

# Epigenetics of metabolism, immunology and aging

**Edited by**

Haitao Wang, Xiaofan Lu and Heng Sun

**Published in**

Frontiers in Genetics

Frontiers in Cell and Developmental Biology



## 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 978-2-83251-538-9  
DOI 10.3389/978-2-83251-538-9

## 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](https://frontiersin.org/about/contact)

# Epigenetics of metabolism, immunology and aging

## Topic editors

Haitao Wang — Center for Cancer Research, National Cancer Institute (NIH), United States

Xiaofan Lu — INSERM U964 Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), France

Heng Sun — University of Macau, China

## Topic coordinator

Hailuan Zeng — Fudan University, China

## Citation

Wang, H., Lu, X., Sun, H., eds. (2023). *Epigenetics of metabolism, immunology and aging*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-538-9

# Table of contents

|     |   |
|-----|---|
| 05  | <b>Editorial: Epigenetics of metabolism, immunology and aging</b><br>Hailuan Zeng, Xiaofan Lu, Heng Sun and Haitao Wang   |
| 07  | <b>Circular RNA as a Potential Biomarker for Forensic Age Prediction</b><br>Junyan Wang, Chunyan Wang, Yangyan Wei, Yanhao Zhao, Can Wang, Chaolong Lu, Jin Feng, Shujin Li and Bin Cong  |
| 17  | <b>Trends in the Contribution of Genetic Susceptibility Loci to Hyperuricemia and Gout and Associated Novel Mechanisms</b><br>Jianan Zhao, Shicheng Guo, Steven J. Schrodi and Dongyi He  |
| 34  | <b>Emerging Roles of Non-proteolytic Ubiquitination in Tumorigenesis</b><br>Xiu Yin, Qingbin Liu, Fen Liu, Xinchun Tian, Tinghao Yan, Jie Han and Shulong Jiang   |
| 64  | <b>Epigenetic Regulation in Knee Osteoarthritis</b><br>Zhengyu Cai, Teng Long, Yaochao Zhao, Ruixin Lin and You Wang  |
| 75  | <b>Glycometabolism-related gene signature of hepatocellular carcinoma predicts prognosis and guides immunotherapy</b><br>Lihua Yu, Xiaoli Liu, Xinhui Wang, Huiwen Yan, Qing Pu, Yuqing Xie, Juan Du and Zhiyun Yang  |
| 91  | <b>Identification of subtypes of clear cell renal cell carcinoma and construction of a prognostic model based on fatty acid metabolism genes</b><br>Shiwen Nie, Youlong Huili, Anliang Yao, Jian Liu, Yong Wang, Lei Wang, Liguang Zhang, Shaosan Kang and Fenghong Cao |
| 103 | <b>Assessment of alterations in histone modification function and guidance for death risk prediction in cervical cancer patients</b><br>Tingting Zhao, Bairong Liu, Mengyuan Zhang, Shiguo Li, Can Zhao and Li Cheng  |
| 120 | <b>Research progress on the relationship between the TOR signaling pathway regulator, epigenetics, and tumor development</b><br>Jiaen Sun, Minglei Yang, Weidi Zhao, Fajiu Wang, Liangwei Yang, Chuntao Tan, Tianjun Hu, Huangkai Zhu and Guofang Zhao                  |
| 129 | <b>Identification of m<sup>6</sup>A-related long non-coding RNAs for predicting prognosis and immune characterizations in gastric cancer</b><br>Xianhui Zhang, Changjing Wang, Zhongxin Liu and Yuan Si   |
| 143 | <b>Effect of metformin on the epigenetic age of peripheral blood in patients with diabetes mellitus</b><br>Man Li, Litao Bao, Ping Zhu and Shuxia Wang  |

- 151 **Epigenome-wide association study of biomarkers of liver function identifies albumin-associated DNA methylation sites among male veterans with HIV**  
Boghuma K. Titanji, Mitch Lee, Zeyuan Wang, Junyu Chen, Qin Hui, Vincent Lo Re III, Kaku So-Armah, Amy C. Justice, Ke Xu, Matthew Freiberg, Marta Gwinn, Vincent C. Marconi and Yan V. Sun
- 164 **Comprehensive analysis of epigenomics and transcriptome data to identify potential target genes associated with obesity**  
Peili Wu, Lei Guo, Xuelin Li, Yuejun Du, Xiaochun Lin, Xiaoqin Ma, Yingbei Lin, Churan Wen, Chuyi Yang, Nannan Liu, Qijian Feng, Yaoming Xue and Meiping Guan
- 176 **CD146 is closely associated with the prognosis and molecular features of osteosarcoma: Guidance for personalized clinical treatment**  
Jingkun Wang, Zhonghan Wu, Meige Zheng, Shuisheng Yu, Xin Zhang and XinZhong Xu
- 191 ***N*<sup>6</sup>-methyladenosine RNA methylation: From regulatory mechanisms to potential clinical applications**  
Peipei Li, Yuntao Wang, Yiwen Sun, Sanjie Jiang and Jingjing Li
- 205 **GOLM1 is related to the inflammatory/immune nature of uveal melanoma and acts as a promising indicator for prognosis and immunotherapy response**  
Xin Liang, Yu Yin and Ning Li
- 219 **Construction of m7G subtype classification on heterogeneity of sepsis**  
Jinru Gong, Jiasheng Yang, Yaowei He, Xiaoxuan Chen, Guangyu Yang and Ruilin Sun



## OPEN ACCESS

EDITED AND REVIEWED BY  
Michael E Symonds,  
University of Nottingham, United Kingdom

## \*CORRESPONDENCE

Hailuan Zeng,  
✉ happyhellen0524@foxmail.com  
Xiaofan Lu,  
✉ xlu.cpu@foxmail.com  
Heng Sun,  
✉ hengsun@um.edu.mo  
Haitao Wang,  
✉ haitao.wang@nih.gov

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 02 January 2023  
ACCEPTED 12 January 2023  
PUBLISHED 18 January 2023

## CITATION

Zeng H, Lu X, Sun H and Wang H (2023),  
Editorial: Epigenetics of metabolism,  
immunology and aging.  
*Front. Genet.* 14:1135889.  
doi: 10.3389/fgene.2023.1135889

## COPYRIGHT

© 2023 Zeng, Lu, Sun and Wang. This is an  
open-access article distributed under the  
terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Editorial: Epigenetics of metabolism, immunology and aging

Hailuan Zeng<sup>1,2\*</sup>, Xiaofan Lu<sup>3,4\*</sup>, Heng Sun<sup>5,6,7\*</sup> and Haitao Wang<sup>8\*</sup>

<sup>1</sup>Department of Endocrinology and Metabolism, Zhongshan Hospital, Fudan University, Shanghai, China, <sup>2</sup>Human Phenome Institute, Fudan University, Shanghai, China, <sup>3</sup>Department of Cancer and Functional Genomics, Institute of Genetics and Molecular and Cellular Biology, CNRS/INSERM/UNISTRA, Illkirch, France, <sup>4</sup>State Key Laboratory of Natural Medicines, Research Center of Biostatistics and Computational Pharmacy, China Pharmaceutical University, Nanjing, China, <sup>5</sup>Cancer Center, Faculty of Health Sciences, University of Macau, Macau, Macau SAR, China, <sup>6</sup>MOE Frontiers Science Center for Precision Oncology, University of Macau, Macau, Macau SAR, China, <sup>7</sup>Zhuhai Research Institute, University of Macau, Zhuhai, China, <sup>8</sup>Thoracic Surgery Branch, Center for Cancer Research, National Cancer Institute (NCI), Bethesda, MD, United States

## KEYWORDS

methylation, histone modification, metabolism, immune, epigenetic age, cancer

## Editorial on the Research Topic Epigenetics of metabolism, immunology and aging

The incapacitation and dysfunction of metabolism and immunity are always coupled with aging, contributing to the development of a series of aging-related diseases, like cancer. Previous studies indicated that many genetic and environmental factors are involved in aging-related diseases. Notably, these diseases are also accompanied by epigenetic changes in DNA methylation, post-translational modifications of histones, RNA-mediated processes, transcription factor binding, and high-order chromatin organization (Booth and Brunet 2016). Besides, increasing evidence has shown that the dysregulation of epigenetic pathways is crucial for coordinating metabolic reprogramming and cellular immunological events in response to extracellular stimuli, leading to and/or fueling the aging-related diseases (Cao and Yan 2020; Sun et al., 2022). Hence, in this Research Topic, we focus on the most recent advances that help understand the roles of epigenetics in aging, metabolism, immunology, and cancer.

The correlation of DNA CpG methylation and aging has been well studied, which serves as the basis for establishment of biological age molecular estimators, such as Horvath's pan-tissue clock and Hannum's clock (Horvath and Raj 2018; Noroozi et al., 2021). These epigenetic clocks might be affected by environmental factors, lifestyles, psychological states, and medical treatments (Noroozi et al., 2021). In the current Research Topic, Li et al. performed a genome-wide methylation study to explore the effect of metformin, a widely used drug for diabetes mellitus, on the epigenetic age of peripheral blood in patients with diabetes mellitus, and they found a strong association between metformin intake and slower epigenetic aging by Horvath's clock and Hannum's clock. In an epigenome-wide association study on biomarkers of liver function in 960 HIV-infected men, Titanji et al. identified nine DNA methylation sites annotating to *TMEM49*, *SOCS3*, *FKBP5*, *ZEB2*, and *SAMD14* genes that were positively associated with serum albumin, and reported that a higher PhenoAge score was significantly associated with a lower level of serum albumin. Moreover, by combining DNA methylation and transcriptome analysis, Wu et al. identified seven key genes that affect the occurrence and development of obesity by influencing the immune microenvironment of adipose tissue. Further, a connective map analysis of drugs suggested that four small-molecule drugs might have therapeutic effects on obesity. The study suggests that epigenetic factors could

mediate the interaction between metabolism and immunity, thus shedding light on identifying novel drugs targeting obesity.

In addition to DNA methylation, other epigenetic mechanisms and technologies have been reported and used in clinical applications. For example, long non-coding RNAs have been shown to regulate gene expressions through recruiting, repelling, or affecting the expression of DNA methyltransferases and TET enzymes (Huang et al., 2022). In this Research Topic, Wang et al. performed a blood circRNA-seq analysis on Chinese participants and identified 5 age-related biomarkers that can be used to develop forensic age prediction models with a mean average error less than 10 years. Zhang et al. developed a risk model based on 11 m6A-related long non-coding RNAs for predicting gastric cancer prognosis and sensitivity to immune checkpoint inhibitors. Gong et al. identified eight N7-Methylguanosine (m7G)-related prognostic genes that can be used to categorize sepsis patients into two molecular subtypes with different metabolic activities and immune status. Wang et al. showed that the expression of CD146, a gene correlated with m5C RNA methylation modification, is linked to poor prognosis in osteosarcoma and can be used to predict response to immunotherapy. Zhao et al. evaluated the activation status of the histone modification pathway in cervical cancer and develop a prognostic model to predict clinical outcomes. These studies may provide valuable insights and potential applications for personalized treatment strategies for heterogeneous diseases.

This Research Topic also includes reviews on different aspects of epigenetic regulations in signaling pathways and pathological processes. For example, Li et al. reviewed the regulatory mechanisms of N6-methyladenosine RNA methylation, the techniques for detecting N6-methyladenosine methylation, the role of N6-methyladenosine modification in cancer and other diseases, and the potential clinical applications. Yin et al. introduced the role of non-proteolytic ubiquitination in tumorigenesis and related signaling pathways. Sun et al. provided an overview of the role of rapamycin (TOR) signaling pathway regulator (TIPRL) in cancer development, as well as the TIPRL/protein phosphatase 2A (PP2A) axis and its epigenetic regulation. In addition, Cai et al. summarized the interactions of genetics, environmental factors, and epigenetics on knee osteoarthritis, with a focus on DNA methylation, histone modification, and non-coding RNA.

Epigenetic changes associated with the onset and progression of aging-related diseases could serve as biomarkers for the diagnosis and prognosis of these diseases. In this Research Topic, several studies have investigated the role of fatty acid metabolism genes (Nie et al.), glycometabolism (Yu et al.), and genetic variants (Zhao et al.) on the prognosis and response to immunotherapy in various types of diseases, including clear cell renal cell carcinoma (ccRCC), hepatocellular carcinoma (HCC), and uveal melanoma (Liang et al.). These studies have shown promising results in using multidisciplinary data analysis to

better understand the relationship between epigenetics and diseases, and have the potential to lead to the development of more effective treatments for these conditions. In particular, the findings on the epigenetic regulation of metabolism, immunology, and cancer may improve our understanding of the underlying causes of these diseases and potentially improve health and longevity in individuals.

## Author contributions

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

## Funding

This work was supported by National Natural Science Foundation of China (No. 81602587; No. 81973145; No. 82273735), Macao Science and Technology Development Fund (FDCT) grants (0006/2021/AGJ and 0065/2021/A), University of Macau, Multi-Year Research Grant – General Research Grant (MYRG2022-00175-FHS), Key R&D Program of Jiangsu Province (Social Development) (BE2020694), and Shanghai Municipal Science and Technology Major Project (Grant No. 2017SHZDZX01).

## Acknowledgments

The editors thank all authors for their contributions to this Research Topic, and we appreciate all reviewers for their time and constructive comments on submitted manuscripts.

## 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.

## References

- Booth, L. N., and Brunet, A. (2016). The aging epigenome. *Mol. Cell* 62 (5), 728–744. doi:10.1016/j.molcel.2016.05.013
- Cao, J., and Yan, Q. (2020). Cancer epigenetics, tumor immunity, and immunotherapy. *Trends Cancer* 6 (7), 580–592. doi:10.1016/j.trecan.2020.02.003
- Horvath, S., and Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat. Rev. Genet.* 19 (6), 371–384. doi:10.1038/s41576-018-0004-3
- Huang, W., Li, H., Yu, Q., Xiao, W., and Wang, D. O. (2022). LncRNA-mediated DNA methylation: An emerging mechanism in cancer and beyond. *J. Exp. Clin. Cancer Res.* 41 (1), 100. doi:10.1186/s13046-022-02319-z
- Noroozi, R., Ghafouri-Fard, S., Pisarek, A., Rudnicka, J., Spólnicka, M., Branicki, W., et al. (2021). DNA methylation-based age clocks: From age prediction to age reversion. *Ageing Res. Rev.* 68, 101314. doi:10.1016/j.arr.2021.101314
- Sun, L., Zhang, H., and Gao, P. (2022). Metabolic reprogramming and epigenetic modifications on the path to cancer. *Protein Cell* 13 (12), 877–919. doi:10.1007/s13238-021-00846-7



# Circular RNA as a Potential Biomarker for Forensic Age Prediction

Junyan Wang<sup>1</sup>, Chunyan Wang<sup>2</sup>, Yangyan Wei<sup>1</sup>, Yanhao Zhao<sup>1</sup>, Can Wang<sup>1</sup>, Chaolong Lu<sup>1</sup>, Jin Feng<sup>2</sup>, Shujin Li<sup>1\*</sup> and Bin Cong<sup>1\*</sup>

<sup>1</sup>Hebei Key Laboratory of Forensic Medicine, Collaborative Innovation Center of Forensic Medical Molecular Identification, Research Unit of Digestive Tract Microecosystem Pharmacology and Toxicology, College of Forensic Medicine, Chinese Academy of Medical Sciences, Hebei Medical University, Shijiazhuang, China, <sup>2</sup>Physical Examination Center of Shijiazhuang First Hospital, Shijiazhuang, China

## OPEN ACCESS

### Edited by:

Lidia Larizza,  
Italian Auxological Institute (IRCCS),  
Italy

### Reviewed by:

Jianhui Xie,  
Fudan University, China  
Hermona Soreq,  
Hebrew University of Jerusalem, Israel

### \*Correspondence:

Shujin Li  
shujinli@163.com  
shujinli@hebmh.edu.cn  
Bin Cong  
hbydcngbin@126.com

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

**Received:** 30 November 2021

**Accepted:** 04 January 2022

**Published:** 07 February 2022

### Citation:

Wang J, Wang C, Wei Y, Zhao Y,  
Wang C, Lu C, Feng J, Li S and Cong B  
(2022) Circular RNA as a Potential  
Biomarker for Forensic Age Prediction.  
Front. Genet. 13:825443.  
doi: 10.3389/fgene.2022.825443

In forensic science, accurate estimation of the age of a victim or suspect can facilitate the investigators to narrow a search and aid in solving a crime. Aging is a complex process associated with various molecular regulations on DNA or RNA levels. Recent studies have shown that circular RNAs (circRNAs) upregulate globally during aging in multiple organisms such as mice and *C.elegans* because of their ability to resist degradation by exonucleases. In the current study, we attempted to investigate circRNAs' potential capability of age prediction. Here, we identified more than 40,000 circRNAs in the blood of thirteen Chinese unrelated healthy individuals with ages of 20–62 years according to their circRNA-seq profiles. Three methods were applied to select age-related circRNA candidates including the false discovery rate, lasso regression, and support vector machine. The analysis uncovered a strong bias for circRNA upregulation during aging in human blood. A total of 28 circRNAs were chosen for further validation in 30 healthy unrelated subjects by RT-qPCR, and finally, 5 age-related circRNAs were chosen for final age prediction models using 100 samples of 19–73 years old. Several different algorithms including multivariate linear regression (MLR), regression tree, bagging regression, random forest regression (RFR), and support vector regression (SVR) were compared based on root mean square error (RMSE) and mean average error (MAE) values. Among five modeling methods, regression tree and RFR performed better than the others with MAE values of 8.767 years (S.rho = 0.6983) and 9.126 years (S.rho = 0.660), respectively. Sex effect analysis showed age prediction models significantly yielded smaller prediction MAE values for males than females (MAE = 6.133 years for males, while 10.923 years for females in the regression tree model). In the current study, we first used circRNAs as additional novel age-related biomarkers for developing forensic age estimation models. We propose that the use of circRNAs to obtain additional clues for forensic investigations and serve as aging indicators for age prediction would become a promising field of interest.

**Keywords:** forensic genetics, biomarkers, circular RNA, age prediction, machine learning

## INTRODUCTION

Age prediction of an unknown individual can facilitate case investigations and disaster victim identification. Estimating the age of known persons with an unclear age can provide important clues in legal affairs (Schmeling et al., 2016). The age of individuals can be determined by techniques that rely on morphological measures of teeth and skeletal remains (Ichikawa et al., 2010; Marquez-Ruiz et al., 2018). But this approach is restricted to samples with a nearly complete skeleton and is influenced by subjective factors (Schmeling et al., 2007).

In most cases, only fragmentary remains are left by the perpetrator after committing a crime. Forensic molecular methods have allowed researchers to obtain genetic information of a person from biological evidence retrieved from crime scenes (Park et al., 2016). In recent years, several molecular methods were proposed, such as telomere shortening (Mensa et al., 2019), mitochondrial DNA deletion (Zapico and Ubelaker, 2016; Yang et al., 2020), signal-joint T-cell receptor excision circle (sjTRECs) (Yamanoi et al., 2018), and DNA methylation (DNAm) (Meng et al., 2019; Dias et al., 2020; Freire-Aradas et al., 2020). Among these biomarkers, the DNAm pattern was considered as the most promising age-predictive biomarkers for forensic and clinical use due to its high prediction accuracy. In particular, Zbiec-Piekarska et al. (Zbiec-Piekarska et al., 2015a) conducted a research based on pyrosequencing data for two CpG sites in the *ELOVL2* gene and achieved relatively high prediction accuracy with a mean absolute deviation (MAD) from the chronological age of 5.03 years. Another model based on 5 CpG sites reported by Zbiec-Piekarska showed high prediction accuracy with a MAD of 3.9 years. Hence, several sites of the gene *ELOVL2* is believed to be the most hopeful locus for age prediction (Zbiec-Piekarska et al., 2015b). DNAm-based age prediction methods can be developed in a way with a relatively high accuracy for individual age estimation. However, it cannot be denied the DNAm method suffers from several challenges: The DNAm change pattern can be greatly influenced by several factors such as smoking, nutrition, and diverse diseases giving rise to the inaccuracy of quantification results, especially when the age increase, the bias bigger (Cho et al., 2017). Furthermore, the problem of a critical level of degradation in the amount of full-length DNA after conventional bisulfite treatment is yet to be addressed (Meng et al., 2019). In light of various restrictions and challenges mentioned above, looking for other appropriate biomarkers in human blood is of great significance for forensic age estimation.

With the advent of next-generation RNA sequencing (RNA-seq) and bioinformatics approaches, circular RNAs (circRNAs) have emerged as an interesting RNA species (Salzman et al., 2012; Jeck et al., 2013). CircRNA is a class of newly discovered noncoding RNA (ncRNA), presenting as a special covalent loop without a 5' cap or 3' tail (Li et al., 2018b). Their unique feature enhances their ability to resist the degradation of exonucleases and thus contributes to their stability compared to their mRNA counterparts (Enuka et al., 2016; Zhang et al., 2016). Recent

studies have reported that circRNAs have a number of binding sites for the microRNA (miRNA) family acting as the miRNA sponge and interact with RNA-binding proteins (RBPs). They play a role in aging and age-related diseases like Alzheimer's disease, through interaction with miRNAs and RBPs (Cai et al., 2019). Apart from the properties of resistance to RNase R digestion, miRNA, and RBP sponges, circRNA also exhibits other biological characteristics such as widespread expression, cell-specificity, tissue-specificity, and developmental stage-specific expression patterns (Abu and Jamal, 2016). Its feature of developmental stage-specific expression has been confirmed across various model organisms such as mice (Gruner et al., 2016), flies (Westholm et al., 2014), and elegans (Cheng et al., 2016). Accumulating evidence indicates the involvement of circRNAs in aging, indicating a potential role as biomarkers of the chronological age. Hall et al. (Hall et al., 2017) identified 38 circRNAs that were differentially expressed between day 10 and 40 from RNA-seq profiles of *Drosophila* photoreceptor neurons. After detecting the global profiles of circRNAs in *C. elegans* from the fourth larval stage (L4) through 10-day-old adults, Cortes-Lopez et al. (Cortes-Lopez et al., 2018) found a massive accumulation of circRNA expression levels during aging. Yu et al. and others identified senescence-associated circRNAs by the whole-transcriptome sequencing and discovered circCCNB1 sponges miRNA miR-449a to inhibit cellular senescence by targeting CCNE2 (Yu et al., 2019). Although the fact that circRNAs expressions are observed to accumulate during aging in various organisms and human senescent cells (Haque et al., 2020), research focusing on circRNAs for age prediction has not been carried out yet in the field of forensic. In our preliminary work, we screened a total of 23 age-related circRNAs using univariate linear regression analysis (Wang et al., 2019). In the current study, we introduced other machine learning algorithms, support vector machine (SVM), and lasso regression to conduct in-depth data mining on circRNA sequencing data and screen out more age-related circRNA candidates. Selected age-related circRNAs were validated in a larger sample size and investigated for modeling for age prediction.

In this study, we analyzed circRNA-seq profiles from 13 blood samples of unrelated Chinese aged between 20 and 62 years, focusing on the potential links between the chronological age and the expression of circular RNAs in human blood. The present study aims to build a novel age prediction model based on a subset of age-related circRNAs for forensic application. Machine learning has gained a place in medicine and captured the interest of medical researchers (Cabitza and Banfi, 2018). We used several machine-learning methods to select age-related circRNAs and build age prediction models. Five different models including multivariate linear regression (MLR), regression tree, bagging regression, random forest regression (RFR), and support vector regression (SVR) were established based on five age-associated circRNAs using 100 samples. To the best of our knowledge, this is the first study that uses circRNAs in the blood as indicators together with machine learning methods to develop prediction models for forensic individual age estimation.

## MATERIALS AND METHODS

### Sample Collection and RNA Isolation

Whole blood samples were collected from 13 healthy volunteers (aged between 20 and 62 years, including 6 females and 7 males) for circRNA sequencing and 30 subjects (aged between 19 and 72 years, including 16 females and 14 males) for RT-qPCR experiment validation (**Supplementary Table S2**). For age prediction model construction, 100 blood samples from 19–73-year olds including 52 females and 48 males were collected. Written informed consent was obtained from all the volunteers, and our study was approved by the Medical Ethics Committee of Hebei Medical University (No. 20190013).

Peripheral whole blood 10 ml were drawn from subjects by venipuncture and collected in an EDTA-containing vacutainer. Total RNA was isolated immediately using TRIzol reagent (Thermo Scientific, United States) according to the manufacturer's instructions after blood collection. RNA quantification was conducted on the NanoDrop 1000 (NanoDrop Technologies). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, United States). Isolated RNAs were preserved on the condition of  $-80^{\circ}\text{C}$  until reverse transcription.

### High-Throughput Sequencing (circRNA-Seq)

A total amount of 5  $\mu\text{g}$  RNA per sample was used as the input material for RNA sample preparation. First, ribosomal RNA was removed by using the Epicentre Ribozero<sup>TM</sup> rRNA Removal Kit (Epicentre, United States), and the rRNA free residue was cleaned up by ethanol precipitation. Subsequently, the linear RNA was digested with 3U of RNase R (Epicentre, United States) per  $\mu\text{g}$  of RNA. The sequencing libraries were generated by NEBNext<sup>®</sup> Ultra<sup>TM</sup> Directional RNA Library recommendations. Barcoded libraries were sequenced at Novogene Co., LTD. (Beijing, China) using the Illumina HiSeq 4000 platform to obtain paired-end 150 nt reads. CircRNA candidates were detected and identified using *find\_circ* (Memczak et al., 2013) and CIRI2 algorithms (Gao et al., 2015). The normalization of contig counts was performed by calculating transcripts per million (TPMs). The normalized expression level = (read counts \* 1,000,000)/libsize (libsize is the sum of circRNA read counts).

### Reverse Transcription and Quantitative PCR (RT-qPCR)

RNA was reverse transcribed into cDNA using PrimeScript Reverse Transcriptase<sup>TM</sup> (Takara Bio Inc., Otsu, Shiga, Japan) according to the manufacturer's protocol. cDNA was obtained using the RT Primer Mix including the oligo dT primer and random 6 mers and was stored at  $-80^{\circ}\text{C}$  waiting for further RT-qPCR tests. qPCR reactions were performed using the QuantiNova<sup>TM</sup> SYBR<sup>®</sup> Green PCR Kit (Qiagen, Germany) on a 7500 System (Applied Biosystems). Circular RNAs are circular in shape and covalently closed. The primer design is of vital importance for PCR quantitation. The use of divergent rather

than convergent primers can selectively detect and quantitate these special RNA molecules (Li et al., 2015). Sequences of primers are available in **Supplementary Table S3**. 18S rRNA was chosen as reference genes which were stably expressed in blood samples in the pre-test. The delta Ct value ( $\text{Ct}_{\text{target}} - \text{Ct}_{\text{reference}}$ ) ( $\Delta\text{Ct}$ ) represented the circRNA expression. These tests were performed using technical duplicates. PCR products were also electrophoresed on agarose gel and recovered to verify the specificity of primers using Sanger sequencing.

### Selection of Potential Age-Associated circRNAs

An enormous amount of data produced by the next generation sequencing makes the pre-selection process very important and challenging. In an attempt to limit the quantity of predictors which may save both time and cost, three statistical methods were conducted to select age-associated markers. According to circRNA-seq data, an adjusted *p*-value lower than 0.01 (the false discovery rate method) after the Spearman's correlation analysis was regarded as a criterion to identify the age-associated circRNAs using IBM SPSS statistics software for Windows, version 21.0. Lasso regression was implemented for feature selection with an alpha value set in a range from 0.0005 to 1 within Python. Another feature selection approach SVM was performed in Python using the 'LinearSVC' command with the 'C' value set to 0.0202. The Spearman correlation coefficients (Srho) were then calculated again for 30 validation samples according to qPCR results to validate age-dependent circRNAs. CircRNA host gene analyses were performed on *string-db.org* and the gene database of [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov).

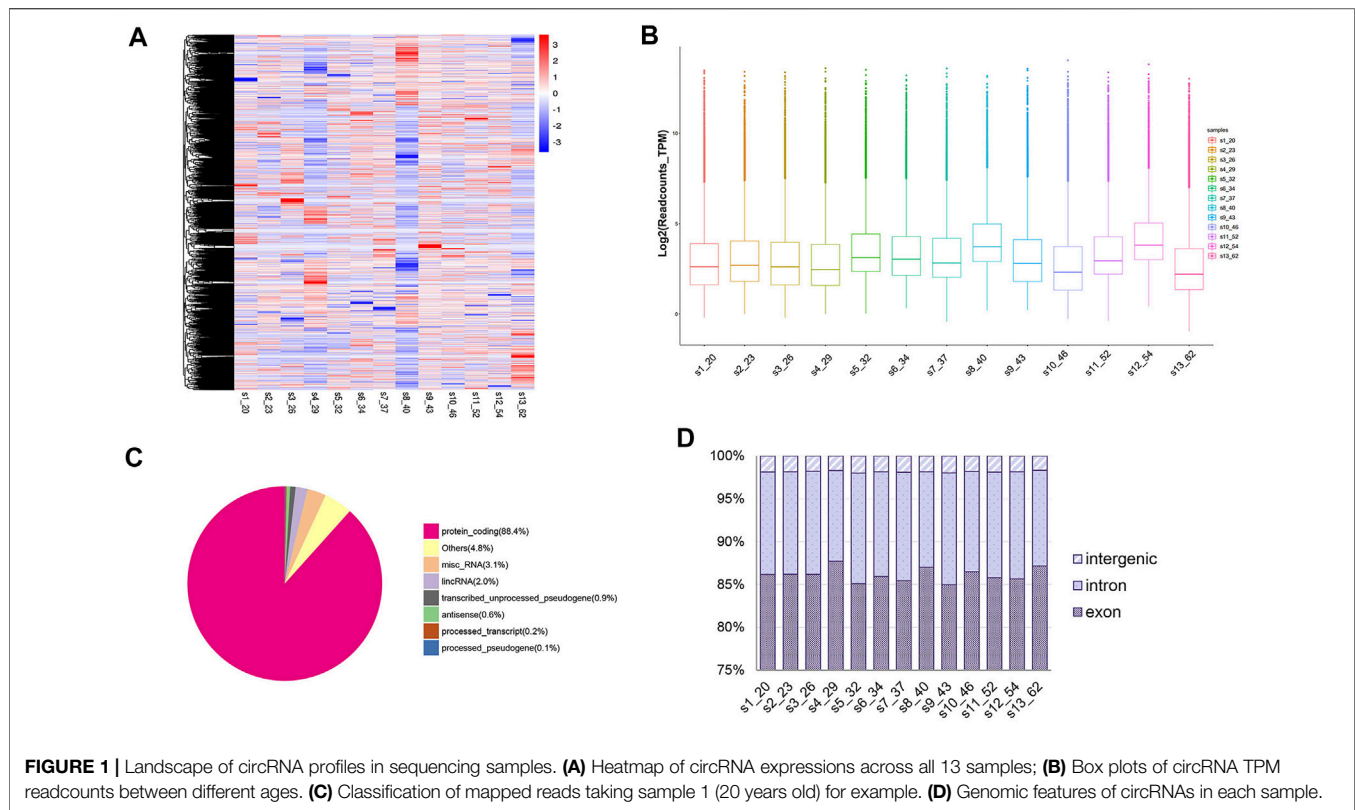
### Age Prediction

The R project for statistical computing software version 4.0.0 was employed to conduct five algorithms: Multivariate Linear Regression (*lm*), regression tree (*rpart*), bagging regression (*ipred*), random forest regression (*randomForest*), and support vector regression (*rminer*). The RMSE and MAE values between chronological and predicted ages were used as performance metrics for the final age prediction models using the five above-mentioned algorithms. Two important parameters *mtry* referring to the number of features to be considered at each split and *ntree* standing for the number of trees in a forest were set to 5 and 300, respectively. VIM is one of the methods in capturing the patterns of dependency between variables and response in the form of a single number. It can essentially be used in many prediction methods but is particularly effective for black-box methods which (in contrast to, say, generalized linear models) yield less interpretable results (Shen et al., 2019). We ranked VIM according to the accuracy, MSE, and standard errors, separately.

## RESULTS

### Selection of circRNAs Differentially Expressed During Aging

A total of 45,697 circRNAs were identified by circRNA high-throughput sequencing [56,330 circRNAs were identified in our



published data that circRNA identification was performed using the find\_circ algorithm alone (Wang et al., 2019)] (**Figure 1A**). Our results demonstrated a high abundance of circRNA in human peripheral whole blood with an average of 26,719 read counts for each individual. To further illustrate distribution of circRNA levels between different ages, we plotted circRNA  $\log_2$  TPM readcounts for all samples (**Figure 1B**). Additionally, circRNAs were distributed across various genomic regions but most commonly from protein coding regions, where over 80% of circRNAs originated (**Figure 1C**). We also observed exons and introns accounting for a higher proportion (more than 95% in each sample) than intergenic regions (**Figure 1D**).

To identify age-correlated circRNA candidates, we introduced three methods of feature dimension reduction, including false discovery rate (FDR), lasso regression (LASSO), and support vector machine (SVM). The Spearman's correlation coefficient ( $S.rho$ ) was calculated to identify the correlation between the age and circRNA expression level (TPM value) for each single circRNA. For FDR-adjusted method after Spearman correlation analysis, 14 circRNAs were considered as age-related markers. Eight circRNAs were selected using the lasso regression method. The SVM method selected 10 circRNAs from more than 40,000 circRNAs (see Methods & Materials). A subset of these 28 circRNAs was chosen for age-associated circRNA candidates for further RT-qPCR validation in a cohort of 30 samples (**Supplementary Table S1**). There were four overlapped circRNAs screened both by LASSO and SVM methods, including hsa\_circ\_0002454, hsa\_circ\_0006117, hsa\_circ\_0014606, and hsa\_circ\_0032800.

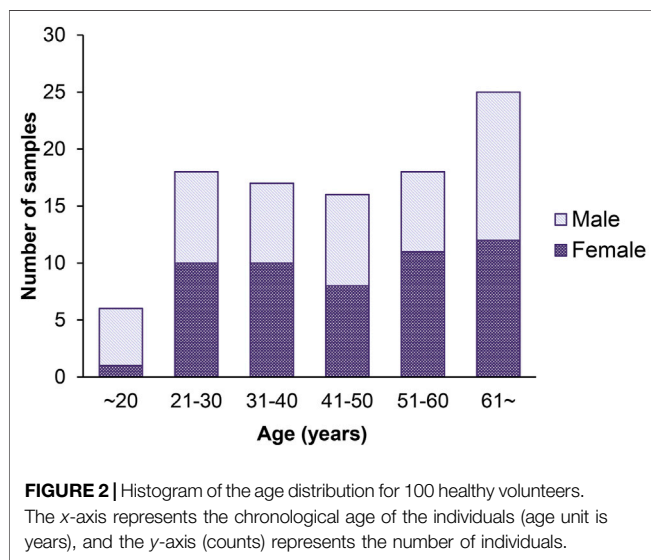
Information of 30 validation samples is listed in **Supplementary Table S2**. Spearman correlation analyses showed five circRNAs selected by RNA-seq were in good agreement with RT-qPCR quantifications. All five age-related circRNAs showing a statistically different change by qPCR were upregulated during aging revealing an overwhelming bias for the upregulation of circRNAs during aging. Note that the lower the  $\Delta Ct$ -value, the higher is the circRNA expression. Among them, circRNA hsa\_circ\_0086306 showed higher Spearman correlation coefficients ( $S.rho = -0.456$ ), followed by hsa\_circ\_0015789 ( $S.rho = -0.453$ ). Derived genes of the 5 age-related circRNAs encode several proteins or involve in multiple pathways. The feature description of 5 age-associated circRNAs is listed in **Table 1**. Finally, 5 age-correlated circRNAs were selected for further age estimation modeling using an independent cohort of 100 samples.

## Development of Age Prediction Models Using Different Algorithms

We detected 5 age-related circRNAs in 100 blood samples aged between 19 and 73 years old using the RT-qPCR strategy (**Figure 2**). The scatter plots of 5 circRNAs in 100 samples are displayed in **Figure 3**. To provide an unbiased estimate of a predictive accuracy for age, various attempts had been made to develop better models to predict the age. We adopted five different algorithms, including Multivariate Linear Regression (MLR), bagging regression, regression tree, random forest regression (RFR), and support vector regression (SVR). Age prediction models were fit based on 5 age-related circRNAs. The whole blood dataset was randomly split

**TABLE 1** | Final circRNA markers for the age prediction models. Description of their features, originated gene name, and tendency during aging.

| circRNA ID       | Gene    | Full gene name                                    | Chr. | Length (nt) | Trend |
|------------------|---------|---|------|-------------|-------|
| hsa_circ_0015789 | DENND1B | DENN domain containing 1B                         | 1    | 477         | Up    |
| hsa_circ_0086306 | UHRF2   | Ubiquitin like with PHD and ring finger domains 2 | 9    | 297         | up    |
| hsa_circ_0002454 | DNAJC6  | DnaJ heat shock protein family (Hsp40) member C6  | 1    | 350         | up    |
| hsa_circ_0000524 | RBM23   | RNA-binding motif protein 23                      | 14   | 189         | up    |
| hsa_circ_0004689 | SWT1    | SWT1 RNA endoribonuclease homolog                 | 1    | 469         | up    |



into the training subset (80%,  $n = 80$ ) and the testing subset (20%,  $n = 20$ ). Results uncovered that the regression tree and the RFR models showed the highest and similar accuracy with MAE values of 8.767 years ( $S.rho = 0.6983$ ) and 9.126 years ( $S.rho = 0.660$ ) in the testing subset, while the SVR model reached the highest accuracy in the training subset (MAE = 8.367,  $S.rho = 0.7423$ ) but performed poorly in the testing subset (MAE = 10.175,  $S.rho = 0.4511$ ). Other models showed medium MAE values in the testing subset: 10.175 years for bagging and 12.190 years for the MLR (Table 2). The relationship between the chronological and the predicted age using different algorithms is displayed in Figure 4. The RFR model fitted 30% (6/20) of individuals within a  $\pm 5$  year error range, while 50% (10/20) within a  $\pm 10$  year error range. Additionally, the variable importance measure (VIM) ranking the variables (i.e., the features) with respect to their relevance for prediction is a byproduct of random forest (Shen et al., 2019). As an illustration, the ranks of variable

importance measures according to accuracy, MSE, and standard errors are displayed in Supplementary Figure S1 separately.

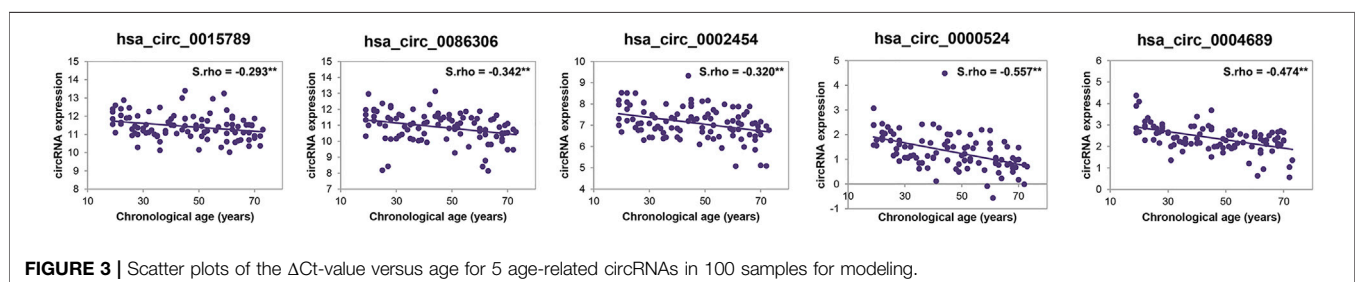
## Age and Sex Analyses

In order to further evaluate the predictive error, samples were divided into five age groups: 20s or less, 30, 40, 50, and 60s or more. In line with previous studies, the MAE increased with age (Naue et al., 2017). In our study, older individuals (more than 60 years old) presented an increased deviation between the chronological and predicted age compared to younger groups (less than 40 years old), and the predicted age of elders was prone to be underestimated. Specifically, the tree model, for example, the MAE was 7.985 years for the youngest age group (20s or less), 7.656 years for 30s, 7.335 years for 40s, 8.604 years for 50s, and 12.493 years for the oldest (60s or more).

The observation that differences between sexes exist was found by the previous study (Fehlmann et al., 2020). Whether sex effects did have an impact on the accuracy of age prediction models in our study, we conducted an analysis on female and male samples for model fitting separately. Similarly, all five models in the current study significantly yielded smaller prediction MAE values for males than females in the training subset, and 3 out of 5 models in the testing subset are as shown in Table 3. For instance, the regression tree model reached an MAE of 6.133 years ( $S.rho = 0.882$ ) in males but 10.923 years in females ( $S.rho = 0.562$ ). Results indicated age prediction models for female samples were less accurate than those for male samples.

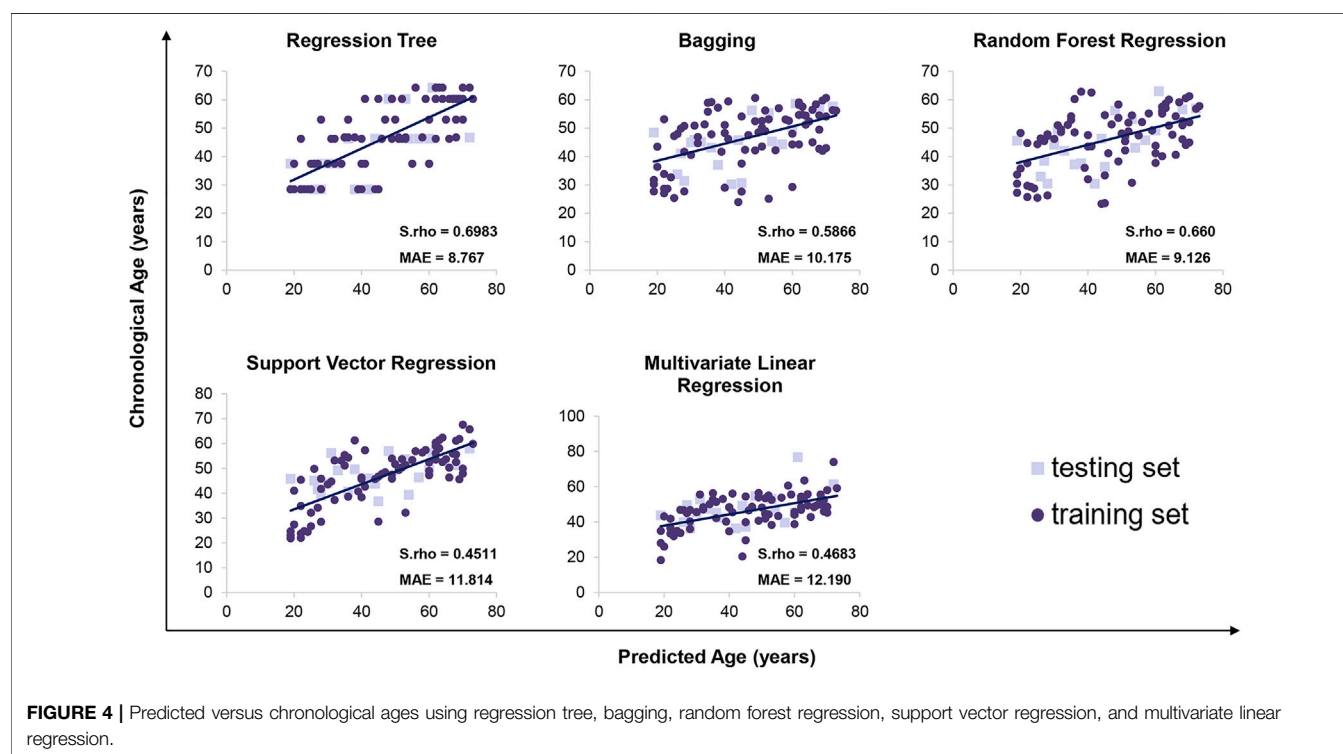
## DISCUSSION

The development of molecular methods for age prediction is valuable when the human specimens such as bloodstains are retrieved from crime scenes without morphological age features.



**TABLE 2** | Comparison of five prediction models.

| Models  | Training set ( $n = 80$ ) |              |        | Testing set ( $n = 20$ ) |              |        |
|---------|---------------------------|--------------|--------|--------------------------|--------------|--------|
|         | MAE (years)               | RMSE (years) | S.rho  | MAE (years)              | RMSE (years) | S.rho  |
| Tree    | 9.343                     | 11.162       | 0.7405 | 8.767                    | 10.584       | 0.6983 |
| Bagging | 12.311                    | 14.650       | 0.4888 | 10.175                   | 12.04        | 0.5866 |
| RFR     | 12.442                    | 14.543       | 0.4975 | 9.126                    | 11.168       | 0.660  |
| SVR     | 8.367                     | 11.187       | 0.7423 | 11.814                   | 13.716       | 0.4511 |
| MLR     | 11.925                    | 13.690       | 0.5662 | 12.190                   | 13.825       | 0.4683 |



Considerable progress of the molecular biology of aging has been made in the last few decades. Using a variety of age-associated biomarkers in blood has emerged as a useful method for age estimation. Nevertheless, estimating age accurately and reliably from molecular biomarkers is still a challenge because identifying minimal numbers of biomarkers that provide maximal age information is extremely tough when analyzing forensic samples as the nature of forensic DNA or RNA being of lower quality and quantity (Zubakov et al., 2016). New insights into discovering novel biomarkers with the property of strong correlation of aging and stability need more attention within the forensic genetic community.

In this study, we provided potential age-associated markers in human blood, namely, circRNA due to its property of highly stable and developmental stage-specific expression patterns. Recently, some analyses conducted by forensic researchers have focused on the capacity of circRNAs to identify body fluids due to their tissue-specificity (Dong et al., 2016). We also noticed that circRNAs have been studied in aging and age-related diseases across different species, revealing a global accumulation of the circRNA expression level during aging. Additionally, the unique feature of circRNAs enhances

their stability and thus might be more suitable for forensic-degraded samples seen commonly. To substantiate this assumption, we identified age-dependent circRNA markers from the human peripheral whole blood by means of circRNA next-generation sequencing and RT-qPCR and developed age-predictive models using different machine learning algorithms. To the best of our knowledge, this is the first one to predict an individual's age using circRNA biomarkers. Our study is non-trivial; we opened a new avenue for forensic age prediction analysis using a novel biomarker in human blood, which would be an important topic for future investigation.

In recent years, sequencing technological advances in monitoring the gene expression have given rise to a dramatic increase in gene expression data. Selecting age-associated circRNAs from a giant gene dataset of RNA-seq profiles (approximately over 40,000 circRNAs were identified) made the current study difficult. Machine learning techniques are known to be excellent at discarding and compacting redundant information from gene expression data. The introduction of machine learning algorithms helped us to select variables from 45,697 down to 28. Five out of 28 age-dependent circRNAs finally

**TABLE 3 |** Age prediction performance of different models between female and male groups.

| Models  | Training set (n = 80) |              |       |               |               |              | Testing set (n = 20) |              |       |               |               |              |
|---------|-----------------------|--------------|-------|---------------|---------------|--------------|----------------------|--------------|-------|---------------|---------------|--------------|
|         | Female (n = 41)       |              |       | Male (n = 39) |               |              | Female (n = 11)      |              |       | Male (n = 9)  |               |              |
|         | MAE (years)           | RMSE (years) | S.rho | MAE (years)   | RMSE (years)  | S.rho        | MAE (years)          | RMSE (years) | S.rho | MAE (years)   | RMSE (years)  | S.rho        |
| Tree    | 10.284                | 12.226       | 0.654 | <b>8.352</b>  | <b>9.975</b>  | <b>0.809</b> | 10.923               | 12.634       | 0.562 | <b>6.133</b>  | <b>7.338</b>  | <b>0.882</b> |
| Bagging | 13.355                | 15.569       | 0.368 | <b>11.213</b> | <b>13.617</b> | <b>0.609</b> | 10.192               | 12.742       | 0.526 | <b>10.154</b> | <b>11.122</b> | <b>0.660</b> |
| RFR     | 13.692                | 15.559       | 0.373 | <b>11.129</b> | <b>13.392</b> | <b>0.624</b> | 9.531                | 11.919       | 0.599 | <b>8.630</b>  | <b>10.175</b> | <b>0.732</b> |
| SVR     | 9.168                 | 11.947       | 0.673 | <b>7.524</b>  | <b>10.327</b> | <b>0.807</b> | 11.380               | 13.649       | 0.497 | 12.345        | 13.798        | 0.409        |
| MLR     | 12.385                | 13.963       | 0.502 | <b>11.441</b> | <b>13.398</b> | <b>0.619</b> | 10.804               | 13.188       | 0.51  | 13.883        | 14.566        | 0.461        |

The bold values means models with a lower MAE value.

selected all showed positive correlations with the chronological age. This expression trend was in accordance with the previous research in other organisms that most circRNA expression levels dramatically increase during aging (Mahmoudi and Cairns, 2019). Though there has been no existing evidence yet indicating that circRNA expression levels of human peripheral whole blood are changing with increasing age, recent reports showed that circRNA expression levels change in the elders' brains with aging and play a vital role in the development of neurodegenerative diseases such as Alzheimer's disease (AD) and Parkinson's disease (PD) (Constantin, 2018; Zhang et al., 2018; Hanan et al., 2020). Furthermore, Shahnaz Haque et al. (Haque et al., 2020) assessed the circRNA expression in aging human blood and found that *circFOXO3* and *circEP300* demonstrated differential expression in one or more human senescent cell types, providing indirect evidence for circRNA as a promising indicator for age estimation.

Selected circRNA candidates in our study were derived from five different genes (*DENND1B*, *UHRF2*, *DNAJC6*, *RBM23*, and *SWT1*) and might involve in several important biological processes. We then undertook a comprehensive analysis. As expected, these five genes code for important proteins or involve in biological metabolic processes. The gene *DENND1B* encodes the guanine nucleotide exchange factor (GEF) for *RAB35* that acts as a regulator of T-cell receptor (TCR) internalization in TH2 cells which functions as a regulator in the process of childhood asthma (Fauna et al., 2017). The gene *UHRF2* encodes a protein that is an intermolecular hub protein in the cell-cycle network (Mori et al., 2002). Through cooperative DNA and histone binding, it may contribute to tighter epigenetic control of the gene expression in differentiated cells. Recent research found that *UHRF2* can promote the DNA damage response (Wang et al., 2018). The gene *DNAJC6* belongs to the evolutionarily conserved *DNAJ/HSP40* family of proteins, which regulate the molecular chaperone activity by stimulating the ATPase activity. This gene plays a role in clathrin-mediated endocytosis in neurons (Yim et al., 2010). The gene *RBM23* encodes probable RNA-binding protein and may be involved in the pre-mRNA splicing process (Dowhan et al., 2005). *SWT1* is a protein coding gene that involves in diseases including kidney sarcoma and Wilms tumor 1. Notably, the function of circRNAs or their host genes can provide some clues for future selection of biomarker candidates. Some circRNAs were confirmed to regulate biogenesis, and some senescence-associated genes might become a clue for the selection of age-dependent

indicators, for instance, gene *Foxo3* circular RNA was highly expressed in heart samples of aged patients and mice reported by William W. Du et al. (Du et al., 2017). *Foxo3* was found with different expressions in one or more human senescent cell types in the study of Shahnaz Haque et al. (Haque et al., 2020). In contrast, some of the aforementioned circRNAs play a key role in the pathogenesis of tumors or diseases, such as *DENND1B* and *SWT1*. The effect of the disease state on age predictions should be considered for further investigation because it is important to build a robust predictive model using biomarkers that would not present differential expression due to the disease status. As models described in the study by Athina Vidaki et al. (Vidaki et al., 2017), they analyzed the effect of different diseases on age prediction that schizophrenia presented the lowest age prediction error, while anemia demonstrated the lower relation with age, indicating that it becomes evident that the error is much higher for blood-related diseases when analyzing separately, samples suffering from blood- vs. non-blood-related diseases.

Among the multiple machine learning algorithms adopted for the present study, the regression tree and RFR models showed the highest and similar accuracy in the testing subset. The SVR model reached the highest accuracy in the training subset but had a lower accuracy in the testing subset, which might be due to overfitting and loss of generalizability. These estimation errors were higher than previous predictive models; therefore, using other biomarkers (Cho et al., 2017; Dias et al., 2020) might be due to the small sample size. Small sample sizes might lead to biased screening of age-related biomarkers, which would explain why we screened age-associated circRNAs with only four overlapped using different methods. The limited sample size might restrict our models' validity to a certain extent. The possibility that these age-related biomarkers might not be able to apply in other groups of different ethnic origins is due to the very small cohorts employed for the initial RNA-sequencing analysis, which limits the value of this study to the Chinese population alone. This also restricts the potential impact of childbearing on women's data as circRNA profiles in human granulosa cells were reported age-related and potentially reflecting decreased oocyte competence during maternal aging (Cheng et al., 2017). Besides, although machine learning algorithms seem to achieve a relatively high accuracy using a subset of weak predictors, the prediction accuracy of models presented here is needed to be further improved to satisfy forensic practice, which means more robust, stable, and higher

age-related circRNA biomarkers need to be found in our future work. Our models performed better in young subjects but poorly for elders over 60 years old. We suppose that the deviation might attribute to the fact that adults and elders suffer from more confounding factors in the aspect of medication, smoking, or alcohol, which are not easily accessible for children and adolescents, as previously reported (Li et al., 2018a). We are aware of the other one weakness in the current study that our sample set did not incorporate children and adolescents (a population sample ranging from 19 to 73 years old was used). Therefore, future studies are needed to explore age-specific biomarkers, namely, under-aged-specific and adult-specific predictive indicators. Focusing on when these biomarkers occur and when they lose functions or their periodical changes during an individual's life span also need to be investigated precisely in further research. Additionally, aging can be also affected by the sex effect as females tend to live longer than males, and the occurrence of certain diseases are sex-specific, for instance, neurodegenerative diseases such as Parkinson's disease are more prevalent in males (Cerri et al., 2019). Sex analysis showed all five models in our study significantly yielded smaller prediction MAE values for males than females in the training subset, indicating age prediction models for female samples were less accurate than male samples. This might be due to female hormones such as estrogen, which can protect females from some diseases, for example, bone loss increases dramatically in women after menopause (Streicher et al., 2017). The presence of a single X chromosome in males (rather than two in females) might also explain why males are more susceptible to genetic diseases linked to the X chromosome such as hemophilia (Shoukat et al., 2020).

It is worth asking what underlying mechanisms might give rise to the increased levels of circRNAs with aging. Some people hold that the age-related increasing trends for circRNAs are reflective of age-accumulation more so than specific regulation. Such an assumption is based on studies of multiple animals discovering that the increased abundance of circRNAs during aging was found to be largely independent of the gene expression from their host due to their lack of 3' or 5' end (Knupp and Miura, 2018). In addition to the contribution of circRNA stability, some researchers suppose that increased circRNA biogenesis due to age-related changes in alternative splicing might play a role (Kramer et al., 2015). These results prompt us that it might be necessary to figure out the functions of circRNAs in aging and mechanisms contributing to circRNA changes during aging.

In the present study, we first constructed age prediction models using novel biomarkers, circular RNAs. Our analysis of blood-derived circRNAs provides insights into changes in circRNA abundance dependent on age and is an exploration for circRNAs as potential biomarkers to be applied in forensic practice in trap future. It is convinced that the accuracy of models can be improved in our future work through newly age-related circRNAs discovered, algorithms optimized, sample sizes enlarged, and a wide age range included. In a word, based on our results, we argue that the prediction of the chronological age utilizing age-dependent changes of specific

circRNAs is a promising application and will become an increasing field of interest.

## 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 Medical Ethics Committee of Hebei Medical University (No. 20190013). The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

SL and BC devised the experiments. CW and JF provided the peripheral whole blood samples. JW, YW, YZ, and CW performed the laboratory work. JW and CL collected, analyzed, and visualized the data. JW wrote the manuscript. SL modified the manuscript. All authors contributed to the manuscript and approved the submitted version.

## FUNDING

This work was supported by the National Natural Science Foundation of China (82072118) and the Hebei Province Science Fund (H2021206451).

## ACKNOWLEDGMENTS

The authors would like to thank voluntary participants who donated blood samples toward this project. The preliminary results of our research with a small sample size had previously appeared online as a preprint at BioRxiv (Wang et al., 2020). Our newest results and more comprehensive analysis are added into the current study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.825443/full#supplementary-material>

## REFERENCES

- Abu, N., and Jamal, R. (2016). Circular RNAs as Promising Biomarkers: A Mini-Review. *Front. Physiol.* 7, 355. doi:10.3389/fphys.2016.00355
- Cabitz, F., and Banfi, G. (2018). Machine Learning in Laboratory Medicine: Waiting for the Flood? *Clin. Chem. Lab. Med.* 56 (4), 516–524. doi:10.1515/cclm-2017-0287
- Cai, H., Li, Y., Niringiyumukiza, J. D., Su, P., and Xiang, W. (2019). Circular RNA Involvement in Aging: An Emerging Player with Great Potential. *Mech. Ageing Dev.* 178, 16–24. doi:10.1016/j.mad.2018.11.002
- Cerri, S., Mus, L., and Blandini, F. (2019). Parkinson's Disease in Women and Men: What's the Difference? *J. Parkinson's Dis.* 9 (3), 501–515. doi:10.3233/JPD-191683
- Cheng, J., Metge, F., and Dieterich, C. (2016). Specific Identification and Quantification of Circular RNAs from Sequencing Data. *Bioinformatics* 32 (7), 1094–1096. doi:10.1093/bioinformatics/btv656
- Cheng, J., Huang, J., Yuan, S., Zhou, S., Yan, W., Shen, W., et al. (2017). Circular RNA Expression Profiling of Human Granulosa Cells during Maternal Aging Reveals Novel Transcripts Associated with Assisted Reproductive Technology Outcomes. *PLoS One* 12 (6), e0177888. doi:10.1371/journal.pone.0177888
- Cho, S., Jung, S.-E., Hong, S. R., Lee, E. H., Lee, J. H., Lee, S. D., et al. (2017). Independent Validation of DNA-Based Approaches for Age Prediction in Blood. *Forensic Sci. Int. Genet.* 29, 250–256. doi:10.1016/j.fsigen.2017.04.020
- Constantin, L. (2018). Circular RNAs and Neuronal Development. *Adv. Exp. Med. Biol.* 1087, 205–213. doi:10.1007/978-981-13-1426-1\_16
- Cortés-López, M., Gruner, M. R., Cooper, D. A., Gruner, H. N., Voda, A.-I., van der Linden, A. M., et al. (2018). Global Accumulation of circRNAs during Aging in *Caenorhabditis elegans*. *BMC Genomics* 19 (1), 8. doi:10.1186/s12864-017-4386-y
- Dias, H. C., Cordeiro, C., Pereira, J., Pinto, C., Real, F. C., Cunha, E., et al. (2020). DNA Methylation Age Estimation in Blood Samples of Living and Deceased Individuals Using a Multiplex SNaPshot Assay. *Forensic Sci. Int.* 311, 110267. doi:10.1016/j.forsciint.2020.110267
- Dong, W.-W., Li, H.-M., Qing, X.-R., Huang, D.-H., and Li, H.-G. (2016). Identification and Characterization of Human Testis Derived Circular RNAs and Their Existence in Seminal Plasma. *Sci. Rep.* 6, 39080. doi:10.1038/srep39080
- Dowhan, D. H., Hong, E. P., Auboeuf, D., Dennis, A. P., Wilson, M. M., Berget, S. M., et al. (2005). Steroid Hormone Receptor Coactivation and Alternative RNA Splicing by U2AF65-Related Proteins CAPERα and CAPERβ. *Mol. Cell* 17 (3), 429–439. doi:10.1016/j.molcel.2004.12.025
- Du, W., Yang, W., Chen, Y., Wu, Z.-K., Foster, F. S., Yang, Z., et al. (2017). Foxo3 Circular RNA Promotes Cardiac Senescence by Modulating Multiple Factors Associated with Stress and Senescence Responses. *Eur. Heart J.* 38 (18), ehw001–1412. doi:10.1093/eurheartj/ehw001
- Enuka, Y., Lauriola, M., Feldman, M. E., Sas-Chen, A., Ulitsky, I., and Yarden, Y. (2016). Circular RNAs Are Long-Lived and Display Only Minimal Early Alterations in Response to a Growth Factor. *Nucleic Acids Res.* 44 (3), 1370–1383. doi:10.1093/nar/gkv1367
- Fehlmann, T., Lehallier, B., Schaum, N., Hahn, O., Kahraman, M., Li, Y., et al. (2020). Common Diseases Alter the Physiological Age-Related Blood microRNA Profile. *Nat. Commun.* 11 (1), 5958. doi:10.1038/s41467-020-19665-1
- Fiuza, B. S. D., Silva, M. d. J., Alcántara-Neves, N. M., Barreto, M. L., Costa, R. d. S., and Figueiredo, C. A. (2017). Polymorphisms in DENND1B Gene Are Associated with Asthma and Atopy Phenotypes in Brazilian Children. *Mol. Immunol.* 90, 33–41. doi:10.1016/j.molimm.2017.06.030
- Freire-Aradas, A., Pośpiech, E., Aliferi, A., Girón-Santamaría, L., Mosquera-Miguel, A., Pisarek, A., et al. (2020). A Comparison of Forensic Age Prediction Models Using Data from Four DNA Methylation Technologies. *Front. Genet.* 11, 932. doi:10.3389/fgene.2020.00932
- Gao, Y., Wang, J., and Zhao, F. (2015). CIRI: an Efficient and Unbiased Algorithm for De Novo Circular RNA Identification. *Genome Biol.* 16, 4. doi:10.1186/s13059-014-0571-3
- Gruner, H., Cortés-López, M., Cooper, D. A., Bauer, M., and Miura, P. (2016). CircRNA Accumulation in the Aging Mouse Brain. *Sci. Rep.* 6, 38907. doi:10.1038/srep38907
- Hall, H., Medina, P., Cooper, D. A., Escobedo, S. E., Rounds, J., Brennan, K. J., et al. (2017). Transcriptome Profiling of Aging *Drosophila* Photoreceptors Reveals Gene Expression Trends that Correlate with Visual Senescence. *BMC Genomics* 18 (1), 894. doi:10.1186/s12864-017-4304-3
- Hanan, M., Simchovitz, A., Yayon, N., Vaknine, S., Cohen-Fultheim, R., Karmon, M., et al. (2020). A Parkinson's Disease Circ RNA S Resource Reveals a Link between Circ SLC 8A1 and Oxidative Stress. *EMBO Mol. Med.* 12 (9), e11942. doi:10.15252/emmm.201911942
- Haque, S., Ames, R. M., Moore, K., Pilling, L. C., Peters, L. L., Bandinelli, S., et al. (2020). circRNAs Expressed in Human Peripheral Blood Are Associated with Human Aging Phenotypes, Cellular Senescence and Mouse Lifespan. *Geroscience* 42 (1), 183–199. doi:10.1007/s11357-019-00120-z
- Ichikawa, S., Koller, D. L., Padgett, L. R., Lai, D., Hui, S. L., Peacock, M., et al. (2010). Replication of Previous Genome-wide Association Studies of Bone mineral Density in Premenopausal American Women. *J. Bone Miner Res.* 25 (8), 1821–1829. doi:10.1002/jbmr.62
- Jeck, W. R., Sorrentino, J. A., Wang, K., Slevin, M. K., Burd, C. E., Liu, J., et al. (2013). Circular RNAs Are Abundant, Conserved, and Associated with ALU Repeats. *RNA* 19 (2), 141–157. doi:10.1261/rna.035667.112
- Knupp, D., and Miura, P. (2018). CircRNA Accumulation: A New Hallmark of Aging? *Mech. Ageing Dev.* 173, 71–79. doi:10.1016/j.mad.2018.05.001
- Kramer, M. C., Liang, D., Tatomer, D. C., Gold, B., March, Z. M., Cherry, S., et al. (2015). Combinatorial Control of *Drosophila* Circular RNA Expression by Intronic Repeats, hnRNPs, and SR Proteins. *Genes Dev.* 29 (20), 2168–2182. doi:10.1101/gad.270421.115
- Li, Z., Huang, C., Bao, C., Chen, L., Lin, M., Wang, X., et al. (2015). Exon-intron Circular RNAs Regulate Transcription in the Nucleus. *Nat. Struct. Mol. Biol.* 22 (3), 256–264. doi:10.1038/nsmb.2959
- Li, C., Gao, W., Yu, C., Lv, J., Lv, R., Duan, J., et al. (2018a). Age Prediction of Children and Adolescents Aged 6–17 Years: an Epigenome-wide Analysis of DNA Methylation. *Aging* 10 (5), 1015–1026. doi:10.18632/aging.101445
- Li, X., Yang, L., and Chen, L.-L. (2018b). The Biogenesis, Functions, and Challenges of Circular RNAs. *Mol. Cell* 71 (3), 428–442. doi:10.1016/j.molcel.2018.06.034
- Mahmoudi, E., and Cairns, M. J. (2019). Circular RNAs Are Temporally Regulated throughout Development and Ageing in the Rat. *Sci. Rep.* 9 (1), 2564. doi:10.1038/s41598-019-38860-9
- Márquez-Ruiz, A. B., González-Herrera, L., and Valenzuela, A. (2018). Usefulness of Telomere Length in DNA from Human Teeth for Age Estimation. *Int. J. Leg. Med.* 132 (2), 353–359. doi:10.1007/s00414-017-1595-5
- Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., et al. (2013). Circular RNAs Are a Large Class of Animal RNAs with Regulatory Potency. *Nature* 495 (7441), 333–338. doi:10.1038/nature11928
- Meng, H., Ma, K. J., Dong, L. M., Li, C. T., Xiao, B., Xu, L. Y., et al. (2019). Research Progress on Age Estimation Based on DNA Methylation. *Fa Yi Xue Za Zhi* 35 (5), 537–544. doi:10.12116/j.issn.1004-5619.2019.05.006
- Mensà, E., Latini, S., Ramini, D., Storci, G., Bonafè, M., and Olivieri, F. (2019). The Telomere World and Aging: Analytical Challenges and Future Perspectives. *Ageing Res. Rev.* 50, 27–42. doi:10.1016/j.arr.2019.01.004
- Mori, T., Li, Y., Hata, H., Ono, K., Kochi, H., and Communications, B. R. (2002). NIRF, a Novel RING finger Protein, Is Involved in Cell-Cycle Regulation. *Biochem. Biophysical Res. Commun.* 296 (3), 530–536. doi:10.1016/s0006-291x(02)00890-2
- Naue, J., Hoefsloot, H. C. J., Mook, O. R. F., Rijlaarsdam-Hoekstra, L., van der Zwalm, M. C. H., Henneman, P., et al. (2017). Chronological Age Prediction Based on DNA Methylation: Massive Parallel Sequencing and Random forest Regression. *Forensic Sci. Int. Genet.* 31, 19–28. doi:10.1016/j.fsigen.2017.07.015
- Park, J.-L., Kim, J. H., Seo, E., Bae, D. H., Kim, S.-Y., Lee, H.-C., et al. (2016). Identification and Evaluation of Age-Related DNA Methylation Markers for Forensic Use. *Forensic Sci. Int. Genet.* 23, 64–70. doi:10.1016/j.fsigen.2016.03.005
- Salzman, J., Gawad, C., Wang, P. L., Lacayo, N., and Brown, P. O. (2012). Circular RNAs Are the Predominant Transcript Isoform from Hundreds of Human Genes in Diverse Cell Types. *PLoS One* 7 (2), e30733. doi:10.1371/journal.pone.0030733
- Schmeling, A., Dettmeyer, R., Rudolf, E., Vieth, V., and Geserick, G. (2016). Forensic Age Estimation: Methods, Certainty, and the Law. *Dtsch Arztebl Int.* 113 (4), 44–50. doi:10.3238/arztebl.2016.0044
- Schmeling, A., Geserick, G., Reisinger, W., and Olze, A. (2007). Age Estimation. *Forensic Sci. Int.* 165 (2–3), 178–181. doi:10.1016/j.forsciint.2006.05.016

- Shen, W., Guo, Y., Wang, Y., Zhao, K., Wang, B., and Yuille, A. (2021). Deep Differentiable Random Forests for Age Estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 404–419. doi:10.1109/TPAMI.2019.2937294
- Shoukat, H. M. H., Ghous, G., Tarar, Z. I., Shoukat, M. M., and Ajmal, N. (2020). Skewed Inactivation of X Chromosome: A Cause of Hemophilia Manifestation in Carrier Females. *Cureus* 12 (10), e11216. doi:10.7759/cureus.11216
- Streicher, C., Heyny, A., Andrukhova, O., Haigl, B., Slavic, S., Schüller, C., et al. (2017). Estrogen Regulates Bone Turnover by Targeting RANKL Expression in Bone Lining Cells. *Sci. Rep.* 7 (1), 6460. doi:10.1038/s41598-017-06614-0
- Vidaki, A., Ballard, D., Aliferi, A., Miller, T. H., Barron, L. P., and Syndercombe Court, D. (2017). DNA Methylation-Based Forensic Age Prediction Using Artificial Neural Networks and Next Generation Sequencing. *Forensic Sci. Int. Genet.* 28, 225–236. doi:10.1016/j.fsigen.2017.02.009
- Wang, Y., Chen, R., Yan, X., Zeng, S., Zhang, T., Cheng, F., et al. (2018). UHRF2 Promotes DNA Damage Response by Decreasing P21 via RING finger Domain. *Biotechnol. Lett.* 40 (8), 1181–1188. doi:10.1007/s10529-018-2577-5
- Wang, J., Fu, G., Wang, Q., Cong, B., and Li, S. (2019). CircRNA: A Novel Biomarker for Forensic Age Estimation? *Forensic Sci. Int. Genet. Suppl. Ser. 7* (1), 504–506. doi:10.1016/j.fsigs.2019.10.068
- Wang, J., Wang, C., Fu, L., Wang, Q., Fu, G., Lu, C., et al. (2020). Circular RNA as a Potential Biomarker for Forensic Age Prediction Using Multiple Machine Learning Models: A Preliminary Study. *bioRxiv*. doi:10.1101/2020.11.10.376418
- Westholm, J. O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., et al. (2014). Genome-wide Analysis of drosophila Circular RNAs Reveals Their Structural and Sequence Properties and Age-dependent Neural Accumulation. *Cel Rep.* 9 (5), 1966–1980. doi:10.1016/j.celrep.2014.10.062
- Yamanai, E., Uchiyama, S., Sakurada, M., and Ueno, Y. (2018). sjTREC Quantification Using SYBR Quantitative PCR for Age Estimation of Bloodstains in a Japanese Population. *Leg. Med.* 32, 71–74. doi:10.1016/j.legalmed.2018.03.003
- Yang, L., Lin, X., Tang, H., Fan, Y., Zeng, S., Jia, L., et al. (2020). Mitochondrial DNA Mutation Exacerbates Female Reproductive Aging via Impairment of the NADH/NAD + Redox. *Aging Cell* 19, e13206. doi:10.1111/acer.13206
- Yim, Y.-I., Sun, T., Wu, L.-G., Raimondi, A., De Camilli, P., Eisenberg, E., et al. (2010). Endocytosis and Clathrin-Uncoating Defects at Synapses of Auxilin Knockout Mice. *Proc. Natl. Acad. Sci.* 107 (9), 4412–4417. doi:10.1073/pnas.1000738107
- Yu, A. Q., Wang, Z. X., Wu, W., Chen, K. Y., Yan, S. R., and Mao, Z. B. (2019). Circular RNA CircCCNB1 Sponges Micro RNA-449a to Inhibit Cellular Senescence by Targeting CCNE2. *Aging* 11 (22), 10220–10241. doi:10.18632/aging.102449
- Zapico, S. C., and Ubelaker, D. H. (2016). Relationship between Mitochondrial DNA Mutations and Aging. Estimation of Age-At-Death. *Gerona* 71 (4), 445–450. doi:10.1093/gerona/glv115
- Zbieć-Piekarska, R., Spólnicka, M., Kupiec, T., Makowska, Ż., Spas, A., Parys-Proszek, A., et al. (2015a). Examination of DNA Methylation Status of the ELOVL2 Marker May Be Useful for Human Age Prediction in Forensic Science. *Forensic Sci. Int. Genet.* 14, 161–167. doi:10.1016/j.fsigen.2014.10.002
- Zbieć-Piekarska, R., Spólnicka, M., Kupiec, T., Parys-Proszek, A., Makowska, Ż., Pałeczka, A., et al. (2015b). Development of a Forensically Useful Age Prediction Method Based on DNA Methylation Analysis. *Forensic Sci. Int. Genet.* 17, 173–179. doi:10.1016/j.fsigen.2015.05.001
- Zhang, Y., Xue, W., Li, X., Zhang, J., Chen, S., Zhang, J.-L., et al. (2016). The Biogenesis of Nascent Circular RNAs. *Cel Rep.* 15 (3), 611–624. doi:10.1016/j.celrep.2016.03.058
- Zhang, Z., Yang, T., and Xiao, J. (2018). Circular RNAs: Promising Biomarkers for Human Diseases. *EBioMedicine* 34, 267–274. doi:10.1016/j.ebiom.2018.07.036
- Zubakov, D., Liu, F., Kokmeijer, I., Choi, Y., van Meurs, J. B. J., van IJcken, W. F. J., et al. (2016). Human Age Estimation from Blood Using mRNA, DNA Methylation, DNA Rearrangement, and Telomere Length. *Forensic Sci. Int. Genet.* 24, 33–43. doi:10.1016/j.fsigen.2016.05.014

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

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

Copyright © 2022 Wang, Wang, Wei, Zhao, Wang, Lu, Feng, Li and Cong. 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.



# Trends in the Contribution of Genetic Susceptibility Loci to Hyperuricemia and Gout and Associated Novel Mechanisms

Jianan Zhao<sup>1,2,6</sup>, Shicheng Guo<sup>3,4\*</sup>, Steven J. Schrodi<sup>3,4\*</sup> and Dongyi He<sup>1,2,5,6\*</sup>

<sup>1</sup>Department of Rheumatology, Shanghai Guanghua Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China, <sup>2</sup>Guanghua Clinical Medical College, Shanghai University of Traditional Chinese Medicine, Shanghai, Shanghai, China, <sup>3</sup>Computation and Informatics in Biology and Medicine, University of WI-Madison, Madison, WI, United States, <sup>4</sup>Department of Medical Genetics, School of Medicine and Public Health, University of WI-Madison, Madison, WI, United States, <sup>5</sup>Arthritis Institute of Integrated Traditional and Western Medicine, Shanghai Chinese Medicine Research Institute, Shanghai, China, <sup>6</sup>Institute of Arthritis Research in Integrative Medicine, Shanghai Academy of Traditional Chinese Medicine, Shanghai, China

## OPEN ACCESS

### Edited by:

Haitao Wang,  
National Cancer Institute,  
United States

### Reviewed by:

Jing Guo,  
Stanford University, United States  
Xin Yin,  
Penn State Milton S. Hershey Medical  
Center, United States  
Yuying Huang,  
University of Texas MD Anderson  
Cancer Center, United States  
Mingdian Tan,  
Stanford University, United States

### \*Correspondence:

Shicheng Guo  
Shicheng.Guo@wisc.edu  
Steven J. Schrodi  
Schrodi@wisc.edu  
Dongyi He  
dongyihe@medmail.com.cn

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

**Received:** 06 May 2022

**Accepted:** 31 May 2022

**Published:** 23 June 2022

### Citation:

Zhao J, Guo S, Schrodi SJ and He D  
(2022) Trends in the Contribution of  
Genetic Susceptibility Loci to  
Hyperuricemia and Gout and  
Associated Novel Mechanisms.  
Front. Cell Dev. Biol. 10:937855.  
doi: 10.3389/fcell.2022.937855

Hyperuricemia and gout are complex diseases mediated by genetic, epigenetic, and environmental exposure interactions. The incidence and medical burden of gout, an inflammatory arthritis caused by hyperuricemia, increase every year, significantly increasing the disease burden. Genetic factors play an essential role in the development of hyperuricemia and gout. Currently, the search on disease-associated genetic variants through large-scale genome-wide scans has primarily improved our understanding of this disease. However, most genome-wide association studies (GWASs) still focus on the basic level, whereas the biological mechanisms underlying the association between genetic variants and the disease are still far from well understood. Therefore, we summarized the latest hyperuricemia- and gout-associated genetic loci identified in the Global Biobank Meta-analysis Initiative (GBMI) and elucidated the comprehensive potential molecular mechanisms underlying the effects of these gene variants in hyperuricemia and gout based on genetic perspectives, in terms of mechanisms affecting uric acid excretion and reabsorption, lipid metabolism, glucose metabolism, and nod-like receptor pyrin domain 3 (NLRP3) inflammasome and inflammatory pathways. Finally, we summarized the potential effect of genetic variants on disease prognosis and drug efficacy. In conclusion, we expect that this summary will increase our understanding of the pathogenesis of hyperuricemia and gout, provide a theoretical basis for the innovative development of new clinical treatment options, and enhance the capabilities of precision medicine for hyperuricemia and gout treatment.

**Keywords:** hyperuricemia, gout, genetic susceptibility loci, novel mechanism, inflammation introduction

## INTRODUCTION

Gout is the leading cause of inflammatory arthritis in males. This is primarily due to multiple mechanisms resulting in the deposition of urate in the synovial fluid and other tissues to form monosodium urate crystals, which are further stimulated by inflammatory irritants, ultimately resulting in gout. The global prevalence of gout is approximately 0.1%–10%, and the incidence ranges

from 0.3 to 6 cases per 1,000 person-years (Kuo et al., 2015; Liu et al., 2015; GBD, 2017). With a worldwide trend of an aging population, the medical disease burden of gout is increasing (Smith et al., 2014). Risk factors for hyperuricemia and gout include the use of medications (thiazides, cyclosporine, low-dose aspirin), insulin resistance, metabolic syndrome, obesity, renal insufficiency, abnormal blood pressure, purine-rich foods, alcohol, and sugary drinks (Neogi, 2011). The role of wine in gout may be contradictory, in any case, a retrospective study said individuals with established gout and pre-existing risk factors should limit all types of alcohol intake to prevent gout episodes (Nieradko-Iwanicka, 2021).

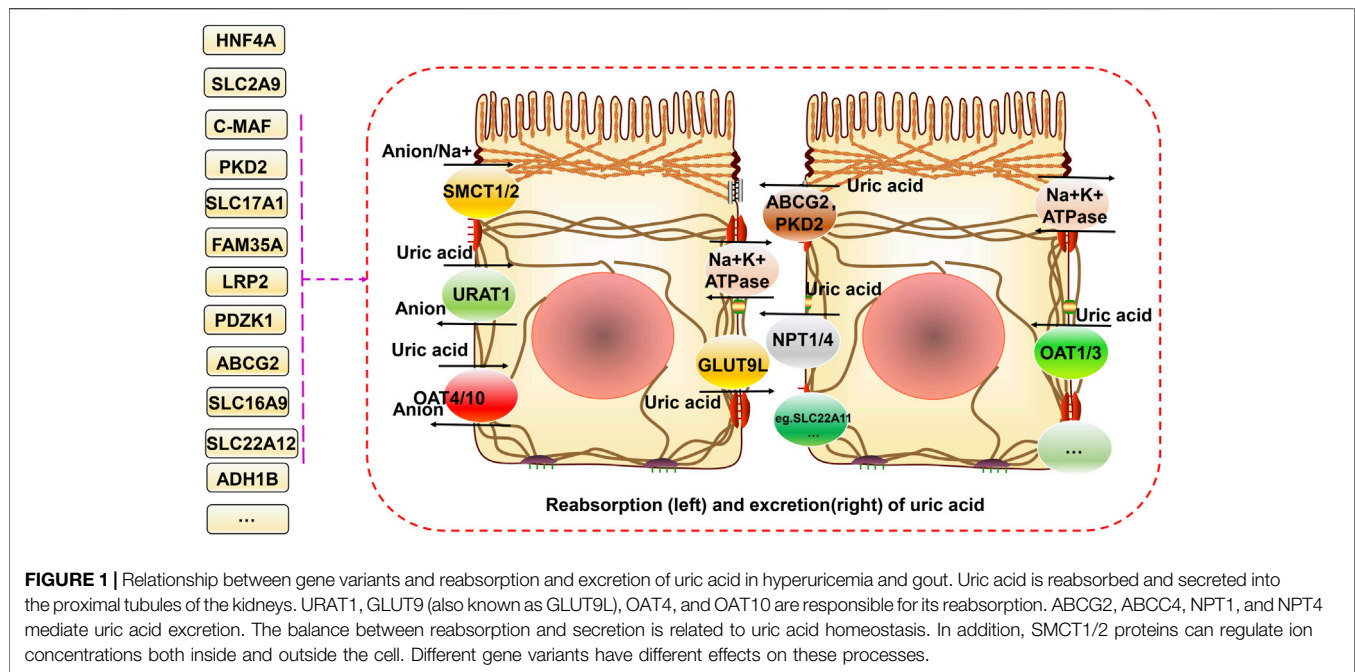
The main source of uric acid is the metabolism of purines and nucleotides in food produced in the liver and excreted by the intestines and kidneys (Köttgen et al., 2013a). Uric acid is reabsorbed and secreted into the proximal tubules of the kidneys, and urate transporter-1 (URAT1), glucose transporter 9 (GLUT9) (also known as GLUT9L), organic anion transporter 4 (OAT4), and organic anion transporter 4 (OAT10) are responsible for its reabsorption. TP-binding cassette superfamily G member 2 (ABCG2), adenosine triphosphate (ATP) binding cassette subfamily C member 4 (ABCC4), and organic anion transporter (NPT1 and NPT4) proteins mediate uric acid excretion. The balance between reabsorption and secretion is related to the homeostasis of uric acid. Otherwise, hyperuricemia and gout can occur. A high serum uric acid concentration is the primary risk factor for gout. Controlling the metabolism of uric acid in circulation at reasonable levels plays a vital role in preventing and improving gout (Chung and Kim, 2021). The progression from high blood uric acid levels to gout occurs in three main steps, hyperuricemia, the deposition of monosodium urate crystals, and inflammatory responses in the joints (Dalbeth et al., 2016). Toll-like receptor (TLR) and NLRP3 inflammasome activation and the associated inflammatory responses are critical factors in the progression of hyperuricemia to gout. This primarily involves activation of the downstream TLR4 and nuclear factor- $\kappa$ B (NF- $\kappa$ B) pathways, activation of the NLRP3 inflammasome, and production of interleukin (IL)-1 $\beta$ , which together regulate immune, metabolic, and inflammatory processes (Qing et al., 2013; McKinney et al., 2015; Rasheed et al., 2016).

Gene variants in related functional proteins can affect uric acid metabolism and inflammation *in vivo* (Reginato et al., 2012). Current research is focused on the heritability of uric acid-associated phenotypes, estimated to be around 40%–70%, implying a clear indication of the importance of its genetic role (Köttgen et al., 2013a). Commonly used drugs to treat acute gout attacks include nonsteroidal anti-inflammatory drugs (NSAIDs), colchicine, and glucocorticoids (Terkeltaub, 2003). Several biologically targeted agents have also been developed, such as IL-1/1 $\beta$  antagonists, anakinra, rilonacept, canakinumab (So et al., 2007; Terkeltaub et al., 2009; Neogi, 2010; So et al., 2010). In addition, patients with gout require a combination of long-term treatments to lower uric acid levels, such as allopurinol, probenecid, and sulfinpyrazone (Terkeltaub, 2003). Although existing gout treatment drugs have achieved some efficacy, the multiple side effects and even poor drug

response in some patients suggest that we should focus, at least in part, on the genetic mechanisms underlying hyperuricemia and gout to identify other effective and well-tolerated clinical treatment options. Large-scale genome-wide association studies have identified many risk loci (Tin et al., 2019a); however, the biological mechanisms underlying hyperuricemia and gout remain unclear. A meta-analysis of the Global Biobank Meta-analysis Initiative improved the understanding of this disease, as well as risk prediction, by integrating GWAS results from six major ancestral groups (African ancestry from African or mixed-race immigrants, mixed-race Americans, Central and South Asians, East Asians, Europeans, and Middle Easterners), while also providing insight into the underlying biology of the traits being studied by integrating gene and protein expression data, enabling the identification of disease-related genes and drug candidates (Zhou et al., 2021). This review discusses the mechanisms through which gene variants affect hyperuricemia and gout by searching *Pubmed* and *GBMI* (<https://www.globalbiobankmeta.org/> and <http://results.globalbiobankmeta.org/>) database. This review further explores and discusses the relationship between multiple biological agents and genetic variants and how they potentially affect gout and hyperuricemia to provide a theoretical reference for further clinical treatment options.

## ASSOCIATION BETWEEN URIC ACID TRANSPORTER-RELATED GENE VARIANTS AND HYPERURICEMIA AND GOUT

A decline in kidney function is a vital cause of hyperuricemia and gout. As the kidney is the main organ that excretes uric acid, when this occurs, uric acid excretion by this organ is also reduced. The increase in uric acid in the blood aggravates hyperuricemia, negatively affecting each other and forming a vicious cycle (Asgari and Hilton, 2021). The proximal tubule excretes most uric acid, and when specific lesions occur there, dysfunction leads to low uric acid excretion, which is often associated with genetic variants of specific uric acid transport proteins. Pleiotropy in genetic variation can underlie the regulation of renal function, hyperuricemia, and gout (García-Nieto et al., 2022). For example, Mendelian randomization analysis can be used to analyze causality in confounding situations. Using this approach to determine the relationship between genetic variants regulating blood uric acid excretion in the kidney and renal function, it was found that the uric acid transporter genetic risk score (mainly comprising solute carrier family 2 member 9 (*SLC2A9*), ABCG2, solute carrier family 22 member 11 (*SLC22A11*), solute carrier family 17 member 1 (*SLC17A1*), and solute carrier family 22 member 12 (*SLC22A12*)) was positively associated with improved renal function in European Caucasian males. The uric acid transporter protein genetic risk score was used as an instrumental variable. Mendelian randomization for renal function using the two-stage least squares method to assess the effect of urate on renal function quantitatively (Hughes et al., 2014). In conclusion, it was found that the variant with



the strongest effect on the protection of renal function was located in *SLC22A11* (Hughes et al., 2014). A meta-analysis further identified four gene loci (*SLC2A9*, *ABCG2*, *SLC22A12*, and *MAF BZIP transcription factor (MAF)*) associated with blood uric acid levels and renal function in an East Asian population (Okada et al., 2012). Another meta-analysis of GWASs on serum urea salt concentrations and gout in African Americans found genome-wide significance at three loci (*SLC2A9*, solute carrier family 2 member 12 (*SLC2A12*), and *SLC22A12*) (Tin et al., 2011). Previous genome-wide significant loci associated with serum urate levels, such as *SLC2A12*, were identified and validated in a meta-analysis by combining GWAS data from more than 14,000 individuals (Köttgen et al., 2013a). Similarly, *ABCG2*, *SLC2A9*, solute carrier family 16 member 9 (*SLC16A9*), glucokinase (hexokinase 4) regulator (*GCKR*), *SLC22A11*, *SLC22A12*, PDZ domain containing 1 (*PDZK1*), and *SLC17A1* were found to be significantly associated with hyperuricemia and gout risk in Asian, native Hawaiian, and Pacific Islander populations estimated using the biospecimens repository at the University of Hawai'i (Alghubayshi et al., 2022). Most gout-related genetic studies have focused on this mechanism and have made some exciting discoveries, but there still could be a need to focus on this and explore it more extensively in the future (Figure 1).

### Hepatocyte Nuclear Factor 4 Alpha (HNF4A), Hepatocyte Nuclear Factor 4 Gamma (HNF4G), and PDZK1

*HNF4A* encodes a nuclear transcription factor that binds DNA and modulates the transcription of multiple genes, mainly in the form of homodimers. A missense variant in *HNF4A* (rs1800961) is probably the most likely leading and causal variant resulting in

better transactivation of the promoter of the urate transporter protein-encoding gene *ABCG2* (Tin et al., 2019a). Additionally, *HNF4A* can also control gene expression in pancreatic islets, potentially further associating with uric acid and gout by affecting insulin secretion (Yoon et al., 2001). The inheritance and expression of different alleles of *HNF4A* might also have potential effects on renal function, but the exact mechanism remains unknown (Leask et al., 2020).

*PDZK1* primarily encodes a scaffolding protein containing the PDZ structural domain. It mediates the localization of cell surface proteins and is linked to cholesterol metabolism through the regulation of multiple receptors. *HNF4A* can also directly regulate *PDZK1*. The T-allele of *PDZK1* single-nucleotide polymorphism (SNP) (rs1967017) enhances *HNF4A* binding to the promoter of *PDZK1*, augmenting its expression, potentially increasing uric acid transport, and regulating uric acid homeostasis, as *PDZK1* is a scaffolding protein for multiple transport proteins (Ketharnathan et al., 2018). A C-MAF BZIP transcription factor-encoding (*C-MAF*) SNP (rs889472) might also be associated with gout susceptibility by affecting uric acid metabolism (Higashino et al., 2018), and part of the mechanism could be related to regulation of the transcription factor *HNF4A* (Leask et al., 2018). *MAF/c-MAF* is mainly expressed in the proximal tubules of the kidney and is a critical factor for maintaining differentiation and functional integrity (Imaki et al., 2004; Tsuchiya et al., 2015). Its b-ZIP structure can form dimers with other b-ZIP proteins and bind to DNA as transcription factors to regulate the functions of various organs, such as the kidney and pancreas (Yang and Cvekl, 2007; Tsuchiya et al., 2015). Leask et al. summarized the genetic mechanisms underlying the detailed regulation of uric acid levels mediated by *MAF* variants, mainly involving the proximal signal cis-expression quantitative trait loci (*cis-eQTL*) of *MAF* (controls

**TABLE 1 |** Gene variants associated with hyperuricemia and gout.

| Items           | SNPs and Its Potential Impact   | Molecular Mechanisms and Associations  | Ref   |
|-----------------|---|--|---|
| <i>HNF4A</i>    | rs1800961 (+)   | Has a stronger activating effect on ABCG2  | Tin et al. (2019a)  |
| <i>PDZK1</i>    | rs1967017 (-)<br>rs112129861 (+)  | Enhances binding to HNF4A to increase uric acid transport and interacts with IGF1R to regulate the inflammatory response   | (Ketharnathan et al., 2018; Fernández-Torres et al., 2019)  |
| <i>C-MAF</i>    | rs889472 (un)   | Can interact with HNF4A and is associated with gout susceptibility   | Higashino et al. (2018)   |
| <i>ABCG2</i>    | rs2231142 (+)<br>rs2231137 (+)<br>rs1481012 (+)<br>rs13120400 (+)<br>rs7672194 (+)                      | Associated with early-onset gout, erythema, and gout stone appearance<br>Variants destabilize the nucleotide-binding structural domain of ABCG2 and inflammatory responses<br>Also interacts with the SNP of PKD2 and serves as a diagnostic and prognostic marker | (Tu et al., 2014; Wong et al., 2016; Dong et al., 2020; Onuora, 2020; Sandoval-Plata et al., 2021)            |
| <i>PKD2</i>     | rs2728121 (+)   | Interacts with the SNP of ABCG2 to increase the risk of gout and increase urate concentration  | (Mejías et al., 1989; Puig et al., 1993; Dong et al., 2020)   |
| <i>SLC2A9</i>   | rs734553 (+)<br>rs16890979 (un)<br>rs16891234 (+)   | Alters protein affinity to increase the risk of hyperuricemia and gout and can be used as a diagnostic and prognostic marker   | (Yi et al., 2018; Sandoval-Plata et al., 2021)  |
| <i>SLC17A1</i>  | rs1183201 (-)   | Protects against disease by enhancing urate excretion and transport and is associated with glucose metabolism  | Kolz et al. (2009)  |
| <i>FAM35A</i>   | rs7903456 (+)   | Reduces the excretion of uric acid in the kidneys  | Nakayama et al. (2017)  |
| <i>LRP2</i>     | rs2390793 (+)<br>rs2544390 (+)<br>rs16856823 (+)  | Mainly affects the renal reabsorption of uric acid, alcohol, and lipid metabolism  | (Kamatani et al., 2010; Rasheed et al., 2013; Kanai et al., 2018; Nakatochi et al., 2019b; Tin et al., 2019b) |
| <i>SLC22A12</i> | rs150255373 (-)<br>rs563239942 (-)<br>rs200104135 (-)<br>rs528619562 (-)<br>rs12800450 (-)              | Protective factor against gout that functions by altering protein function   | Tin et al. (2018)   |
| <i>ADH1B</i>    | rs129984 (+)  | It mainly affects the renal reabsorption of uric acid and acts synergistically with transporter protein URAT1 and can be used to predict the transition from asymptomatic hyperuricemia to gout  | (Lieber et al., 1962; Edenberg, 2007; Macgregor et al., 2009; Sandoval-Plata et al., 2021)                    |
| <i>HNF4G</i>    | rs2941484 (+)   | Associated with gout by regulating endogenous fatty acid metabolism  | Wisely et al. (2002)  |
| <i>PNPLA3</i>   | rs738409 (-)  | Affects gout susceptibility by influencing lipid metabolism and oxidative stress processes   | Diogo et al. (2018)   |
| <i>IGF1R</i>    | rs12908437 (un)<br>rs659854 (un)<br>rs659854 (un)<br>rs1291127 (un)<br>rs4966024 (un)<br>rs7193778 (un) | Affects gout susceptibility by influencing lipid metabolism and oxidative stress processes and modulates the inflammatory response by interacting with PDZK1   | Park et al. (2021)  |
| <i>GCKR</i>     | rs780094 (+)<br>rs1260326 (+)   | Regulates uric acid levels by modulating glucolipid metabolism, promotes an inflammatory response by interacting with STC1, and can be used as a diagnostic and prognostic marker  | (Köttgen et al., 2013b; Fernández-Torres et al., 2019; Sandoval-Plata et al., 2021)                           |
| <i>A1CF</i>     | rs10821905 (+)  | Regulates uric acid levels by modulating dyslipidemia and alcohol metabolism   | Köttgen et al. (2013b)  |
| <i>MLXIP</i>    | rs7953704 (un)  | Transcription factor that might regulate serum uric acid through the pentose phosphate pathway   | Boocock et al. (2020)   |
| <i>MLXIPL</i>   | rs1178977 (un)  | Responsible for regulating glucose flux and potentially affects the renal clearance of urate   | (Hutton et al., 2018; Boocock et al., 2020)   |
| <i>STC1</i>     | rs17786744 (+)  | Promotes the precipitation of monosodium urate crystals to activate the inflammatory response  | (Köttgen et al., 2013b; Fernández-Torres et al., 2019)  |
| <i>CLNK</i>     | rs16869924 (+)<br>rs2041215 (+)<br>rs1686947 (+)  | Regulates B-cell development and activation and mediates the formation of immune complexes through the STAT signaling pathway and might serve as a diagnostic and prognostic marker  | (Siniachenko et al., 1984; Wang et al., 2002; Marrero et al., 2006; Jin et al., 2015)                         |
| <i>SLC22A6</i>  | rs3017670 (un)<br>rs2276300 (un)<br>rs4149171 (un)<br>rs4149170 (un)                                    | Might be associated with the regulation of tryptophan metabolism   | Granados et al. (2021)  |
| <i>BCAS3</i>    | rs11653176 (+)  | Activates estrogen receptor alpha to regulate sex hormone levels affecting uric acid levels  | Sakiyama et al. (2018)  |
| <i>SLC16A9</i>  | rs12356193 (un)   | Might be related to sex hormone regulation   | Köttgen et al. (2013a)  |
| <i>HCRTR2</i>   | rs4715517 (un)  | Might affect the immune system by regulating sleep rhythms   | (Lane et al., 2017; Dashti et al., 2019)  |
| <i>SLC22A11</i> | rs2078267 (+)   |  | Sandoval-Plata et al. (2021)  |

(Continued on following page)

**TABLE 1 |** (Continued) Gene variants associated with hyperuricemia and gout.

| Items               | SNPs and Its Potential Impact                    | Molecular Mechanisms and Associations   | Ref                          |
|---------------------|--|---|------------------------------|
| <i>MEPE</i>         | rs114580333 (+)                                  | Can be used to predict the transition from asymptomatic hyperuricemia to gout | Sandoval-Plata et al. (2021) |
| <i>PPM1K-DT</i>     | rs4693211 (+)<br>rs28793136 (+)<br>rs1545207 (+) | Can be used to predict the transition from asymptomatic hyperuricemia to gout | Sandoval-Plata et al. (2021) |
| <i>LOC105377323</i> | rs114791459 (+)                                  | Can be used to predict the transition from asymptomatic hyperuricemia to gout | Sandoval-Plata et al. (2021) |

(+), The SNP promotes hyperuricemia or gout (-), The SNP inhibits hyperuricemia or gout (un), The SNP has unknown or contradictory effects on hyperuricemia or gout.

the expression of MAF transcriptional regulator RNA (*MAFTRR*) and the distal signal *cis-eQTL* (controls the expression of LINC01229). The *MAFTRR* lncRNA region binds to the MAF promoter and recruits the histone imprint H3K27me3 to repress MAF transcription, whereas the removal of both LINC01229 and *MAFTRR* promotes MAF expression (Leask and Merriman, 2021).

*HNF4G* also encodes a transcription factor involved in the positive regulation of transcription by RNA polymerase II. It has a lower transcriptional activation potential than that of *HNF4A*. An *HNF4G* SNP (rs2941484) can increase gout susceptibility in the Chinese population and mainly affects serum uric acid concentration and gout risk in men (Dong et al., 2017). In Chinese Han men, the TT genotype of the *HNF4G* rs2941484 may represent a gender-specific genetic marker of hyperuricemia. The distribution frequency of TT and CC+CT alleles in hyperuricemic and normokalemic males differed considerably ( $p = 0.011$ ) in the rs2941484 recessive model (Chen et al., 2017). miR-34a can regulate *HNF4G* to control the survival, proliferation, and invasion of bladder cancer cells (Sun et al., 2015) and might bind to endogenous fatty acids to regulate fatty acid metabolic pathways affecting gout (Wisely et al., 2002).

## ABCG2 and Polycystin 2 (PKD2)

*ABCG2* is a multispecific heterotrimeric and endogenous transporter protein expressed mainly in the kidney, liver, and gastrointestinal tract that affects drug metabolism and plays a key role in uric acid excretion. Its variants can lead to destabilization of the nucleotide-binding structural domain of *ABCG2*, resulting in its reduced expression and dysfunction, leading to the inadequate renal excretion of urate, causing hyperuricemia and gout (Wong et al., 2016). *ABCG2* variants (rs2231142) are variants associated with gout and an increased frequency of erythema (Onuora, 2020). Individuals carrying the *ABCG2* SNP (rs2231142) have a nearly 2-fold increased susceptibility to gout (Lee et al., 2019), and alcohol consumption independently increases the risk of gout stones in the Han Chinese population in Taiwan (Tu et al., 2014). The *alpha kinase 1* variant in combination with the *ABCG2* SNP (rs2231142), the *SLC2A9* SNP (rs1014290), or the *SLC22A12* SNP (rs475688 and rs3825016) is linked to gout in the recessive model (Tu et al.,

2018). *ABCG2* SNP (rs2231142) significantly increased the risk of gout in Asians (dominant model: OR = 2.64, 95% CI = 2.04–3.43,  $p = 0.02$  for heterogeneity; recessive model: OR = 3.19, 95% CI = 2.56–3.97,  $p = 0.28$  for heterogeneity; co-dominant model: OR = 1.37, 95% CI = 1.18–1.59,  $p = 0.09$  for heterogeneity) as well as other populations (dominant model: OR = 1.85, 95% CI = 1.20–2.85,  $p < 0.0001$  for heterogeneity; recessive model: OR = 3.78, 95% CI = 2.28–6.27,  $p = 0.19$  for heterogeneity; co-dominant model: OR = 1.48, 95% CI = 1.26–1.74,  $p = 0.19$  for heterogeneity) (Li et al., 2015a). The *ABCG2* SNP (rs72552713) also significantly increased the risk of gout in Asians (dominant model: OR = 3.87, 95% CI = 2.07–7.24,  $p = 0.06$  for heterogeneity) (Li et al., 2015a). *ABCG2* and *PKD2* were found to have epistatic interactions, and two SNP pairs (rs2728121:rs1481012 and rs2728121:rs2231137) were mainly identified as associated with the serum urate concentration or risk of hyperuricemia (Dong et al., 2020). *ABCG2* variants might affect disease progression through inflammatory pathways, in addition to lowering uric acid excretion. The knockdown of *ABCG2* in endothelial cells leads to higher IL-8 release, which further leads to inflammation (Chen et al., 2018). *ABCG2* deficiency in hepatocytes leads to mitochondrial dysfunction and dynamics. Owing to increased intracellular protoporphyrin IX/DRP-1-mediated mitochondrial fission, abnormal protein function results in aggregate formation, leading to excessive reactive oxygen species activation of the NLRP3 inflammasome, which plays a role in the development of gout (Lin et al., 2013). Mitochondrial dysfunction can induce the NLRP3 inflammasome in gout to promote IL-1 $\beta$  and inflammation (Gosling et al., 2018). In addition, monosodium urate crystals also disrupt proteasomal degradation, leading to increased P62 expression, impaired cellular autophagy, and the inability to clear dysfunctional proteins, thus leading to aggregates formation. An *ABCG2* SNP (rs2231142) enhances this autophagic impairment, diminishes the formation of neutrophil extracellular traps, and aggravates gout via the overactive release of the NLRP3 inflammasome and IL-1 $\beta$ . Neutrophil extracellular traps can degrade cytokines and chemokines to limit inflammation (Luciani et al., 2010; Shi et al., 2012; Choe et al., 2014; Schauer et al., 2014). *PKD2* is localized near *ABCG2* and encodes a urate transporter protein. *PKD2* variants in autosomal dominant polycystic kidney disease

result in PKD2 transporter dysfunction and elevated serum urate concentrations, which are associated with hyperuricemia and gout (Mejías et al., 1989; Puig et al., 1993). A transcript assay revealed that *PKD2* and *ABCG2* gene expression levels are positively correlated; thus, the regulators of *PKD2* interact with *ABCG2* to indirectly influence gout incidence (Dong et al., 2020).

### SLC16A9, SLC17A1, and Shieldin Complex Subunit 2 (FAM35A)

A *SLC16A9* SNP (rs2242206) can affect the function of its encoded monocarboxylate transporter 9 (MCT9) protein, resulting in inadequate urate excretion in the kidney (Kolz et al., 2009; Nakayama et al., 2013). *SLC2A9* is expressed in the liver, kidney, and bone cells and transports various substances, including urates and sugars. The SNP rs734553 alters protein affinity to increase the risk of hyperuricemia and gout (Yi et al., 2018).

*SLC17A1* encodes the NPT1 protein. The SNP rs1183201 appears to have a protective effect against diseases by enhancing urate excretion and transport (Kolz et al., 2009). A meta-analysis of GWASs on serum uric acid and gout in 28,283 Caucasian individuals found genome-wide significance for the *SLC17A1* SNP with serum urate levels (Yang et al., 2010).

*FAM35A* variants are associated with gout and hyperuricemia via a mechanism that might involve a reduction in uric acid excretion during renal excretion (Nakayama et al., 2017). *FAM35A* encodes a DNA repair protein expressed mainly in the distal tubules of the kidney and has not been directly linked to uric acid metabolism in functional assays. Therefore, there might be other indirect mechanisms and the potential involvement of kidney function in the regulation of uric acid excretion (Nakayama et al., 2017; García-Nieto et al., 2022).

### SLC22A12

*SLC22A12* encodes the transporter protein URAT1, which is primarily responsible for urate reabsorption following urine filtration. Tin et al. identified 97 rare variants of *SLC22A12*, of which functional validation of *p. Trp325*, *p. Cys405*, and *p. Met467* variants revealed that they cause loss of function of the encoded protein affecting serum uric acid levels. Individuals carrying *SLC22A12* variants have a lower risk of developing gout (Tin et al., 2018). Linkage disequilibrium between *SLC22A12* and *SLC22A11* might be associated with uric acid in Caucasian individuals (Yang et al., 2010). Novel G65W variants of *SLC22A12* (rs12800450) are characterized as functional alleles with an approximately 6–10-fold greater effect on uric acid than that observed for common variants in *SLC22A12* (Tin et al., 2011). Existing drugs have been developed to target URAT1, such as probenecid and benzbromarone. In addition, a new URAT1 inhibitor for the treatment of chronic gout, lesinurad (Zurampic®; RDEA594), was approved in the United States and Europe in 2016 (Miner et al., 2016). However, lesinurad alone appears to impair renal function and should be used in

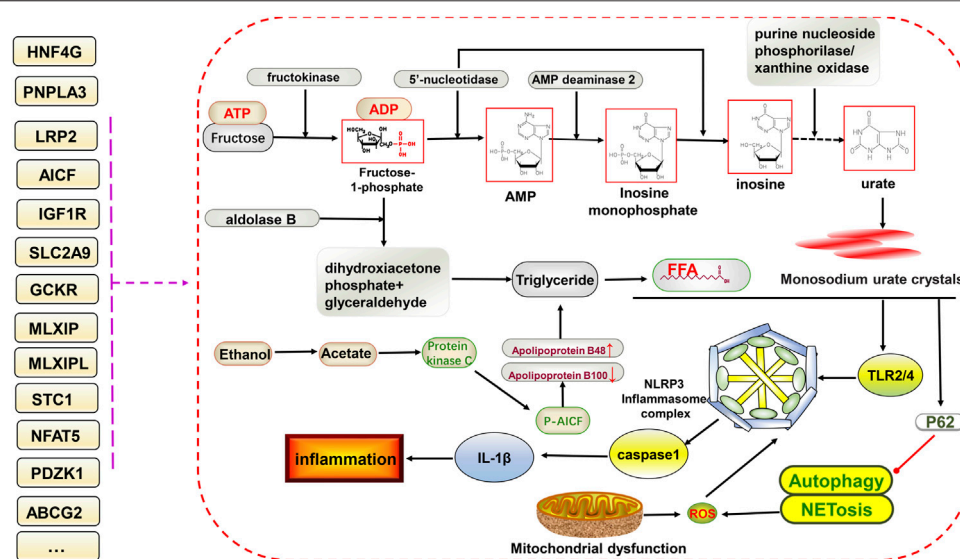
combination with xanthine oxidase inhibitors, and recipients should be closely monitored for renal function (Narang and Dalbeth, 2018). The alcohol dehydrogenase 1B (Class I), beta polypeptide (*ADH1B*) SNP (rs129984) might increase the NADH/NAD ratio to promote lactate production by facilitating ethanol conversion to highly reactive acetaldehyde, thereby increasing uric acid reabsorption in synergy with the *SLC22A12*-encoded transporter protein URAT1 (Lieber et al., 1962; Edenberg, 2007; Macgregor et al., 2009; Sandoval-Plata et al., 2021).

### Solute Carrier Family 22 Member 6 (SLC22A6)

*SLC22A6* primarily encodes organic anion transporter 1 (OAT1) involved in eliminating endogenous and exogenous organic anions from the kidney. Tanner et al. identified multiple SNPs in *SLC22A6* associated with hyperuricemia, including rs3017670, rs2276300, rs4149171, and rs4149170. Strong association studies with gout have been performed; however, there is potential evidence linking it to gout (Tanner et al., 2017). Granados et al. found altered tryptophan metabolite profiles in *SLC22A6*-knockout mice, including several gut microbiota metabolites that are thought to be deleterious for chronic kidney disease. Probenecid, a gout treatment drug, elevates the levels of circulating tryptophan metabolites. Different variants affect the ability of OAT1 to regulate tryptophan metabolism, thus potentially causing gout. Therefore, based on the relationship between OAT1 and tryptophan metabolism, it might be a potential future direction for targets of drug development (Granados et al., 2021). Liu et al. also demonstrated that OAT1 is associated with various metabolic processes, including the tricarboxylic acid cycle, tryptophan metabolism, and other amino acids, fatty acids, and prostaglandins (Liu et al., 2016).

### SLC2A9

*SLC2A9* mainly encodes the GLUT9 protein. The missense variants (rs16890979) of *SLC2A9* showed an association with uric acid and gout (Dehghan et al., 2008). On the one hand, *SLC2A9* is related to regulation by the transcription factor HNF4A. HNF4A overexpression enhances the activity of *SLC2A9*. The mRNA expression levels of *HNF4A* and *SLC2A9* are significantly correlated, and there is an interaction between them (Prestin et al., 2014). The contribution of the coding sequence variants of *SLC2A9* to overall uric acid metabolism is still unknown because of the presence of linkage disequilibrium and heterogeneity, but 24 annotated nonsynonymous variants have been identified (Reginato et al., 2012). The effects of variants in *SLC2A9* (Val253Ile, and Arg265His) are also inconsistent based on studies on gout and hyperuricemia, and further studies are required (Hollis-Moffatt et al., 2009; Tu et al., 2010; Urano et al., 2010; Reginato et al., 2012). On the other hand, *SLC2A9* can exchange uric acid with glucose and fructose, which are involved in gluconeogenesis. This may also have a



**FIGURE 2 |** Potential relationship between gene variants and glucolipid metabolism and NLRP3 inflammasome-mediated inflammatory pathways in hyperuricemia and gout. The ingestion of fructose increases uric acid formation via the gluconeogenic pathway. In addition, triglycerides are also produced, increasing free fatty acid contents via the lipid metabolic pathway. Free fatty acids and monosodium urate crystals together stimulate downstream TLR2/4 and NLRP3 inflammasome formation, facilitate the release of IL-1 $\beta$ , and inhibit P62-mediated activation of autophagy and NETosis, ultimately promoting inflammation. Gene variants have different effects on different processes.

potential impact on hyperuricemia and gout (Batt et al., 2014).

## EFFECT OF GENE VARIANTS RELATED TO GLUCOLIPID METABOLISM ON HYPERURICEMIA AND GOUT

Multiple metabolic factors, including abnormalities in glucose regulation, lipid levels, obesity, and arterial hypertension, are associated with primary gout and hyperuricemia (González-Senac et al., 2014). The glycolytic pathway leads to increased serum uric acid levels through various mechanisms. In addition, insulin resistance and high blood glucose levels can directly affect uric acid clearance in the kidneys (Padova et al., 1964). Hyperinsulinemia increases urate reabsorption in the kidney and decreases renal uric acid and sodium excretion, and this effect can also occur at sites other than the proximal tubule (Quinones Galvan et al., 1995; Maaten et al., 1997). Studies have demonstrated that a high intake of fructose or other high-calorie foods can dramatically increase serum uric acid levels beyond what the body can typically handle, resulting in urate deposition with a severe disruption in hepatocyte metabolism. The rapid intake of fructose can also cause an increase in blood lactate, probably via a mechanism involving blockage of the gluconeogenic pathway caused by the inhibition of glucose-phosphate isomerase mediated by fructose 1-phosphate, leading to the excessive production of lactic acid in hepatocytes (Perheentupa and Raivio, 1967). In addition, fructose phosphorylation in the liver can increase serum uric acid levels by interacting with aldolase B, ATP, and adenosine monophosphate deaminase 2 (AMPD2) (Lanaspa et al., 2011).

Dyslipidemia, insulin resistance, hyperuricemia, and gout are interrelated (Schmidt et al., 1996). Excessive alcohol intake, but not including wine, has been shown to increase serum uric acid levels in several studies (Choi and Curhan, 2004; Yu et al., 2008). Some alcohols, such as beer, contain high levels of purines, and excessive intake can increase uric acid synthesis, leading to hyperuricemia and resistance to some of the antioxidant components of the plasma (van der Gaag et al., 2000; Nishioka et al., 2002). The specific underlying mechanism might involve the degradation of adenosine triphosphate to monophosphate during alcohol metabolism, thereby increasing adenosine and uric acid synthesis. The oxidation of alcohol (ethanol) increases blood lactate, further decreases uric acid excretion, and potentially affects lipid metabolism, thereby increasing the risk of hyperuricemia and gout (Nakamura et al., 2012). Multiple gene variants are associated with glucose metabolism and are potentially associated with hyperuricemia and gout.

## GCKR, MLX Interacting Protein (MLXIP), and MLX Interacting Protein-Like (MLXIPL)

GCKR encodes the GCKR subfamily of proteins that are regulatory proteins that inhibit glucokinase in the liver and pancreatic islet cells by binding noncovalently with the enzyme to form inactive complexes. GCKR variants (rs1260326) are missense variants that serve as possible candidate causal variants for which the leucine allele leads to increased glucokinase GCK activity, resulting in increased glycolytic flux, which facilitates hepatic glucose metabolism (Beer et al., 2009). A GCKR SNP (rs780094) is strongly

**TABLE 2 |** Gene tests in clinical trials.

| Items           | ClinicalTrials.gov identifier | Condition or Disease  | Intervention/Treatment                                 | Aims   |
|-----------------|-------------------------------|---|--|--|
| <i>HNF4A</i>    | NCT01181505                   | Genotype guided (HNF4a), healthy subjects   | Tolterodine  | To study the effect of the <i>HNF-4A</i> G60D variant on the CYP2D6 activity <i>in vivo</i>  |
|                 | NCT04239586                   | Maturity onset diabetes in the young (MODY)   | Sulfonylurea   | To detect the association of the <i>HNF4A</i> variant with insulin secretion in MODY.  |
| <i>ABCG2</i>    | NCT03710395                   | Hypertensive breastfeeding women  | Nifedipine   | The present study aimed to evaluate the effect of <i>ABCG2</i> c.421C>A on nifedipine breast milk/plasma concentration ratio in hypertensive breastfeeding women   |
|                 | NCT04410965                   | Multiple sclerosis  | Teriflunomide  | To evaluate the relationship between <i>ABCG2</i> mutation (rs2231142) and teriflunomide exposure  |
|                 | NCT04608344                   | Rheumatoid arthritis  | Atorvastatin, pravastatin, rosuvastatin, filgotinib    | To evaluate the effect of filgotinib on a mixed organic anion transporting polypeptide/cytochrome P450 3A (OATP/CYP3A), OATP/breast cancer resistance protein (BCRP), and OATP substrates  |
| <i>PKD2</i>     | NCT02112136                   | Autosomal dominant polycystic kidney disease (ADPKD)  | Blood collection                                       | To identify families with ADPKD, characterize the phenotype, and screen for mutations in <i>PKD2</i> genes   |
| <i>SLC22A12</i> | NCT04040907                   | Gout  | XNW3009, placebo                                       | XNW3009 is a small molecule hURAT1 inhibitor. The study investigates the safety, tolerability, pharmacokinetics, and pharmacodynamics of XNW3009   |
|                 | NCT02815839                   | Gout, hyperuricemia   | SHR4640, placebo                                       | SHR4640 is a URAT1 inhibitor. The study assesses the safety, tolerance, food effect, and pharmacokinetic and pharmacodynamic properties of single-dose administration of SHR4640 in healthy volunteers   |
|                 | NCT01883167                   | Healthy   | RDEA3170, febuxostat, placebo                          | To evaluate the potential pharmacokinetic and pharmacodynamic interaction between the xanthine oxidase inhibitor febuxostat and the investigational URAT1 inhibitor RDEA3170   |
|                 | NCT03316131                   | Asymptomatic hyperuricemia  | RDEA3170, febuxostat, dapagliflozin, placebo           | To assess the effect of intensive uric acid lowering therapy with RDEA3170, febuxostat, and dapagliflozin on urinary excretion of uric acid, in asymptomatic hyperuricemic patients  |
| <i>GCKR</i>     | NCT01023750                   | Hypertriglyceridemia, insulin resistance  | Fenofibrate  | To study the pretreatment genotyping at <i>GCKR</i> loci and response to fenofibrate therapy   |
| <i>PNPLA3</i>   | NCT04640324                   | Non-alcoholic fatty liver disease, insulin resistance   | Nutraceutical therapy                                  | To explore the effect of 303 mg of silybin-phospholipids complex, 10 mg of vitamin D, and 15 mg of vitamin E twice a day for 6 months in NAFLD patients carrying <i>PNPLA3</i> -rs738409, <i>TM6SF2</i> -rs58542926, and <i>MBOAT7</i> -rs641738 genetic variants  |
|                 | NCT04483947                   | Non-alcoholic steatohepatitis (NASH)  | AZD2693, placebo                                       | AZD2693 is a <i>PNPLA3</i> inhibitor. This study is intended to investigate the safety and tolerability, pharmacokinetics, and pharmacodynamics of AZD2693 in NASH patients carrying <i>PNPLA3</i> -rs738409   |
| <i>GNAS</i>     | NCT04671719                   | Fibrous dysplasia, albright syndrome, adult children hypoparathyroidism hyperparathyroidism pseudo hypoparathyroidism | blood sample   | To study the full spectrum of PTH and <i>GNAS</i> pathologies  |
| <i>SLC22A6</i>  | NCT02743260                   | Healthy   | Pitavastatin, metformin, digoxin, Adefovir sitagliptin | To establish <i>in vivo</i> phenotyping procedures for organic anionic transporter polypeptide 1B1 (OATP1B1), organic cation transporters 1 and 2 (OCT1/2), multidrug and toxic compound extrusion transporters 1 and 2, kidney splice variant (MATE1/2K), organic anion transporters 1 and 3 (OAT1/3), and p-glycoprotein (P-gp) transporters via a cocktail approach |

associated with gout in Polynesian, European, Japanese, and Chinese populations (Wang et al., 2012; Köttgen et al., 2013a; Urano et al., 2013). In the recessive model, *GCKR* SNP (rs780094) was shown to be associated with the risk of hyperuricemia in men in the Uyghur population of Xinjiang in China ( $p = 0.015$ , OR = 1.311) (Wang et al., 2018). *GCKR* and *NFAT5* are associated with glucose metabolism or the insulin response, and *GCKR* increases

the metabolites that cause gout-related factors through glycolysis (Köttgen et al., 2013a; Rasheed et al., 2017). *MLXIP* encodes a protein that forms a heterodimer with MAX dimerization protein. It regulates the genes that moderate cellular glucose levels. *MLXIPL* encodes a Myc/Max/Mad superfamily basic helix-loop-helix leucine zipper transcription factor that forms a heterodimeric complex. That binds and activates the

carbohydrate response element-binding protein motif within the triglyceride synthesis gene promoter in a glucose-dependent manner. *MLXIP* and *MLXIPL* variants can also correlate with serum urate concentrations (Boocock et al., 2020). *MLXIPL* is primarily associated with cellular carbohydrate metabolism and glycolytic processes. It is directly responsible for regulating glucose flux, interpreted as the pentose phosphate pathway producing ribose 5-phosphate, an essential precursor of *de novo* purine synthesis, and is involved in the production of uric acid (Köttgen et al., 2013a). In addition, the overproduction of lactate affects the transmembrane transport of urate, leading to impaired clearance of urate by the kidney (Luo et al., 2005; Tong et al., 2009; Levine and Puzio-Kuter, 2010).

### Patatin-Like Phospholipase Domain-Containing 3 (*PNPLA3*) and Insulin Like Growth Factor 1 Receptor (*IGF1R*)

*PNPLA3* encodes an active lipase that hydrolyzes various lipids and is associated with oxidative stress (Huang et al., 2011). A *PNPLA3* SNP (rs738409) is associated with hyperuricemia in a Japanese population (Nakatochi et al., 2019a). The rs738409-G allele was found to be associated with a reduced risk of gout in phenome-wide association studies (Diogo et al., 2018). This study indicated that the *PNPLA3* SNP (rs738409) enhances susceptibility to metabolism-related fatty liver disease (MAFLD) and is involved in the pathology of liver fibrosis (Kawaguchi et al., 2018; Namjou et al., 2019). In the recessive model, the *PNPLA3* SNP (rs738409) was associated with NAFLD in different ethnic groups in China: Han (OR = 1.84, 95% CI: 1.03–3.27,  $p = 0.036$ ), Uyghur (OR = 2.25, 95% CI: 1.23–4.09,  $p = 0.006$ ) (Zhang et al., 2014). *IGF1R* encodes the insulin-like growth factor I receptor. This receptor binds insulin-like growth factor with a high affinity. It has tyrosine kinase activity. The insulin-like growth factor I receptor plays a critical role in transformation events. *IGF1R* SNPs (rs12908437, rs659854, rs1291127, and rs4966024) might correlate with blood uric acid levels by affecting the body mass index (BMI) (Park et al., 2021). An abnormal BMI is indicative of abnormal lipid metabolism, and plasma uric acid is a powerful antioxidant (Ames et al., 1981). Thus, *PNPLA3* and *IGF1R* variants might be linked to hyperuricemia and gout by affecting lipid metabolism and oxidative stress.

### APOBEC1 Complementation Factor (*A1CF*)

*A1CF* encodes a protein that may primarily act as an RNA binding subunit and be involved in RNA editing or processing. Rasheed et al. found that both a *GCKR* SNP (rs780094) and *A1CF* SNP (rs10821905) interact with alcohol exposure to increase the risk of gout in a European population under alcohol exposure conditions, suggesting that the involvement of *GCKR* and *A1CF* in alcohol metabolism promotes the development of gout (Rasheed et al., 2017). The *A1CF* SNP has been previously associated with hyperuricemia (Köttgen et al., 2013b). Makoto et al. further investigated the association between the *A1CF* SNP (rs10821905) and gout in Japanese individuals. They found that it was significantly

associated with elevated serum uric acid and gout via a mechanism that might involve the regulation of dyslipidemia and uric acid metabolism (Kawaguchi et al., 2021). Further investigation of the mechanism of interaction between alcohol and *A1CF* could suggest that the metabolite acetate of alcohol (ethanol) leads to the increased production of diacylglycerol and further activates protein kinase C and *A1CF* phosphorylation in the nucleus. This leads to the increased production of apolipoprotein B (ApoB)-48 and decreased production of ApoB-100 at the transcriptional level, further causing increased free fatty acid production from very-low-density lipoproteins/triglycerides, which stimulates downstream TLR, the NLRP3 inflammasome, and IL-1 $\beta$  to activate the inflammatory response and produce more monosodium urate crystals (Joosten et al., 2010; Jump et al., 2013; Rasheed et al., 2017).

### Lipoprotein Receptor-Related Protein 2 (*LRP2*)

*LRP2* encodes an endocytic receptor protein, low-density lipoprotein-related protein 2, which is associated with multiple ligands such as ApoB, lipoprotein lipase, and lactoferrin. It is expressed in numerous tissues, such as proximal renal tubules (Christensen and Birn, 2002). *LRP2* SNPs (rs2390793, rs2544390, and rs16856823) are associated with blood uric acid (Kamatani et al., 2010; Kanai et al., 2018; Nakatochi et al., 2019b; Tin et al., 2019b) and increased gout susceptibility in Japanese (Akashi et al., 2020) and Chinese populations (Dong et al., 2015). However, its variants might lead to renal tubular dysfunction, affecting the renal reabsorption of uric acid (Kamatani et al., 2010; Kanai et al., 2018; Nakatochi et al., 2019b; Tin et al., 2019b). In contrast, rs2544390 was shown to have a non-additive interaction effect with alcohol consumption (beer or spirits), which can increase the risk of serum urate accumulation and gout in alcohol drinkers (Rasheed et al., 2013). However, there are additional contradictory results showing that *LRP2* is not associated with gout susceptibility (Nakayama et al., 2014). *LRP2* can also regulate the activity of lipoprotein lipase to modulate lipid metabolism, which is associated with uric acid metabolism (Rasheed et al., 2013). Additional experiments are needed to clarify the potential biological mechanisms and links between *LRP2* and hyperuricemia and gout.

## NLRP3 INFLAMMASOME AND INFLAMMATION-ASSOCIATED GENES PROMOTE THE PROGRESSION OF HYPERURICEMIA TO GOUT

Pyroptosis, involving the NLRP3 inflammasome, can lead to cell destruction and the release of the pro-inflammatory factors IL-18 and IL-1 $\beta$ , thus promoting inflammation, which has been discussed in rheumatoid arthritis and MAFLD. Both have similarities to gouty arthritis in terms of disease mechanisms (Zhao et al., 2021a; Zhao et al., 2021b). As mentioned earlier, the excessive deposition of uric acid leads to the appearance of monosodium urate crystals, which are stimulated by the

NLRP3 inflammasome and inflammatory factors to progress further toward inflammation. Many genetic variants could be involved in this (Figure 2).

## IGF1R

IGF1R might be associated with activation of the NLRP3 inflammasome in gout. Spadaro et al. found that macrophages lacking the IGF1R have reduced NLRP3 activation and a controlled inflammatory response (Spadaro et al., 2016). Liang et al. found that IGF1R primarily regulates vascular homeostasis and precise endothelial functions and that IGF1R deficiency impairs endothelial function in experimental mice and increases the degree of fibrosis in renal disease, which is associated with a poor wound healing response owing to repeated irritation from inflammation (Liang et al., 2015). Thus, both studies suggest that IGF1R variants might influence gout by regulating inflammation. In addition, a *IGF1R* SNP (rs7193778) and *PDZK1* SNP (rs112129861) could interact with each other, further enriching our understanding of the genetic and biological mechanisms underlying uric acid accumulation and gout (Fernández-Torres et al., 2019).

## Stanniocalcin 1 (STC1)

*STC1* encodes stanniocalcin-1, a glycoprotein that plays a role in multiple biological responses, including bone development, angiogenesis, and inflammatory responses (Yeung et al., 2012). Studies have reported that *STC1* is associated with elevated serum uric acid levels (Köttgen et al., 2013b). An *STC1* SNP (rs17786744) might cause the crystalline precipitation of sodium urate to trigger the inflammatory process, further exacerbating cartilage damage and promoting knee osteoarthritis, which could be associated with the inflammatory response in gouty arthritis. In addition, an interaction between an *STC1* SNP (rs17786744) and *GCKR* SNP (rs1260326) synergistically promotes crystalline precipitation with urate-promoting gout (Fernández-Torres et al., 2019).

## THE ASSOCIATION BETWEEN GENETIC VARIANTS INVOLVED IN OTHER MECHANISMS AND HYPERURICEMIA AND GOUT

Various factors, such as coffee intake, tryptophan metabolism, B-cell development and activation, and sex hormones, are interlinked with genetic variants that play a role in hyperuricemia and gout. Hutton et al. found a negative association between coffee intake and gout. *ABCG2*, *GCKR*, *MLXIPL*, and cytochrome P450 family 1 subfamily A member 2 (*CYP1A2*) are variants associated with coffee consumption habits, and *GCKR* and *ABCG2* are associated with low coffee intake and a high gout risk. Coffee consumption habits indirectly affect the association between gene variants and gout. In contrast, the direct effect of these gene variants on gout is still possible through other mechanisms, as described previously herein (Hutton et al., 2018). Evidence from studies involving genetic

variants associated with other mechanisms is relatively scarce, and further research is needed in the future. Therefore, in this section, we summarize briefly the association of other mechanisms with gout, including sleep rhythm, immune response and B-cell activation (hypocretin receptor 2 (*HCRTR2*), cytokine-dependent hematopoietic cell linker (*CLNK*), guanine nucleotide-binding protein  $\alpha$ -stimulating polypeptide (*GNAS*)), sex hormones (breast cancer-amplified sequence 3 (*BCAS3*)).

## HCRTR2, CLNK and GNAS

The protein encoded by *HCRTR2* is a G protein-coupled receptor involved in the regulation of feeding. The encoded proteins bind to orexin A and orexin B. A *HCRTR2* SNP (rs4715517), a variant associated with serum uric acid, appears to be specific to Asian populations with significantly higher allele frequencies than those in European populations. Differences in allele frequencies might contribute to interethnic differences in serum uric acid levels (Park et al., 2021). *HCRTR2* is mainly involved in the sleep rhythm of the body (Lane et al., 2017; Dashti et al., 2019). With the accelerated pace of life in modern society, irregular sleep affects the immune system and the function of multiple organs, including the kidneys and liver. Therefore, genes regulating sleep rhythms might be potentially associated with hyperuricemia and gout. *CLNK*, a member of the SLP76 family, plays an essential role in integrating immunotyrosine-based activation motif-bearing receptors and integrins and is a positive regulator of immune response signaling (Yu et al., 2001; Wu and Koretzky, 2004). The allele “G” of *CLNK* SNP (rs2041215 and rs1686947) was identified as susceptibility genes for gout in the Chinese population by using dominant model (OR 1.66; 95% CI 1.04–2.63;  $p = 0.031$ ) (OR 2.19; 95% CI 1.38–3.46;  $p = 0.001$ ) and additive model (OR 1.39; 95% CI 1.00–1.93;  $p = 0.049$ ) (OR 1.67; 95% CI 1.19–2.32;  $p = 0.003$ ), respectively (Jin et al., 2015). A *CLNK* SNP (rs16869924) within the established *SLC2A9* gout-associated locus was shown to increase the risk of gout in Polynesian and Chinese Tibetan individuals, genetically independent on the *SLC2A9* association signal (Lan et al., 2016; Ji et al., 2021). It is hypothesized that *CLNK* mainly regulates B-cell development and activation and co-mediate the formation of immune complexes through the STAT signaling pathway to promote gout, as suggested by a combination of related studies (Siniachenko et al., 1984; Wang et al., 2002; Marrero et al., 2006). *GNAS* encodes G $\alpha$ s protein, which activates downstream cyclic AMP (cAMP) production and promotes signaling (Turan and Bastepe, 2015; Tafaj and Jüppner, 2017). *GNAS* variants predispose patients to an abnormal synovial environment and the deposition of uric acid crystals, promoting the formation of gout and related osteoarthritis (Rhyu and Bhat, 2021).

## BCAS3

*BCAS3* encodes proteins that are associated with several functions, such as angiogenesis, activation and recruitment of cell division cycle 42, reorganization of the actin cytoskeleton at the leading edge, regulation of cell polarity, the endothelial cell migration, filopodia formation, estrogen receptor response, and autophagy.

Rs11653176 in *BCAS3* is significantly associated with uric acid levels and gout in Japanese and Chinese Han populations (Li et al., 2015b; Sakiyama et al., 2018). *BCAS3* can activate estrogen receptor alpha (Sakiyama et al., 2018). Studies have shown that sex hormones can affect uric acid levels (Adamopoulos et al., 1977). Postmenopausal women might have elevated uric acid levels owing to a decrease in estrogen, especially estradiol, because estrogen is more effective in promoting urate clearance by the kidneys (Hak and Choi, 2008). Similarly, the effects of some gene variants on serum uric acid levels and gout appear to be sex-specific. For example, variants in *SLC2A9* and *ABCG2*, which are associated with urate concentrations, are sex-specific (Köttgen et al., 2013a). An *SLC16A9* SNP (rs12356193) was found to be weakly associated with gout but strongly associated with blood uric acid and showed a sex-specific difference (Köttgen et al., 2013a). It is thus possible that sex hormones primarily contribute to sex differences in disease or drug efficacy.

## Gene Variants as Potential Diagnostic Markers of Drug Efficacy and Prognosis

A case-control association study of gout in Chinese populations revealed that *CLNK* SNPs (rs2041215 and rs1686947) are associated with various clinicopathological parameters and might have potential as diagnostic and prognostic markers for patients with gout (Jin et al., 2015). Patients with gout carrying an *ABCG2* SNP (rs2231142) respond poorly to allopurine therapy (Wen et al., 2015; Roberts et al., 2017; Wallace et al., 2018). A GWAS and polygenic risk scores in patients with asymptomatic hyperuricemia and gout revealed that *ABCG2* (rs2231142, rs13120400, and rs7672194), *SLC2A9* (rs16890979 and rs16891234), *SLC22A11* (rs2078267), *GCKR* (rs1260326), matrix extracellular phosphoglycoprotein (*MEPE*) (rs114580333), protein phosphatase, Mg<sup>2+</sup>/Mn<sup>2+</sup> dependent 1 K-divergent transcript (*PPM1K-DT*) (rs4693211, rs28793136, and rs1545207), *LOC105377323* (rs114791459), and alcohol dehydrogenase 1B (*Class I*), *beta* polypeptide (*ADH1B*) (rs1229984) SNPs can be used as markers of asymptomatic hyperuricemia to identify transition predictors (Sandoval-Plata et al., 2021). However, little research has been conducted on genetic variants as markers to predict disease progression and drug efficacy, as serum uric acid levels can effectively predict the risk of gout. However, this might provide more relevant results that could be uncovered through in-depth studies in the future.

## CONCLUSION

Gout is a form of arthritis that damages patients' physical and mental health and causes severe pain during acute attacks. Identifying individuals at risk in the early stages of the disease is essential to prevent and reduce hyperuricemia and gout and to provide pharmacological and lifestyle interventions to better treat patients with clinically diagnosed gout. The identification of genetic variants might help in disease prevention and intervention. Many GWASs have performed to uncover loci related to hyperuricemia and gout, mostly linking it to uric

acid transporter proteins, such as the widely studied *URAT1* and *GLUT9*. Some drugs have been used as targets for drug development (see **Table 1**). We also summarize the latest clinical trials of these genes, and some of these were conducted in the context of gout, which is certainly instructive. Although some trials were not investigated in the context of gout, they have some informative implications for the clinical management of gout, which urgently needs to be studied in depth in the context of gout in the future (see **Table 2**). Currently, the most elucidated is the effect of variants in the uric acid transporter protein gene on hyperuricemia and gout. We aim to increase our understanding of the genetic mechanisms behind the disease by adding descriptions of other genes of potential clinical value. All of these genes are undoubtedly promising and essential. Associations between genetic variants and traits are often located in regions of strong linkage disequilibrium and aided by eQTL analysis and fine localization studies, these can be exploited to identify true causal variants of gout in complex genetic backgrounds. Genetic-related issues in multiple disease contexts still need attention and elucidation, such as disease-specific genetic variants in different ethnic backgrounds, genetic variants based on sex differences, rare and low frequency variants, functional polymorphisms in genetic susceptibility genes, and epigenetic mechanisms. With the rapid development of modern molecular biotechnologies and multi-omics techniques, these issues require further clarification. In addition, attention should be paid to the interconnection between hyperuricemia/gout and other diseases, such as metabolic syndrome and cardiovascular diseases, as well as the role of genetic factors in these diseases. Elucidating these genetic issues will contribute to the improvement of clinical outcomes and precision medicine.

## AUTHOR CONTRIBUTIONS

JZ is responsible for the collection, collation, and writing of the original manuscript. SG, SS, and DH are responsible for the concept development, revision, and manuscript review. All authors reviewed and accepted the final version.

## FUNDING

This work was funded by the National Natural Science Funds of China (82074234 and 82071756), National Key Research and Development Project (2018YFC1705200 and 2018YFC1705203), Shanghai Chinese Medicine Development Office, National Administration of Traditional Chinese Medicine, Regional Chinese Medicine (Specialist) Diagnosis and Treatment Center Construction Project-Rheumatology, State Administration of Traditional Chinese Medicine, National TCM Evidence-Based Medicine Research and Construction Project, Basic TCM Evidence-Based Capacity Development Program, Shanghai Municipal Health Commission, and East China Region-based Chinese and Western Medicine Joint Disease Specialist Alliance.

## REFERENCES

- Adamopoulos, D., Vlassopoulos, C., Seitanides, B., Contoyiannis, P., and Vassilopoulos, P. (1977). The Relationship of Sex Steroids to Uric Acid Levels in Plasma and Urine. *Eur. J. Endocrinol.* 85 (1), 198–208. doi:10.1530/acta.0.0850198
- Akashi, A., Nakayama, A., Kamatani, Y., Higashino, T., Shimizu, S., Kawamura, Y., et al. (2020). A Common Variant of LDL Receptor Related Protein 2 (LRP2) Gene Is Associated with Gout Susceptibility: a Meta-Analysis in a Japanese Population. *Hum. Cell* 33 (2), 303–307. doi:10.1007/s13577-019-00318-5
- Alghubayshi, A., Edelman, A., Alrajeh, K., and Roman, Y. (2022). Genetic Assessment of Hyperuricemia and Gout in Asian, Native Hawaiian, and Pacific Islander Subgroups of Pregnant Women: Biospecimens Repository Cross-Sectional Study. *BMC Rheumatol.* 6 (1), 1. doi:10.1186/s41927-021-00239-7
- Ames, B. N., Cathcart, R., Schwiers, E., and Hochstein, P. (1981). Uric Acid Provides an Antioxidant Defense in Humans against Oxidant- and Radical-Caused Aging and Cancer: a Hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 78 (11), 6858–6862. doi:10.1073/pnas.78.11.6858
- Asgari, E., and Hilton, R. M. (2021). One Size Does Not Fit All: Understanding Individual Living Kidney Donor Risk. *Pediatr. Nephrol.* 36 (2), 259–269. doi:10.1007/s00467-019-04456-8
- Batt, C., Phipps-Green, A. J., Black, M. A., Cadzow, M., Merriman, M. E., Topless, R., et al. (2014). Sugar-sweetened Beverage Consumption: a Risk Factor for Prevalent Gout with SLC2A9 genotype-specific Effects on Serum Urate and Risk of Gout. *Ann. Rheum. Dis.* 73 (12), 2101–2106. doi:10.1136/annrheumdis-2013-203600
- Beer, N. L., Tribble, N. D., McCulloch, L. J., Roos, C., Johnson, P. R. V., Orholm-Melander, M., et al. (2009). The P446L Variant in GCKR Associated with Fasting Plasma Glucose and Triglyceride Levels Exerts its Effect through Increased Glucokinase Activity in Liver. *Hum. Mol. Genet.* 18 (21), 4081–4088. doi:10.1093/hmg/ddp357
- Boocock, J., Leask, M., Okada, Y., Matsuo, H., Kawamura, Y., Shi, Y., et al. (2020). Genomic Dissection of 43 Serum Urate-Associated Loci Provides Multiple Insights into Molecular Mechanisms of Urate Control. *Hum. Mol. Genet.* 29 (6), 923–943. doi:10.1093/hmg/ddaa013
- Chen, B.-D., Chen, X.-C., Pan, S., Yang, Y.-N., He, C.-H., Liu, F., et al. (2017). TT Genotype of Rs2941484 in the Human HNF4G Gene Is Associated with Hyperuricemia in Chinese Han Men. *Oncotarget* 8 (16), 26918–26926. doi:10.18632/oncotarget.15851
- Chen, C.-J., Tseng, C.-C., Yen, J.-H., Chang, J.-G., Chou, W.-C., Chu, H.-W., et al. (2018). ABCG2 Contributes to the Development of Gout and Hyperuricemia in a Genome-wide Association Study. *Sci. Rep.* 8 (1), 3137. doi:10.1038/s41598-018-21425-7
- Choe, J.-Y., Jung, H.-Y., Park, K.-Y., and Kim, S.-K. (2014). Enhanced P62 Expression through Impaired Proteasomal Degradation Is Involved in Caspase-1 Activation in Monosodium Urate Crystal-Induced Interleukin-1 Expression. *Rheumatology* 53 (6), 1043–1053. doi:10.1093/rheumatology/ket474
- Choi, H. K., and Curhan, G. (2004). Beer, Liquor, and Wine Consumption and Serum Uric Acid Level: the Third National Health and Nutrition Examination Survey. *Arthritis & Rheumatism* 51 (6), 1023–1029. doi:10.1002/art.20821
- Christensen, E. I., and Birn, H. (2002). Megalin and Cubilin: Multifunctional Endocytic Receptors. *Nat. Rev. Mol. Cell Biol.* 3 (4), 258–267. doi:10.1038/nrm778
- Chung, S., and Kim, G.-H. (2021). Urate Transporters in the Kidney: What Clinicians Need to Know. *Electrolyte Blood Press* 19 (1), 1–9. doi:10.5049/ebp.2021.19.1.1
- Dalbeth, N., Merriman, T. R., and Stamp, L. K. (2016). Gout. *Lancet* 388 (10055), 2039–2052. doi:10.1016/s0140-6736(16)00346-9
- Dashti, H. S., Jones, S. E., Wood, A. R., Lane, J. M., van Hees, V. T., Wang, H., et al. (2019). Genome-wide Association Study Identifies Genetic Loci for Self-Reported Habitual Sleep Duration Supported by Accelerometer-Derived Estimates. *Nat. Commun.* 10 (1), 1100. doi:10.1038/s41467-019-08917-4
- Dehghan, A., Köttgen, A., Yang, Q., Hwang, S.-J., Kao, W. L., Rivadeneira, F., et al. (2008). Association of Three Genetic Loci with Uric Acid Concentration and Risk of Gout: a Genome-wide Association Study. *Lancet* 372 (9654), 1953–1961. doi:10.1016/s0140-6736(08)61343-4
- Diogo, D., Tian, C., Franklin, C. S., Alanne-Kinnunen, M., March, M., Spencer, C. C. A., et al. (2018). Phenome-wide Association Studies across Large Population Cohorts Support Drug Target Validation. *Nat. Commun.* 9 (1), 4285. doi:10.1038/s41467-018-06540-3
- Dong, Z., Zhao, D., Yang, C., Zhou, J., Qian, Q., Ma, Y., et al. (2015). Common Variants in LRP2 and COMT Genes Affect the Susceptibility of Gout in a Chinese Population. *PLoS One* 10 (7), e0131302. doi:10.1371/journal.pone.0131302
- Dong, Z., Zhou, J., Jiang, S., Li, Y., Zhao, D., Yang, C., et al. (2020). Epistatic Interaction between PKD2 and ABCG2 Influences the Pathogenesis of Hyperuricemia and Gout. *Heredity* 157 (1), 2. doi:10.1186/s41065-020-0116-6
- Dong, Z., Zhou, J., Jiang, S., Li, Y., Zhao, D., Yang, C., et al. (2017). Effects of Multiple Genetic Loci on the Pathogenesis from Serum Urate to Gout. *Sci. Rep.* 7, 43614. doi:10.1038/srep43614
- Edenberg, H. J. (2007). The Genetics of Alcohol Metabolism: Role of Alcohol Dehydrogenase and Aldehyde Dehydrogenase Variants. *Alcohol Res. Health* 30 (1), 5–13.
- Fernández-Torres, J., Martínez-Nava, G. A., Oliviero, F., López-Reyes, A. G., Martínez-Flores, K., Garrido-Rodríguez, D., et al. (2019). Common Gene Variants Interactions Related to Uric Acid Transport Are Associated with Knee Osteoarthritis Susceptibility. *Connect. tissue Res.* 60 (3), 219–229. doi:10.1080/03008207.2018.1483359
- García-Nieto, V. M., Claverie-Martín, F., Moraleda-Mesa, T., Perdomo-Ramírez, A., Tejera-Carreño, P., Córdoba-Lanus, E., et al. (2022). La gota asociada a reducción de la excreción renal de ácido úrico. Esa tubulopatía que no tratamos los nefrólogos. *Nefrología* 42, 273–279. doi:10.1016/j.nefro.2021.03.013
- Gbd (2017). Global, Regional, and National Incidence, Prevalence, and Years Lived with Disability for 328 Diseases and Injuries for 195 Countries, 1990–2016: a Systematic Analysis for the Global Burden of Disease Study 2016. *Lancet* 390 (10100), 1211–1259. doi:10.1016/s0140-6736(17)32154-2
- González-Senac, N. M., Bailén, R., Torres, R. J., de Miguel, E., and Puig, J. G. (2014). Metabolic Syndrome in Primary Gout. *Nucleosides, Nucleotides Nucleic Acids* 33 (4–6), 185–191. doi:10.1080/15257770.2013.853785
- Gosling, A. L., Boocock, J., Dalbeth, N., Harré Hindmarsh, J., Stamp, L. K., Stahl, E. A., et al. (2018). Mitochondrial Genetic Variation and Gout in Māori and Pacific People Living in Aotearoa New Zealand. *Ann. Rheum. Dis.* 77 (4), 571–578. doi:10.1136/annrheumdis-2017-212416
- Granados, J. C., Richelle, A., Gutierrez, J. M., Zhang, P., Zhang, X., Bhatnagar, V., et al. (2021). Coordinate Regulation of Systemic and Kidney Tryptophan Metabolism by the Drug Transporters OAT1 and OAT3. *J. Biol. Chem.* 296, 100575. doi:10.1016/j.jbc.2021.100575
- Hak, A. E., and Choi, H. K. (2008). Menopause, Postmenopausal Hormone Use and Serum Uric Acid Levels in US Women - the Third National Health and Nutrition Examination Survey. *Arthritis Res. Ther.* 10 (5), R116. doi:10.1186/ar2519
- Higashino, T., Matsuo, H., Okada, Y., Nakashima, H., Shimizu, S., Sakiyama, M., et al. (2018). A Common Variant of MAF/c-MAF, Transcriptional Factor Gene in the Kidney, Is Associated with Gout Susceptibility. *Hum. cell* 31 (1), 10–13. doi:10.1007/s13577-017-0186-6
- Hollis-Moffatt, J. E., Xu, X., Dalbeth, N., Merriman, M. E., Topless, R., Waddell, C., et al. (2009). Role of the Urate transporter SLC2A9 gene in Susceptibility to Gout in New Zealand Māori, Pacific Island, and Caucasian Case-Control Sample Sets. *Arthritis & Rheumatism* 60 (11), 3485–3492. doi:10.1002/art.24938
- Huang, Y., Cohen, J. C., and Hobbs, H. H. (2011). Expression and Characterization of a PNPLA3 Protein Isoform (I148M) Associated with Nonalcoholic Fatty Liver Disease. *J. Biol. Chem.* 286 (43), 37085–37093. doi:10.1074/jbc.M111.290114
- Hughes, K., Flynn, T., de Zoysa, J., Dalbeth, N., and Merriman, T. R. (2014). Mendelian Randomization Analysis Associates Increased Serum Urate, Due to Genetic Variation in Uric Acid Transporters, with Improved Renal Function. *Kidney Int.* 85 (2), 344–351. doi:10.1038/ki.2013.353
- Hutton, J., Fatima, T., Major, T. J., Topless, R., Stamp, L. K., Merriman, T. R., et al. (2018). Mediation Analysis to Understand Genetic Relationships between

- Habitual Coffee Intake and Gout. *Arthritis Res. Ther.* 20 (1), 135. doi:10.1186/s13075-018-1629-5
- Imaki, J., Tsuchiya, K., Mishima, T., Onodera, H., Kim, J. I., Yoshida, K., et al. (2004). Developmental Contribution of C-Maf in the Kidney: Distribution and Developmental Study of C-Maf mRNA in Normal Mice Kidney and Histological Study of C-Maf Knockout Mice Kidney and Liver. *Biochem. Biophysical Res. Commun.* 320 (4), 1323–1327. doi:10.1016/j.bbrc.2004.05.222
- Ji, A., Shaikat, A., Takei, R., Bixley, M., Cadzow, M., Topless, R. K., et al. (2021). Aotearoa New Zealand Māori and Pacific Population-Amplified Gout Risk Variants: CLNK Is a Separate Risk Gene at the SLC2A9 Locus. *J. Rheumatol.* 48 (11), 1736–1744. doi:10.3899/jrheum.201684
- Jin, T.-b., Ren, Y., Shi, X., Jiri, M., He, N., Feng, T., et al. (2015). Genetic Variations in the CLNK Gene and ZNF518B Gene Are Associated with Gout in Case-Control Sample Sets. *Rheumatol. Int.* 35 (7), 1141–1147. doi:10.1007/s00296-015-3215-3
- Joosten, L. A. B., Netea, M. G., Mylona, E., Koenders, M. I., Malireddi, R. K. S., Oosting, M., et al. (2010). Engagement of Fatty Acids with Toll-like Receptor 2 Drives Interleukin-1 $\beta$  Production via the ASC/caspase 1 Pathway in Monosodium Urate Monohydrate Crystal-Induced Gouty Arthritis. *Arthritis & Rheumatism* 62 (11), 3237–3248. doi:10.1002/art.27667
- Jump, D. B., Tripathy, S., and Depner, C. M. (2013). Fatty Acid-Regulated Transcription Factors in the Liver. *Annu. Rev. Nutr.* 33, 249–269. doi:10.1146/annurev-nutr-071812-161139
- Kamatani, Y., Matsuda, K., Okada, Y., Kubo, M., Hosono, N., Daigo, Y., et al. (2010). Genome-wide Association Study of Hematological and Biochemical Traits in a Japanese Population. *Nat. Genet.* 42 (3), 210–215. doi:10.1038/ng.531
- Kanai, M., Akiyama, M., Takahashi, A., Matoba, N., Momozawa, Y., Ikeda, M., et al. (2018). Genetic Analysis of Quantitative Traits in the Japanese Population Links Cell Types to Complex Human Diseases. *Nat. Genet.* 50 (3), 390–400. doi:10.1038/s41588-018-0047-6
- Kawaguchi, M., Nakayama, A., Aoyagi, Y., Nakamura, T., Shimizu, S., Kawamura, Y., et al. (2021). Both variants of A1CF and BAZ1B Genes Are Associated with Gout Susceptibility: a Replication Study and Meta-Analysis in a Japanese Population. *Hum. Cell* 34 (2), 293–299. doi:10.1007/s13577-021-00485-4
- Kawaguchi, T., Shima, T., Mizuno, M., Mitsumoto, Y., Umemura, A., Kanbara, Y., et al. (2018). Risk Estimation Model for Nonalcoholic Fatty Liver Disease in the Japanese Using Multiple Genetic Markers. *PLoS One* 13 (1), e0185490. doi:10.1371/journal.pone.0185490
- Ketharnathan, S., Leask, M., Boock, J., Phipps-Green, A. J., Antony, J., O'Sullivan, J. M., et al. (2018). A Non-coding Genetic Variant Maximally Associated with Serum Urate Levels Is Functionally Linked to HNF4A-dependent PDZK1 Expression. *Hum. Mol. Genet.* 27 (22), 3964–3973. doi:10.1093/hmg/ddy295
- Kolz, M., Johnson, T., Sanna, S., Teumer, A., Vitart, V., Perola, M., et al. (2009). Meta-analysis of 28,141 Individuals Identifies Common Variants within Five New Loci that Influence Uric Acid Concentrations. *PLoS Genet.* 5 (6), e1000504. doi:10.1371/journal.pgen.1000504
- Köttgen, A., Albrecht, E., Teumer, A., Vitart, V., Krumsiek, J., Hundertmark, C., et al. (2013). Genome-wide Association Analyses Identify 18 New Loci Associated with Serum Urate Concentrations. *Nat. Genet.* 45 (2), 145–154. doi:10.1038/ng.2500
- Köttgen, A., Albrecht, E., Teumer, A., Vitart, V., Krumsiek, J., Hundertmark, C., et al. (2013). Genome-wide Association Analyses Identify 18 New Loci Associated with Serum Urate Concentrations. *Nat. Genet.* 45 (2), 145–154. doi:10.1038/ng.2500
- Kuo, C.-F., Grainge, M. J., Zhang, W., and Doherty, M. (2015). Global Epidemiology of Gout: Prevalence, Incidence and Risk Factors. *Nat. Rev. Rheumatol.* 11 (11), 649–662. doi:10.1038/nrrheum.2015.91
- Lan, B., Chen, P., Jiri, M., He, N., Feng, T., Liu, K., et al. (2016). WDR1 and CLNK Gene Polymorphisms Correlate with Serum Glucose and High-Density Lipoprotein Levels in Tibetan Gout Patients. *Rheumatol. Int.* 36 (3), 405–412. doi:10.1007/s00296-015-3378-y
- Lanaspa, M. A., Tapia, E., Soto, V., Sautin, Y., and Sánchez-Lozada, L. G. (2011). Uric Acid and Fructose: Potential Biological Mechanisms. *Seminars Nephrol.* 31 (5), 426–432. doi:10.1016/j.semnephrol.2011.08.006
- Lane, J. M., Liang, J., Vlasac, I., Anderson, S. G., Bechtold, D. A., Bowden, J., et al. (2017). Genome-wide Association Analyses of Sleep Disturbance Traits Identify New Loci and Highlight Shared Genetics with Neuropsychiatric and Metabolic Traits. *Nat. Genet.* 49 (2), 274–281. doi:10.1038/ng.3749
- Leask, M., Dowdle, A., Salvesen, H., Topless, R., Fadason, T., Wei, W., et al. (2018). Functional Urate-Associated Genetic Variants Influence Expression of lincRNAs LINC01229 and MAFTRR. *Front. Genet.* 9, 733. doi:10.3389/fgene.2018.00733
- Leask, M. P., and Merriman, T. R. (2021). The Genetic Basis of Urate Control and Gout: Insights into Molecular Pathogenesis from Follow-Up Study of Genome-wide Association Study Loci. *Best Pract. Res. Clin. Rheumatology* 35 (4), 101721. doi:10.1016/j.berh.2021.101721
- Leask, M. P., Sumpter, N. A., Lupi, A. S., Vazquez, A. I., Reynolds, R. J., Mount, D. B., et al. (2020). The Shared Genetic Basis of Hyperuricemia, Gout, and Kidney Function. *Seminars Nephrol.* 40 (6), 586–599. doi:10.1016/j.semnephrol.2020.12.002
- Lee, M.-t. G., Hsu, T.-C., Chen, S.-C., Lee, Y.-C., Kuo, P.-H., Yang, J.-H., et al. (2019). Integrative Genome-wide Association Studies of eQTL and GWAS Data for Gout Disease Susceptibility. *Sci. Rep.* 9 (1), 4981. doi:10.1038/s41598-019-41434-4
- Levine, A. J., and Puzio-Kuter, A. M. (2010). The Control of the Metabolic Switch in Cancers by Oncogenes and Tumor Suppressor Genes. *Science* 330 (6009), 1340–1344. doi:10.1126/science.1193494
- Li, C., Li, Z., Liu, S., Wang, C., Han, L., Cui, L., et al. (2015). Genome-wide Association Analysis Identifies Three New Risk Loci for Gout Arthritis in Han Chinese. *Nat. Commun.* 6, 7041. doi:10.1038/ncomms8041
- Li, R., Miao, L., Qin, L., Xiang, Y., Zhang, X., Peng, H., et al. (2015). A Meta-Analysis of the Associations between the Q141K and Q126X ABCG2 Gene Variants and Gout Risk. *Int. J. Clin. Exp. Pathol.* 8 (9), 9812–9823.
- Liang, M., Woodard, L. E., Liang, A., Luo, J., Wilson, M. H., Mitch, W. E., et al. (2015). Protective Role of Insulin-like Growth Factor-1 Receptor in Endothelial Cells against Unilateral Ureteral Obstruction-Induced Renal Fibrosis. *Am. J. Pathology* 185 (5), 1234–1250. doi:10.1016/j.ajpath.2015.01.027
- Lieber, C. S., Jones, D. P., Losowsky, M. S., and Davidson, C. S. (1962). Interrelation of Uric Acid and Ethanol Metabolism in Man\*. *J. Clin. Invest.* 41 (10), 1863–1870. doi:10.1172/jci104643
- Lin, Y.-H., Chang, H.-M., Chang, F.-P., Shen, C.-R., Liu, C.-L., Mao, W.-Y., et al. (2013). Protoporphyrin IX Accumulation Disrupts Mitochondrial Dynamics and Function in ABCG2-Deficient Hepatocytes. *FEBS Lett.* 587 (19), 3202–3209. doi:10.1016/j.febslet.2013.08.011
- Liu, H. C., Jamshidi, N., Chen, Y., Eraly, S. A., Cho, S. Y., Bhatnagar, V., et al. (2016). An Organic Anion Transporter 1 (OAT1)-Centered Metabolic Network. *J. Biol. Chem.* 291 (37), 19474–19486. doi:10.1074/jbc.M116.745216
- Liu, R., Han, C., Wu, D., Xia, X., Gu, J., Guan, H., et al. (2015). Prevalence of Hyperuricemia and Gout in Mainland China from 2000 to 2014: A Systematic Review and Meta-Analysis. *BioMed Res. Int.* 2015, 1–12. doi:10.1155/2015/762820
- Luciani, A., Villella, V. R., Esposito, S., Brunetti-Pierri, N., Medina, D., Settembre, C., et al. (2010). Defective CFTR Induces Aggresome Formation and Lung Inflammation in Cystic Fibrosis through ROS-Mediated Autophagy Inhibition. *Nat. Cell Biol.* 12 (9), 863–875. doi:10.1038/ncb2090
- Luo, Z., Saha, A. K., Xiang, X., and Ruderman, N. B. (2005). AMPK, the Metabolic Syndrome and Cancer. *Trends Pharmacol. Sci.* 26 (2), 69–76. doi:10.1016/j.tips.2004.12.011
- Maaten, J. C. T., Voorburg, A., Heine, R. J., Wee, P. M. T., Donker, A. J. M., and Gans, R. O. B. (1997). Renal Handling of Urate and Sodium during Acute Physiological Hyperinsulinaemia in Healthy Subjects. *Clin. Sci.* 92 (1), 51–58. doi:10.1042/cs0920051
- Macgregor, S., Lind, P. A., Bucholz, K. K., Hansell, N. K., Madden, P. A. F., Richter, M. M., et al. (2009). Associations of ADH and ALDH2 Gene Variation with Self Report Alcohol Reactions, Consumption and Dependence: an Integrated Analysis. *Hum. Mol. Genet.* 18 (3), 580–593. doi:10.1093/hmg/ddn372
- Marrero, M. B., Baner-Berceli, A. K., Stern, D. M., and Eaton, D. C. (2006). Role of the JAK/STAT Signaling Pathway in Diabetic Nephropathy. *Am. J. Physiology-Renal Physiology* 290 (4), F762–F768. doi:10.1152/ajprenal.00181.2005
- McKinney, C., Stamp, L. K., Dalbeth, N., Topless, R. K., Day, R. O., Kannangara, D. R., et al. (2015). Multiplicative Interaction of Functional Inflammation Genetic Variants in Determining the Risk of Gout. *Arthritis Res. Ther.* 17, 288. doi:10.1186/s13075-015-0802-3

- Mejias, E., Navas, J., Lluberes, R., and Martínez-Maldonado, M. (1989). Hyperuricemia, Gout, and Autosomal Dominant Polycystic Kidney Disease. *Am. J. Med. Sci.* 297 (3), 145–148. doi:10.1097/00000441-198903000-00002
- Miner, J. N., Tan, P. K., Hyndman, D., Liu, S., Iverson, C., Nanavati, P., et al. (2016). Lesinurad, a Novel, Oral Compound for Gout, Acts to Decrease Serum Uric Acid through Inhibition of Urate Transporters in the Kidney. *Arthritis Res. Ther.* 18 (1), 214. doi:10.1186/s13075-016-1107-x
- Nakamura, K., Sakurai, M., Miura, K., Morikawa, Y., Yoshita, K., Ishizaki, M., et al. (2012). Alcohol Intake and the Risk of Hyperuricaemia: A 6-year Prospective Study in Japanese Men. *Nutr. Metabolism Cardiovasc. Dis.* 22 (11), 989–996. doi:10.1016/j.numecd.2011.01.003
- Nakatochi, M., Kanai, M., Nakayama, A., Hishida, A., Kawamura, Y., Ichihara, S., et al. (2019). Genome-wide Meta-Analysis Identifies Multiple Novel Loci Associated with Serum Uric Acid Levels in Japanese Individuals. *Commun. Biol.* 2, 115. doi:10.1038/s42003-019-0339-0
- Nakatochi, M., Kanai, M., Nakayama, A., Hishida, A., Kawamura, Y., Ichihara, S., et al. (2019). Genome-wide Meta-Analysis Identifies Multiple Novel Loci Associated with Serum Uric Acid Levels in Japanese Individuals. *Commun. Biol.* 2, 115. doi:10.1038/s42003-019-0339-0
- Nakayama, A., Matsuo, H., Shimizu, T., Ogata, H., Takada, Y., Nakashima, H., et al. (2013). A Common Missense Variant of Monocarboxylate Transporter 9 (MCT9/SLC16A9) Gene Is Associated with Renal Overload Gout, but Not with All Gout Susceptibility. *Hum. Cell* 26 (4), 133–136. doi:10.1007/s13577-013-0073-8
- Nakayama, A., Matsuo, H., Shimizu, T., Takada, Y., Nakamura, T., Shimizu, S., et al. (2014). Common Variants of a Urate-Associated Gene LRP2 Are Not Associated with Gout Susceptibility. *Rheumatol. Int.* 34 (4), 473–476. doi:10.1007/s00296-013-2924-8
- Nakayama, A., Nakaoka, H., Yamamoto, K., Sakiyama, M., Shaukat, A., Toyoda, Y., et al. (2017). GWAS of Clinically Defined Gout and Subtypes Identifies Multiple Susceptibility Loci that Include Urate Transporter Genes. *Ann. Rheum. Dis.* 76 (5), 869–877. doi:10.1136/annrheumdis-2016-209632
- Namjou, B., Lingren, T., Lingren, T., Huang, Y., Parameswaran, S., Cobb, B. L., et al. (2019). GWAS and Enrichment Analyses of Non-alcoholic Fatty Liver Disease Identify New Trait-Associated Genes and Pathways across eMERGE Network. *BMC Med.* 17 (1), 135. doi:10.1186/s12916-019-1364-z
- Narang, R. K., and Dalbeth, N. (2018). Management of Complex Gout in Clinical Practice: Update on Therapeutic Approaches. *Best Pract. Res. Clin. Rheumatology* 32 (6), 813–834. doi:10.1016/j.berh.2019.03.010
- Neogi, T. (2011). Gout. *N. Engl. J. Med.* 364 (5), 443–452. doi:10.1056/NEJMcpl001124
- Neogi, T. (2010). Interleukin-1 Antagonism in Acute Gout: Is Targeting a Single Cytokine the Answer? *Arthritis & Rheumatism* 62 (10), 2845–2849. doi:10.1002/art.27635
- Nieradko-Iwanicka, B. (2021). The Role of Alcohol Consumption in Pathogenesis of Gout. *Crit. Rev. Food Sci. Nutr.* 1–9. doi:10.1080/10408398.2021.1911928
- Nishioka, K., Sumida, T., Iwatani, M., Kusumoto, A., Ishikura, Y., Hatanaka, H., et al. (2002). Influence of Moderate Drinking on Purine and Carbohydrate Metabolism. *Alcohol Clin. Exp. Res.* 26 (8 Suppl. 1), 20s–25S. doi:10.1097/01.Alc.0000026829.60802.67
- Okada, Y., Sim, X., Sim, X., Go, M. J., Wu, J.-Y., Gu, D., et al. (2012). Meta-analysis Identifies Multiple Loci Associated with Kidney Function-Related Traits in East Asian Populations. *Nat. Genet.* 44 (8), 904–909. doi:10.1038/ng.2352
- Onuora, S. (2020). ABCG2 SNP Associated with Early-Onset Gout. *Nat. Rev. Rheumatol.* 16 (4), 186. doi:10.1038/s41584-020-0393-5
- Padova, J., Onesti, G., Faludi, G., Bendersky, G., and Bendersky, G. (1964). THE EFFECT OF GLUCOSE LOADS ON RENAL URIC ACID EXCRETION IN DIABETIC PATIENTS. *Metabolism* 13, 507–512. doi:10.1016/0026-0495(64)90137-4
- Park, J. S., Kim, Y., and Kang, J. (2021). Genome-wide Meta-Analysis Revealed Several Genetic Loci Associated with Serum Uric Acid Levels in Korean Population: an Analysis of Korea Biobank Data. *J. Hum. Genet.* 67, 231–237. doi:10.1038/s10038-021-00991-1
- Perheentupa, J., and Raivio, K. (1967). Fructose-induced Hyperuricaemia. *Lancet* 290 (7515), 528–531. doi:10.1016/s0140-6736(67)90494-1
- Prestin, K., Wolf, S., Feldtmann, R., Hussner, J., Geissler, I., Rimmbach, C., et al. (2014). Transcriptional Regulation of Urate Transportosome Member SLC2A9 by Nuclear Receptor HNF4a. *Am. J. Physiology-Renal Physiology* 307 (9), F1041–F1051. doi:10.1152/ajprenal.00640.2013
- Puig, J. G., Miranda, M. E., Mateos, F. A., Picazo, M. L., Jiménez, M. L., Calvin, T. S., et al. (1993). Hereditary Nephropathy Associated with Hyperuricemia and Gout. *Arch. Intern. Med.* 153 (3), 357–365. doi:10.1001/archinte.1993.00410030063009
- Qing, Y.-F., Zhou, J.-G., Zhang, Q.-B., Wang, D.-S., Li, M., Yang, Q.-B., et al. (2013). Association of TLR4 Gene Rs2149356 Polymorphism with Primary Gouty Arthritis in a Case-Control Study. *PLoS One* 8 (5), e64845. doi:10.1371/journal.pone.0064845
- Quinones Galvan, A., Natali, A., Baldi, S., Frascerra, S., Sanna, G., Ciociaro, D., et al. (1995). Effect of Insulin on Uric Acid Excretion in Humans. *Am. J. Physiology-Endocrinology Metabolism* 268, E1–E5. doi:10.1152/ajpendo.1995.268.1.E1
- Rasheed, H., McKinney, C., Stamp, L. K., Dalbeth, N., Topless, R. K., Day, R., et al. (2016). The Toll-like Receptor 4 (TLR4) Variant Rs2149356 and Risk of Gout in European and Polynesian Sample Sets. *PLoS One* 11 (1), e0147939. doi:10.1371/journal.pone.0147939
- Rasheed, H., Phipps-Green, A., Topless, R., Hollis-Moffatt, J. E., Hindmarsh, J., Franklin, C., et al. (2013). Association of the Lipoprotein Receptor-Related Protein 2 Gene with Gout and Non-additive Interaction with Alcohol Consumption. *Arthritis Res. Ther.* 15 (6), R177. doi:10.1186/ar4366
- Rasheed, H., Stamp, L. K., Dalbeth, N., and Merriman, T. R. (2017). Interaction of the GCKR and AICF Loci with Alcohol Consumption to Influence the Risk of Gout. *Arthritis Res. Ther.* 19 (1), 161. doi:10.1186/s13075-017-1369-y
- Reginato, A. M., Mount, D. B., Yang, I., and Choi, H. K. (2012). The Genetics of Hyperuricaemia and Gout. *Nat. Rev. Rheumatol.* 8 (10), 610–621. doi:10.1038/nrrheum.2012.144
- Rhyu, J., and Bhat, S. P. (2021). Skeletal Complications with GNAS Mutation: An Unusual Case with Osteoma Cutis, Gout, and Synovial Chondromatosis in a Patient with Pseudopseudohypoparathyroidism. *AACE Clin. Case Rep.* 7 (3), 180–183. doi:10.1016/j.aace.2020.11.036
- Roberts, R. L., Wallace, M. C., Phipps-Green, A. J., Topless, R., Drake, J. M., Tan, P., et al. (2017). ABCG2 Loss-Of-Function Polymorphism Predicts Poor Response to Allopurinol in Patients with Gout. *Pharmacogenomics J.* 17 (2), 201–203. doi:10.1038/tpj.2015.101
- Sakiyama, M., Matsuo, H., Nakaoka, H., Kawamura, Y., Kawaguchi, M., Higashino, T., et al. (2018). Common Variant of BCAS3 Is Associated with Gout Risk in Japanese Population: the First Replication Study after Gout GWAS in Han Chinese. *BMC Med. Genet.* 19 (1), 96. doi:10.1186/s12881-018-0583-z
- Sandoval-Plata, G., Morgan, K., and Abhishek, A. (2021). Variants in Urate Transporters, ADH1B, GCKR and MEPE Genes Associate with Transition from Asymptomatic Hyperuricaemia to Gout: Results of the First Gout versus Asymptomatic Hyperuricaemia GWAS in Caucasians Using Data from the UK Biobank. *Ann. Rheum. Dis.* 80 (9), 1220–1226. doi:10.1136/annrheumdis-2020-219796
- Schauer, C., Janko, C., Munoz, L. E., Zhao, Y., Kienhöfer, D., Frey, B., et al. (2014). Aggregated Neutrophil Extracellular Traps Limit Inflammation by Degrading Cytokines and Chemokines. *Nat. Med.* 20 (5), 511–517. doi:10.1038/nm.3547
- Schmidt, M. I., Watson, R. L., Duncan, B. B., Metcalf, P., Brancati, F. L., Richey Sharrett, A., et al. (1996). Clustering of Dyslipidemia, Hyperuricemia, Diabetes, and Hypertension and its Association with Fasting Insulin and Central and Overall Obesity in a General Population. *Metabolism* 45 (6), 699–706. doi:10.1016/s0026-0495(96)90134-1
- Shi, C.-S., Shenderov, K., Huang, N.-N., Kabat, J., Abu-Asab, M., Fitzgerald, K. A., et al. (2012). Activation of Autophagy by Inflammatory Signals Limits IL-1 $\beta$  Production by Targeting Ubiquitinated Inflammasomes for Destruction. *Nat. Immunol.* 13 (3), 255–263. doi:10.1038/ni.2215
- Siniachenko, O. V., Diadyk, A. I., Nikolenko, Iu. I., and Khomenko, M. V. (1984). Circulating Immune Complexes in Gout. *Vrach Delo* (9), 61–63.
- Smith, E., Hoy, D., Cross, M., Merriman, T. R., Vos, T., Buchbinder, R., et al. (2014). The Global Burden of Gout: Estimates from the Global Burden of Disease 2010 Study. *Ann. Rheum. Dis.* 73 (8), 1470–1476. doi:10.1136/annrheumdis-2013-204647
- So, A., De Meulemeester, M., Pikhak, A., Yücel, A. E., Richard, D., Murphy, V., et al. (2010). Canakinumab for the Treatment of Acute Flares in Difficult-To-Treat Gouty Arthritis: Results of a Multicenter, Phase II, Dose-Ranging Study. *Arthritis & Rheumatism* 62 (10), 3064–3076. doi:10.1002/art.27600

- So, A., De Smedt, T., Revaz, S., and Tschopp, J. (2007). A Pilot Study of IL-1 Inhibition by Anakinra in Acute Gout. *Arthritis Res. Ther.* 9 (2), R28. doi:10.1186/ar2143
- Spadaro, O., Goldberg, E. L., Camell, C. D., Youm, Y.-H., Kopchick, J. J., Nguyen, K. Y., et al. (2016). Growth Hormone Receptor Deficiency Protects against Age-Related NLRP3 Inflammasome Activation and Immune Senescence. *Cell Rep.* 14 (7), 1571–1580. doi:10.1016/j.celrep.2016.01.044
- Sun, H., Tian, J., Xian, W., Xie, T., and Yang, X. (2015). miR-34a Inhibits Proliferation and Invasion of Bladder Cancer Cells by Targeting Orphan Nuclear Receptor HNF4G. *Dis. markers* 2015, 1–8. doi:10.1155/2015/879254
- Tafaj, O., and Jüppner, H. (2017). Pseudohypoparathyroidism: One Gene, Several Syndromes. *J. Endocrinol. Invest* 40 (4), 347–356. doi:10.1007/s40618-016-0588-4
- Tanner, C., Boocock, J., Stahl, E. A., Dobbryn, A., Mandal, A. K., Cadzow, M., et al. (2017). Population-Specific Resequencing Associates the ATP-Binding Cassette Subfamily C Member 4 Gene with Gout in New Zealand Māori and Pacific Men. *Arthritis & Rheumatology* 69 (7), 1461–1469. doi:10.1002/art.40110
- Terkeltaub, R. A. (2003). Gout. *N. Engl. J. Med.* 349 (17), 1647–1655. doi:10.1056/NEJMcp030733
- Terkeltaub, R., Sundy, J. S., Schumacher, H. R., Murphy, F., Bookbinder, S., Biedermann, S., et al. (2009). The Interleukin 1 Inhibitor Rilonacept in Treatment of Chronic Gouty Arthritis: Results of a Placebo-Controlled, Monosequence Crossover, Non-randomised, Single-Blind Pilot Study. *Ann. Rheumatic Dis.* 68 (10), 1613–1617. doi:10.1136/ard.2009.108936
- Tin, A., Marten, J., Halperin Kuhns, V. L., Li, Y., Wuttke, M., Kirsten, H., et al. (2019). Target Genes, Variants, Tissues and Transcriptional Pathways Influencing Human Serum Urate Levels. *Nat. Genet.* 51 (10), 1459–1474. doi:10.1038/s41588-019-0504-x
- Tin, A., Marten, J., Halperin Kuhns, V. L., Li, Y., Wuttke, M., Kirsten, H., et al. (2019). Target Genes, Variants, Tissues and Transcriptional Pathways Influencing Human Serum Urate Levels. *Nat. Genet.* 51 (10), 1459–1474. doi:10.1038/s41588-019-0504-x
- Tin, A., Li, Y., Brody, J. A., Nutile, T., Chu, A. Y., Huffman, J. E., et al. (2018). Large-scale Whole-Exome Sequencing Association Studies Identify Rare Functional Variants Influencing Serum Urate Levels. *Nat. Commun.* 9 (1), 4228. doi:10.1038/s41467-018-06620-4
- Tin, A., Woodward, O. M., Kao, W. H. L., Liu, C.-T., Lu, X., Nalls, M. A., et al. (2011). Genome-wide Association Study for Serum Urate Concentrations and Gout Among African Americans Identifies Genomic Risk Loci and a Novel URAT1 Loss-Of-Function Allele. *Hum. Mol. Genet.* 20 (20), 4056–4068. doi:10.1093/hmg/ddr307
- Tong, X., Zhao, F., Mancuso, A., Gruber, J. J., and Thompson, C. B. (2009). The Glucose-Responsive Transcription Factor ChREBP Contributes to Glucose-dependent Anabolic Synthesis and Cell Proliferation. *Proc. Natl. Acad. Sci. U.S.A.* 106 (51), 21660–21665. doi:10.1073/pnas.0911316106
- Tsuchiya, M., Misaka, R., Nitta, K., and Tsuchiya, K. (2015). Transcriptional Factors, Mafs and Their Biological Roles. *Wjd* 6 (1), 175–183. doi:10.4239/wjd.v6.i1.175
- Tu, H.-P., Chen, C.-J., Tovosia, S., Ko, A. M.-S., Lee, C.-H., Ou, T.-T., et al. (2010). Associations of a Non-synonymous Variant in SLC2A9 with Gouty Arthritis and Uric Acid Levels in Han Chinese Subjects and Solomon Islanders. *Ann. Rheumatic Dis.* 69 (5), 887–890. doi:10.1136/ard.2009.113357
- Tu, H.-P., Ko, A. M.-S., Chiang, S.-L., Lee, S.-S., Lai, H.-M., Chung, C.-M., et al. (2014). Joint Effects of Alcohol Consumption and ABCG2 Q141K on Chronic Tophaceous Gout Risk. *J. Rheumatol.* 41 (4), 749–758. doi:10.3899/jrheum.130870
- Tu, H.-P., Min-Shan Ko, A., Lee, S.-S., Lee, C.-P., Kuo, T.-M., Huang, C.-M., et al. (2018). Variants of ALPK1 with ABCG2, SLC2A9, and SLC22A12 Increased the Positive Predictive Value for Gout. *J. Hum. Genet.* 63 (1), 63–70. doi:10.1038/s10038-017-0368-9
- Turan, S., and Bastepe, M. (2015). GNAS Spectrum of Disorders. *Curr. Osteoporos. Rep.* 13 (3), 146–158. doi:10.1007/s11914-015-0268-x
- Urano, W., Taniguchi, A., Anzai, N., Inoue, E., Sekita, C., Endou, H., et al. (2010). Association between GLUT9 and Gout in Japanese Men. *Ann. Rheumatic Dis.* 69 (5), 932–933. doi:10.1136/ard.2009.111096
- Urano, W., Taniguchi, A., Inoue, E., Sekita, C., Ichikawa, N., Koseki, Y., et al. (2013). Effect of Genetic Polymorphisms on Development of Gout. *J. Rheumatol.* 40 (8), 1374–1378. doi:10.3899/jrheum.121244
- van der Gaag, M., van den Berg, R., van den Berg, H., Schaafsma, G., and Hendriks, H. (2000). Moderate Consumption of Beer, Red Wine and Spirits Has Counteracting Effects on Plasma Antioxidants in Middle-Aged Men. *Eur. J. Clin. Nutr.* 54 (7), 586–591. doi:10.1038/sj.ejcn.1601061
- Wallace, M. C., Roberts, R. L., Nanavati, P., Miner, J. N., Dalbeth, N., Topless, R., et al. (2018). Association between ABCG2 Rs2231142 and Poor Response to Allopurinol: Replication and Meta-Analysis. *Rheumatol. Oxf.* 57 (4), 656–660. doi:10.1093/rheumatology/kex467
- Wang, J., Liu, S., Wang, B., Miao, Z., Han, L., Chu, N., et al. (2012). Association between Gout and Polymorphisms in GCKR in Male Han Chinese. *Hum. Genet.* 131 (7), 1261–1265. doi:10.1007/s00439-012-1151-9
- Wang, L., Ma, Q., Yao, H., He, L. J., Fang, B. B., Cai, W., et al. (2018). Association of GCKR Rs780094 Polymorphism with Circulating Lipid Levels in Type 2 Diabetes and Hyperuricemia in Uyghur Chinese. *Int. J. Clin. Exp. Pathol.* 11 (9), 4684–4694.
- Wang, X., Shaw, S., Amiri, F., Eaton, D. C., and Marrero, M. B. (2002). Inhibition of the JAK/STAT Signaling Pathway Prevents the High Glucose-Induced Increase in TGF- $\beta$  and Fibronectin Synthesis in Mesangial Cells. *Diabetes* 51 (12), 3505–3509. doi:10.2337/diabetes.51.12.3505
- Wen, C., Yee, S., Liang, X., Hoffmann, T., Kvale, M., Banda, Y., et al. (2015). Genome-wide Association Study Identifies ABCG2 (BCRP) as an Allopurinol Transporter and a Determinant of Drug Response. *Clin. Pharmacol. Ther.* 97 (5), 518–525. doi:10.1002/cpt.89
- Wisely, G. B., Miller, A. B., Davis, R. G., Thornquest, A. D., Jr., Johnson, R., Spitzer, T., et al. (2002). Hepatocyte Nuclear Factor 4 Is a Transcription Factor that Constitutively Binds Fatty Acids. *Structure* 10 (9), 1225–1234. doi:10.1016/s0969-2126(02)00829-8
- Wong, K., Briddon, S. J., Holliday, N. D., and Kerr, I. D. (2016). Plasma Membrane Dynamics and Tetrameric Organisation of ABCG2 Transporters in Mammalian Cells Revealed by Single Particle Imaging Techniques. *Biochimica Biophysica Acta (BBA) - Mol. Cell Res.* 1863 (1), 19–29. doi:10.1016/j.bbamer.2015.10.002
- Wu, J. N., and Koretzky, G. A. (2004). The SLP-76 Family of Adapter Proteins. *Winnars Immunol.* 16 (6), 379–393. doi:10.1016/j.smim.2004.08.018
- Yang, Q., Köttgen, A., Dehghan, A., Smith, A. V., Glazer, N. L., Chen, M.-H., et al. (2010). Multiple Genetic Loci Influence Serum Urate Levels and Their Relationship with Gout and Cardiovascular Disease Risk Factors. *Circ. Cardiovasc. Genet.* 3 (6), 523–530. doi:10.1161/circgenetics.109.934455
- Yang, Y., and Cvekl, A. (2007). Large Maf Transcription Factors: Cousins of AP-1 Proteins and Important Regulators of Cellular Differentiation. *Ejbm* 23 (1), 2–11. doi:10.23861/ejbm20072347
- Yeung, B. H. Y., Law, A. Y. S., and Wong, C. K. C. (2012). Evolution and Roles of Stanniocalcin. *Mol. Cell. Endocrinol.* 349 (2), 272–280. doi:10.1016/j.mce.2011.11.007
- Yi, X. L., Li, J., Meng, D. M., Liu, Y. J., Liu, Y. H., Ma, H. M., et al. (2018). An Intron Variant of SLC2A9 Increases the Risk for Type 2 Diabetes Mellitus Complicated with Hyperuricemia in Chinese Male Population. *Iran. J. Public Health* 47 (6), 844–851.
- Yoon, J. C., Puigserver, P., Chen, G., Donovan, J., Wu, Z., Rhee, J., et al. (2001). Control of Hepatic Gluconeogenesis through the Transcriptional Coactivator PGC-1. *Nature* 413 (6852), 131–138. doi:10.1038/35093050
- Yu, J., Riou, C., Davidson, D., Minhas, R., Robson, J. D., Julius, M., et al. (2001). Synergistic Regulation of Immunoreceptor Signaling by SLP-76-Related Adaptor Clnk and Serine/threonine Protein Kinase HPK-1. *Mol. Cell Biol.* 21 (18), 6102–6112. doi:10.1128/mcb.21.18.6102-6112.2001
- Yu, K.-H., See, L.-C., Huang, Y.-C., Yang, C.-H., and Sun, J.-H. (2008). Dietary Factors Associated with Hyperuricemia in Adults. *Seminars Arthritis Rheumatism* 37 (4), 243–250. doi:10.1016/j.semarthrit.2007.04.007
- Zhang, Y., Cai, W., Song, J., Miao, L., Zhang, B., Xu, Q., et al. (2014). Association between the PNPLA3 I148M Polymorphism and Non-alcoholic Fatty Liver Disease in the Uyghur and Han Ethnic Groups of Northwestern China. *PLoS One* 9 (10), e108381. doi:10.1371/journal.pone.0108381
- Zhao, J., Hu, Y., and Peng, J. (2021). Targeting Programmed Cell Death in Metabolic Dysfunction-Associated Fatty Liver Disease (MAFLD): a Promising New Therapy. *Cell Mol. Biol. Lett.* 26 (1), 17. doi:10.1186/s11658-021-00254-z

- Zhao, J., Jiang, P., Guo, S., Schrodi, S. J., and He, D. (2021). Apoptosis, Autophagy, NETosis, Necroptosis, and Pyroptosis Mediated Programmed Cell Death as Targets for Innovative Therapy in Rheumatoid Arthritis. *Front. Immunol.* 12, 809806. doi:10.3389/fimmu.2021.809806
- Zhou, W., Kanai, M., Wu, K-H. H., Humaira, R., Tsuo, K., Hirbo, J. B., et al. (2021). Global Biobank Meta-Analysis Initiative: Powering Genetic Discovery across Human Diseases. *medRxiv*. 2021.11.19.21266436. doi:10.1101/2021.11.19.21266436

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

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

Copyright © 2022 Zhao, Guo, Schrodi and He. 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.

## GLOSSARY

|                                |  |                 |  |
|--------------------------------|--|-----------------|--|
| <b>GWAS</b>                    | genome-wide association studies            | <b>HNF4G</b>    | hepatocyte nuclear factor 4 gamma  |
| <b>GBMI</b>                    | Global Biobank Meta-analysis Initiative    | <b>PKD2</b>     | polycystin 2   |
| <b>URAT1</b>                   | urate transporter-1                        | <b>ADH1B</b>    | alcohol dehydrogenase 1B (Class I), beta polypeptide                                       |
| <b>GLUT9</b>                   | glucose transporter 9                      | <b>MCT9</b>     | monocarboxylate transporter 9  |
| <b>OAT4</b>                    | organic anion transporter 4                | <b>FAM35A</b>   | shieldin complex subunit 2   |
| <b>OAT10</b>                   | organic anion transporter 4                | <b>AMPD2</b>    | adenosine monophosphate deaminase 2  |
| <b>ABCG2</b>                   | TP-binding cassette superfamily G member 2 | <b>MLXIP</b>    | MLX interacting protein  |
| <b>ABCC4</b>                   | ATP Binding Cassette Subfamily C Member 4  | <b>MLXIPL</b>   | MLX interacting protein-like   |
| <b>ATP</b>                     | adenosine triphosphate                     | <b>PNPLA3</b>   | patatin-like phospholipase domain containing 3   |
| <b>TLR</b>                     | Toll-like receptor                         | <b>MAFLD</b>    | metabolism-related fatty liver disease   |
| <b>NLRP3</b>                   | nod-like receptor pyrin domain 3           | <b>IGF1R</b>    | insulin like growth factor 1 receptor  |
| <b>NF-<math>\kappa</math>B</b> | nuclear factor- $\kappa$ B                 | <b>BMI</b>      | body mass index  |
| <b>IL</b>                      | interleukin                                | <b>A1CF</b>     | APOBEC1 complementation factor   |
| <b>NSAIDs</b>                  | non-steroidal anti-inflammation drugs      | <b>ApoB</b>     | apolipoprotein B   |
| <b>SLC2A9</b>                  | solute carrier family 2 member 9           | <b>LRP2</b>     | lipoprotein receptor-related protein 2   |
| <b>SLC22A11</b>                | solute carrier family 22 member 11         | <b>STC1</b>     | stanniocalcin 1  |
| <b>SLC17A1</b>                 | solute carrier family 17 member 1          | <b>CYP1A2</b>   | cytochrome P450 family 1 subfamily A member 2  |
| <b>SLC22A12</b>                | solute carrier family 22 member 12         | <b>HCRT2</b>    | hypocretin receptor 2  |
| <b>MAF</b>                     | MAF BZIP transcription factor              | <b>CLNK</b>     | cytokine-dependent hematopoietic cell linker   |
| <b>SLC2A12</b>                 | solute carrier family 2 member 12          | <b>GNAS</b>     | guanine nucleotide-binding protein a-stimulating polypeptide                               |
| <b>SLC16A9</b>                 | solute carrier family 16 member 9          | <b>SCL22A6</b>  | solute carrier family 22 member 6  |
| <b>GCKR</b>                    | glucokinase (hexokinase 4) regulator       | <b>BCAS3</b>    | breast cancer-amplified sequence 3   |
| <b>PDZK1</b>                   | PDZ domain-containing 1                    | <b>cAMP</b>     | cyclic AMP   |
| <b>HNF4A</b>                   | hepatocyte nuclear factor 4 alpha          | <b>OAT1</b>     | organic anion transporter 1  |
| <b>SNP</b>                     | single-nucleotide polymorphism             | <b>MEPE</b>     | matrix extracellular phosphoglycoprotein   |
| <b>C-MAF</b>                   | C-MAF BZIP transcription factor            | <b>PPM1K-DT</b> | protein phosphatase, Mg <sup>2+</sup> /Mn <sup>2+</sup> dependent 1K- divergent transcript |
| <b>cis-eQTL</b>                | cis-expression quantitative trait loci     | <b>ADH1B</b>    | alcohol dehydrogenase 1B (Class I), beta polypeptide                                       |
| <b>MAFTRR</b>                  | MAF transcriptional regulator RNA          |                 |  |



# Emerging Roles of Non-proteolytic Ubiquitination in Tumorigenesis

Xiu Yin<sup>1</sup>, Qingbin Liu<sup>1</sup>, Fen Liu<sup>1</sup>, Xinchun Tian<sup>1,2</sup>, Tinghao Yan<sup>1,2</sup>, Jie Han<sup>3\*</sup> and Shulong Jiang<sup>1\*</sup>

<sup>1</sup>Clinical Medical Laboratory Center, Jining First People's Hospital, Jining Medical University, Jining, China, <sup>2</sup>Cheeloo College of Medicine, Shandong University, Jinan, China, <sup>3</sup>Department of Thyroid and Breast Surgery, Jining First People's Hospital, Jining Medical University, Jining, China

## OPEN ACCESS

### Edited by:

Haitao Wang,  
Center for Cancer Research (NIH),  
United States

### Reviewed by:

Jianping Guo,  
Sun Yat-sen University, China  
Petro Starokadomskyy,  
University of Texas Southwestern  
Medical Center, United States  
Jianlin Liu,  
Dana-Farber Cancer Institute,  
United States  
Shivani Dixit,  
National Cancer Institute (NIH),  
United States

### \*Correspondence:

Jie Han  
hanj3669@163.com  
Shulong Jiang  
jnslijiang@163.com

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

**Received:** 15 May 2022

**Accepted:** 15 June 2022

**Published:** 06 July 2022

### Citation:

Yin X, Liu Q, Liu F, Tian X, Yan T, Han J  
and Jiang S (2022) Emerging Roles of  
Non-proteolytic Ubiquitination  
in Tumorigenesis.  
Front. Cell Dev. Biol. 10:944460.  
doi: 10.3389/fcell.2022.944460

Ubiquitination is a critical type of protein post-translational modification playing an essential role in many cellular processes. To date, more than eight types of ubiquitination exist, all of which are involved in distinct cellular processes based on their structural differences. Studies have indicated that activation of the ubiquitination pathway is tightly connected with inflammation-related diseases as well as cancer, especially in the non-proteolytic canonical pathway, highlighting the vital roles of ubiquitination in metabolic programming. Studies relating degradable ubiquitination through lys48 or lys11-linked pathways to cellular signaling have been well-characterized. However, emerging evidence shows that non-degradable ubiquitination (linked to lys6, lys27, lys29, lys33, lys63, and Met1) remains to be defined. In this review, we summarize the non-proteolytic ubiquitination involved in tumorigenesis and related signaling pathways, with the aim of providing a reference for future exploration of ubiquitination and the potential targets for cancer therapies.

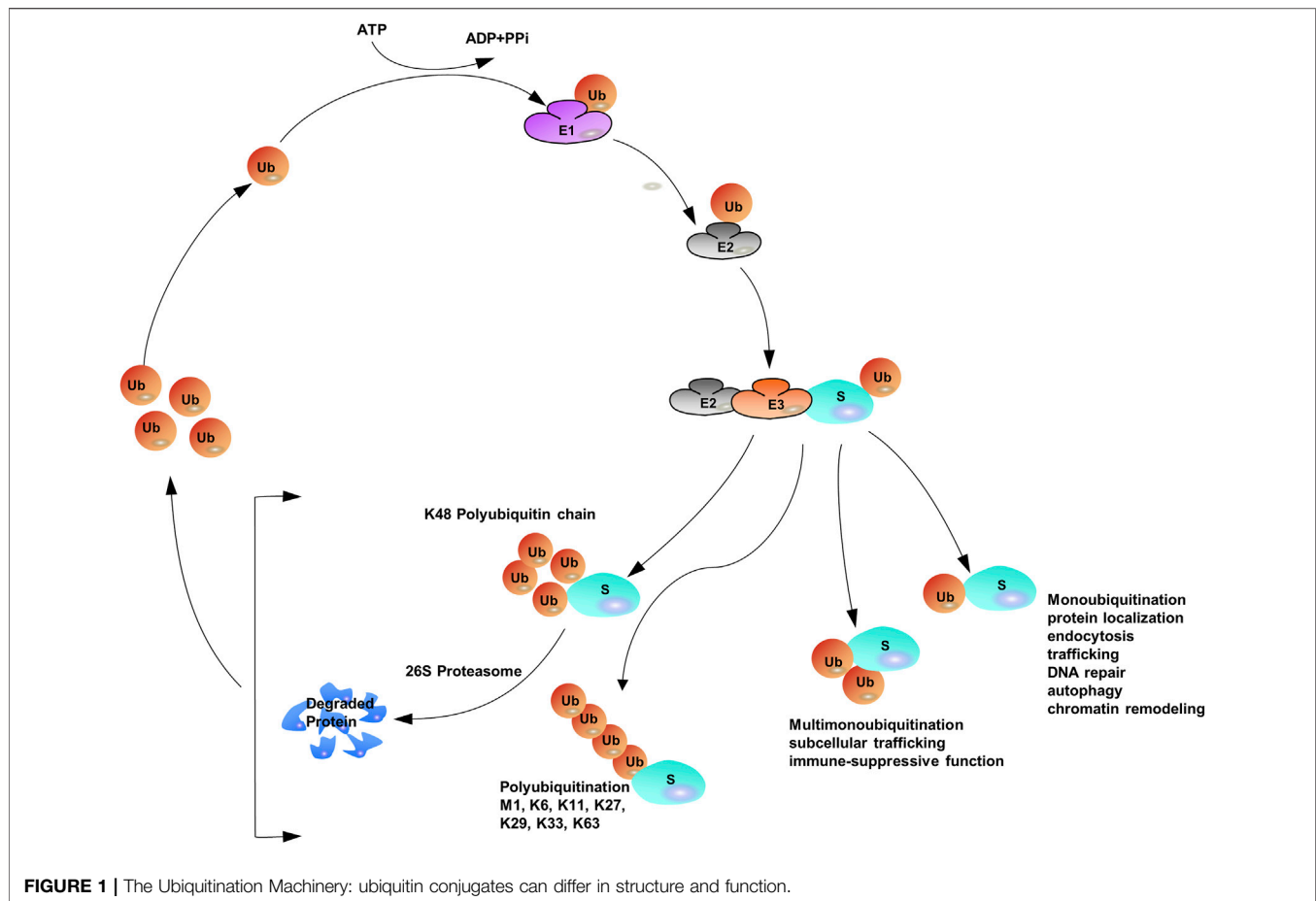
**Keywords:** ubiquitin, atypical ubiquitination, ubiquitin E2 conjugating enzyme, ubiquitin E3 ligase, tumorigenesis, ubiquitin-proteasome system

## 1 INTRODUCTION

### 1.1 The Ubiquitin-Proteasome System

Ubiquitination, also known as ubiquitylation, refers to the process by which ubiquitin (Ub, a small and highly conserved protein), with the help of a series of special enzymes, classifies proteins in cells, selects target proteins, and modifies those proteins (McDowell and Philpott, 2016; Seeler and Dejean, 2017; Rape, 2018). Ubiquitination plays fundamental roles in many cellular events such as cell proliferation (Kwon and Ciechanover, 2017; Werner et al., 2017; Senft et al., 2018; Song and Luo, 2019), cell cycle (Teixeira and Reed, 2013; Darling et al., 2017; Gilberto and Peter, 2017), DNA repair (Alpi and Patel, 2009; Abbas and Dutta, 2011), immune response (Heaton et al., 2016; Rudnicka and Yamauchi, 2016; Manthiram et al., 2017), transcription (Zhou et al., 2018; Imam et al., 2019; Sun et al., 2019), angiogenesis (Zhang et al., 2022), metastasis (Rossi and Rossi, 2022), and apoptosis (Xu et al., 2017; Zhou et al., 2017).

Protein ubiquitination requires three different enzymes: E1 ubiquitin-activating enzymes, E2 ubiquitin-conjugating enzymes, and E3 ubiquitin ligases (Heap et al., 2017; O'Connor and Huibregtse, 2017). E1 enzymes activate the ubiquitin polypeptide in an ATP-dependent manner, and the activated forms are then conjugated to E2 enzymes through the formation of thioester bonds (Wang et al., 2018). Finally, E3 ubiquitin ligases recognize both E2 enzymes and specific target substrates that confer specificity to the system, and then the Ub can be transferred from E2



**FIGURE 1 |** The Ubiquitination Machinery: ubiquitin conjugates can differ in structure and function.

enzymes to the target substrates to complete the ubiquitination process (Rittinger and Ikeda, 2017; Yau et al., 2017). The coordination of E1, E2, and E3 enzymes earmarks target proteins with a wide variety of ubiquitin modifications such that distinct ubiquitin modifications transmit different cellular signals (Khaminets et al., 2016; Venuto and Merla, 2019). Specific ubiquitin-binding domains (UBDs) can identify ubiquitylated substrate proteins (Kristariyanto et al., 2015; Kung et al., 2019; Polykratis et al., 2019; Spiliotopoulos et al., 2019). UBDs utilize diverse mechanisms to interact with various surface patches on ubiquitin molecules or different ubiquitin linkages (Hicke et al., 2005; Dikic et al., 2009; Komander and Rape, 2012). The hydrophobic surfaces of ubiquitin molecules, such as isoleucine 36 (Ile36) and isoleucine 44 (Ile44), are the structural basis for the recognition of ubiquitination signals. Different ubiquitin chains have various spatial structures and thus expose different hydrophobic surfaces, which can be recognized by specific UBDs. Proteins containing UBDs recognize and transmit the functional signals represented by ubiquitin chains (Buetow and Huang, 2016). Similar to many other protein post-translational modifications (PTMs), ubiquitination can be preserved through cleavage by deubiquitylating enzymes (DUBs) (Herhaus and Sapkota, 2014; Nishi et al., 2014; Clague et al., 2019) (**Figure 1**).

The versatility of ubiquitination is determined by the complex assembly pattern of ubiquitin molecules on the target protein. Ubiquitins can be attached to one or multiple lysine residues with either a single ubiquitin molecule (mono- and multi-mono-ubiquitination, respectively) or ubiquitin polymers (poly-ubiquitination) (Eger et al., 2010; Akutsu et al., 2016; Chitale and Richly, 2017; van der Heden van Noort et al., 2017). Poly-ubiquitin chains comprising only one single linkage are often assumed to be homotypic (Jeusset and McManus, 2017; Leestemaker and Ovaa, 2017), whereas heterotypic chains adopt multiple linkages within the same polymer (branched or non-branched) (Akturk et al., 2018). In a poly-ubiquitin chain, ubiquitin moieties can be linked through any of the seven lysine residues (K6, K11, K27, K29, K33, K48, and K63) or N-terminal methionine (Met1) (Dittmar and Winklhofer, 2019; Spit et al., 2019; Liao et al., 2022), resulting in an almost unlimited number of poly-ubiquitin chain topologies (Varadan et al., 2004; Sadowski and Sarcevic, 2010; Dondelinger et al., 2016; O'Connor and Huibregtse, 2017; Padala et al., 2017). Further complexity is added to ubiquitination when the ubiquitin polypeptide is modified by phosphorylation (Dong et al., 2017) or acetylation (Choudhary et al., 2009; Ohtake et al., 2015). Given the sophisticated assembly of protein ubiquitination, it has been often referred to as the “ubiquitin code” (Komander and Rape, 2012). Proteomics studies have shown that Lys48-linked chains

are predominant in cells (>50% of all linkages), and Lys63-linked chains rank the second abundant chain form, however, researchers have begun to characterize the remain chain types, which were considered to be “atypical” ubiquitin modifications (linked through Lys6, Lys11, Lys27, Lys29, Lys33 and Met1) (Xu et al., 2009; Dammer et al., 2011; Kim et al., 2011; Wagner et al., 2011; Ziv et al., 2011).

## 1.2 Structural Features of Poly-Ubiquitin Chains

Poly-ubiquitin chains occur when a single ubiquitin molecule is repeatedly connected in series with another ubiquitin lysine residue. Substrate proteins can be distinguished by poly-ubiquitin chains by attaching between different types of deubiquitination formations, single mono-ubiquitination events, multiple mono-ubiquitinations events, homotypic ubiquitination events and heterotypic ubiquitination events (branched and non-branched ubiquitination) (Sokratous et al., 2014; Yau and Rape, 2016; Zhao et al., 2017). This also led to the formation of eight different homotypic chains. The key distinguishing feature of how this can be achieved is the specific combination of E2/E3 enzymes, thereby triggering distinct cellular fates of substrate proteins. However, some of the reported E2/E3 enzyme combinations were not 100% specific for targeted linkage. It has been reported that two E1 enzymes were selected for ubiquitin in humans: UBA1 and UBA6 (Barghout and Schimmer, 2021). Humans also encoded 40 E2 conjugation enzymes cooperate with approximately 600 E3 ligase enzymes (Hodson et al., 2014). The E3s were categorized into three groups: RING/U-box, HECT, and RING between RING (RBR) (Wang et al., 2020).

With the assistance of E1, E2, and E3, mono-ubiquitination occurs when a single ubiquitin is attached to its target proteins, then Ub molecules are added to the model in linear ways one by one (Torres et al., 2009; Tang et al., 2011). The sequential addition model of Ub on the substrate contributes to the elongation of the Ub lines. When secondary Ub molecules are connected to specific lysine residues, they are called homotypic chains. If any of the attached adjacent Ub molecules are linked to each other by different lysine residues (mixed or branched model), a heteropic structure is formed. For homotypic chains, reports have found that different chain types are closely related to the confirmation of the structure, either “compact” or “open”. Generally, non-proteolytic ubiquitination Lys63 linkages and Met1 linkages, adopt “open” ones. Contrast to the aforementioned linkages, internal structural molecules interact with each other among degradable linkages, like Lys6, Lys11 and Lys48 linkages, and those linkages display “open” conformation. Furthermore, all ubiquitin moieties can be modified by acetylation or phosphorylation to add additional layers of complexity (Kane et al., 2014; Kazlauskaitė et al., 2014; Koyano et al., 2014; Ordureau et al., 2014; Ohtake et al., 2015; Swaney et al., 2015) (Figure 1).

Numerous studies have found that ubiquitin acetylation inhibits poly-ubiquitination elongation, and phosphoubiquitin leads to mitophagy. Any PTM chain can be changed to ubiquitin

chains, which may prevent or facilitate ubiquitin interactions. Protein phosphorylation is linked to ubiquitination for proteasomal degradation. Reports have shown that the phosphorylation of ULK1 by MAPK1/3 kinase interacts with BTRC, which leads to subsequent proteasomal degradation and attenuates breast cancer bone metastasis (Deng et al., 2020). However, the stability of some proteins is also regulated by phosphorylation. Reports have also shown that Aurora B-mediated phosphorylation of ubiquitin specific protease 13 (USP13) at Serine 114 promoted the stability of Aurora B.

## 1.3 Encoding and Decoding the Ubiquitin Code

Encoding and decoding the ubiquitin code is performed by factors that recognize Ub chains and connect the substrate proteins to the downstream response (Ji and Kwon, 2017; Kwon and Ciechanover, 2017). Recognition of chains occurs through discrete domains and affinity binders with specificity for a particular Ub substrate and chain type (Fu et al., 2012; Suryadinata et al., 2014; Kniss et al., 2018; Michel et al., 2018). This complex system consists of the conjugation of diverse mono, multi-mono and polymeric chains (Rösner et al., 2015; Ji and Kwon, 2017). The interpretation of how, when, and why the ubiquitin codes are written, read, and erased emerged to be characterized. Ubiquitination is a powerful decoration process of proteins and is typically actualized by “ubiquitinase” (Zientara-Rytter and Subramani, 2019). Ubiquitin can be successfully linked to one of the seven lysine residues, all of which can be characterized as poly-ubiquitin chains (Laplanche et al., 2009; Regev et al., 2015; Bax et al., 2019). Poly-ubiquitin chains with different topologies depended upon the lysine residues (which were chosen to be attached) and the substantial chain length, determine the lucky chance of the target proteins and regulate diverse cellular processes, known as “ubiquitin code” (Dittmar and Selbach, 2017; Chatr-Aryamontri et al., 2018; Fottner et al., 2019; O'Donnell, 2019; Song and Luo, 2019). Recent discoveries have deepened our understanding of the whole picture of the ubiquitin code, the interplay between “writers” (E1/E2/E3s), “erasers” (DUBs) and “readers” (ubiquitin binding domain containing proteins) (Tanno and Komada, 2013; Heride et al., 2014; Rogerson et al., 2015; Di Lello and Hymowitz, 2016; Smeenk and Mailand, 2016). Studies have highlighted that mono-ubiquitination can be catalyzed by different E2 and E3 enzymes, acting either individually or together to determine specific substrates. Notably, the linkage specificity of E3s containing RING or U-box domains is likely dictated by E2. As for the HECT E3s class of enzymes, HECT domain swaps can activate the acceptor lysine and are sufficient to determine the linkage specificity. RBR E3s, however, are somewhat complex. RBR E3s display linkage specificity in multiple chains, Met1-, Lys63-, Lys48-, and Lys27-linked chains, as well as mono-ubiquitination, while cooperating with E2 to synthesize Lys-linked or Met1-linked chains. The Ub tag attached to a certain substrate, which is preferably achieved through the cooperation of specific E2/E3 enzyme pairs, represents a complex yet specific message encoded by the cell (Kim and Huibregtse, 2009; David

et al., 2011; Turek et al., 2018). Interestingly, this is achieved by Ub receptors which are equipped with one or more UBDs (Mattern et al., 2019). Ubiquitin recognition achieved by UBDs can translate written code into specific outcomes. There are more than 20 families of UBDs that bind to different patches on Ub surrounding hydrophobic patches, either Ile44 or Ile36 (Dikic et al., 2009; Hendriks et al., 2018). Moreover, UBDs are able to sense unique 3D conformations of distinct chain types, which are associated with diverse biological activities, meaning that this small motif-containing protein can help recognize versatile signals and affect the desired effect (Lee et al., 2006; Penengo et al., 2006; Reyes-Turcu et al., 2006; Bomar et al., 2010). In addition, ubiquitin-interacting proteins which served as “decoders” of the ubiquitin message, participated in the downstream regulation of the ubiquitinated substrate (Heideker and Wertz, 2015; Leznicki and Kulathu, 2017; Mevissen and Komander, 2017). These “decoders” may specifically reverse the ubiquitination process or function as receptors for the transfer of the targeted substrate toward downstream signaling components and/or subcellular compartments (Hu et al., 2002; Lin et al., 2008; Komander et al., 2009). To date, 55 ubiquitin specific proteases (USPs), 14 ovarian tumor DUBs (OTUs), 10 JAMM family DUBs, 4 ubiquitin C-terminal hydrolases (UCHs) and 4 Josephin domain DUBs have been identified. Encoding and decoding ubiquitin codes are responsible for all levels of epigenetic changes, and by changing substrate protein activities, they can also activate and repress effects on gene transcription depending on their target proteins and the ubiquitin chain types, all of which are connectively related to the process of tumor proliferation (Kim and Baek, 2006; Zheng et al., 2008; Fradet-Turcotte et al., 2013; Yeh et al., 2018). Ubiquitin code signaling is frequently dysregulated in numerous cancer types and can function as a tumor suppressor or tumor promoter, suggesting a potential target for cancer therapy (Loch and Strickler, 2012; Choudhry et al., 2018; Emanuelli et al., 2019).

## 1.4 Physiological Functions of Non-proteolytic Poly-Ubiquitin Chains

The simplest version of ubiquitination or mono-ubiquitination confers non-degradative activities including protein localization (Yang et al., 2017), endocytosis (Shih et al., 2002; Windheim et al., 2008), trafficking, DNA repair (Xie et al., 2014; Whiteaker et al., 2018), autophagy (Chen S. et al., 2017; Zheng et al., 2020; Leng et al., 2021) and chromatin remodeling (Cole et al., 2021). When mono-ubiquitination is further modified, multiple lysine residues of the substrate are yielded to induce multi-mono-ubiquitination. Emerging investigations have found that this process connects ubiquitin with subcellular trafficking (Yin et al., 2010; Cooray et al., 2011) and immune-suppressive functions (Zhu et al., 2018). As for poly-ubiquitination, Lys48-linked poly-ubiquitination leads to the degradation of substrates (Komander and Rape, 2012). In contrast, Lys63-linked poly-ubiquitination exerts critical signaling functions in regulating protein stability, including nuclear factor  $\kappa$ B (NF- $\kappa$ B) signaling (Syed et al., 2006; Wu Z. et al., 2014; Gallo et al., 2014), endocytosis

(Galan and Haguenauer-Tsapis, 1997), DNA damage responses, and immune responses (Wu and Karin, 2015; Hrdinka et al., 2016; Liu et al., 2017; Paul and Wang, 2017). Emerging experiments have shown that incorrect regulation of cellular processes (either tumour inhibitors or promoters) contributes to cancer pathogenesis and progression (Shirane et al., 1999). Additionally, the cellular functions of atypical ubiquitin linkages (except Lys11-linked ubiquitin) are supposed to be non-degradable (Kulathu and Komander, 2012; Iwai, 2014, 2015; Meza Gutierrez et al., 2018) and in most cases, activities involved in Lys63-linked poly-ubiquitin chains are also considered to be non-degradable (Liu et al., 2015).

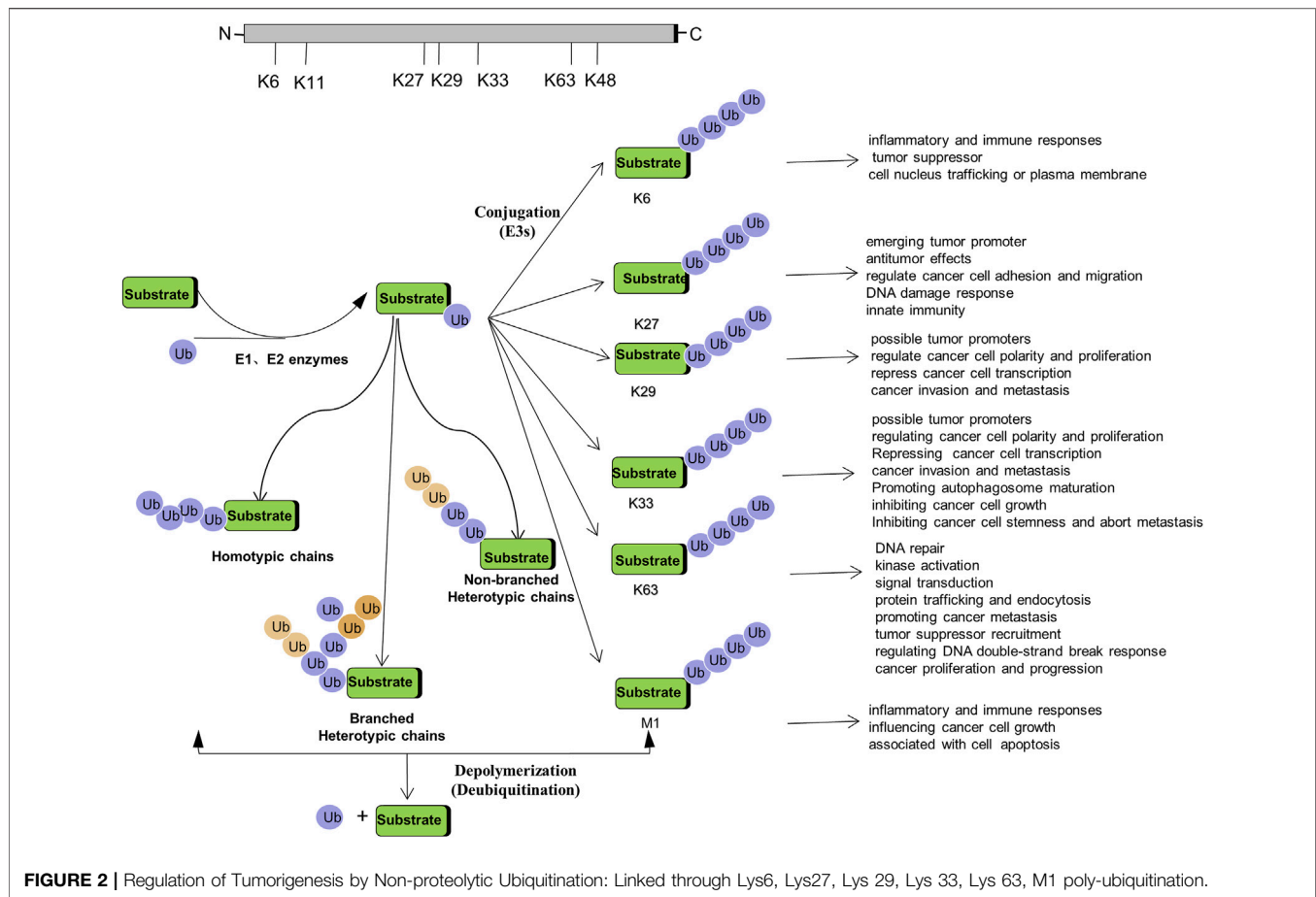
Of note, other chains also regulate specific physiological functions: Lys6-linked poly-ubiquitin was supposed to be indirectly linked to DNA damage response with the help of heterodimeric ubiquitin E3 ligase BRCA1-BARD1 (Wu-Baer et al., 2003; Wu-Baer et al., 2010). Linear (M1) chains can regulate NF- $\kappa$ B activation (Behrends and Harper, 2011; Chen et al., 2015; Borghi et al., 2018), whereas Lys11 linkage acts as a powerful degradation signal in heterotypic ubiquitin conjugates (Locke et al., 2014; Mevissen et al., 2016) and Lys27 linkage can prompt mitochondrial depolarization and mediate translocation of the E3 ligase Parkin, which accumulates Lys27-linked linkages on mitochondrial protein voltage-dependent anion-selective channel protein1 (VDAC1). This exact Lys27-linked translocation leads to Parkinson's disease in the presence of Parkin (Geisler et al., 2010; Glauser et al., 2011). Lys27 linkage has also been demonstrated to be related with the DNA damage response and innate immune response (Xue et al., 2018). The propagation of Wnt/ $\beta$ -catenin signaling through Lys29- or Lys11-linked ubiquitin chains is closely associated with cancer pathogenesis and is involved in protein ubiquitination at multiple levels (Hay-Koren et al., 2011). Several studies have reported the functions of both Lys29- and Lys33-linked chains in the regulation of AMPK-related protein kinases (Al-Hakim et al., 2008), and the Lys33 linkage is negatively regulated by T-cell antigen receptors (TCRs), which indirectly affect cellular activities in tumors (Huang et al., 2010). In this review, we will focus on recent progresses in non-degradable ubiquitin chains in tumorigenesis and the essential proteins involved in non-proteolytic ubiquitination.

## 2 THE ROLES OF NON-PROTEOLYTIC UBIQUITINATION IN TUMORIGENESIS

Non-proteolytic ubiquitination, including both mono-ubiquitination and poly-ubiquitination (mainly K6-, K27-, K29-, K33-, K63-, and M1-linked poly-ubiquitination), has become key regulators in a variety of cancers. How they function as signaling entities in the pathogenesis and progression of cancer will be beneficial to gain an in-depth understanding of ubiquitination (Figure 2).

### 2.1 Mono-Ubiquitination

Mono-ubiquitination has been suggested to be even more dynamic than previously thought, and its functions have been



deciphered by various ubiquitin-binding proteins. It has been demonstrated that the cellular functionality of ubiquitin is mediated by mono-ubiquitin and/or poly-ubiquitin. In fact, mono-ubiquitination is associated with tumorigenesis, not limited to membrane transportation, endocytosis, receptor internalization, degradation in lysosomes, and protein reprocessing (Haglund et al., 2003). BMI1 interacts with histone H2A through mono-ubiquitination, repressing multiple genes, such as INK4A/ARF, which function in the pRb and p53 pathways, thereby facilitating cancer progression (Lin et al., 2015). Interestingly, in the presence of UBE2E1, PRC1 catalyzes the mono-ubiquitination of H2A, contributing to cancer cell proliferation (Wheaton et al., 2017). Mono-ubiquitination is also involved in the process of USP22 regulating histone H2B and exhibits both oncogenic and tumor-suppressor roles in cancer development (Jeusset and McManus, 2017). In addition, reversible mono-ubiquitination activity plays an essential role in balancing TGF- $\beta$ /SMAD signaling, which is involved in cancer initiation and progression (Xie et al., 2014). Mono-ubiquitination also regulates forkhead box O (FOXO) proteins, which control specific gene expression programs that are vital for slowing the onset of cancer in aging individuals (Greer and Brunet, 2008). Strikingly, FANCL cooperates with, UBE2T and catalyzes mono-ubiquitination, which participated in the regulation of Fanconi Anemia pathway, leading to chromosome instability and

promoting tumorigenesis (Machida et al., 2006; Hodson et al., 2014; Miles et al., 2015; Sun et al., 2020; Wang S. et al., 2021) (Table 1, 2).

## 2.2 Linear (M1) Linkage

**2.2.1 The LUBAC Complex Encodes the M1 Linkage**  
Emerging evidences connect Met1-linked ubiquitin chains to NF- $\kappa$ B signaling, which enables physiological regulation of inflammation and immune responses (Emmerich et al., 2011; Borghi et al., 2018). Indeed, mutations and deficiencies involved in the formation and dissolution of Met1-linked poly-ubiquitin chains have been extensively illustrated in immune-related disorders (Tokunaga and Iwai, 2012; Fiil et al., 2013; Fiil and Gyrd-Hansen, 2014). Until now, linear ubiquitin chain assembly complex (LUBAC) is the only known E3 ubiquitin ligase assembling this type of chain (Kirisako et al., 2006; Walczak et al., 2012). The multi-subunit E3 ligase comprises catalytically active hybrid organic-inorganic perovskite (HOIP, also known as RNF31) and two adaptor proteins, HOIPL (also known as RBCK1) and SHANK-associated RH domain-interacting protein (SHARPIN) (Tokunaga et al., 2011). Both HOIL-1 and SHARPIN can co-operate with HOIP, which might be a central part of the LUBAC complex (Ikeda et al., 2011; Elton et al., 2015). The HOIP orthologue, linear ubiquitin E3 ligase (LUBEL), modifies Kenny with M1-linked linear ubiquitin chains in *Drosophila* and is

**TABLE 1** | Summary of the combined E2/E3 enzymes in Tumorigenesis.

| E2                 | Alias               | Accompanied E3 | Linkage | Phenotypic Characteristics                      | Neoplastic Implications  | Substrate                       | Mechanism Summary   |
|--------------------|---------------------|----------------|---------|---|--|---------------------------------|---|
| UBE2N              | UBC13               | TRAF2/TRAF6?   | K63     | preventing tumor formation and metastasis       | modulating breast cancer metastasis                            | NEMO?                           | UEV1A, together with Ubc13, promote breast cancer metastasis through Lys63-linked polyubiquitination of target proteins and NF- $\kappa$ B-mediated MMP1 expression (Wu et al., 2014b)  |
| UBE2N              | UBC13               | TRAF6          | K63     | DNA damage repair and protein kinase activation | metastatic spread and lung colonization by breast cancer cells | p38                             | Ubc13 catalyzes K63-linked proteins, accompanied by TAK1-p38 activation, whose activity is essential for breast cancer metastasis (Wu et al., 2014a)  |
| Ubc13: Uev1A/ Uev2 | UBE2N: UBE2V1/ Uev2 | RNF8           | K63     | DNA damage repair and cytokinesis               | genomic instability in adult T-cell leukemia (ATL)             | TAK1?IKK?                       | Ubc13/Uev1A and RNF8 interact with each other and generate K63-pUb, which is recognized by Tax, stimulating TAK1 and IKK activation (Ho et al., 2015)   |
| Ubc13              | UBE2N               | TRAF6          | K63     | activating NF- $\kappa$ B signaling             | elicit anti-tumour responses                                   | RANK?                           | STAT3 negatively regulates Ubc1 involving K63-linked ubiquitination, and suppress pro-inflammatory cytokines by modulating NF- $\kappa$ B signaling (Zhang et al., 2014)  |
| Ubc13              | UBE2N               | RNF8           | K63     | DNA double-strand break (DSB) responses         | BRCA1 Tumor Suppressor Recruitment                             | histone                         | RNF8 stimulates Ubc13-dependent Lys-63 ubiquitination, which is pivotal for DNA damage response and recruitment of BRCA1 (Hodge et al., 2016)   |
| Ubc13              | UBE2N               | TRAF6          | K63     | innate and adaptive immunity                    | osteoclast differentiation                                     | TRAF6 (autoubiquitination)      | Ubc13/Uev1A interacts and binds to the active RING domain of TRAF6, which is essential for the formation of Lys63-linked pUb, thus triggering NF- $\kappa$ B activation and osteoclast differentiation (Lamothe et al., 2008) |
| Ubc13/ Uev1A       | UBE2N/ UBE2V1       | TRAF6          | K63     | activating NF- $\kappa$ B signaling             | osteoclast differentiation                                     | TRAF6 (autoubiquitination)      | TRAF6 in combination with Ubc/Uev1A catalyzes TRAF6 auto-ubiquitination through Lys63-linked poly-Ub chains, which controls NF- $\kappa$ B signaling and osteoclast differentiation (Lamothe et al., 2007b)                   |
| Ubc13/ Uev1A       | UBE2N/ UBE2V1       | TRAF6          | K63     | spontaneous osteoclast differentiation          |  | TRAF6 (auto-ubiquitination)     | TRAF6 interacts with Ubc13/ Uev1A, facilitating Lys-linked auto-ubiquitination of TRAF in a RING domain-dependent fashion, and modulating downstream NF- $\kappa$ B signaling (Lamothe et al., 2007a)                         |
| UBE2O              | UBE2O               | TRAF6          | K63     | activating NF- $\kappa$ B signaling             | modulating NF- $\kappa$ B signaling associated cancers         | TRAF6 (auto-polyubiquitination) | UBE2O negatively regulates the recruitment of TRAF6, inducing TRAF6 auto-ubiquitination through binding to K63 residue, and subsequently prevents NF- $\kappa$ B activation by the IL-1R/TLR complex (Zhang et al., 2013b)    |

(Continued on following page)

**TABLE 1 |** (Continued) Summary of the combined E2/E3 enzymes in Tumorigenesis.

| E2    | Alias | Accompanied E3   | Linkage                | Phenotypic Characteristics                           | Neoplastic Implications                               | Substrate        | Mechanism Summary  |
|-------|-------|------------------|------------------------|--|---|------------------|--|
| UBE2T | UBE2T | FANCL            | monoubiquitination     | maintenance of chromosome stability                  | disrupting DNA repair pathways                        | FANCD2           | In the presence of FANCL, UBE2T stimulates monoubiquitination of FANCD2, which is vital for disrupting abnormal chromosomes and efficient DNA damage repair (Machida et al., 2006)     |
| UBE2T | UBE2T | FANCL            | automonoubiquitination | maintenance of chromosome stability                  | disrupting DNA repair pathways                        | UBE2T            | Automonoubiquitination of UBE2T inhibits own conjugation activity (Machida et al., 2006)   |
| UBE2T | UBE2T | FANCL            | monoubiquitination     | DNA repair   | leading to leukemia and bone marrow failure           | FANCD2           | FANCL interacts with UBE2T in an ELF-domain-dependent fashion, which regulates DNA damage-induced FANCD2 monoubiquitination (Miles et al., 2015)                                       |
| UBE2T | UBE2T | FANCL            | monoubiquitination     | DNA interstrand crosslink repair                     | genomic instabilities                                 | FANCD2           | FANCL specifically interacts with UBE2T, leading to FANCD2 ubiquitination, which is involved in Fanconi Anemia pathway (Hodson et al., 2014)   |
| RAD6  | UBE2B | in absence of E3 | K63?                   | promoting DNA repair                                 | promoting recurrence and metastasis in ovarian cancer | $\beta$ -catenin | Rad6 facilitates DNA repair and stem cell gene transcription, through mediating H2B ubiquitination (Somasagara et al., 2017)   |
| RAD6B | UBE2B | in absence of E3 | K63                    | DNA repair and mutagenesis                           | h-catenin modification in breast cancer               | $\beta$ -catenin | Rad6B ubiquitinates b-catenin through K63-linked ubiquitination, which regulates transcriptional activity in breast cancer (Shekhar et al., 2008)                                      |
| UBE2B | UBE2B | BRE1             | monoubiquitination     | promoting the G1-S transition and cell proliferation | promoting G1-S transition and cell proliferation      | H2B              | UBE2B modulates CCND1 transcription level by regulating the levels of H2B monoubiquitination, promoting cell cycle progression and proliferation (Cai et al., 2014)                    |
| Ube2w | UBE2W | RNF4             | monoubiquitination     | DNA damage repair                                    | potential prostate, breast and lung cancer target     | SUMO             | Ube2w associated with RNF4, mediating mono-ubiquitination of SUMO chains. Those chains can be further ubiquitinated through K63 chains in response to DNA damager (Maure et al., 2016) |
| Ubc13 | UBE2N | TRAF6?           | K63                    | Autoimmunity and aberrant T cell activation          | modulating NF- $\kappa$ B associated cancer           | IKK              | Ubc13 conjugates K63-linked ubiquitin chains involving Ubc13-IKK signaling axis, which have a robust evidence in regulating T cell function (Chang et al., 2012)                       |
| UBC13 | UBE2N | Bcl10            | K63                    | activating NF-kappaB pathway                         | modulating NF- $\kappa$ B associated cancer           | NEMO             | UBC13 is dependent in Bcl10 modulating NEMO lysine-63-linked ubiquitination, and subsequent NF-kappaB activation (Zhou et al., 2004)   |

(Continued on following page)

**TABLE 1 |** (Continued) Summary of the combined E2/E3 enzymes in Tumorigenesis.

| E2                 | Alias               | Accompanied E3 | Linkage            | Phenotypic Characteristics                                  | Neoplastic Implications                                  | Substrate       | Mechanism Summary   |
|--------------------|---------------------|----------------|--------------------|---|--|-----------------|---|
| UBE2N              | UBE2N               | MEKK1          | K63                | embryonic survival  | promoting ES-cell differentiation and tumour formation   | TAB1            | Together with UBE2N, MEKK1 could tag TAB1 with Lys63-linked poly-Ub, promoting ES-cell differentiation and tumourigenesis (Charlaftis et al., 2014)   |
| UbcH6              | UBE2E1              | TRAF4          | K63                | DNA damage  | overcome chemoresistance in colorectal cancer            | CHK1            | UbcH6 combined with TRAF4, which is critical for CHK1 K63-linked ubiquitination and essential for cell proliferation, colony formation (Yu et al., 2020b)   |
| UBE2T              | UBE2T               | RNF8           | monoubiquitination | facilitating cell cycle arrest activation                   | conferring HCC radioresistance                           | H2AX            | UBE2T/RNF8 complex, monoubiquitinated H2AX/γH2AX, facilitating cell cycle arrest activation, thus inducing HCC radioresistance (Sun et al., 2020)   |
| Ubc13: Uev1A/ Uev2 | UBE2N: UBE2V1/ UEV2 | RNF8           | K63                | DNA damage repair and cytokinesis                           | genomic instability of ATL cells                         | NEMO and TAB2/3 | RNF8 and Ubc13:Uev1A/ Uev2 assemble K63-pUb chains on NEMO and TAB2/3 respectively, allowing TAK1 and IKK activation (Ho et al., 2015)  |
| Ube2E1             |                     | PRC1           | monoubiquitination | maintenance of stem cell proliferation                      | promoting cancer cell proliferation                      | H2A             | Ube2E1 interacts with PRC1 complex, catalyzing monoubiquitination of H2A (Wheaton et al., 2017)   |
| Ubc13              | UBE2N               | TRAF6          | K63                | regulating immune signaling                                 | NF-κB signaling related cancer                           |                 | Ubc13, together with TRAF6, mediates K63-linked polyubiquitin signaling pathway, including NF-κB signaling (Lenoir et al., 2021)  |
| Uev1A/ Ubc13       | UBE2N/ UBE2V1       | TRAF6          | K63                | regulating AKT signaling pathway                            | promoting breast cancer cell migration and EMT signaling | AKT             | Uev1A/Ubc13 interact with TRAF6, ubiquitinates AKT with K63-linked ubiquitination, which is required for AKT activation, promoting cell migration and EMT in breast cancer (Niu et al., 2021)                         |
| UBE2T              |                     | FANCL          | monoubiquitination | involving in FA pathway-induced chromosome instability      | functions in cancer predisposition                       | ID              | UBE2T/FANCI-FANCD2 complex remodeling the ID-DNA complex, preventing clamp opening after monoubiquitination (Wang et al., 2021a)  |
| UbcH6              |                     | NEDD4          | K63                | regulating cell-cell adhesion, mechanosensing and autophagy | involving in angiogenesis and tumor growth               | IGPR-1          | NEDD4 and UbcH6 are involved in the K63-linked ubiquitination of IGPR-1, regulating different cellular activities, such as cell adhesion, autophagy, mechanosensing, angiogenesis and tumor growth (Sun et al., 2021) |
| Ubc13              | UBE2N               | RNF213         | K63                | angiogenic activity   | regulating cell mobility and invasion                    |                 | RNF213 interacts with Ubc13 and promotes its own autoubiquitination, controlling inflammatory responses and angiogenic activities (Habu and Harada, 2021)   |

**TABLE 2 |** Summary of the Identified DUBs Involving in Tumorigenesis.

| Type | DUBs            | Linkage            | Phenotypic Characteristics   | Neoplastic Implications   | Substrate                   | Mechanism Summary  |
|------|-----------------|--------------------|--|---|-----------------------------|--|
| OTU  | TRABID (ZRANB1) | K63                | stem cell self-renewal or differentiation  | Wnt-induced transcription in colorectal cancer cell                               | APC                         | Trabid preferentially binding to K63-linked ubiquitination chains, which is required for Wnt-induced transcription (Tran et al., 2008)   |
| OTU  | TRABID (ZRANB1) | K29, K33           | inhibiting autophagy flux  | Promoting autophagosome maturation and inhibiting hepatocellular carcinoma growth | UVRAG                       | ZRANB1 precisely cleaves K29 and K33-linked polyubiquitin chains from UVRAG, regulating autophagy system (Feng et al., 2019)   |
| OTU  | OTUD1           | K33                | restricting the TGF- $\beta$ signaling   | inhibiting breast cancer stemness and metastasis                                  | SMAD7                       | OTUD1 directly deubiquitinates the SMAD7, shuts off TGF- $\beta$ signals, thereby suppressing metastasis in breast cancer (Zhang et al., 2017)   |
| OTU  | OTUD1           | K63                | regulating organ growth, tissue regeneration   | Regulating tumorigenesis, cancer cell survival and chemoresistance                | YAP                         | OTUD1 cleaves K63-linked poly Ub from YAP, which is assembled by SKP2 E3 ligase, regulating tumorigenesis (Yao et al., 2018)   |
| OTU  | OTUD1           | K63?               | suppressing colony formation and increasing apoptosis  | arresting cell growth and inducing apoptosis                                      | p53                         | OTUD1 interacts with and stabilizes p53. Its overexpression significantly suppress colony formation, and increases apoptosis (Piao et al., 2017)   |
| OTU  | OTULIN          | M1                 | activating NF- $\kappa$ B and promoting pro-inflammatory cytokines and restricting bacterial proliferation | NF- $\kappa$ B signaling associated cancers                                       | cytosolic <i>salmonella</i> | OTULIN dissolves linear Ub chains on cytosolic S, resulting in ultimately NF- $\kappa$ B activation (van Wijk et al., 2017)  |
| OTU  | OTULIN          | M1                 | regulating NF- $\kappa$ B signaling and sensitizing cell death   | NF- $\kappa$ B signaling associated cancers                                       | RIPK1                       | OTULIN interacts with LUBAC, balancing Met1-polyUb chains, thereby regulating NF- $\kappa$ B signaling (Keusekotten et al., 2013)  |
| OTU  | OTUD1           | K63?               | decreases cell proliferation and increases cell apoptosis  | regulating tumor-suppressor p53   | p53                         | OTUD1 interacts with and deubiquitinates p53, regulates p53 stability and activity (Piao et al., 2017)   |
| OTU  | A20             | K63                | NF- $\kappa$ B transcriptional activity-mediated cell death and chronic inflammation                       | NF- $\kappa$ B signaling-related cancers  | RIP                         | A20 erases K63-linked ubiquitin chains from RIP, and it also polyubiquitinate RIP with K48-linked ubiquitin chains in a carboxy-terminal-domain-dependent manner, which downregulating NF- $\kappa$ B signalling, (Wertz et al., 2004) |
| OTU  | A20             | K63                | downregulating NF- $\kappa$ B pathway  | NF- $\kappa$ B signaling-related cancers  | TRAF6/RIP                   | A20 display dual ubiquitin-editing functions, mediating both non-proteolytic Lys63-linked ubiquitin chains and degradative Lys48-linked ubiquitin chains, thus regulating NF- $\kappa$ B activities (Lin et al., 2008)                 |
| OTU  | OTUD7B          | K63                | regulating mTORC2 signalling, thus relating to cell growth and metabolic disorders                         | activates Akt signaling and Kras-driven lung tumorigenesis <i>in vivo</i>         | G $\beta$ L                 | OTUD7B and TRAF2 regulate stability of G $\beta$ L, which plays critical roles in mTORC2 signaling (Wang et al., 2017a)  |
| UCH  | BAP1            | monoubiquitination | inhibiting cell apoptosis  | mediating tumor suppression in both mice and humans                               | H2A                         | Inactivation of BAP1 causes apoptosis through regulating H2A monoubiquitination, regulating tumor suppression (He et al., 2019)  |
| USP  | USP4            | K63                | activating inflammation and immune response  | inhibiting TNF $\alpha$ -induced cancer cell migration                            | TRAF2/ TRAF6                | USP4 negatively regulates the TRAF2- and TRAF6-stimulated NF- $\kappa$ B activation, and inhibits cancer cell migration (Xiao et al., 2012)  |
| USP  | CYLD            | K63                | regulating NF- $\kappa$ B-mediated inflammation  | associating the development of head and neck squamous cell carcinomas             | NEMO                        | TRAF3/CYLD complex regulate NF- $\kappa$ B transcriptional level, which is associated with head and neck squamous cell carcinomas with HPV infection (Chen et al., 2017b)  |
| USP  | CYLD            | K63/M1             | regulating Innate Immune Signaling   | tumor suppressor  | RIPK2                       | CYLD counteracts Met1-Ub and Lys63-Ub conjugated to Ripk2, and this deubiquitinase activity plays an important role in innate immune regulation (Hrdinka et al., 2016)   |

(Continued on following page)

**TABLE 2 |** (Continued) Summary of the Identified DUBs Involving in Tumorigenesis.

| Type | DUBs  | Linkage | Phenotypic Characteristics  | Neoplastic Implications                                       | Substrate | Mechanism Summary  |
|------|-------|---------|---|---|-----------|--|
| USP  | USP8  | K63     | DNA damage response   | genomic instability in cancer                                 | BRIT1     | USP8 rescues BRIT1 from K63 ubiquitin and regulates its recruitment to DNA double-strand break sites (Ge et al., 2015)   |
| USP  | USP10 | K63     | controlling cell cycle  | promoting proliferation of t chronic myeloid leukemia cells   | Bcr-Abl   | SKP2 acts as co-regulator of K63-linked ubiquitination of Bcr-Abl for its activation. While USP10 deubiquitinates and stabilizes SKP protein levels and amplifies Bcr-Abl activation in chronic myeloid leukemia cells (Liao et al., 2019) |
| USP  | USP20 | K63     | negatively regulating inflammation, cell proliferation, and apoptosis | promoting adult T cell leukemia (ATL) development             | Tax       | USP20 targets and deubiquitinates TRAF6 and TAX, negatively regulating NF- $\kappa$ B signaling (Yasunaga et al., 2011)  |
| USP  | USP17 | K63     | enhancing inflammation and promoting macrophage recruitment           | promoting lung cancer growth                                  | cIAP1/2   | USP17 interacted with and disrupted the TRAF2/TRAF3 complex through reducing K63-linked ubiquitination of TRAF2 and TRAF6. This activity positively drives stemness and inflammation in lung cancer (Lu et al., 2018)                      |
| USP  | CYLD  | K63     | regulating inflammation   | promoting tumor growth  | TAK1      | Itch-Cyld complex sequentially cleaving K63-linked ubiquitin chain on Tak1 thus terminating the inflammatory response (Ahmed et al., 2011)   |
| USP  | CYLD  | K63     | negative regulate the NF- $\kappa$ B pathway                          | tumor suppressor  | E6        | HPV E6 suppresses the CYLD under hypoxic conditions, promoting unrestricted NF- $\kappa$ B activation and allowing for malignant progression of tumors (An et al., 2008)   |
| USP  | CYLD  | K63     | controlling inflammation  | inhibiting tumor formation                                    | Bcl-3     | Cyld erases K63-linked polyubiquitin chains from Bcl-3, inactivating NF- $\kappa$ B signaling (Massoumi et al., 2006)  |
| USP  | CYLD  | K63     | controls survival and inflammation                                    | inhibiting tumor cell Proliferation                           | TRAF2     | Cyld regulates inflammation through deubiquitinating TRAF2 and blocking NF- $\kappa$ B pathway (Massoumi et al., 2006)   |
| USP  | USP14 | K63     | inflammation  | acute colitis and colitis-associated colon cancer development | p100/p52  | TRAM14 recruits USP14 to cleave K63-linked ubiquitin chains of p100/p52, regulating NF- $\kappa$ B-mediated autophagy and innate immunity. (Chen et al., 2020)   |
| USP  | USP1  | K63     | regulating macroautophagy/ autophagy                                  | affecting breast cancer cell growth                           | ULK1      | USP1 modulates ULK1 K63-linked deubiquitination, and regulates autophagy, also affects breast cancer cell growth relying on autophagy (Raimondi et al., 2019)  |
| USP  | USP1  | K63     | Double-strand breaks (DSBs)   | potential tumour suppressor                                   | histones  | USP1 actively destorys K63-linked poly-ubiquitin chains on histones. And its recruitment to damage sites has a close link with genome stability and double-strand breaks (Ha et al., 2017)   |
| USP  | USP34 | K63     | genome stability maintenance  | promoting ES-cell differentiation and tumour formation        | H2A       | USP34 stabilizes RNF168, recruiting repair proteins at DSBs, which, is critical for genome stability (Sy et al., 2013)   |
| USP  | CYLD  | K63/M1  | DNA damage-induced apoptosis  | enhancing sensitivity to chemodrug in cancer cells            | NEMO      | CYLD downregulates K63-linked and linear ubiquitination of NEMO, promoting apoptosis (Niu et al., 2011)  |
| USP  | CYLD  | K63     | activating the cell death pathway                                     | regulating ATLL cell death                                    | RIPK1     | CYLD erases k63-linked ubiquitin chains from RIPK1, which activates the cell death pathway and activates CYLD and RIPK1-dependent tumor cell death in ATLL (Xu et al., 2020)   |

(Continued on following page)

**TABLE 2 |** (Continued) Summary of the Identified DUBs Involving in Tumorigenesis.

| Type  | DUBs   | Linkage | Phenotypic Characteristics   | Neoplastic Implications  | Substrate | Mechanism Summary   |
|-------|--------|---------|--|--|-----------|---|
| USP   | USP38  | K63     | restoring genome stability   | regulating cancer cell response to genotoxic insults                 | HDAC1     | USP38 preferentially removed the K63-linked ubiquitin chains from HDAC1, regulating genomic stability (Yang et al., 2020)   |
|       | UBQLN4 | K63/K48 | homologous recombination repair  | predictor of poor survival in various cancer entities                | MRE11     | Overexpression of UBQLN4 represses homologous recombination activity through inhibiting MRE11 ubiquitination, thus presenting close relationship with survival rates in various cancer (Jachimowicz et al., 2019) |
| OTU   | ZRANB1 | K29/K33 | promoting autophagosome maturation                                     | inhibiting cell growth in hepatocellular carcinoma                   | UVRAG     | ZRANB1 specifically cleaves SMURF1-induced K29 and K33-linked polyubiquitin chains from UVRAG, regulating autophagosome maturation and HCC growth (Feng et al., 2019)   |
| JAMM  | POH1   | K63     | double-strand DNA break responses, embryonic stem cell differentiation | promoting tumour formation in human hepatocellular carcinomas (HCCs) | E2F1      | POH1 binds to and stabilizes the E2F1, upregulating Survivin and FOXM1 protein levels, accompanied by accelerating tumor growth (Wang et al., 2015)   |
| USP   | USP30  | K6      | regulating mitophagy and neurodegenerative disease                     | functions in hepatocellular carcinoma                                |           | USP30 specifically cleaves the Lys6 linked ubiquitin chains, regulating mitophagy, apoptosis and tumorigenesis  |
| OTUD1 | OTUD1  | K63     | suppressing intestinal inflammation                                    | NF- $\kappa$ B signaling-related cancers                             | RIPK1     | OTUD1 preferentially cleaves K63-linked polyubiquitin chains from RIPK1, inhibiting colonic inflammation and NF- $\kappa$ B signaling   |
| USP   | CYLD   | K63     | regulating ERK activation  | regulating cancer cell growth, proliferation, migration              | ERK1/2    | CYLD cleaves K63-linked ubiquitination mediated by TRIM1, regulating ERK signaling and the associated cancer development  |
| USP   | USP10  | K63     |  | inhibiting NSCLC cell proliferation and migration                    | PTEN      | USP10 suppress NSCLC cell proliferation and migration through abolishing PTEN from K63-linked polyubiquitination mediated by TRIM25   |
| OTU   | A20    | K63     | anti-inflammatory effects  | tumor suppressor   | TBK1      | A20 inhibits TBK1 activation through reducing K63-linked ubiquitination of Nrdp1, regulating inflammation   |

indispensable for inflammatory responses by activating Imd pathway (Aalto et al., 2019). Additionally, HOIP, with the help of cIAP1, is recruited to the linear ubiquitination of the TNFR2 signaling complex and activates canonical NF- $\kappa$ B, thereby facilitating cancer progression (Borghi et al., 2018). Multiple investigations found that HOIP, presumably through M1-linked ubiquitination, was proved to be connected with sorts of malignancies, including breast and prostate cancer (Guo et al., 2015; Zhu et al., 2016). Furthermore, previous experiments highlight that HOIL-1 interacts with HOIP, which adds a M1-linked poly-ubiquitin chain to specific NF- $\kappa$ B signaling proteins, suggesting its link to a diversity of immune disorders, antiviral signaling (Elton et al., 2015), iron and xenobiotic metabolism (Elton et al., 2015), apoptosis (Emmerich et al., 2011; Lewis et al., 2015; Sasaki and Iwai, 2015), and cancer (Queisser et al., 2014; Taminiau et al., 2016). Another member of the LUBAC

family, SHARPIN, is a novel component of the LUBAC complex. Spontaneous mutation of this tiny gene led to dysregulation of the NF- $\kappa$ B signaling pathway through linear ubiquitination of NEMO (the key modulator of NF- $\kappa$ B). This systematic linear ubiquitination is obvious and can induce immune system disorders in SHARPIN-deficient mice (Tokunaga et al., 2011). Interestingly, SHARPIN-containing complexes can also interact with NEMO to activate NF- $\kappa$ B pathway (Ikeda et al., 2011). Moreover, NEMO was identified to be modified by LUBAC, generating M1-linked chains that were recognized by the UBAN domain of NEMO, causing conformational changes in the intertwined helices of NEMO dimers (Haas et al., 2009; Belgnaoui et al., 2012; Tokunaga and Iwai, 2012; Noad et al., 2017). NF- $\kappa$ B signaling regulates human cellular activities in different ways and is balanced by ubiquitination and deubiquitination (Adhikari et al., 2007; Lork et al., 2017) (Table 2).

**TABLE 3 |** Summary of the E3 enzymes in Tumorigenesis.

| E3     | Linkage  | Phenotypic Characteristics   | Neoplastic Implications                                   | Substrate                          | Mechanism Summary   |
|--------|----------|--|---|------------------------------------|---|
| RNF8   | K63      | regulating DNA double-strand break responses                                 | BRCA1 Tumor Suppressor Recruitment                        | histone                            | RNF8 interact with Ubc13, generating K63-linked ubiquitin chains on histone, which positively regulate DNA double-strand break and BRCA1 recruitment (Hodge et al., 2016)   |
| RNF8   | K63      | DNA damage repair and cytokinesis  | genomic instability of ATL cells                          | Tax                                | Stimulated by Tax, RNF8 and Ubc13:Uev1A function together, generating K63-pUb chains, which activated TAK and NF- $\kappa$ B signaling (Zhi et al., 2020)   |
| HOIL-1 | M1       | immune regulation  | NF- $\kappa$ B activation in cancers                      | NEMO                               | HOIL-1 modifies NF- $\kappa$ B core proteins with linear ubiquitin chains, regulating NF- $\kappa$ B pathway signaling (Elton et al., 2015)   |
| CHIP   | K6       | suppressing of cell death  |   | FADD                               | CHIP triggers K6-linked polyubiquitylation of FADD, leading to the suppression of cell death (Seo et al., 2018)   |
| BRCA1  | K6       | DNA damage response  | tumor suppressor  | BRCA1 autoubiquitination           | UBXN1 binds to BRCA1 active site and decorate it with K6-linked polyubiquitin chains in a UBA-domain-dependent manner and BRCA1/BARD1 complex is regulated by the ubiquitination status of BRCA1 (Wu-Baer et al., 2010) |
| BRCA1  | K6       | DNA repair, transcriptional regulation, and cell cycle checkpoint control    | tumor suppressor  | BRCA1 autoubiquitination           | BRCA1 mediates autoubiquitination by conjugating to K6-linked polymers, which impart cellular properties (Wu-Baer et al., 2003)   |
| BRCA1  | K6       | DNA double-stranded breaks repair  | tumor suppressor  | might be BRCA1 autoubiquitination? | BRCA1 recruits its autoubiquitination at DNA damage sites, which is dependent on K6-linked linkage. BRCA1:BARD1 enzyme activity is regulated by BRCA1 ubiquitin status. (Morris and Solomon, 2004)                      |
| BRCA1  | K6       | regulating DNA repair, transcriptional levels, cell cycle and cell apoptosis | tumor suppressor  | BRCA1 autoubiquitination           | BRCA1-BARD1 regulate BRCA1 autoubiquitination by preferentially mediating K6-linked polyubiquitin chains (Nishikawa et al., 2004)   |
| Hectd3 | K27, K29 | leading to NF- $\kappa$ B activation   | NF- $\kappa$ B associated cancer                          | Malt1                              | Hectd3 promotes K27 and K29 polyubiquitination on Malt1, regulating autoimmunity and other Th17-related diseases (Cho et al., 2019)   |
| WWP1   | K27      | suppressing the dimerization, membrane recruitment                           | restoring tumor-suppressive activity                      | PTEN                               | WWP1 triggers K27-linked polyubiquitination of PTEN to regulate subcellular localization cancer susceptibility syndromes (Lee et al., 2019c)  |
| TRAF4  | K27, K29 | facilitating immune cell migration   | promoting cancer cell invasion                            | TrkA                               | TRAF4 promotes K27 and K29-linked ubiquitin linkages on TrkA, facilitating prostate cell invasion (Singh et al., 2018)  |
| RNF4   | K63      | DNA damage repair  | potential prostate, breast and lung cancer target         | Trim5 $\alpha$                     | Ube2w interacts with RNF4, promoting monoubiquitination of SUMO chains, which are further modified to form K63-linked ubiquitin chains (Maure et al., 2016)   |
| RNF8   | K63      | DNA damage response  | breast cancer predisposition                              | H2A/H2AX                           | RNF8 activated with Ubc13, promoting K63-linked polyubiquitin conjugation to histones H2A/H2AX, then contributing to breast cancer predisposition (Vuorela et al., 2011)  |
| Skp2   | K63      | promoting survival and Akt-mediated glycolysis                               | restricting cancer cell progression                       | Akt                                | Skp2/SCF complex catalyzes K63-linked ubiquitination chains on Akt, which is required for glycolysis and cancer development (Chan et al., 2013)   |
| Skp2   | K63      | controlling cell cycle   | promoting proliferation of chronic myeloid leukemia cells | Bcr-Abl                            | SKP2 triggers K63-linked ubiquitination of Bcr-Abl, regulating downstream signaling, and is vital for chronic myeloid leukemia development and progression (Liao et al., 2019)  |

(Continued on following page)

**TABLE 3 |** (Continued) Summary of the E3 enzymes in Tumorigenesis.

| E3               | Linkage             | Phenotypic Characteristics                                 | Neoplastic Implications  | Substrate           | Mechanism Summary   |
|------------------|---------------------|--|--|---------------------|---|
| RNF113A          | K63                 | DNA repair   | potentially associating with tumor progression                   | BRR2                | RNF113A interacts with BRR2 through K63-linked polyubiquitin, mediating repairment of DNA alkylation damage (Brickner et al., 2017)   |
| TRAF2            | K63                 | mediating several cell growth and metabolic pathways       | facilitating tumorigenesis                                       | GβL                 | TRAF2 promotes K63-linked polyubiquitination of GβL, and regulates mTORC2 signalling, thus mediating several cell growth and metabolic pathways (Wang et al., 2017a)  |
| TRIM37           | mono-ubiquitination | regulating transcriptional repression                      | promoting transformation in breast cancer                        | H2A                 | TRIM37 mono-ubiquitinates histone H2A, thus associating with transcriptional repression (Bhatnagar et al., 2014)  |
| RNF8/<br>RNF168  | K63                 | DNA double-strand breaks (DSBs)                            | mediating ATM-dependent carcinogenesis                           | H2A/H2AX            | RNF8 and RNF168 combined together to catalyze K63-linked poly-Ub chains on H2A/H2AX, which is important for transcription and DNA double-strand breaks (Shanbhag et al., 2010)  |
| HectH9           | K63                 | regulating transcriptional activation and repression       | tumor cell Proliferation   | Myc                 | HectH9 recruits 63-linked polyubiquitin chains to Myc, modulating cell proliferation in various tumor cells (Adhikary et al., 2005)   |
| Bcl10            | K63                 | activating the NF-κB pathway                               | NF-κB associated cancer  | NEMO                | UBC13 and Bcl10 function together inducing NEMO ubiquitination through lysine-63-linked ubiquitination, and subsequent NF-κB activation (Zhou et al., 2004)   |
| RNF8             | K63                 | DNA repair   | tumour-promoting   | probably histone H1 | In p97-ATX3 activated conditions, RNF8 mediates K63-Ub at sites of DNA lesions, regulating genome instability, cell invasion and metastasis (Singh et al., 2019)  |
| PARK2/<br>Parkin | K33                 | fine-tune necroptosis and inflammation                     | tumor suppressor/inflammation-associated tumorigenesis           | RIPK3               | AMPK activated Parkin/RIPK3 complex through K33-linked polyubiquitination, which negatively regulates necroptosis and inflammation-associated tumorigenesis (Lee et al., 2019b)   |
| TRAF6            | K63/K27             | maintaining nuclear genome integrity                       | promoting cancer progression                                     | hDNA2               | hTRAF6 catalyzes the K27- and K63-linked polyubiquitination of hDNA2, maintaining nuclear genome integrity and the associated cancer biology (Meng et al., 2019)  |
| HectH9           | K63                 | integrating glycolysis activation and apoptosis resilience | regulating tumor metabolism and cancer stem cell expansion       | HK2                 | HectH9 catalyzes HK2's K63-linked ubiquitination, regulating stem cell expansion and CSC-induced chemoresistance in prostate cancer (Lee et al., 2019a)   |
| TRAF6            | K63                 | regulating inflammation and immunity                       | promoting liver tumorigenesis and correlates with poor prognosis | HDAC3               | TRAF6 ubiquitinates HDAC3 with K63-linked ubiquitin chains, regulating inflammation and malignant transformation and progression in HCC (Wu et al., 2020)   |
| ITCH             | K27                 | immune response  | promoting proliferation and invasion of melanoma cells           | BRAF                | Activated ITCH maintains BRAF activity and subsequent MEK/ERK signaling through Lysine 27-linked ubiquitination, enhancing proliferation and invasion of melanoma cells (Yin et al., 2019)  |
| TRAF6            | K63                 | immunity   | anti-tumor immunity in the cancer setting                        | FOXP3               | TRAF6 bind to and facilitates Regulatory T cells (Tregs) activities through K63-linked ubiquitination at lysine 262, acting as a Treg-stabilizing regulator and playing crucial roles in immune control and anti-tumor immunity (Ni et al., 2019) |
| SMURF1           | K29/K33             | promoting autophagosome maturation                         | inhibiting cell growth in hepatocellular carcinoma               | UVRAG               | SMURF1 mediates K29 and K33-linked polyubiquitin chains on UVRAG, promoting autophagosome maturation and inhibiting hepatocellular carcinoma growth (Feng et al., 2019)   |

(Continued on following page)

**TABLE 3 |** (Continued) Summary of the E3 enzymes in Tumorigenesis.

| E3      | Linkage | Phenotypic Characteristics  | Neoplastic Implications                           | Substrate   | Mechanism Summary  |
|---------|---------|---|---|-------------|--|
| TRAF6   | K63     | NF- $\kappa$ B activation and autophagy activation                        | cancer cell migration, cell invasion              | BECN1/TRAF6 | PRDX1 negatively regulates TRAF6 ubiquitin-ligase activity, leading to NF- $\kappa$ B inactivation and autophagy activation (Min et al., 2018)   |
| RNF8    | K63     | DNA double-strand break repair  | regulating L3MBTL2 mutation in leukemia           | L3MBTL2     | MDC1 recruits L3MBTL2 to sites of DNA lesion and is then ubiquitinated by RNF8, promoting DNA DSB repair and regulating L3MBTL2-induced cancers (Nowshien et al., 2018)  |
| Itch    | K63     | regulating tissue patterning, stem cell maintenance                       | modulating medulloblastoma tumorigenesis          | SuFu        | Itch/ $\beta$ -arrestin2 complex mediates Lys63-linked polyubiquitylation on SuFu, thus controlling Hedgehog signalling and medulloblastoma tumorigenesis (Infante et al., 2018)   |
| FBXO32  | K63     | brain development   | promoting tumorigenicity and metastasis in humans | CtBP1       | FBXO32 directly ubiquitinates CtBP1 with K63-linked ubiquitin chains, and this interaction activity regulates downstream EMT signaling and is essential for tumor metastasis and brain development (Sahu et al., 2017)                   |
| Cullins | K29     | promoting cell motility   | modulating cell migration                         | hnRNP A1    | SPSB1 catalyzes K29-linked polyUb chains on hnRNP A1, modulating cell migration and cell motility in EGF signaling (Wang et al., 2017b)  |
| HUWE1   | K63     | preventing DNA damage accumulation  | colonic tumour suppressor                         | Myc         | Huwe1 mediates MYC transactivation activity via K63-linked ubiquitination, inhibiting accumulation of DNA damage and preventing tumour initiation especially in colonic cancers (Myant et al., 2017)                                     |
| HectH9  | K63     | relating to embryonic lethal  | promoting hypoxia-induced tumour progression      | HAUSP(USP7) | HectH9 mediates K63-polyubiquitin chains conjugated to HAUSP. HAUSP then deubiquitinates HIF-1 $\alpha$ , promoting hypoxia-induced tumour progression (Wu et al., 2016)   |
| RNF8    | K63     | conferring chemoresistance  | tumor-promoting function                          | Twist       | RNF8 activates and ubiquitinate Twist, leading to subsequent EMT and CSC functions, thus exerting tumor-promoting functions such as cell migration and invasion (Lee et al., 2016)   |
| FBXW7   | K63     | genome integrity  | tumor suppressor                                  | XRCC4       | FBXW7 firstly phosphorylated by ATM and then it ubiquitylates XRCC4 via K63-linkage, promoting NHEJ repair, which is closely related to DSB and genomic stability (Zhang et al., 2016)   |
| Trim7   | K63     | regulating proliferation and apoptosis                                    | promoting Ras-mediated lung adenocarcinoma        | RACO-1      | Trim7 catalyzes Lys63-linked ubiquitination of RACO-1 in response to RAS signaling, and Trim7 overexpression increases lung tumour burden while knockdown of Trim7 reduces tumour growth in xenografts models (Chakraborty et al., 2015) |
| Skp2    | K63     | regulating energy metabolism, proliferation, apoptosis, and cell polarity | tumor growth in HCC                               | LKB1        | Skp2-dependent activation of LKB1 through K63-linked Ubiquitination is essential for HCC tumor growth and related to poor survival outcomes (Lee et al., 2015)   |
| TRAF6   | K63     | enhancing chemotherapeutic efficacy                                       | promoting T-ALL progression                       | MCL1        | IRAK1/4 signaling activated TRAF6, mediating K63-linked ubiquitination of MCL1, promoting T-ALL progression (Li et al., 2015)  |
| PELI1   | K63     | maintenance of autoimmunity   | promoting lymphomagenesis                         | BCL6        | PELI1 specifically binds to BCL6 and induces lysine 63-linked ubiquitin chains on BCL6, promoting lymphomagenesis, modulating the maintenance of autoimmunity through TLR and TCR signaling (Park et al., 2014)                          |

(Continued on following page)

**TABLE 3 |** (Continued) Summary of the E3 enzymes in Tumorigenesis.

| E3               | Linkage             | Phenotypic Characteristics   | Neoplastic Implications                                | Substrate   | Mechanism Summary  |
|------------------|---------------------|--|--|-------------|--|
| MEKK1            | K63                 | embryonic survival   | promoting ES-cell differentiation and tumour formation | TAB1        | MEKK1 ubiquitylates TAB1 with Lys63-linked polyubiquitin in a PHD motif-dependent manner, inducing TAK1 and MAPK activation, which are crucial for ES-cell differentiation and tumorigenesis (Charlaftis et al., 2014)                   |
| TRAF4            | K63                 | regulating immunity  | driving Breast Cancer Metastasis                       | TβRI        | TβRI-receptor TRAF4 interacts with each other, triggering Lys 63-linked TRAF4 polyubiquitylation and TAK1 activation, promoting cell migration and metastasis in breast cancer (Zhang et al., 2013a)                                     |
| RNF8/<br>RNF168  | K63                 | maintaining genome stability                                       | suppressing tumorigenesis                              | BLM         | RNF8/RNF168, triggers BLM activation, leading to BLM recruiting to the ubiquitin-interacting motifs of RAP80, which is vital to maintain genome stability and suppressing tumorigenesis (Tikoo et al., 2013)                             |
| RFP              | K27                 | inhibiting apoptosis and promoting cell survival and proliferation | tumor suppression                                      | PTEN        | E3 ubiquitin ligase RFP interacts with PTEN through K27-linked ubiquitination diminishing the effect of AKT signaling, involving in tumor suppression regulation (Lee et al., 2013)  |
| FAAP20           | K63                 | DNA damage repair and genome maintenance                           | leading to hematologic defects and cancer in patients  | FANCA       | FAAP20 binds K-63-linked ubiquitin chains in a UBZ domain-dependent manner, modulating DNA damage repair and genome maintenance (Ali et al., 2012)   |
| LUBAC            | M1                  | DNA damage-induced apoptosis                                       | enhancing sensitivity to chemodrug in cancer cells     | NEMO        | LUBAC-mediated NEMO linear ubiquitination promotes TAK1 and IKK activation, protecting cells from DNA damage-induced apoptosis (Niu et al., 2011)  |
| TRAF4<br>(RNF83) | K63                 | DNA damage   | overcome chemoresistance in colorectal cancer          | CHK1        | CHK1 K63-linked ubiquitination is mediated by TRAF4, which is essential for CHK1 phosphorylation and activation during DNA damage response, and is close to cell proliferation, colony formation in colorectal cancer (Yu et al., 2020b) |
| LUBAC            | M1                  | regulating cell activation and death                               | promoting breast cancer                                | NEMO        | Epsins 1 and 2 interact with LUBAC, promoting NEMO linear ubiquitination and resulting in breast cancer development (Song et al., 2021)  |
| DZIP3            | K63                 | regulating cell cycle  | promoting cancer cell growth, migration, and invasion  | Cyclin D1   | DZIP3/hRUL138 stabilizes and ubiquitinated Cyclin D1 protein through K63-linked ubiquitination, and closely related with cell cycle progression, cancer cell growth, invasion, migration (Kolapalli et al., 2021)                        |
| RNF138           | K63                 | driving NF-κB activation and innate immunity                       | promoting NF-κB activation in lymphomas                | MYD88L265P  | RNF138 triggers K63-linked polyubiquitination of MYD88L265P, resulting in constitutive activation of NF-κB and the associated lymphomagenesis (Yu et al., 2020a)   |
| RNF181           | K63                 | endocrine resistance   | facilitating breast cancer progression                 | ERα protein | RNF181 functions as E3 enzymes through K63-linked ubiquitination, which facilitates breast cancer progression (Zhu et al., 2020)   |
| TRIM11           | mono-ubiquitination | regulating estrogen-dependent gene expression                      | promoting cell growth and migration                    | ERα         | TRIM11 interacts with the N terminal of ERα and maintains ERα stability through mono-ubiquitination, thus promoting cell growth and proliferation in breast cancer (Tang et al., 2020)   |
| HUWE1            | K27-, K29-          | regulating DNA damage response                                     | promoting radio-resistance of prostate cancer cells    | JMJD1A      | HUWE1 mediates the K27-/K29-linked ubiquitination of JMJD1A, enhancing c-Myc activity, promoting DSB repair and<br>(Continued on following page)   |

**TABLE 3 |** (Continued) Summary of the E3 enzymes in Tumorigenesis.

| E3     | Linkage | Phenotypic Characteristics                                  | Neoplastic Implications   | Substrate                    | Mechanism Summary  |
|--------|---------|---|---|------------------------------|--|
| RNF6   | K63     | maintaining nuclear receptors                               | promoting cell proliferation  | glucocorticoid receptor (GR) | sensitizing the response of prostate cancer (Fan et al., 2020)<br>RNF6 stabilize GR genes and enhances its transcriptional activity by catalyzing its K63-linked polyubiquitination, promoting MM cell proliferation and survival (Ren et al., 2020) |
| TRIM27 |         | suppressing cell senescence                                 | cell cycle dysregulation, tumor cell proliferation and migration        | p21                          | TRIM27 regulates cell apoptosis, cell senescence through mediating the ubiquitination of p21 in breast cancer (Xing et al., 2020)  |
| SPOP   | K27     | increasing DNA replication stress                           | sensitizing cancer cells to ATR inhibition                              | Geminin                      | SPOP binding Geminin catalyzes K27-linked poly-ubiquitination of Geminin, preventing DNA replication over-firing and sensitizing cancer cells to ATR inhibition (Ma et al., 2021)  |
| NF-X1  | K33     | regulating glycine metabolism                               | preventing glioma tumor growth  | GLDC                         | Acetylation of GLDC inhibits its enzymes activity, and facilitates K33-linked ubiquitination by NF-X1, regulating glycine metabolism and tumorigenesis (Liu et al., 2021)  |
| HUWE1  | K63     | promoting c-Myc activity                                    | promoting retinoblastoma cell proliferation                             | c-Myc                        | HEL22 triggers K63-linked ubiquitination activity of c-Myc by HUWE1 to mediate retinoblastoma tumorigenesis (Dai et al., 2021)   |
| RNF8   | K63     | activating AKT pathway                                      | promoting lung cancer cell proliferation and resistance to chemotherapy | Akt                          | RNF8 mediates K63-linked ubiquitination of Akt, promoting lung cancer cell proliferation and resistance to DNA damage (Xu et al., 2021)  |
| TRIM31 | K63     | stabilizing and activating p53                              | inhibiting breast cancer progression                                    | p53                          | TRIM31 directly ubiquitinates p53 with K63-linked ubiquitination through its RING domain, activating p53 pathway, suppressing breast cancer progression (Guo et al., 2021)   |
| TRIM15 | K63     | activating NF- $\kappa$ B and Akt signaling pathway         | regulating cancer cell growth, proliferation, migration                 | ERK1/2                       | TRIM15 mediates K63-linked polyubiquitination of ERK1/2, then activating ERK signaling, leading to cell proliferation, migration and differentiation (Zhu et al., 2021)  |
| NEDD4  | K63     | regulating cell-cell adhesion, mechanosensing and autophagy | involving in angiogenesis and tumor growth                              | IGPR-1                       | NEDD4 and UbcH6 are involved in the K63-linked ubiquitination of IGPR-1, regulating different cellular activities (Sun et al., 2021)   |
| TRIM25 | K63     | activating AKT/mTOR signaling                               | promoting NSCLC cell survival and tumor growth                          | PTEN                         | TRIM25 directly interacts with PTEN and catalyzes its K63-linked ubiquitination, modulating PTEN signaling and involving in cell survival and tumor growth in NSCLC (He et al., 2022)  |
| TRIM41 | K63     | innate antiviral response                                   | NF- $\kappa$ B associated cancer  | BCL10                        | TRIM41 modifies K63-linked polyubiquitination of BCL10, activating NF- $\kappa$ B and TBK1 signaling pathway (Yu et al., 2021)   |
| HECTD3 | K63     | regulating inflammation                                     | NF- $\kappa$ B associated cancer  | TRAF3                        | HECTD3 interacts with TRAF3 via K63-linked polyubiquitination, reducing inflammation and facilitating NF- $\kappa$ B inflammation pathway (Zhou et al., 2021)  |
| DZIP3  | K63     | driving cell cycle  | promoting cancer progression  | Cyclin D1                    | DZIP3 stabilizes Cyclin D1 by promoting K63-linked ubiquitination of Cyclin D1, driving cell cycle and cancer progression (Kolapalli et al., 2021)   |
| TRIM22 | K63     | activating NF- $\kappa$ B signaling                         | promoting glioblastoma tumor growth                                     | IKK $\gamma$                 | TRIM22 promotes K63-linked ubiquitination of IKK $\gamma$ , leading to degradation of I $\kappa$ B $\alpha$ and NF- $\kappa$ B activation (Ji et al., 2021)  |

### 2.2.2 OTULIN Disassembles M1-Linked Poly-Ubiquitin Chains

OTULIN, a methionine 1 (M1)-specific deubiquitinase (DUB), is a rare member of the OTU family of DUBs. Its proximal ubiquitin moiety cannot break down isopeptide linkages of ubiquitin chains, but can efficiently cleave peptide bonds present in the linear chains (Keusekotten et al., 2013; Rivkin et al., 2013). OTULIN presents negative regulation in the cellular process of immune homeostasis and inflammation (Damgaard et al., 2019). Depletion of OTULIN resulted in an increase in the formation of linear Ub chains and demonstrated proteasome dysregulation as the cause of NF- $\kappa$ B positive activation, which in turn restricts bacterial proliferation (Takiuchi et al., 2014; van Wijk et al., 2017). Moreover, the deficiency of OTULIN led to the inability to remove M1-linked poly-ubiquitin signals, which are typically conjugated by the LUBAC, resulting in LUBAC degradation and dysregulation of TNF signaling and cell death (Tokunaga, 2013; Damgaard et al., 2016). Notably, the function of LUBAC is controlled by cylindromatosis (CYLD), which interacts with the LUBAC core subunit HOIP to generate Met1-linked ubiquitin. However, this interaction can be weakened by the Met1-Ub-specific deubiquitinase OTULIN. In addition, through the deubiquitinase function of OTULIN, LUBAC can regulate Met1-Ub to ensure an advisable response to innate immune activity (Elliott et al., 2016; Hrdinka et al., 2016). The CYLD/TRAF3 complex has also been reported to regulate NF- $\kappa$ B-mediated inflammation and interferon signaling, which defines a subset of head and neck cancers that harbor human papillomavirus (Chen T. et al., 2017). Notably, accumulating evidence has manifested that linear ubiquitin chains play essential roles in ensuring appropriate activity of inflammatory responses and innate immune signaling (Tokunaga and Iwai, 2012; Tokunaga, 2013; Jing et al., 2017). The linear ubiquitination of cFLIP is directly induced by RNF31, a catalytic subunit of LUBAC, at Lys-351 and Lys-353, contributing to TNF $\alpha$ -induced apoptosis, thereby protecting cells from apoptosis (Tang et al., 2018) (Table 3).

### 2.2.3 Linear Ubiquitination in Tumorigenesis

M1-linked ubiquitination, specifically N-terminal Met1-linked ubiquitination, is able to form eight different inter-ubiquitin linkages via its N-terminal methionine (M1). It can be specifically catalyzed by LUBAC. For instance, EGFR recruits PKP2 to the plasma membrane and cooperates with LUBAC (HOIP), activating linear ubiquitination of NEMO, which is critical for tumor cell proliferation (Hua et al., 2021). Also, Epsins 1/2 promotes NEMO linear ubiquitination via LUBAC, driving breast cancer development (Song et al., 2021). Moreover, LUBAC (SHARPIN) regulated  $\beta$ -catenin activity through linear ubiquitination, promoting gastric tumorigenesis (Zhang et al., 2021). Interestingly, OTULIN exclusively cleaves M1-linked ubiquitination and exhibits a high affinity for linear ubiquitination. LUBAC and linear ubiquitination have been found to be relevant to TNF signaling. Interestingly, OTULIN was shown to remove linear ubiquitination from LUBAC-modified proteins, which is critical for various cellular activities (Draber et al., 2015). Moreover, the deubiquitinating

enzyme, CYLD, was also identified for disassembly Met1-linked-Ub (mostly the immune system). A previous report demonstrated that modification of proliferating cell nuclear antigen (PCNA) induces apoptosis and inhibits tumor growth through the linear ubiquitin chain (Qin et al., 2018).

## 2.3 Lys63 Linkage

### 2.3.1 The Writer Enzymes for the Lys63 Linkage

It has been well established that Lys63 chains (ubiquitin chains topology lysine 63 poly-ubiquitin linkages) regulate and trigger distinct cellular signaling, including kinase activation, signal transduction, protein trafficking, endocytosis and DNA repair (Spence et al., 1995; Hofmann and Pickart, 1999; Komander and Rape, 2012; Wu and Karin, 2015; Hrdinka et al., 2016). The E2 conjugating enzyme complex Ubc13/Uev1A preferentially assembles the K63-pUb chain (Smith et al., 2013; Zhang et al., 2018). Furthermore, Tax can be recruited to K63-pUb by E3 ligase RNF8 and Ubc13/Uev1A, which allows the activation of TGF $\beta$ -activating kinase 1 (TAK1), followed by multiple downstream signaling pathways such as the IKK and JNK pathways. These ultimately lead to DNA damage repair, cytokinesis, and the genomic instability in ATL cells (Ho et al., 2015; Lee et al., 2017). Similarly, tumor necrosis factor receptor associated factor 6 (TRAF6) can interact with the E2 conjugation enzyme Ubc13/Uev1A in a RING-dependent manner, catalyzing Lys63-linked TRAF6 auto-ubiquitination. This activates IKK and NF- $\kappa$ B, thereby promoting TAK1 and IKK to trigger spontaneous osteoclast differentiation (Lamothe et al., 2007a; Lenoir et al., 2021). Another report also showed that TRAF6, in a RING-dependent fashion, catalyzed auto-ubiquitination by conjugating with Ubc13/Uev1A, activating the AKT pathway, and promoting cell migration in breast cancer (Lamothe et al., 2007b; Niu et al., 2021).

The Ubc13/Uev1A complex has been shown to conjugate Lys63-linked poly-ubiquitination of substrate proteins, which contribute to breast cancer metastasis via NF- $\kappa$ B signaling regulation (Wu Z. et al., 2014; 2017). It is also reported that Ubc13 can catalyze K63-linked protein poly-ubiquitination, which is indispensable for the activation of non-SMAD signaling by TAK1 and p38, whose activity controls breast cancer metastatic spread and lung colonization (Wu X. et al., 2014). Interacting with Ubc13, RNF213 mediates autoubiquitination and controls inflammatory responses and angiogenic activities (Habu and Harada, 2021). RNF8 was demonstrated to activate Ubc13 and recruit K63-linked poly-ubiquitin conjugation to histones H2A/H2AX, thus contributing to breast cancer predisposition (Vuorela et al., 2011). Moreover, RNF8 utilizes the RING domain, mediating Lys63-linked Ub chains, which is required for DNA double-strand break (DSB) signaling and the downstream BRCA1 tumor suppressor recruitment or lung cancer cell proliferation (Hodge et al., 2016; Xu et al., 2021). Inhibiting the Ubc13/Uev1A complex specifically, which is critical for Lys63-linked ubiquitination, promotes ubiquitin conjugation at the Lys147 site, thereby upregulating NF- $\kappa$ B signaling in multiple myeloma and other

cancers (Gallo et al., 2014). UbcH6 and NEDD4 regulate angiogenesis and tumor growth (Sun et al., 2021). Ube2w accompanies the E3 ligase RNF4 function in distinct DNA repair pathways through Lys63-linked chains and BRIC6 (also named BRUCE) acting on K63-linked ubiquitinylate in unstimulated cells, which regulates the DNA double-strand break response through bridging USP8 and BRCT-repeat inhibitor of hTERT expression (BRIT1) in a deubiquitination manner (Ge et al., 2015; Maure et al., 2016). Poly-ubiquitination of histone H1 depends on Ubc13 and RNF8, which prolong pre-existing ubiquitin modifications to K63-linked chains, thereby stimulating RNF8-Ubc13 mediated DNA damage response (Mandemaker et al., 2017). SKP2 triggers non-proteolytic K63-linked ubiquitination, which is crucial for cancer initiation and progression by positively regulating cancer cell survival and glycolysis. The depletion of SKP2 restricts cancer stem cell proliferation and survival (Chan et al., 2013). However, the non-proteolytic K63-linked ubiquitination triggered by SKP2 can be reversed and modulated by deubiquitinase ovarian tumor domain-containing protein 1 (OTUD1) (Yao et al., 2018) and USP10 (Liao et al., 2019). USP10 has been identified as a novel deubiquitinase of SKP2 that modulates and stabilizes SKP2. Indeed, USP10 can recognize and remove Lys63-linked ubiquitin chains from Bcr-Abl, leading to positive activities in chronic myeloid leukemia cells. OTUD1 interacts with p53 and is essential for constant stabilization of p53. Its overexpression dramatically induces the cell cycle and apoptosis (Mevisen et al., 2013; Piao et al., 2017). In addition, several more DUBs have been defined as having linkage specificity for Lys63-linked ubiquitination. CYLD, a tumor suppressor, inhibits NF- $\kappa$ B signaling by cleaving K63-linked ubiquitination of NEMO/IKK $\gamma$ , thus reducing its stability and averting the IKK complex from phosphorylation of I $\kappa$ B (Chen T. et al., 2017). Furthermore, CYLD was associated with the catalytic LUBAC subunit HOIP to counteract Lys63-Ub and Met1-Ub conjugation to receptor-interacting protein kinase 2 (RIPK2), leading to restriction of innate immune signaling and cytokine production (Hrdinka et al., 2016). TRAF-binding protein domain (TRABID) was demonstrated to bind and cleave Lys63-linked ubiquitin moieties on APC tumor suppressor substrates, which led to the disruption of APC and activation of Wnt signaling in colorectal cancer cell lines (Tran et al., 2008). The USP17/TRAF2/TRAF3 complex acts to stabilize its client proteins, enhancing inflammatory responses and stemness in lung cancer cells (Lu et al., 2018). USP20 deubiquitinates TRAF6 and Tax, and may function as a key regulator of adult T cell leukemia (ATL) leukemogenesis through suppressing NF- $\kappa$ B activation (Yasunaga et al., 2011). In terms of actively removing Lys63-linked poly-ubiquitin chains on G $\beta$ L, OTUD7B appears to be the primary regulator in governing mTOR2 integrity, which is essential for NF- $\kappa$ B activation in cancer biology (Wang B. et al., 2017). Further detailed analysis of Lys63-linked ubiquitination is required to better understand how ubiquitin chains function in numerous cellular events. However, identification of this linkage will provide an ideal point to understand the potential mechanism in the cellular regulation of the incoming ubiquitin research (Table 1, 2).

### 2.3.2 Disassembly of Lys63 Poly-Ubiquitin Chains by DUBs

Although a number of DUBs have been identified, the disassembly ability of the substrate and its physiological role, especially for Lys63 poly-ubiquitin chains, is poorly defined. Previous studies have shown that K63-specific DUBs, A20, and CYLD present the anti-tumor effects by regulating NF- $\kappa$ B signaling (Heyninck and Beyaert, 2005; Sun, 2010; Meng et al., 2021; Zhu et al., 2021). USP10 reduces K63-linked ubiquitination chains and functions as a tumor suppressor (He et al., 2021). Stimulated by the pro-inflammatory cytokine TNF, A20 exerts two major functions: sequential de-ubiquitination and catalyzing receptor-interacting protein (RIP), resulting in downregulation of NF- $\kappa$ B signaling. Absence of A20 led to hyper-accumulation of Lys63 poly-ubiquitin in NF- $\kappa$ B signaling. Like A20, CYLD disassembles the Lys63 poly-ubiquitin recruiting signal and degrades IKK $\kappa$ . In this way, USP4 and OTUD1 cleaves K63-linked ubiquitination chains, and regulates downstream NF- $\kappa$ B signaling (Fan et al., 2011; Liang et al., 2013; Wu et al., 2022) (Table 3).

### 2.3.3 The Multifaceted Functions of Lys63 Poly-Ubiquitin Chains in Cancer

Lys63-linked chains have been described as non-degradative signals and are the core of inflammatory system and the NF- $\kappa$ B pathway, which is involved in either mono-ubiquitination or poly-ubiquitination. TGF $\beta$  type I receptor (TBRI) was shown to interact with TRAF6 through Lys63-dependent poly-ubiquitination in order to promote tumor invasion (Mu et al., 2011). It has also been reported that polo-like kinase 1 (PLK1) phosphorylates kruppel-like factor 4 (KLF4), leading to the recruitment of TRAF6, resulting in Lys63-linked ubiquitination and promoting nasopharyngeal tumorigenesis (Mai et al., 2019). Moreover, TRAF6 functions in hepatocarcinogenesis by regulating TRAF6/HDAC3/c-Myc signal axis. Mechanistically, TRAF6 interacts with histone deacetylase 3 (HDAC3) via Lys63-linked ubiquitin chains, which enhances c-Myc stability, and its overexpression has been proposed to facilitate HCC progression (Wu et al., 2020). Similarly, HUWE1 promotes K63-linked ubiquitination of c-Myc, thereby promoting retinoblastoma cell proliferation (Dai et al., 2021). It was also demonstrated that ASK1-dependent phosphorylation and lys63-linked poly-ubiquitination of LKB1 is critical for its activation. Tankyrase and RNF146 act as upstream regulators of the LKB1/AMPK pathway to promote Lys63-linked ubiquitination resulting in AMPK activation and tumor suppression (Li et al., 2019). Lys63-linked non-proteolytic ubiquitination of Hexokinase 2 by HectH9 effectively disrupts glycolysis activation and facilitates apoptosis in prostate cancer cells (Lee H. J. et al., 2019). SET domain bifurcated histone lysine methyltransferase 1 (SETDB1) methylates Akt and functions as a scaffold to recruit JMJD2A, which then binds TRAF6 and Skp2-SCF to the Akt complex, thereby promoting tumor development in lung cancer (Wang et al., 2019). Notably, numbers of E3 ligase have been discovered to catalyze K63-linked ubiquitination chains, which are involved in activating NF- $\kappa$ B or Akt signaling pathway, acting as tumor promoters and facilitating

tumor progression (Ji et al., 2021; Xu et al., 2021; Yu et al., 2021; Zhou et al., 2021; Zhu et al., 2021; He et al., 2022). Lys63 poly-ubiquitin chains were also found to be involved in driving the cell cycle and promoting cancer progression (Kolapalli et al., 2021). In contrast, tripartite motif-containing 31 (TRIM31) directly ubiquitinates p53 via K63-linked ubiquitination, resulting in tumor-suppressing effects (Guo et al., 2021) (Table 2).

## 2.4 Lys6 Linkages–Tumor Suppressors

Lys6-linked ubiquitin chains are less abundant in resting cells and their functional implications are unclear (Durcan et al., 2014; Michel et al., 2017). It has been reported that BRCA1 can be autoubiquitinated and then bound by UBX domain containing protein 1 (UBXN1) through K6-linked poly-ubiquitin chains. Interestingly, UBXN1 regulates BRCA1 expression upon ubiquitination. Autoubiquitinated forms of BRCA1 act as tumor suppressors and inhibit their enzymatic function (Wu-Baer et al., 2010). Moreover, BRCA1 auto-ubiquitination occurs in a way that the BRCA1/BARD1 complex conducts polymerization by conjugation with K6-linked polymers, which imparts cellular properties to its natural enzymatic substrates (Wu-Baer et al., 2003), further linking BRCA1 auto-ubiquitination to the tumor suppressor. C terminus HSC70-interacting protein (CHIP) was reported to bind to fas-associated protein with death domain (FADD) to induce K6-linked poly-ubiquitination of FADD, which was demonstrated to be essential for the prevention of cell death (Seo et al., 2018).

It has also been revealed that community-based learning collaborative (CBLC) assembles K6- and K11-linked poly-ubiquitin on EGFR and positively regulates its stability. The sustained activation of EGFR is largely dependent on CBLC-mediated ubiquitination, and its dysregulation is preferentially destined for membrane recycling, which plays an important role in non-small-cell lung carcinoma (NSCLC) progression (Hong et al., 2018).

Structural findings indicated that USP30 efficiently cleaves Lys6-linked Ub chains, which abrogates parkin-mediated Ub-chain formation in mitochondria. Dysfunction of this novel distal Ub-recognition mechanism is associated with physiological disorders such as hepatocellular carcinoma (Sato et al., 2017; Wang et al., 2022). Although an extraordinary progress has been made over the last 2 decades, the detailed functional consequences of Lys6 modifications require further investigation (Tables 2, 3).

## 2.5 Lys27 linkages—emerging Tumor Promoter

Emerging investigations have demonstrated that K27-linked poly-ubiquitination is crucial for promoting cancer development, such as facilitating cell proliferation, invasion, and metastasis (Peng et al., 2011; Yin et al., 2019). ITCH, an E3 ubiquitin ligase, was shown to generate Lys27-linked poly-ubiquitination of the transcription factor TIEG1, which inhibits TIEG1 nuclear translocation and subsequent Treg development (Peng et al., 2011). Additionally, TGF- $\beta$ , in the presence of cytokine IL-6, can efficiently promote mono- and poly-ubiquitination of TIEG1 and modulate Treg/Th17

differentiation (Peng et al., 2011). Yet, it seems that tumor immunity in TIEG1<sup>-/-</sup> mice were apparently enhanced by hampering Treg development and increasing Th17 response, suggesting its pro-tumor effects (Peng et al., 2011). In the presence of proinflammatory cytokines, ITCH can catalyze BRAF to disrupt 14–3–3-mediated inhibition of BRAF kinase activity, resulting in MEK/ERK signaling activation. This ubiquitin function plays an essential role in supporting the proliferation and invasion abilities of melanoma cells (Yin et al., 2019). Another report connected Lys27-linked ubiquitination to melanoma cell invasive properties, in which HECT domain and ankyrin repeat containing E3 ubiquitin protein ligase 1 (HACE1) decorated fibronectin with Lys27 Ub moieties (El-Hachem et al., 2018). As mentioned in this study, upregulation of fibronectin in turn modulates the transcription levels of integrin subunit alpha V (ITGAV) and integrin  $\beta$ 1 (ITGB1), leading to an increased invasive power of melanoma cells (El-Hachem et al., 2018). TRAF4 was found to increase TrkA kinase activity through K27- and K29-linked ubiquitination upon nerve growth factor (NGF) stimulation, followed by the recruitment of downstream adaptor proteins and increased metastasis in prostate cancer (Singh et al., 2018).

Strikingly, reports have also revealed that Lys27-linked ubiquitination is implicated in the DNA damage response (DDR) and innate immunity. A previous study demonstrated that RNF168 can target the N-terminal tail of histones H2A/H2A.X, generating Lys27-linked ubiquitin chains. In response to DNA damage, K27 ubiquitination is the major source of PTMs that mark chromatin. Meanwhile, the activation of DDR can be inhibited by mutation of K27, which hinders the localization of 53BP1 and BRCA1 to DDR foci (Gatti et al., 2015). DDR is responsible for the recognition, signal transduction and repair of DNA damage. The inactivation of DDR will result in the accumulation of cell mutations and the increase of genomic instability, which play an essential role in cancer initiation (Klinakis et al., 2020). It is also indicated that K27 and the related DDR mediators may be the potential targets for the development of anti-tumor drugs. cGAS is subjected to K27 poly-ubiquitination by RNF185, facilitating cGAS-mediated innate immune response. TRIM-mediated ubiquitination by binding to residue K27 activates TBK1 recruitment to mitochondrial antiviral signaling (MAVS) and promotes innate immunity (Wang Q. et al., 2017). TRIM31 and TRIM40 can also mediate K27-linked ubiquitination, thereby regulating innate and adaptive immunity (Wang X. et al., 2021; Shen et al., 2021). Specifically, TRIM40 interacts with Rik3, resulting in RIG-I and MDA5 degradation through K27-linked ubiquitination, and negatively regulates innate immunity (Shen et al., 2021). Nevertheless, TRIM31 catalyzes K27-linked ubiquitination of SYK, facilitating antifungal immunity (Wang X. et al., 2021). Additionally, auto-ubiquitination of TRIM23 through K27-linked ubiquitination was found to mediate autophagy via activation of TBK1 (Sparrer et al., 2017). Moreover, E3 ubiquitin ligase Hectd3 decorates Stata3 with non-degradative K27-linked poly-ubiquitin chains and Malt1 with K27/K29-linked poly-ubiquitin chains, leading to signaling-related ubiquitination in neuroinflammation (Cho

et al., 2019) (Table 3). In fact, the innate immune system plays a key role in the formation of tumor. The response of it is generally affected by a variety of immune cells and cytokines in tumor microenvironment (Wenbo and Wang, 2017). In this way, it is beneficial to clarify the regulatory effects of K27 on tumor innate immunity for understanding the mechanisms of cancer.

## 2.6 Lys29 and Lys33 linkage—possible Tumor Promoters

Recent proteomic data have identified the role of Lys29- and Lys33-linked ubiquitin chains in various biological processes, including the control of AMPK-mediated mitochondrial function and Wnt-induced transcription signaling (Jiang et al., 2015; Nusse and Clevers, 2017). AMPK-related kinases, AMPK-related kinase 5 (ARK5, also known as NUA1) and microtubule-affinity-regulating kinase 4 (MARK4), which are mediated by unconventional Lys29/Lys33 linkage, are involved in cell polarity and proliferation (Al-Hakim et al., 2008). One of the underlying mechanisms is that USP9X specifically identifies Lys29/Lys33-conjugated ubiquitin chains on NUA1 and MARK4 (Al-Hakim et al., 2008). However, it should be noted that the ubiquitination of NUA1 and MARK4 suppresses their phosphorylations rather than restores their stability and facilitates LKB1 activation. Interestingly, the TRABID core domain N-terminal Npl4-like zinc finger (NAZF1) preferentially hydrolyzes K29/K33-linked diUb, and this novel AnkUBD displays TRABID linkage specificity (Licchesi et al., 2011). Additionally, TRABID interacts with APC tumor suppressor protein, recruits TCF target genes, and activates their transcription in colorectal cancer cells (Tran et al., 2008). The OUT family DUB TRABID preferentially abolishes Smad ubiquitination regulatory factor 1-induced K29/K33-linked poly-ubiquitin chains from UV radiation resistance associated gene (UVRAG), thereby promoting autophagosome maturation and inhibiting cell proliferation in hepatocellular cancer (Feng et al., 2019). Considering its linkage specificity, it would be insightful to explore the potential positive regulation of TRABID in cancer tumorigenesis via these two pathways.

K29-linked ubiquitin chains play important roles in driving cancer invasion and metastasis, as well as in the positive regulation of immunity (Singh et al., 2018; Cho et al., 2019; Gao et al., 2021). Several studies have also manifested that, with the assistance of Cbl-b and ITCH, T cell receptor-zeta (TCR-zeta) was decorated with a K33 linkage, accompanied by positively activated T cells and immune responses (Huang et al., 2010). In addition, OTUD1 directly deubiquitinates the inhibitor SMAD7 of TGF- $\beta$  pathway and aborates Lys33-linked poly-ubiquitin chains, which inhibits cell stemness and suppresses metastasis (Zhang et al., 2017). Acetylation of GLDC inhibits its enzymatic activity and facilitates K33-linked ubiquitination by NF-X1, thereby suppressing glioma tumor growth (Liu et al., 2021). Ultimately, Lys29- and Lys33-linked ubiquitin chains appear to have more complicated cellular functions that remain to be characterized (Table 2, 3).

## 2.7 Mixed Linkage and Branched Poly-Ubiquitination

Mixed linkage chains send mixed signaling messages that can be identified by different linkage-specific receptors. Our understanding of the mixed linkages and branched poly-ubiquitin is limited. Previous reports have elucidated that Tax in combination with UbcH2, Ubc5c, or UbcH7, can catalyze the construction of free mixed-linkage poly-ubiquitin chains, which are responsible for IKK-NF- $\kappa$ B activation and induction of T cell transformation (Wang et al., 2016). Furthermore, Brcc 36 isopeptidase complex (BRISC), the JAMM/MPN + family of DUBs, preferentially cleaves K63 linkages within mixed-linkage chains. In addition, RING1B was found to generate atypical mixed poly-ubiquitin chains and mediate mono-ubiquitination of H2A (Ben-Saadon et al., 2006). A recent study showed that SPOP triggered mixed-linkage ubiquitination of MyD88 in human lymphoma cells and mouse HSCs, suggesting that the SPOP-MyD88 pathway plays a critical role in hematopoietic neoplasms (Jin et al., 2020).

## 3 CONCLUSION AND FUTURE PERSPECTIVES

Non-proteolytic ubiquitination, the molecular switch in cell fate regulation, plays a crucial role in post-translational protein modifications. Diverse ubiquitination enzymes are essential for ubiquitination linkages that are necessary for normal metabolism and physiological functions. Meanwhile, it is also the root cause of physiological disorders, such as cancer. The aberrant regulation of the UPS is typically achieved by ubiquitination enzymes, DUBs, 20S proteasome catalytic core particles and 19S proteasome regulatory particles (Rape, 2018). In general, specific ubiquitination enzymes determine ubiquitin linking with one of the seven lysine residues to form distinctive styles of poly-ubiquitin chains, deciding the fate of substrate proteins. It is noteworthy that the ubiquitin proteasome system functions as a theoretical target for drug screening, and the study of ubiquitination will provide more insights into the development of anti-tumor drugs. Recently, the quantities of E3s inhibitors have already been processed in preclinical models of cancer immunotherapy. It has been shown that E3 enzyme Smac mimetics (SMs) were the promising immune modulators for cancer therapy as the antagonists targeting E3 ligases IAPs (Cossu et al., 2019). The small molecule inhibitor AMG-232, targeting another E3-ubiquitin ligase oncogenic mouse double minute 2 homolog (MDM2), was shown to strengthen T cell killing of cancer cells, especially when combined with an anti-PD-1 monoclonal antibody (Sahin et al., 2020). Several other proteasome inhibitors have been confirmed to be clinically effective against malignancies as well. The neddylation (NAE) inhibitor pevonedistat has already been tested in multiple clinical trials and has shown positive effects in patients with AML or advanced solid tumors (Barghout and Schimmer, 2021). Small molecule inhibitors based on deubiquitinase have been widely used in experimental anti-tumor therapy, most of which are still

in the preclinical research stage. Due to dose-limiting toxicity, the first deubiquitinase inhibitor VLX1570 was terminated in the clinical trial phase. No other deubiquitinase inhibitors have been approved for clinical studies since then. Also, the discovered deubiquitinase inhibitors related to tumor treatment are mainly concentrated in the USP family, and the relationship between the inhibitors of non-USP family members and the treatment of malignant tumors needs to be further studied (Zhang et al., 2022). Interestingly, the innovative approaches of proteolysis targeting chimeras (PROTACs) and molecular glues might facilitate clinical cancer therapy (Cruz Walma et al., 2022; Kung and Weber, 2022; Zhou and Xu, 2022). Consequently, it is essential further to investigate the role of ubiquitination enzymes in tumorigenesis. Meanwhile, targeting different ubiquitination, including K6-, K27-, K29-, K33-, K63-, and M1-linked poly-ubiquitination, may also be one of the directions for cancer drug discovery. Although great progresses have been achieved with the development of anti-cancer drugs aimed at the UPS, numerous challenges still stand in the way. Some candidate inhibitors have emerged as drug resistant or have limited efficacy in patients. Despite these

questions, the discovery of new drugs targeting single or multiple segments of the UPS is still worthy of future research.

## AUTHOR CONTRIBUTIONS

Conceptualization, SJ and JH; writing original draft preparation, XY; writing review and editing, QL, FL, SJ and JH; visualization: XT and TY; supervision: SJ and JH; funding acquisition: SJ and XY. All authors have read and agreed to the published version of the manuscript.

## FUNDING

This work was supported in part by the National Natural Science Foundation of China (grant nos. 81873249 and 82074360) and the Young Taishan Scholars Program of Shandong Province (grant no. tsqn201909200) and Shandong Province Medicine and Health Project (grant no2019WS091).

## REFERENCES

- Aalto, A. L., Mohan, A. K., Schwintzer, L., Kupka, S., Kietz, C., and Walczak, H. (2019). M1-linked Ubiquitination by LUBEL Is Required for Inflammatory Responses to Oral Infection in *Drosophila*. *Cell Death Differ.* 26 (5), 860–876. doi:10.1038/s41418-018-0164-x
- Abbas, T., and Dutta, A. (2011). CRL4Cdt2: Master Coordinator of Cell Cycle Progression and Genome Stability. *Cell Cycle* 10 (2), 241–249. doi:10.4161/cc.10.2.14530
- Adhikari, A., Xu, M., and Chen, Z. J. (2007). Ubiquitin-mediated Activation of TAK1 and IKK. *Oncogene* 26 (22), 3214–3226. doi:10.1038/sj.onc.1210413
- Adhikary, S., Marinoni, F., Hock, A., Hulleman, E., Popov, N., Beier, R., et al. (2005). The Ubiquitin Ligase HectH9 Regulates Transcriptional Activation by Myc and Is Essential for Tumor Cell Proliferation. *Cell* 123 (3), 409–421. doi:10.1016/j.cell.2005.08.016
- Ahmed, N., Zeng, M., Sinha, I., Polin, L., Wei, W. Z., Rathinam, C., et al. (2011). The E3 Ligase Itch and Deubiquitinase Cyld Act Together to Regulate Tak1 and Inflammation. *Nat. Immunol.* 12 (12), 1176–1183. doi:10.1038/ni.2157
- Akturk, A., Wasilko, D. J., Wu, X., Liu, Y., Zhang, Y., Qiu, J., et al. (2018). Mechanism of Phosphoribosyl-Ubiquitination Mediated by a Single Legionella Effector. *Nature* 557 (7707), 729–733. doi:10.1038/s41586-018-0147-6
- Akutsu, M., Dikic, I., and Bremm, A. (2016). Ubiquitin Chain Diversity at a Glance. *J. Cell Sci.* 129 (5), 875–880. doi:10.1242/jcs.183954
- Al-Hakim, A. K., Zagorska, A., Chapman, L., Deak, M., Pegg, M., and Alessi, D. R. (2008). Control of AMPK-Related Kinases by USP9X and Atypical Lys(29)/Lys(33)-Linked Polyubiquitin Chains. *Biochem. J.* 411 (2), 249–260. doi:10.1042/bj20080067
- Ali, A. M., Pradhan, A., Singh, T. R., Du, C., Li, J., Wahengbam, K., et al. (2012). FAAP20: a Novel Ubiquitin-Binding FA Nuclear Core-Complex Protein Required for Functional Integrity of the FA-BRCA DNA Repair Pathway. *Blood* 119 (14), 3285–3294. doi:10.1182/blood-2011-10-385963
- Alpi, A. F., and Patel, K. J. (2009). Monoubiquitylation in the Fanconi Anemia DNA Damage Response Pathway. *DNA Repair (Amst)* 8 (4), 430–435. doi:10.1016/j.dnarep.2009.01.019
- An, J., Mo, D., Liu, H., Vena, M. S., Srivatsan, E. S., Massoumi, R., et al. (2008). Inactivation of the CYLD Deubiquitinase by HPV E6 Mediates Hypoxia-Induced NF- $\kappa$ B Activation. *Cancer Cell* 14 (5), 394–407. doi:10.1016/j.ccr.2008.10.007
- Barghout, S. H., and Schimmer, A. D. (2021). E1 Enzymes as Therapeutic Targets in Cancer. *Pharmacol. Rev.* 73 (1), 1–58. doi:10.1124/pharmrev.120.000053
- Bax, M., McKenna, J., Do-Ha, D., Stevens, C. H., Higginbottom, S., Balez, R., et al. (2019). The Ubiquitin Proteasome System Is a Key Regulator of Pluripotent Stem Cell Survival and Motor Neuron Differentiation. *Cells* 8 (6). doi:10.3390/cells8060581
- Behrends, C., and Harper, J. W. (2011). Constructing and Decoding Unconventional Ubiquitin Chains. *Nat. Struct. Mol. Biol.* 18 (5), 520–528. doi:10.1038/nsmb.2066
- Belgnaoui, S. M., Paz, S., Samuel, S., Goulet, M. L., Sun, Q., Kikkert, M., et al. (2012). Linear Ubiquitination of NEMO Negatively Regulates the Interferon Antiviral Response through Disruption of the MAVS-TRAF3 Complex. *Cell Host Microbe* 12 (2), 211–222. doi:10.1016/j.chom.2012.06.009
- Ben-Saadon, R., Zaaroor, D., Ziv, T., and Ciechanover, A. (2006). The Polycomb Protein Ring1B Generates Self Atypical Mixed Ubiquitin Chains Required for its *In Vitro* Histone H2A Ligase Activity. *Mol. Cell* 24 (5), 701–711. doi:10.1016/j.molcel.2006.10.022
- Bhatnagar, S., Gazin, C., Chamberlain, L., Ou, J., Zhu, X., Tushir, J. S., et al. (2014). TRIM37 Is a New Histone H2A Ubiquitin Ligase and Breast Cancer Oncoprotein. *Nature* 516 (7529), 116–120. doi:10.1038/nature13955
- Bomar, M. G., D'Souza, S., Bienko, M., Dikic, I., Walker, G. C., and Zhou, P. (2010). Unconventional Ubiquitin Recognition by the Ubiquitin-Binding Motif within the Y Family DNA Polymerases Iota and Rev1. *Mol. Cell* 37 (3), 408–417. doi:10.1016/j.molcel.2009.12.038
- Borghi, A., Haegman, M., Fischer, R., Carpentier, I., Bertrand, M. J. M., Libert, C., et al. (2018). The E3 Ubiquitin Ligases HOIP and cIAP1 Are Recruited to the TNFR2 Signaling Complex and Mediate TNFR2-Induced Canonical NF- $\kappa$ B Signaling. *Biochem. Pharmacol.* 153, 292–298. doi:10.1016/j.bcp.2018.01.039
- Brickner, J. R., Soll, J. M., Lombardi, P. M., Vagbo, C. B., Mudge, M. C., Oyeniran, C., et al. (2017). A Ubiquitin-dependent Signalling axis Specific for ALKBH-Mediated DNA Dealkylation Repair. *Nature* 551 (7680), 389–393. doi:10.1038/nature24484
- Buetow, L., and Huang, D. T. (2016). Structural Insights into the Catalysis and Regulation of E3 Ubiquitin Ligases. *Nat. Rev. Mol. Cell Biol.* 17 (10), 626–642. doi:10.1038/nrm.2016.91
- Cai, F., Chen, P., Chen, L., Biskup, E., Liu, Y., Chen, P. C., et al. (2014). Human RAD6 Promotes G1-S Transition and Cell Proliferation through Upregulation of Cyclin D1 Expression. *PLoS One* 9 (11), e113727. doi:10.1371/journal.pone.0113727
- Chakraborty, A., Diefenbacher, M. E., Mylona, A., Kassel, O., and Behrens, A. (2015). The E3 Ubiquitin Ligase Trim7 Mediates C-Jun/AP-1 Activation by Ras Signalling. *Nat. Commun.* 6, 6782. doi:10.1038/ncomms7782

- Chan, C. H., Morrow, J. K., Li, C. F., Gao, Y., Jin, G., Moten, A., et al. (2013). Pharmacological Inactivation of Skp2 SCF Ubiquitin Ligase Restricts Cancer Stem Cell Traits and Cancer Progression. *Cell* 154 (3), 556–568. doi:10.1016/j.cell.2013.06.048
- Chang, J. H., Xiao, Y., Hu, H., Jin, J., Yu, J., Zhou, X., et al. (2012). Ubc13 Maintains the Suppressive Function of Regulatory T Cells and Prevents Their Conversion into Effector-like T Cells. *Nat. Immunol.* 13 (5), 481–490. doi:10.1038/ni.2267
- Charlaftis, N., Suddason, T., Wu, X., Anwar, S., Karin, M., and Gallagher, E. (2014). The MEKK1 PHD Ubiquitinates TAB1 to Activate MAPKs in Response to Cytokines. *EMBO J.* 33 (21), 2581–2596. doi:10.15252/embj.201488351
- Chatr-Aryamontri, A., van der Sloot, A., and Tyers, M. (2018). At Long Last, a C-Terminal Bookend for the Ubiquitin Code. *Mol. Cell* 70 (4), 568–571. doi:10.1016/j.molcel.2018.05.006
- Chen, M., Zhao, Z., Meng, Q., Liang, P., Su, Z., Wu, Y., et al. (2020). TRIM14 Promotes Noncanonical NF-kappaB Activation by Modulating P100/p52 Stability via Selective Autophagy. *Adv. Sci. (Weinh)* 7 (1), 1901261. doi:10.1002/advs.201901261
- Chen, S., Jing, Y., Kang, X., Yang, L., Wang, D. L., Zhang, W., et al. (2017a). Histone H2B Monoubiquitination Is a Critical Epigenetic Switch for the Regulation of Autophagy. *Nucleic Acids Res.* 45 (3), 1144–1158. doi:10.1093/nar/gkw1025
- Chen, T., Zhang, J., Chen, Z., and Van Waes, C. (2017b). Genetic Alterations in TRAF3 and CYLD that Regulate Nuclear Factor kappaB and Interferon Signaling Define Head and Neck Cancer Subsets Harboring Human Papillomavirus. *Cancer* 123 (10), 1695–1698. doi:10.1002/cncr.30659
- Chen, Y., He, L., Peng, Y., Shi, X., Chen, J., Zhong, J., et al. (2015). The Hepatitis C Virus Protein NS3 Suppresses TNF- $\alpha$ -Stimulated Activation of NF-Kb by Targeting LUBAC. *Sci. Signal* 8 (403), ra118. doi:10.1126/scisignal.aab2159
- Chitale, S., and Richly, H. (2017). Timing of DNA Lesion Recognition: Ubiquitin Signaling in the NER Pathway. *Cell Cycle* 16 (2), 163–171. doi:10.1080/15384101.2016.1261227
- Cho, J. J., Xu, Z., Parthasarathy, U., Drashansky, T. T., Helm, E. Y., Zuniga, A. N., et al. (2019). Hctd3 Promotes Pathogenic Th17 Lineage through Stat3 Activation and Malt1 Signaling in Neuroinflammation. *Nat. Commun.* 10 (1), 701. doi:10.1038/s41467-019-08605-3
- Choudhary, C., Kumar, C., Gnäd, F., Nielsen, M. L., Rehman, M., Walther, T. C., et al. (2009). Lysine Acetylation Targets Protein Complexes and Co-regulates Major Cellular Functions. *Science* 325 (5942), 834–840. doi:10.1126/science.1175371
- Choudhry, H., Zamzami, M. A., Omran, Z., Wu, W., Mousli, M., Bronner, C., et al. (2018). Targeting microRNA/UHRF1 Pathways as a Novel Strategy for Cancer Therapy. *Oncol. Lett.* 15 (1), 3–10. doi:10.3892/ol.2017.7290
- Clague, M. J., Urbé, S., and Komander, D. (2019). Breaking the Chains: Deubiquitylating Enzyme Specificity Begets Function. *Nat. Rev. Mol. Cell Biol.* 20 (6), 338–352. doi:10.1038/s41580-019-0099-1
- Cole, A. J., Dickson, K. A., Liddle, C., Stitzaker, C., Shah, J. S., Clifton-Bligh, R., et al. (2021). Ubiquitin Chromatin Remodelling after DNA Damage Is Associated with the Expression of Key Cancer Genes and Pathways. *Cell Mol. Life Sci.* 78 (3), 1011–1027. doi:10.1007/s00018-020-03552-5
- Cooray, S. N., Guasti, L., and Clark, A. J. (2011). The E3 Ubiquitin Ligase Mahogunin Ubiquitinates the Melanocortin 2 Receptor. *Endocrinology* 152 (11), 4224–4231. doi:10.1210/en.2011-0147
- Cossu, F., Milani, M., Mastrangelo, E., and Lecis, D. (2019). Targeting the BIR Domains of Inhibitor of Apoptosis (IAP) Proteins in Cancer Treatment. *Comput. Struct. Biotechnol. J.* 17, 142–150. doi:10.1016/j.csbj.2019.01.009
- Cruz Walma, D. A., Chen, Z., Bullock, A. N., and Yamada, K. M. (2022). Ubiquitin Ligases: Guardians of Mammalian Development. *Nat. Rev. Mol. Cell Biol.* 23 (5), 350–367. doi:10.1038/s41580-021-00448-5
- Dai, H., Zeng, W., Zeng, W., Yan, M., Jiang, P., Li, Y., et al. (2021). HELZ2 Promotes K63-Linked Polyubiquitination of C-Myc to Induce Retinoblastoma Tumorigenesis. *Med. Oncol.* 39 (1), 11. doi:10.1007/s12032-021-01603-w
- Damgaard, R. B., Elliott, P. R., Swatek, K. N., Maher, E. R., Stepensky, P., Elpeleg, O., et al. (2019). OTULIN Deficiency in ORAS Causes Cell Type-specific LUBAC Degradation, Dysregulated TNF Signalling and Cell Death. *EMBO Mol. Med.* 11 (3). doi:10.15252/emmm.201809324
- Damgaard, R. B., Walker, J. A., Marco-Casanova, P., Morgan, N. V., Titheradge, H. L., Elliott, P. R., et al. (2016). The Deubiquitinase OTULIN Is an Essential Negative Regulator of Inflammation and Autoimmunity. *Cell* 166 (5), 1215–1230. doi:10.1016/j.cell.2016.07.019
- Dammer, E. B., Na, C. H., Xu, P., Seyfried, N. T., Duong, D. M., Cheng, D., et al. (2011). Polyubiquitin Linkage Profiles in Three Models of Proteolytic Stress Suggest the Etiology of Alzheimer Disease. *J. Biol. Chem.* 286 (12), 10457–10465. doi:10.1074/jbc.M110.149633
- Darling, S., Fielding, A. B., Sabat-Pospiech, D., Prior, I. A., and Coulson, J. M. (2017). Regulation of the Cell Cycle and Centrosome Biology by Deubiquitylases. *Biochem. Soc. Trans.* 45 (5), 1125–1136. doi:10.1042/bst20170087
- David, Y., Ternette, N., Edelmänn, M. J., Ziv, T., Gayer, B., Sertchook, R., et al. (2011). E3 Ligases Determine Ubiquitination Site and Conjugate Type by Enforcing Specificity on E2 Enzymes. *J. Biol. Chem.* 286 (51), 44104–44115. doi:10.1074/jbc.M111.234559
- Deng, R., Zhang, H. L., Huang, J. H., Cai, R. Z., Wang, Y., Chen, Y. H., et al. (2020). MAPK1/3 Kinase-dependent ULK1 Degradation Attenuates Mitophagy and Promotes Breast Cancer Bone Metastasis. *Autophagy*, 1–19. doi:10.1080/15548627.2020.1850609
- Di Lello, P., and Hymowitz, S. G. (2016). Unveiling the Structural and Dynamic Nature of the Ubiquitin Code. *Structure* 24 (4), 498–499. doi:10.1016/j.str.2016.03.013
- Dikic, I., Wakatsuki, S., and Walters, K. J. (2009). Ubiquitin-binding Domains - from Structures to Functions. *Nat. Rev. Mol. Cell Biol.* 10 (10), 659–671. doi:10.1038/nrm2767
- Dittmar, G., and Selbach, M. (2017). Deciphering the Ubiquitin Code. *Mol. Cell* 65 (5), 779–780. doi:10.1016/j.molcel.2017.02.011
- Dittmar, G., and Winkhofer, K. F. (2019). Linear Ubiquitin Chains: Cellular Functions and Strategies for Detection and Quantification. *Front. Chem.* 7, 915. doi:10.3389/fchem.2019.00915
- Dondelinger, Y., Darding, M., Bertrand, M. J., and Walczak, H. (2016). Poly-ubiquitination in TNFR1-Mediated Necroptosis. *Cell Mol. Life Sci.* 73 (11–12), 2165–2176. doi:10.1007/s00018-016-2191-4
- Dong, X., Gong, Z., Lu, Y. B., Liu, K., Qin, L. Y., Ran, M. L., et al. (2017). Ubiquitin S65 Phosphorylation Engenders a pH-Sensitive Conformational Switch. *Proc. Natl. Acad. Sci. U. S. A.* 114 (26), 6770–6775. doi:10.1073/pnas.1705718114
- Draber, P., Kupka, S., Reichert, M., Draberova, H., Lafont, E., de Miguel, D., et al. (2015). LUBAC-recruited CYLD and A20 Regulate Gene Activation and Cell Death by Exerting Opposing Effects on Linear Ubiquitin in Signaling Complexes. *Cell Rep.* 13 (10), 2258–2272. doi:10.1016/j.celrep.2015.11.009
- Durcan, T. M., Tang, M. Y., Pérusse, J. R., Dashti, E. A., Aguilera, M. A., McLelland, G. L., et al. (2014). USP8 Regulates Mitophagy by Removing K6-Linked Ubiquitin Conjugates from Parkin. *Embo J.* 33 (21), 2473–2491. doi:10.15252/embj.201489729
- Eger, S., Scheffner, M., Marx, A., and Rubini, M. (2010). Synthesis of Defined Ubiquitin Dimers. *J. Am. Chem. Soc.* 132 (46), 16337–16339. doi:10.1021/ja1072838
- El-Hachem, N., Habel, N., Naiken, T., Bziouche, H., Cheli, Y., Beranger, G. E., et al. (2018). Uncovering and Deciphering the Pro-invasive Role of HACE1 in Melanoma Cells. *Cell Death Differ.* 25 (11), 2010–2022. doi:10.1038/s41418-018-0090-y
- Elliott, P. R., Leske, D., Hrdinka, M., Bagola, K., Fiil, B. K., McLaughlin, S. H., et al. (2016). SPATA2 Links CYLD to LUBAC, Activates CYLD, and Controls LUBAC Signaling. *Mol. Cell* 63 (6), 990–1005. doi:10.1016/j.molcel.2016.08.001
- Elton, L., Carpentier, I., Verhelst, K., Staal, J., and Beyaert, R. (2015). The Multifaceted Role of the E3 Ubiquitin Ligase HOIL-1: beyond Linear Ubiquitination. *Immunol. Rev.* 266 (1), 208–221. doi:10.1111/immr.12307
- Emanuelli, A., Manikoth Ayyathan, D., Koganti, P., Shah, P. A., Apel-Sarid, L., Paolini, B., et al. (2019). Altered Expression and Localization of Tumor Suppressive E3 Ubiquitin Ligase SMURF2 in Human Prostate and Breast Cancer. *Cancers (Basel)* 11 (4). doi:10.3390/cancers11040556
- Emmerich, C. H., Schmukle, A. C., and Walczak, H. (2011). The Emerging Role of Linear Ubiquitination in Cell Signaling. *Sci. Signal* 4 (204), re5. doi:10.1126/scisignal.2002187
- Fan, L., Xu, S., Zhang, F., Cui, X., Fazli, L., Gleave, M., et al. (2020). Histone Demethylase JMJD1A Promotes Expression of DNA Repair Factors and Radio-Resistance of Prostate Cancer Cells. *Cell Death Dis.* 11 (4), 214. doi:10.1038/s41419-020-2405-4

- Fan, Y. H., Yu, Y., Mao, R. F., Tan, X. J., Xu, G. F., Zhang, H., et al. (2011). USP4 Targets TAK1 to Downregulate TNF $\alpha$ -Induced NF- $\kappa$ B Activation. *Cell Death Differ.* 18 (10), 1547–1560. doi:10.1038/cdd.2011.11
- Feng, X., Jia, Y., Zhang, Y., Ma, F., Zhu, Y., Hong, X., et al. (2019). Ubiquitination of UVRAG by SMURF1 Promotes Autophagosome Maturation and Inhibits Hepatocellular Carcinoma Growth. *Autophagy* 15 (7), 1130–1149. doi:10.1080/15548627.2019.1570063
- Fiil, B. K., Damgaard, R. B., Wagner, S. A., Keusekotten, K., Fritsch, M., Bekker-Jensen, S., et al. (2013). OTULIN Restricts Met1-Linked Ubiquitination to Control Innate Immune Signaling. *Mol. Cell* 50 (6), 818–830. doi:10.1016/j.molcel.2013.06.004
- Fiil, B. K., and Gyrd-Hansen, M. (2014). Met1-linked Ubiquitination in Immune Signalling. *Febs J.* 281 (19), 4337–4350. doi:10.1111/febs.12944
- Fottner, M., Brunner, A. D., Bittl, V., Horn-Ghetko, D., Jussupow, A., Kaila, V. R. I., et al. (2019). Site-specific Ubiquitylation and SUMOylation Using Genetic-Code Expansion and Sortase. *Nat. Chem. Biol.* 15 (3), 276–284. doi:10.1038/s41589-019-0227-4
- Fradet-Turcotte, A., Canny, M. D., Escibano-Díaz, C., Orthwein, A., Leung, C. C., Huang, H., et al. (2013). 53BP1 Is a Reader of the DNA-Damage-Induced H2A Lys 15 Ubiquitin Mark. *Nature* 499 (7456), 50–54. doi:10.1038/nature12318
- Fu, Q. S., Song, A. X., and Hu, H. Y. (2012). Structural Aspects of Ubiquitin Binding Specificities. *Curr. Protein Pept. Sci.* 13 (5), 482–489. doi:10.2174/138920312802430581
- Galan, J. M., and Haguenaer-Tsapis, R. (1997). Ubiquitin Lys63 Is Involved in Ubiquitination of a Yeast Plasma Membrane Protein. *Embo J.* 16 (19), 5847–5854. doi:10.1093/emboj/16.19.5847
- Gallo, L. H., Meyer, A. N., Motamedchaboki, K., Nelson, K. N., Haas, M., and Donoghue, D. J. (2014). Novel Lys63-Linked Ubiquitination of IKK $\beta$  Induces STAT3 Signaling. *Cell Cycle* 13 (24), 3964–3976. doi:10.4161/15384101.2014.988026
- Gao, P., Ma, X., Yuan, M., Yi, Y., Liu, G., Wen, M., et al. (2021). E3 Ligase Nedd4l Promotes Antiviral Innate Immunity by Catalyzing K29-Linked Cysteine Ubiquitination of TRAF3. *Nat. Commun.* 12 (1), 1194. doi:10.1038/s41467-021-21456-1
- Gatti, M., Pinato, S., Maiolica, A., Rocchio, F., Prato, M. G., Aebersold, R., et al. (2015). RNF168 Promotes Noncanonical K27 Ubiquitination to Signal DNA Damage. *Cell Rep.* 10 (2), 226–238. doi:10.1016/j.celrep.2014.12.021
- Ge, C., Che, L., Ren, J., Pandita, R. K., Lu, J., Li, K., et al. (2015). BRUCE Regulates DNA Double-Strand Break Response by Promoting USP8 Deubiquitination of BRIT1. *Proc. Natl. Acad. Sci. U. S. A.* 112 (11), E1210–E1219. doi:10.1073/pnas.1418335112
- Geisler, S., Holmström, K. M., Skujat, D., Fiesel, F. C., Rothfuss, O. C., Kahle, P. J., et al. (2010). PINK1/Parkin-mediated Mitophagy Is Dependent on VDAC1 and p62/SQSTM1. *Nat. Cell Biol.* 12 (2), 119–131. doi:10.1038/ncb2012
- Gilberto, S., and Peter, M. (2017). Dynamic Ubiquitin Signaling in Cell Cycle Regulation. *J. Cell Biol.* 216 (8), 2259–2271. doi:10.1083/jcb.201703170
- Glauser, L., Sonnay, S., Stafa, K., and Moore, D. J. (2011). Parkin Promotes the Ubiquitination and Degradation of the Mitochondrial Fusion Factor Mitofusin 1. *J. Neurochem.* 118 (4), 636–645. doi:10.1111/j.1471-4159.2011.07318.x
- Greer, E. L., and Brunet, A. (2008). FOXO Transcription Factors in Ageing and Cancer. *Acta Physiol. (Oxf)* 192 (1), 19–28. doi:10.1111/j.1748-1716.2007.01780.x
- Guo, J., Liu, X., and Wang, M. (2015). miR-503 Suppresses Tumor Cell Proliferation and Metastasis by Directly Targeting RNF31 in Prostate Cancer. *Biochem. Biophys. Res. Commun.* 464 (4), 1302–1308. doi:10.1016/j.bbrc.2015.07.127
- Guo, Y., Li, Q., Zhao, G., Zhang, J., Yuan, H., Feng, T., et al. (2021). Loss of TRIM31 Promotes Breast Cancer Progression through Regulating K48- and K63-Linked Ubiquitination of P53. *Cell Death Dis.* 12 (10), 945. doi:10.1038/s41419-021-04208-3
- Ha, K., Ma, C., Lin, H., Tang, L., Lian, Z., Zhao, F., et al. (2017). The Anaphase Promoting Complex Impacts Repair Choice by Protecting Ubiquitin Signalling at DNA Damage Sites. *Nat. Commun.* 8, 15751. doi:10.1038/ncomms15751
- Haas, T. L., Emmerich, C. H., Gerlach, B., Schmukle, A. C., Cordier, S. M., Rieser, E., et al. (2009). Recruitment of the Linear Ubiquitin Chain Assembly Complex Stabilizes the TNF-R1 Signaling Complex and Is Required for TNF-Mediated Gene Induction. *Mol. Cell* 36 (5), 831–844. doi:10.1016/j.molcel.2009.10.013
- Habu, T., and Harada, K. H. (2021). UBC13 Is an RNF213-Associated E2 Ubiquitin-Conjugating Enzyme, and Lysine 63-linked Ubiquitination by the RNF213-UBC13 axis Is Responsible for Angiogenic Activity. *FASEB Bioadv* 3 (4), 243–258. doi:10.1096/fba.2019-00092
- Haglund, K., Sigismund, S., Polo, S., Szymkiewicz, I., Di Fiore, P. P., and Dikic, I. (2003). Multiple Monoubiquitination of RTKs Is Sufficient for Their Endocytosis and Degradation. *Nat. Cell Biol.* 5 (5), 461–466. doi:10.1038/ncb983
- Hay-Koren, A., Caspi, M., Zilberberg, A., and Rosin-Arbesfeld, R. (2011). The EDD E3 Ubiquitin Ligase Ubiquitinates and Up-Regulates Beta-Catenin. *Mol. Biol. Cell* 22 (3), 399–411. doi:10.1091/mbc.E10-05-0440
- He, M., Chaurushiya, M. S., Webster, J. D., Kummerfeld, S., Reja, R., Chaudhuri, S., et al. (2019). Intrinsic Apoptosis Shapes the Tumor Spectrum Linked to Inactivation of the Deubiquitinase BAP1. *Science* 364 (6437), 283–285. doi:10.1126/science.aav4902
- He, Y., Jiang, S., Mao, C., Zheng, H., Cao, B., Zhang, Z., et al. (2021). The Deubiquitinase USP10 Restores PTEN Activity and Inhibits Non-small Cell Lung Cancer Cell Proliferation. *J. Biol. Chem.* 297 (3), 101088. doi:10.1016/j.jbc.2021.101088
- He, Y. M., Zhou, X. M., Jiang, S. Y., Zhang, Z. B., Cao, B. Y., Liu, J. B., et al. (2022). TRIM25 Activates AKT/mTOR by Inhibiting PTEN via K63-Linked Polyubiquitination in Non-small Cell Lung Cancer. *Acta Pharmacol. Sin.* 43 (3), 681–691. doi:10.1038/s41401-021-00662-z
- Heap, R. E., Gant, M. S., Lamoliatte, F., Peltier, J., and Trost, M. (2017). Mass Spectrometry Techniques for Studying the Ubiquitin System. *Biochem. Soc. Trans.* 45 (5), 1137–1148. doi:10.1042/bst20170091
- Heaton, S. M., Borg, N. A., and Dixit, V. M. (2016). Ubiquitin in the Activation and Attenuation of Innate Antiviral Immunity. *J. Exp. Med.* 213 (1), 1–13. doi:10.1084/jem.20151531
- Heideker, J., and Wertz, I. E. (2015). DUBs, the Regulation of Cell Identity and Disease. *Biochem. J.* 465 (1), 1–26. doi:10.1042/bj20140496
- Hendriks, I. A., Lyon, D., Su, D., Skotte, N. H., Daniel, J. A., Jensen, L. J., et al. (2018). Site-specific Characterization of Endogenous SUMOylation across Species and Organs. *Nat. Commun.* 9 (1), 2456. doi:10.1038/s41467-018-04957-4
- Herhaus, L., and Sapkota, G. P. (2014). The Emerging Roles of Deubiquitylating Enzymes (DUBs) in the TGF $\beta$  and BMP Pathways. *Cell Signal* 26 (10), 2186–2192. doi:10.1016/j.cellsig.2014.06.012
- Heride, C., Urbé, S., and Clague, M. J. (2014). Ubiquitin Code Assembly and Disassembly. *Curr. Biol.* 24 (6), R215–R220. doi:10.1016/j.cub.2014.02.002
- Heyninck, K., and Beyaert, R. (2005). A20 Inhibits NF- $\kappa$ B Activation by Dual Ubiquitin-Editing Functions. *Trends Biochem. Sci.* 30 (1), 1–4. doi:10.1016/j.tibs.2004.11.001
- Hicke, L., Schubert, H. L., and Hill, C. P. (2005). Ubiquitin-binding Domains. *Nat. Rev. Mol. Cell Biol.* 6 (8), 610–621. doi:10.1038/nrm1701
- Ho, Y. K., Zhi, H., Bowlin, T., Dorjbal, B., Philip, S., Zahoor, M. A., et al. (2015). HTLV-1 Tax Stimulates Ubiquitin E3 Ligase, Ring Finger Protein 8, to Assemble Lysine 63-Linked Polyubiquitin Chains for TAK1 and IKK Activation. *PLoS Pathog.* 11 (8), e1005102. doi:10.1371/journal.ppat.1005102
- Hodge, C. D., Ismail, I. H., Edwards, R. A., Hura, G. L., Xiao, A. T., Tainer, J. A., et al. (2016). RNF8 E3 Ubiquitin Ligase Stimulates Ubc13 E2 Conjugating Activity that Is Essential for DNA Double Strand Break Signaling and BRCA1 Tumor Suppressor Recruitment. *J. Biol. Chem.* 291 (18), 9396–9410. doi:10.1074/jbc.M116.715698
- Hodson, C., Purkiss, A., Miles, J. A., and Walden, H. (2014). Structure of the Human FANCL RING-Ube2T Complex Reveals Determinants of Cognate E3-E2 Selection. *Structure* 22 (2), 337–344. doi:10.1016/j.str.2013.12.004
- Hofmann, R. M., and Pickart, C. M. (1999). Noncanonical MMS2-Encoded Ubiquitin-Conjugating Enzyme Functions in Assembly of Novel Polyubiquitin Chains for DNA Repair. *Cell* 96 (5), 645–653. doi:10.1016/s0092-8674(00)80575-9
- Hong, S. Y., Kao, Y. R., Lee, T. C., and Wu, C. W. (2018). Upregulation of E3 Ubiquitin Ligase CBLC Enhances EGFR Dysregulation and Signaling in Lung Adenocarcinoma. *Cancer Res.* 78 (17), 4984–4996. doi:10.1158/0008-5472.Can-17-3858
- Hrdinka, M., Fiil, B. K., Zucca, M., Leske, D., Bagola, K., Yabal, M., et al. (2016). CYLD Limits Lys63- and Met1-Linked Ubiquitin at Receptor Complexes to

- Regulate Innate Immune Signaling. *Cell Rep.* 14 (12), 2846–2858. doi:10.1016/j.celrep.2016.02.062
- Hu, M., Li, P., Li, M., Li, W., Yao, T., Wu, J. W., et al. (2002). Crystal Structure of a UBP-Family Deubiquitinating Enzyme in Isolation and in Complex with Ubiquitin Aldehyde. *Cell* 111 (7), 1041–1054. doi:10.1016/s0092-8674(02)01199-6
- Hua, F., Hao, W., Wang, L., and Li, S. (2021). Linear Ubiquitination Mediates EGFR-Induced NF-kappaB Pathway and Tumor Development. *Int. J. Mol. Sci.* 22 (21). doi:10.3390/ijms22211875
- Huang, H., Jeon, M. S., Liao, L., Yang, C., Elly, C., Yates, J. R., 3rd, et al. (2010). K33-linked Polyubiquitination of T Cell Receptor-Zeta Regulates Proteolysis-independent T Cell Signaling. *Immunity* 33 (1), 60–70. doi:10.1016/j.immuni.2010.07.002
- Ikeda, F., Deribe, Y. L., Skånland, S. S., Stieglitz, B., Grabbe, C., Franz-Wachtel, M., et al. (2011). SHARPIN Forms a Linear Ubiquitin Ligase Complex Regulating NF-Kb Activity and Apoptosis. *Nature* 471 (7340), 637–641. doi:10.1038/nature09814
- Imam, S., Kömürlü, S., Mattick, J., Selyutina, A., Talley, S., Eddins, A., et al. (2019). K63-Linked Ubiquitin Is Required for Restriction of HIV-1 Reverse Transcription and Capsid Destabilization by Rhesus TRIM5a. *J. Virol.* 93 (14). doi:10.1128/jvi.00558-19
- Infante, P., Faedda, R., Bernardi, F., Bufalieri, F., Lospinoso Severini, L., Alfonsi, R., et al. (2018). Itch/beta-arrestin2-dependent Non-proteolytic Ubiquitylation of SuFu Controls Hedgehog Signalling and Medulloblastoma Tumorigenesis. *Nat. Commun.* 9 (1), 976. doi:10.1038/s41467-018-03339-0
- Iwai, K. (2014). Diverse Roles of the Ubiquitin System in NF-Kb Activation. *Biochim. Biophys. Acta* 1843 (1), 129–136. doi:10.1016/j.bbamer.2013.03.011
- Iwai, K. (2015). “Linear Polyubiquitination: A Crucial Regulator of NF-Kb Activation,” in *Innovative Medicine: Basic Research and Development*. Editors K. Nakao, N. Minato, and S. Uemoto (Tokyo: Springer Copyright), 51–59. doi:10.1007/978-4-431-55651-0\_4
- Jachimowicz, R. D., Beleggia, F., Isensee, J., Velpula, B. B., Goergens, J., Bustos, M. A., et al. (2019). UBQLN4 Represses Homologous Recombination and Is Overexpressed in Aggressive Tumors. *Cell* 176 (3), 505–519. e522. doi:10.1016/j.cell.2018.11.024
- Jeusset, L. M., and McManus, K. J. (2017). Ubiquitin Specific Peptidase 22 Regulates Histone H2B Mono-Ubiquitination and Exhibits Both Oncogenic and Tumor Suppressor Roles in Cancer. *Cancers (Basel)* 9 (12). doi:10.3390/cancers9120167
- Ji, C. H., and Kwon, Y. T. (2017). Crosstalk and Interplay between the Ubiquitin-Proteasome System and Autophagy. *Mol. Cells* 40 (7), 441–449. doi:10.14348/molcells.2017.0115
- Ji, J., Ding, K., Luo, T., Zhang, X., Chen, A., Zhang, D., et al. (2021). TRIM22 Activates NF-kappaB Signaling in Glioblastoma by Accelerating the Degradation of IkappaBalpha. *Cell Death Differ.* 28 (1), 367–381. doi:10.1038/s41418-020-00606-y
- Jiang, S., Park, D. W., Gao, Y., Ravi, S., Darley-Usmar, V., Abraham, E., et al. (2015). Participation of Proteasome-Ubiquitin Protein Degradation in Autophagy and the Activation of AMP-Activated Protein Kinase. *Cell Signal* 27 (6), 1186–1197. doi:10.1016/j.cellsig.2015.02.024
- Jin, X., Shi, Q., Li, Q., Zhou, L., Wang, J., Jiang, L., et al. (2020). CRL3-SPOP Ubiquitin Ligase Complex Suppresses the Growth of Diffuse Large B-Cell Lymphoma by Negatively Regulating the MyD88/NF-kappaB Signaling. *Leukemia* 34 (5), 1305–1314. doi:10.1038/s41375-019-0661-z
- Jing, H., Fang, L., Ding, Z., Wang, D., Hao, W., Gao, L., et al. (2017). Porcine Reproductive and Respiratory Syndrome Virus Nsp1a Inhibits NF-Kb Activation by Targeting the Linear Ubiquitin Chain Assembly Complex. *J. Virol.* 91 (3). doi:10.1128/jvi.01911-16
- Kane, L. A., Lazarou, M., Fogel, A. I., Li, Y., Yamano, K., Sarraf, S. A., et al. (2014). PINK1 Phosphorylates Ubiquitin to Activate Parkin E3 Ubiquitin Ligase Activity. *J. Cell Biol.* 205 (2), 143–153. doi:10.1083/jcb.201402104
- Kazlauskaitė, A., Kondapalli, C., Gourlay, R., Campbell, D. G., Ritorto, M. S., Hofmann, K., et al. (2014). Parkin Is Activated by PINK1-dependent Phosphorylation of Ubiquitin at Ser65. *Biochem. J.* 460 (1), 127–139. doi:10.1042/bj20140334
- Keusekotten, K., Elliott, P. R., Glockner, L., Fiil, B. K., Damgaard, R. B., Kulathu, Y., et al. (2013). OTULIN Antagonizes LUBAC Signaling by Specifically Hydrolyzing Met1-Linked Polyubiquitin. *Cell* 153 (6), 1312–1326. doi:10.1016/j.cell.2013.05.014
- Khaminets, A., Behl, C., and Dikic, I. (2016). Ubiquitin-Dependent and Independent Signals in Selective Autophagy. *Trends Cell Biol.* 26 (1), 6–16. doi:10.1016/j.tcb.2015.08.010
- Kim, H. C., and Huibregtse, J. M. (2009). Polyubiquitination by HECT E3s and the Determinants of Chain Type Specificity. *Mol. Cell Biol.* 29 (12), 3307–3318. doi:10.1128/mcb.00240-09
- Kim, K. I., and Baek, S. H. (2006). SUMOylation Code in Cancer Development and Metastasis. *Mol. Cells* 22 (3), 247–253.
- Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., et al. (2011). Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome. *Mol. Cell* 44 (2), 325–340. doi:10.1016/j.molcel.2011.08.025
- Kirisako, T., Kamei, K., Murata, S., Kato, M., Fukumoto, H., Kanie, M., et al. (2006). A Ubiquitin Ligase Complex Assembles Linear Polyubiquitin Chains. *Embo J.* 25 (20), 4877–4887. doi:10.1038/sj.emboj.7601360
- Klinakis, A., Karagiannis, D., and Rampias, T. (2020). Targeting DNA Repair in Cancer: Current State and Novel Approaches. *Cell Mol. Life Sci.* 77 (4), 677–703. doi:10.1007/s00018-019-03299-8
- Kniss, A., Schuetz, D., Kazemi, S., Pluska, L., Spindler, P. E., Rogov, V. V., et al. (2018). Chain Assembly and Disassembly Processes Differently Affect the Conformational Space of Ubiquitin Chains. *Structure* 26 (2), 249–258. e244. doi:10.1016/j.str.2017.12.011
- Kolapalli, S. P., Sahu, R., Chauhan, N. R., Jena, K. K., Mehto, S., Das, S. K., et al. (2021). RNA-binding RING E3-Ligase DZIP3/hRUL138 Stabilizes Cyclin D1 to Drive Cell-Cycle and Cancer Progression. *Cancer Res.* 81 (2), 315–331. doi:10.1158/0008-5472.CAN-20-1871
- Komander, D., Clague, M. J., and Urbé, S. (2009). Breaking the Chains: Structure and Function of the Deubiquitinases. *Nat. Rev. Mol. Cell Biol.* 10 (8), 550–563. doi:10.1038/nrm2731
- Komander, D., and Rape, M. (2012). The Ubiquitin Code. *Annu. Rev. Biochem.* 81, 203–229. doi:10.1146/annurev-biochem-060310-170328
- Koyano, F., Okatsu, K., Kosako, H., Tamura, Y., Go, E., Kimura, M., et al. (2014). Ubiquitin Is Phosphorylated by PINK1 to Activate Parkin. *Nature* 510 (7503), 162–166. doi:10.1038/nature13392
- Kristariyanto, Y. A., Abdul Rehman, S. A., Campbell, D. G., Morrice, N. A., Johnson, C., Toth, R., et al. (2015). K29-selective Ubiquitin Binding Domain Reveals Structural Basis of Specificity and Heterotypic Nature of K29 Polyubiquitin. *Mol. Cell* 58 (1), 83–94. doi:10.1016/j.molcel.2015.01.041
- Kulathu, Y., and Komander, D. (2012). Atypical Ubiquitylation - the Unexplored World of Polyubiquitin beyond Lys48 and Lys63 Linkages. *Nat. Rev. Mol. Cell Biol.* 13 (8), 508–523. doi:10.1038/nrm3394
- Kung, C. P., and Weber, J. D. (2022). It's Getting Complicated-A Fresh Look at P53-MDM2-ARF Triangle in Tumorigenesis and Cancer Therapy. *Front. Cell Dev. Biol.* 10, 818744. doi:10.3389/fcell.2022.818744
- Kung, W. W., Ramachandran, S., Makukhin, N., Bruno, E., and Ciulli, A. (2019). Structural Insights into Substrate Recognition by the SOCS2 E3 Ubiquitin Ligase. *Nat. Commun.* 10 (1), 2534. doi:10.1038/s41467-019-10190-4
- Kwon, Y. T., and Ciechanover, A. (2017). The Ubiquitin Code in the Ubiquitin-Proteasome System and Autophagy. *Trends Biochem. Sci.* 42 (11), 873–886. doi:10.1016/j.tibs.2017.09.002
- Lamothe, B., Besse, A., Campos, A. D., Webster, W. K., Wu, H., and Darnay, B. G. (2007a). Site-specific Lys-63-Linked Tumor Necrosis Factor Receptor-Associated Factor 6 Auto-Ubiquitination Is a Critical Determinant of I Kappa B Kinase Activation. *J. Biol. Chem.* 282 (6), 4102–4112. doi:10.1074/jbc.M609503200
- Lamothe, B., Campos, A. D., Webster, W. K., Gopinathan, A., Hur, L., and Darnay, B. G. (2008). The RING Domain and First Zinc Finger of TRAF6 Coordinate Signaling by Interleukin-1, Lipopolysaccharide, and RANKL. *J. Biol. Chem.* 283 (36), 24871–24880. doi:10.1074/jbc.M802749200
- Lamothe, B., Webster, W. K., Gopinathan, A., Besse, A., Campos, A. D., and Darnay, B. G. (2007b). TRAF6 Ubiquitin Ligase Is Essential for RANKL Signaling and Osteoclast Differentiation. *Biochem. Biophys. Res. Commun.* 359 (4), 1044–1049. doi:10.1016/j.bbrc.2007.06.017
- Laplantine, E., Fontan, E., Chiaravalli, J., Lopez, T., Lakisic, G., Véron, M., et al. (2009). NEMO Specifically Recognizes K63-Linked Poly-Ubiquitin Chains

- through a New Bipartite Ubiquitin-Binding Domain. *Embo J.* 28 (19), 2885–2895. doi:10.1038/emboj.2009.241
- Lee, B. L., Singh, A., Mark Glover, J. N., Hendzel, M. J., and Spyropoulos, L. (2017). Molecular Basis for K63-Linked Ubiquitination Processes in Double-Strand DNA Break Repair: A Focus on Kinetics and Dynamics. *J. Mol. Biol.* 429 (22), 3409–3429. doi:10.1016/j.jmb.2017.05.029
- Lee, H. J., Li, C. F., Ruan, D., He, J., Montal, E. D., Lorenz, S., et al. (2019a). Non-proteolytic Ubiquitination of Hexokinase 2 by HectH9 Controls Tumor Metabolism and Cancer Stem Cell Expansion. *Nat. Commun.* 10 (1), 2625. doi:10.1038/s41467-019-10374-y
- Lee, H. J., Li, C. F., Ruan, D., Powers, S., Thompson, P. A., Frohman, M. A., et al. (2016). The DNA Damage Transducer RNF8 Facilitates Cancer Chemoresistance and Progression through Twist Activation. *Mol. Cell* 63 (6), 1021–1033. doi:10.1016/j.molcel.2016.08.009
- Lee, J. T., Shan, J., Zhong, J., Li, M., Zhou, B., Zhou, A., et al. (2013). RFP-Mediated Ubiquitination of PTEN Modulates its Effect on AKT Activation. *Cell Res.* 23 (4), 552–564. doi:10.1038/cr.2013.27
- Lee, S. B., Kim, J. J., Han, S. A., Fan, Y., Guo, L. S., Aziz, K., et al. (2019b). The AMPK-Parkin axis Negatively Regulates Necroptosis and Tumorigenesis by Inhibiting the Necrosome. *Nat. Cell Biol.* 21 (8), 940–951. doi:10.1038/s41556-019-0356-8
- Lee, S., Tsai, Y. C., Mattera, R., Smith, W. J., Kostelansky, M. S., Weissman, A. M., et al. (2006). Structural Basis for Ubiquitin Recognition and Autoubiquitination by Rabex-5. *Nat. Struct. Mol. Biol.* 13 (3), 264–271. doi:10.1038/nsmb1064
- Lee, S. W., Li, C. F., Jin, G., Cai, Z., Han, F., Chan, C. H., et al. (2015). Skp2-dependent Ubiquitination and Activation of LKB1 Is Essential for Cancer Cell Survival under Energy Stress. *Mol. Cell* 57 (6), 1022–1033. doi:10.1016/j.molcel.2015.01.015
- Lee, Y. R., Chen, M., Lee, J. D., Zhang, J., Lin, S. Y., Fu, T. M., et al. (2019c). Reactivation of PTEN Tumor Suppressor for Cancer Treatment through Inhibition of a MYC-WWP1 Inhibitory Pathway. *Science* 364 (6441). doi:10.1126/science.aau0159
- Leestemaker, Y., and Ovaa, H. (2017). Tools to Investigate the Ubiquitin Proteasome System. *Drug Discov. Today Technol.* 26, 25–31. doi:10.1016/j.ddtec.2017.11.006
- Leng, S., Huang, W., Chen, Y., Yang, Y., Feng, D., Liu, W., et al. (2021). SIRT1 Coordinates with the CRL4B Complex to Regulate Pancreatic Cancer Stem Cells to Promote Tumorigenesis. *Cell Death Differ.* 28 (12), 3329–3343. doi:10.1038/s41418-021-00821-z
- Lenoir, J. J., Parisien, J. P., and Horvath, C. M. (2021). Immune Regulator LGP2 Targets Ubc13/UBE2N to Mediate Widespread Interference with K63 Polyubiquitination and NF-kappaB Activation. *Cell Rep.* 37 (13), 110175. doi:10.1016/j.celrep.2021.110175
- Lewis, M. J., Vyse, S., Shields, A. M., Boeltz, S., Gordon, P. A., Spector, T. D., et al. (2015). UBE2L3 Polymorphism Amplifies NF-Kb Activation and Promotes Plasma Cell Development, Linking Linear Ubiquitination to Multiple Autoimmune Diseases. *Am. J. Hum. Genet.* 96 (2), 221–234. doi:10.1016/j.ajhg.2014.12.024
- Leznicki, P., and Kulathu, Y. (2017). Mechanisms of Regulation and Diversification of Deubiquitylating Enzyme Function. *J. Cell Sci.* 130 (12), 1997–2006. doi:10.1242/jcs.201855
- Li, N., Wang, Y., Neri, S., Zhen, Y., Fong, L. W. R., Qiao, Y., et al. (2019). Tankyrase Disrupts Metabolic Homeostasis and Promotes Tumorigenesis by Inhibiting LKB1-AMPK Signalling. *Nat. Commun.* 10 (1), 4363. doi:10.1038/s41467-019-12377-1
- Li, Z., Younger, K., Gartenhaus, R., Joseph, A. M., Hu, F., Baer, M. R., et al. (2015). Inhibition of IRAK1/4 Sensitizes T Cell Acute Lymphoblastic Leukemia to Chemotherapies. *J. Clin. Invest* 125 (3), 1081–1097. doi:10.1172/JCI75821
- Liang, L., Fan, Y., Cheng, J., Cheng, D., Zhao, Y., Cao, B., et al. (2013). TAK1 Ubiquitination Regulates Doxorubicin-Induced NF-kappaB Activation. *Cell Signal* 25 (1), 247–254. doi:10.1016/j.cellsig.2012.09.003
- Liao, Y., Liu, N., Xia, X., Guo, Z., Li, Y., Jiang, L., et al. (2019). USP10 Modulates the SKP2/Bcr-Abl axis via Stabilizing SKP2 in Chronic Myeloid Leukemia. *Cell Discov.* 5, 24. doi:10.1038/s41421-019-0092-z
- Liao, Y., Sumara, I., and Pangou, E. (2022). Non-proteolytic Ubiquitylation in Cellular Signaling and Human Disease. *Commun. Biol.* 5 (1), 114. doi:10.1038/s42003-022-03060-1
- Licchesi, J. D., Mieszanek, J., Mevissen, T. E., Rutherford, T. J., Akutsu, M., Virdee, S., et al. (2011). An Ankyrin-Repeat Ubiquitin-Binding Domain Determines TRABID's Specificity for Atypical Ubiquitin Chains. *Nat. Struct. Mol. Biol.* 19 (1), 62–71. doi:10.1038/nsmb.2169
- Lin, S. C., Chung, J. Y., Lamothe, B., Rajashankar, K., Lu, M., Lo, Y. C., et al. (2008). Molecular Basis for the Unique Deubiquitinating Activity of the NF-kappaB Inhibitor A20. *J. Mol. Biol.* 376 (2), 526–540. doi:10.1016/j.jmb.2007.11.092
- Lin, X., Ojo, D., Wei, F., Wong, N., Gu, Y., and Tang, D. (2015). A Novel Aspect of Tumorigenesis-BMI1 Functions in Regulating DNA Damage Response. *Biomolecules* 5 (4), 3396–3415. doi:10.3390/biom5043396
- Liu, B., Zhang, M., Chu, H., Zhang, H., Wu, H., Song, G., et al. (2017). The Ubiquitin E3 Ligase TRIM31 Promotes Aggregation and Activation of the Signaling Adaptor MAVS through Lys63-Linked Polyubiquitination. *Nat. Immunol.* 18 (2), 214–224. doi:10.1038/ni.3641
- Liu, R., Zeng, L. W., Gong, R., Yuan, F., Shu, H. B., and Li, S. (2021). mTORC1 Activity Regulates Post-translational Modifications of glycine Decarboxylase to Modulate glycine Metabolism and Tumorigenesis. *Nat. Commun.* 12 (1), 4227. doi:10.1038/s41467-021-24321-3
- Liu, Z., Gong, Z., Jiang, W. X., Yang, J., Zhu, W. K., Guo, D. C., et al. (2015). Lys63-linked Ubiquitin Chain Adopts Multiple Conformational States for Specific Target Recognition. *Elife* 4. doi:10.7554/eLife.05767
- Loch, C. M., and Strickler, J. E. (2012). A Microarray of Ubiquitylated Proteins for Profiling Deubiquitylase Activity Reveals the Critical Roles of Both Chain and Substrate. *Biochim. Biophys. Acta* 1823 (11), 2069–2078. doi:10.1016/j.bbamcr.2012.05.006
- Locke, M., Toth, J. I., and Petroski, M. D. (2014). Lys11- and Lys48-Linked Ubiquitin Chains Interact with P97 during Endoplasmic-Reticulum-Associated Degradation. *Biochem. J.* 459 (1), 205–216. doi:10.1042/bj20120662
- Lork, M., Verhelst, K., and Beyaert, R. (2017). CYLD, A20 and OTULIN Deubiquitinases in NF-Kb Signaling and Cell Death: So Similar, yet So Different. *Cell Death Differ.* 24 (7), 1172–1183. doi:10.1038/cdd.2017.46
- Lu, C. H., Yeh, D. W., Lai, C. Y., Liu, Y. L., Huang, L. R., Lee, A. Y., et al. (2018). USP17 Mediates Macrophage-Promoted Inflammation and Stemness in Lung Cancer Cells by Regulating TRAF2/TRAF3 Complex Formation. *Oncogene* 37 (49), 6327–6340. doi:10.1038/s41388-018-0411-0
- Ma, J., Shi, Q., Cui, G., Sheng, H., Botuyan, M. V., Zhou, Y., et al. (2021). SPOP Mutation Induces Replication Over-firing by Impairing Geminin Ubiquitination and Triggers Replication Catastrophe upon ATR Inhibition. *Nat. Commun.* 12 (1), 5779. doi:10.1038/s41467-021-26049-6
- Machida, Y. J., Machida, Y., Chen, Y., Gurtan, A. M., Kupfer, G. M., D'Andrea, A. D., et al. (2006). UBE2T Is the E2 in the Fanconi Anemia Pathway and Undergoes Negative Autoregulation. *Mol. Cell* 23 (4), 589–596. doi:10.1016/j.molcel.2006.06.024
- Mai, J., Zhong, Z. Y., Guo, G. F., Chen, X. X., Xiang, Y. Q., Li, X., et al. (2019). Polo-like Kinase 1 Phosphorylates and Stabilizes KLF4 to Promote Tumorigenesis in Nasopharyngeal Carcinoma. *Theranostics* 9 (12), 3541–3554. doi:10.7150/thno.32908
- Mandemaker, I. K., van Cuijk, L., Janssens, R. C., Lans, H., Bezstarosti, K., Hoeijmakers, J. H., et al. (2017). DNA Damage-Induced Histone H1 Ubiquitylation Is Mediated by HUWE1 and Stimulates the RNF8-Rnf168 Pathway. *Sci. Rep.* 7 (1), 15353. doi:10.1038/s41598-017-15194-y
- Manthiram, K., Zhou, Q., Aksentjevich, I., and Kastner, D. L. (2017). The Monogenic Autoinflammatory Diseases Define New Pathways in Human Innate Immunity and Inflammation. *Nat. Immunol.* 18 (8), 832–842. doi:10.1038/ni.3777
- Massoumi, R., Chmielarska, K., Hennecke, K., Pfeifer, A., and Fassler, R. (2006). Cyld Inhibits Tumor Cell Proliferation by Blocking Bcl-3-dependent NF-kappaB Signaling. *Cell* 125 (4), 665–677. doi:10.1016/j.cell.2006.03.041
- Mattern, M., Sutherland, J., Kadimisetty, K., Barrio, R., and Rodriguez, M. S. (2019). Using Ubiquitin Binders to Decipher the Ubiquitin Code. *Trends Biochem. Sci.* 44 (7), 599–615. doi:10.1016/j.tibs.2019.01.011
- Maure, J. F., Moser, S. C., Jaffray, E. G., Alpi, A., and Hay, R. T. (2016). Loss of Ubiquitin E2 Ube2w Rescues Hypersensitivity of Rnf4 Mutant Cells to DNA Damage. *Sci. Rep.* 6, 26178. doi:10.1038/srep26178
- McDowell, G. S., and Philpott, A. (2016). New Insights into the Role of Ubiquitylation of Proteins. *Int. Rev. Cell Mol. Biol.* 325, 35–88. doi:10.1016/bbs.ircmb.2016.02.002

- Meng, Y., Liu, C., Shen, L., Zhou, M., Liu, W., Kowolik, C., et al. (2019). TRAF6 Mediates Human DNA2 Polyubiquitination and Nuclear Localization to Maintain Nuclear Genome Integrity. *Nucleic Acids Res.* 47 (14), 7564–7579. doi:10.1093/nar/gkz537
- Meng, Z., Xu, R., Xie, L., Wu, Y., He, Q., Gao, P., et al. (2021). A20/Nrdp1 Interaction Alters the Inflammatory Signaling Profile by Mediating K48- and K63-Linked Polyubiquitination of Effectors MyD88 and TBK1. *J. Biol. Chem.* 297 (1), 100811. doi:10.1016/j.jbc.2021.100811
- Mevissen, T. E., Hospenthal, M. K., Geurink, P. P., Elliott, P. R., Akutsu, M., Arnaudo, N., et al. (2013). OTU Deubiquitinases Reveal Mechanisms of Linkage Specificity and Enable Ubiquitin Chain Restriction Analysis. *Cell* 154 (1), 169–184. doi:10.1016/j.cell.2013.05.046
- Mevissen, T. E. T., and Komander, D. (2017). Mechanisms of Deubiquitinase Specificity and Regulation. *Annu. Rev. Biochem.* 86, 159–192. doi:10.1146/annurev-biochem-061516-044916
- Mevissen, T. E. T., Kulathu, Y., Mulder, M. P. C., Geurink, P. P., Maslen, S. L., Gersch, M., et al. (2016). Molecular Basis of Lys11-Polyubiquitin Specificity in the Deubiquitinase Cezanne. *Nature* 538 (7625), 402–405. doi:10.1038/nature19836
- Meza Gutierrez, F., Simsek, D., Mizrak, A., Deutschbauer, A., Braberg, H., Johnson, J., et al. (2018). Genetic Analysis Reveals Functions of Atypical Polyubiquitin Chains. *Elife* 7. doi:10.7554/eLife.42955
- Michel, M. A., Komander, D., and Elliott, P. R. (2018). Enzymatic Assembly of Ubiquitin Chains. *Methods Mol. Biol.* 1844, 73–84. doi:10.1007/978-1-4939-8706-1\_6
- Michel, M. A., Swatek, K. N., Hospenthal, M. K., and Komander, D. (2017). Ubiquitin Linkage-specific Affimers Reveal Insights into K6-Linked Ubiquitin Signaling. *Mol. Cell* 68 (1), 233–246. e235. doi:10.1016/j.molcel.2017.08.020
- Miles, J. A., Frost, M. G., Carroll, E., Rowe, M. L., Howard, M. J., Sidhu, A., et al. (2015). The Fanconi Anemia DNA Repair Pathway Is Regulated by an Interaction between Ubiquitin and the E2-like Fold Domain of FANCL. *J. Biol. Chem.* 290 (34), 20995–21006. doi:10.1074/jbc.M115.675835
- Min, Y., Kim, M. J., Lee, S., Chun, E., and Lee, K. Y. (2018). Inhibition of TRAF6 Ubiquitin-Ligase Activity by PRDX1 Leads to Inhibition of NFkB Activation and Autophagy Activation. *Autophagy* 14 (8), 1347–1358. doi:10.1080/15548627.2018.1474995
- Morris, J. R., and Solomon, E. (2004). BRCA1 : BARD1 Induces the Formation of Conjugated Ubiquitin Structures, Dependent on K6 of Ubiquitin, in Cells during DNA Replication and Repair. *Hum. Mol. Genet.* 13 (8), 807–817. doi:10.1093/hmg/ddh095
- Mu, Y., Sundar, R., Thakur, N., Ekman, M., Gudey, S. K., Yakymovych, M., et al. (2011). TRAF6 Ubiquitinates TGFbeta Type I Receptor to Promote its Cleavage and Nuclear Translocation in Cancer. *Nat. Commun.* 2, 330. doi:10.1038/ncomms1332
- Myant, K. B., Cammareri, P., Hodder, M. C., Wills, J., Von Kriegsheim, A., Gyorffy, B., et al. (2017). HUWE1 is a Critical Colonic Tumour Suppressor Gene that Prevents MYC Signalling, DNA Damage Accumulation and Tumour Initiation. *EMBO Mol. Med.* 9 (2), 181–197. doi:10.15252/emmm.201606684
- Ni, X., Kou, W., Gu, J., Wei, P., Wu, X., Peng, H., et al. (2019). TRAF6 Directs FOXP3 Localization and Facilitates Regulatory T-Cell Function through K63-Linked Ubiquitination. *EMBO J.* 38 (9). doi:10.15252/emboj.201899766
- Nishi, R., Wijnhoven, P., le Sage, C., Tjeertes, J., Galanty, Y., Forment, J. V., et al. (2014). Systematic Characterization of Deubiquitylating Enzymes for Roles in Maintaining Genome Integrity. *Nat. Cell Biol.* 16 (10), 10161011–10261018. doi:10.1038/ncb3028
- Nishikawa, H., Ooka, S., Sato, K., Arima, K., Okamoto, J., Klevit, R. E., et al. (2004). Mass Spectrometric and Mutational Analyses Reveal Lys-6-Linked Polyubiquitin Chains Catalyzed by BRCA1-BARD1 Ubiquitin Ligase. *J. Biol. Chem.* 279 (6), 3916–3924. doi:10.1074/jbc.M308540200
- Niu, J., Shi, Y., Iwai, K., and Wu, Z. H. (2011). LUBAC Regulates NF-kappaB Activation upon Genotoxic Stress by Promoting Linear Ubiquitination of NEMO. *EMBO J.* 30 (18), 3741–3753. doi:10.1038/emboj.2011.264
- Niu, T., Wu, Z., and Xiao, W. (2021). Uev1A Promotes Breast Cancer Cell Migration by Up-Regulating CT45A Expression via the AKT Pathway. *BMC Cancer* 21 (1), 1012. doi:10.1186/s12885-021-08750-3
- Noad, J., von der Malsburg, A., Pathe, C., Michel, M. A., Komander, D., and Randow, F. (2017). LUBAC-Synthesized Linear Ubiquitin Chains Restrict Cytosol-Invasive Bacteria by Activating Autophagy and NF-Kb. *Nat. Microbiol.* 2, 17063. doi:10.1038/nmicrobiol.2017.63
- Nowsheen, S., Aziz, K., Aziz, A., Deng, M., Qin, B., Luo, K., et al. (2018). L3MBTL2 Orchestrates Ubiquitin Signalling by Dictating the Sequential Recruitment of RNF8 and RNF168 after DNA Damage. *Nat. Cell Biol.* 20 (4), 455–464. doi:10.1038/s41556-018-0071-x
- Nusse, R., and Clevers, H. (2017). Wnt/ $\beta$ -Catenin Signaling, Disease, and Emerging Therapeutic Modalities. *Cell* 169 (6), 985–999. doi:10.1016/j.cell.2017.05.016
- O'Connor, H. F., and Huibregtse, J. M. (2017). Enzyme-substrate Relationships in the Ubiquitin System: Approaches for Identifying Substrates of Ubiquitin Ligases. *Cell Mol. Life Sci.* 74 (18), 3363–3375. doi:10.1007/s00018-017-2529-6
- O'Donnell, M. A. (2019). Gustavo Silva: Translating the Ubiquitin Code. *J. Cell Biol.* 218 (1), 3–4. doi:10.1083/jcb.201812083
- Ohtake, F., Saeki, Y., Sakamoto, K., Ohtake, K., Nishikawa, H., Tsuchiya, H., et al. (2015). Ubiquitin Acetylation Inhibits Polyubiquitin Chain Elongation. *EMBO Rep.* 16 (2), 192–201. doi:10.15252/embr.201439152
- Ordureau, A., Sarraf, S. A., Duda, D. M., Heo, J. M., Jedrychowski, M. P., Svendskiy, V. O., et al. (2014). Quantitative Proteomics Reveal a Feedforward Mechanism for Mitochondrial PARKIN Translocation and Ubiquitin Chain Synthesis. *Mol. Cell* 56 (3), 360–375. doi:10.1016/j.molcel.2014.09.007
- Padala, P., Soudah, N., Giladi, M., Haitin, Y., Isupov, M. N., and Wiener, R. (2017). The Crystal Structure and Conformations of an Unbranched Mixed Tri-ubiquitin Chain Containing K48 and K63 Linkages. *J. Mol. Biol.* 429 (24), 3801–3813. doi:10.1016/j.jmb.2017.10.027
- Park, H. Y., Go, H., Song, H. R., Kim, S., Ha, G. H., Jeon, Y. K., et al. (2014). Pellino 1 Promotes Lymphomagenesis by Deregulating BCL6 Polyubiquitination. *J. Clin. Invest* 124 (11), 4976–4988. doi:10.1172/JCI75667
- Paul, A., and Wang, B. (2017). RNF8- and Ube2S-dependent Ubiquitin Lysine 11-Linkage Modification in Response to DNA Damage. *Mol. Cell* 66 (4), 458–472. e455. doi:10.1016/j.molcel.2017.04.013
- Penengo, L., Mapelli, M., Murachelli, A. G., Confalonieri, S., Magri, L., Musacchio, A., et al. (2006). Crystal Structure of the Ubiquitin Binding Domains of Rabex-5 Reveals Two Modes of Interaction with Ubiquitin. *Cell* 124 (6), 1183–1195. doi:10.1016/j.cell.2006.02.020
- Peng, D. J., Zeng, M., Muromoto, R., Matsuda, T., Shimoda, K., Subramaniam, M., et al. (2011). Noncanonical K27-Linked Polyubiquitination of TIEG1 Regulates Foxp3 Expression and Tumor Growth. *J. Immunol.* 186 (10), 5638–5647. doi:10.4049/jimmunol.1003801
- Piao, S., Pei, H. Z., Huang, B., and Baek, S. H. (2017). Ovarian Tumor Domain-Containing Protein 1 Deubiquitinates and Stabilizes P53. *Cell Signal* 33, 22–29. doi:10.1016/j.cellsig.2017.02.011
- Polykratis, A., Martens, A., Eren, R. O., Shirasaki, Y., Yamagishi, M., Yamaguchi, Y., et al. (2019). A20 Prevents Inflammasome-dependent Arthritis by Inhibiting Macrophage Necroptosis through its ZnF7 Ubiquitin-Binding Domain. *Nat. Cell Biol.* 21 (6), 731–742. doi:10.1038/s41556-019-0324-3
- Qin, Z., Jiang, W., Wang, G., Sun, Y., and Xiao, W. (2018). Linear Ubiquitin Chain Induces Apoptosis and Inhibits Tumor Growth. *Apoptosis* 23 (1), 16–26. doi:10.1007/s10495-017-1433-8
- Queisser, M. A., Dada, L. A., Deiss-Yehiely, N., Angulo, M., Zhou, G., Kouri, F. M., et al. (2014). HOIL-1L Functions as the PKC $\zeta$  Ubiquitin Ligase to Promote Lung Tumor Growth. *Am. J. Respir. Crit. Care Med.* 190 (6), 688–698. doi:10.1164/rccm.201403-0463OC
- Raimondi, M., Cesselli, D., Di Loreto, C., La Marra, F., Schneider, C., and Demarchi, F. (2019). USP1 (Ubiquitin Specific Peptidase 1) Targets ULK1 and Regulates its Cellular Compartmentalization and Autophagy. *Autophagy* 15 (4), 613–630. doi:10.1080/15548627.2018.1535291
- Rape, M. (2018). Ubiquitylation at the Crossroads of Development and Disease. *Nat. Rev. Mol. Cell Biol.* 19 (1), 59–70. doi:10.1038/nrm.2017.83
- Regev, O., Roth, Z., Korman, M., Khalaila, I., and Gur, E. (2015). A Kinetic Model for the Prevalence of Mono- over Poly-Pupylation. *Febs J.* 282 (21), 4176–4186. doi:10.1111/febs.13413
- Ren, Y., Xu, X., Mao, C. Y., Han, K. K., Xu, Y. J., Cao, B. Y., et al. (2020). RNF6 Promotes Myeloma Cell Proliferation and Survival by Inducing Glucocorticoid Receptor Polyubiquitination. *Acta Pharmacol. Sin.* 41 (3), 394–403. doi:10.1038/s41401-019-0309-6
- Reyes-Turcu, F. E., Horton, J. R., Mullally, J. E., Heroux, A., Cheng, X., and Wilkinson, K. D. (2006). The Ubiquitin Binding Domain ZnF UBP Recognizes

- the C-Terminal Diglycine Motif of Unanchored Ubiquitin. *Cell* 124 (6), 1197–1208. doi:10.1016/j.cell.2006.02.038
- Rittinger, K., and Ikeda, F. (2017). Linear Ubiquitin Chains: Enzymes, Mechanisms and Biology. *Open Biol.* 7 (4). doi:10.1098/rsob.170026
- Rivkin, E., Almeida, S. M., Ceccarelli, D. F., Juang, Y. C., MacLean, T. A., Srikumar, T., et al. (2013). The Linear Ubiquitin-specific Deubiquitinase Gumbi Regulates Angiogenesis. *Nature* 498 (7454), 318–324. doi:10.1038/nature12296
- Rogerson, D. T., Sachdeva, A., Wang, K., Haq, T., Kazlauskaitė, A., Hancock, S. M., et al. (2015). Efficient Genetic Encoding of Phosphoserine and its Nonhydrolyzable Analog. *Nat. Chem. Biol.* 11 (7), 496–503. doi:10.1038/nchembio.1823
- Rösner, D., Schneider, T., Schneider, M., and Marx, A. (2015). Click Chemistry for Targeted Protein Ubiquitylation and Ubiquitin Chain Formation. *Nat. Protoc.* 10 (10), 1594–1611. doi:10.1038/nprot.2015.106
- Rossi, F. A., and Rossi, M. (2022). Emerging Role of Ubiquitin-specific Protease 19 in Oncogenesis and Cancer Development. *Front. Cell Dev. Biol.* 10, 889166. doi:10.3389/fcell.2022.889166
- Rudnicka, A., and Yamauchi, Y. (2016). Ubiquitin in Influenza Virus Entry and Innate Immunity. *Viruses* 8 (10). doi:10.3390/v8100293
- Sadowski, M., and Sarcevic, B. (2010). Mechanisms of Mono- and Poly-Ubiquitination: Ubiquitination Specificity Depends on Compatibility between the E2 Catalytic Core and Amino Acid Residues Proximal to the Lysine. *Cell Div.* 5, 19. doi:10.1186/1747-1028-5-19
- Sahin, I., Zhang, S., Navaraj, A., Zhou, L., Dizon, D., Safran, H., et al. (2020). AMG-232 Sensitizes High MDM2-Expressing Tumor Cells to T-Cell-Mediated Killing. *Cell Death Discov.* 6, 57. doi:10.1038/s41420-020-0292-1
- Sahu, S. K., Tiwari, N., Pataskar, A., Zhuang, Y., Borisova, M., Diken, M., et al. (2017). FBXO32 Promotes Microenvironment Underlying Epithelial-Mesenchymal Transition via CtBP1 during Tumour Metastasis and Brain Development. *Nat. Commun.* 8 (1), 1523. doi:10.1038/s41467-017-01366-x
- Sasaki, K., and Iwai, K. (2015). Roles of Linear Ubiquitylation, a Crucial Regulator of NF- $\kappa$ B and Cell Death, in the Immune System. *Immunol. Rev.* 266 (1), 175–189. doi:10.1111/immr.12308
- Sato, Y., Okatsu, K., Saeki, Y., Yamano, K., Matsuda, N., Kaiho, A., et al. (2017). Structural Basis for Specific Cleavage of Lys6-Linked Polyubiquitin Chains by USP30. *Nat. Struct. Mol. Biol.* 24 (11), 911–919. doi:10.1038/nsmb.3469
- Seeler, J. S., and Dejean, A. (2017). SUMO and the Robustness of Cancer. *Nat. Rev. Cancer* 17 (3), 184–197. doi:10.1038/nrc.2016.143
- Senft, D., Qi, J., and Ronai, Z. A. (2018). Ubiquitin Ligases in Oncogenic Transformation and Cancer Therapy. *Nat. Rev. Cancer* 18 (2), 69–88. doi:10.1038/nrc.2017.105
- Seo, J., Lee, E. W., Shin, J., Seong, D., Nam, Y. W., Jeong, M., et al. (2018). K6 Linked Polyubiquitylation of FADD by CHIP Prevents Death Inducing Signaling Complex Formation Suppressing Cell Death. *Oncogene* 37 (36), 4994–5006. doi:10.1038/s41388-018-0323-z
- Shanbhag, N. M., Rafalska-Metcalf, I. U., Balane-Bolivar, C., Janicki, S. M., and Greenberg, R. A. (2010). ATM-Dependent Chromatin Changes Silence Transcription in Cis to DNA Double-Strand Breaks. *Cell* 141 (6), 970–981. doi:10.1016/j.cell.2010.04.038
- Shekhar, M. P., Gerard, B., Pauley, R. J., Williams, B. O., and Tait, L. (2008). Rad6B Is a Positive Regulator of Beta-Catenin Stabilization. *Cancer Res.* 68 (6), 1741–1750. doi:10.1158/0008-5472.CAN-07-2111
- Shen, Y., Tang, K., Chen, D., Hong, M., Sun, F., Wang, S., et al. (2021). Riok3 Inhibits the Antiviral Immune Response by Facilitating TRIM40-Mediated RIG-I and MDA5 Degradation. *Cell Rep.* 35 (12), 109272. doi:10.1016/j.celrep.2021.109272
- Shih, S. C., Katzmman, D. J., Schnell, J. D., Sutanto, M., Emr, S. D., and Hicke, L. (2002). Epsins and Vps27p/Hrs Contain Ubiquitin-Binding Domains that Function in Receptor Endocytosis. *Nat. Cell Biol.* 4 (5), 389–393. doi:10.1038/ncb790
- Shirane, M., Hatakeyama, S., Hattori, K., Nakayama, K., and Nakayama, K. (1999). Common Pathway for the Ubiquitination of IkappaBalpha, IkappaBbeta, and IkappaBepsilon Mediated by the F-Box Protein FWD1. *J. Biol. Chem.* 274 (40), 28169–28174. doi:10.1074/jbc.274.40.28169
- Singh, A. N., Oehler, J., Torrecilla, I., Kilgas, S., Li, S., Vaz, B., et al. (2019). The P97-Ataxin 3 Complex Regulates Homeostasis of the DNA Damage Response E3 Ubiquitin Ligase RNF8. *EMBO J.* 38 (21), e102361. doi:10.15252/embj.2019102361
- Singh, R., Karri, D., Shen, H., Shao, J., Dasgupta, S., Huang, S., et al. (2018). TRAF4-mediated Ubiquitination of NGF Receptor TrkA Regulates Prostate Cancer Metastasis. *J. Clin. Invest.* 128 (7), 3129–3143. doi:10.1172/JCI96060
- Smeenk, G., and Mailand, N. (2016). Writers, Readers, and Erasers of Histone Ubiquitylation in DNA Double-Strand Break Repair. *Front. Genet.* 7, 122. doi:10.3389/fgene.2016.00122
- Smith, C. J., Berry, D. M., and McGlade, C. J. (2013). The E3 Ubiquitin Ligases RNF126 and Rabr7 Regulate Endosomal Sorting of the Epidermal Growth Factor Receptor. *J. Cell Sci.* 126 (Pt 6), 1366–1380. doi:10.1242/jcs.116129
- Sokratous, K., Hadjisavvas, A., Diamandis, E. P., and Kyriacou, K. (2014). The Role of Ubiquitin-Binding Domains in Human Pathophysiology. *Crit. Rev. Clin. Lab. Sci.* 51 (5), 280–290. doi:10.3109/10408363.2014.915287
- Somasagara, R. R., Spencer, S. M., Tripathi, K., Clark, D. W., Mani, C., Madeira da Silva, L., et al. (2017). RAD6 Promotes DNA Repair and Stem Cell Signaling in Ovarian Cancer and Is a Promising Therapeutic Target to Prevent and Treat Acquired Chemoresistance. *Oncogene* 36 (48), 6680–6690. doi:10.1038/onc.2017.279
- Song, K., Cai, X., Dong, Y., Wu, H., Wei, Y., Shankavaram, U. T., et al. (2021). Epsins 1 and 2 Promote NEMO Linear Ubiquitination via LUBAC to Drive Breast Cancer Development. *J. Clin. Invest.* 131 (1). doi:10.1172/JCI129374
- Song, L., and Luo, Z. Q. (2019). Post-Translational Regulation of Ubiquitin Signaling. *J. Cell Biol.* 218 (6), 1776–1786. doi:10.1083/jcb.201902074
- Sparrer, K. M. J., Gableske, S., Zurenski, M. A., Parker, Z. M., Full, F., Baumgart, G. J., et al. (2017). TRIM23 Mediates Virus-Induced Autophagy via Activation of TBK1. *Nat. Microbiol.* 2 (11), 1543–1557. doi:10.1038/s41564-017-0017-2
- Spence, J., Sadis, S., Haas, A. L., and Finley, D. (1995). A Ubiquitin Mutant with Specific Defects in DNA Repair and Multiubiquitination. *Mol. Cell Biol.* 15 (3), 1265–1273. doi:10.1128/mcb.15.3.1265
- Spiliotopoulos, A., Blokpoel Ferreras, L., Densham, R. M., Caulton, S. G., Maddison, B. C., Morris, J. R., et al. (2019). Discovery of Peptide Ligands Targeting a Specific Ubiquitin-like Domain-Binding Site in the Deubiquitinase USP11. *J. Biol. Chem.* 294 (2), 424–436. doi:10.1074/jbc.RA118.004469
- Spit, M., Rieser, E., and Walczak, H. (2019). Linear Ubiquitination at a Glance. *J. Cell Sci.* 132 (2). doi:10.1242/jcs.208512
- Sun, J., Zhu, Z., Li, W., Shen, M., Cao, C., Sun, Q., et al. (2020). UBE2T-regulated H2AX Monoubiquitination Induces Hepatocellular Carcinoma Radioresistance by Facilitating CHK1 Activation. *J. Exp. Clin. Cancer Res.* 39 (1), 222. doi:10.1186/s13046-020-01734-4
- Sun, L., Amraei, R., and Rahimi, N. (2021). NEDD4 Regulates Ubiquitination and Stability of the Cell Adhesion Molecule IGPR-1 via Lysosomal Pathway. *J. Biomed. Sci.* 28 (1), 35. doi:10.1186/s12929-021-00731-9
- Sun, S. C. (2010). CYLD: a Tumor Suppressor Deubiquitinase Regulating NF- $\kappa$ B Activation and Diverse Biological Processes. *Cell Death Differ.* 17 (1), 25–34. doi:10.1038/cdd.2009.43
- Sun, X., Ding, Y., Zhan, M., Li, Y., Gao, D., Wang, G., et al. (2019). Usp7 Regulates Hippo Pathway through Deubiquitinating the Transcriptional Coactivator Yorkie. *Nat. Commun.* 10 (1), 411. doi:10.1038/s41467-019-08334-7
- Suryadinata, R., Roesley, S. N., Yang, G., and Sarčević, B. (2014). Mechanisms of Generating Polyubiquitin Chains of Different Topology. *Cells* 3 (3), 674–689. doi:10.3390/cells3030674
- Swaney, D. L., Rodríguez-Mias, R. A., and Villén, J. (2015). Phosphorylation of Ubiquitin at Ser65 Affects its Polymerization, Targets, and Proteome-wide Turnover. *EMBO Rep.* 16 (9), 1131–1144. doi:10.15252/embr.201540298
- Sy, S. M., Jiang, J., O, W. S., Deng, Y., and Huen, M. S. (2013). The Ubiquitin Specific Protease USP34 Promotes Ubiquitin Signaling at DNA Double-Strand Breaks. *Nucleic Acids Res.* 41 (18), 8572–8580. doi:10.1093/nar/gkt622
- Syed, N. A., Andersen, P. L., Warrington, R. C., and Xiao, W. (2006). Uev1A, a Ubiquitin Conjugating Enzyme Variant, Inhibits Stress-Induced Apoptosis through NF- $\kappa$ B Activation. *Apoptosis* 11 (12), 2147–2157. doi:10.1007/s10495-006-0197-3
- Takiuchi, T., Nakagawa, T., Tamiya, H., Fujita, H., Sasaki, Y., Saeki, Y., et al. (2014). Suppression of LUBAC-Mediated Linear Ubiquitination by a Specific Interaction between LUBAC and the Deubiquitinases CYLD and OTULIN. *Genes cells.* 19 (3), 254–272. doi:10.1111/gtc.12128
- Taminiau, A., Draime, A., Tys, J., Lambert, B., Vandeputte, J., Nguyen, N., et al. (2016). HOXA1 Binds RBCK1/HOIL-1 and TRAF2 and Modulates the TNF/

- NF- $\kappa$ B Pathway in a Transcription-independent Manner. *Nucleic Acids Res.* 44 (15), 7331–7349. doi:10.1093/nar/gkw606
- Tang, J., Luo, Y., Tian, Z., Liao, X., Cui, Q., Yang, Q., et al. (2020). TRIM11 Promotes Breast Cancer Cell Proliferation by Stabilizing Estrogen Receptor Alpha. *Neoplasia* 22 (9), 343–351. doi:10.1016/j.neo.2020.06.003
- Tang, L. Y., Yamashita, M., Coussens, N. P., Tang, Y., Wang, X., Li, C., et al. (2011). Ablation of Smurf2 Reveals an Inhibition in TGF- $\beta$  Signalling through Multiple Mono-Ubiquitination of Smad3. *Embo J.* 30 (23), 4777–4789. doi:10.1038/emboj.2011.393
- Tang, Y., Joo, D., Liu, G., Tu, H., You, J., Jin, J., et al. (2018). Linear Ubiquitination of cFLIP Induced by LUBAC Contributes to TNF $\alpha$ -Induced Apoptosis. *J. Biol. Chem.* 293 (52), 20062–20072. doi:10.1074/jbc.RA118.005449
- Tanno, H., and Komada, M. (2013). The Ubiquitin Code and its Decoding Machinery in the Endocytic Pathway. *J. Biochem.* 153 (6), 497–504. doi:10.1093/jb/mvt028
- Teixeira, L. K., and Reed, S. I. (2013). Ubiquitin Ligases and Cell Cycle Control. *Annu. Rev. Biochem.* 82, 387–414. doi:10.1146/annurev-biochem-060410-105307
- Tikoo, S., Madhavan, V., Hussain, M., Miller, E. S., Arora, P., Zlatanou, A., et al. (2013). Ubiquitin-dependent Recruitment of the Bloom Syndrome Helicase upon Replication Stress Is Required to Suppress Homologous Recombination. *EMBO J.* 32 (12), 1778–1792. doi:10.1038/emboj.2013.117
- Tokunaga, F., and Iwai, K. (2012). Linear Ubiquitination: a Novel NF-Kb Regulatory Mechanism for Inflammatory and Immune Responses by the LUBAC Ubiquitin Ligase Complex. *Endocr. J.* 59 (8), 641–652. doi:10.1507/endocrj.ej12-0148
- Tokunaga, F. (2013). Linear Ubiquitination-Mediated NF-Kb Regulation and its Related Disorders. *J. Biochem.* 154 (4), 313–323. doi:10.1093/jb/mvt079
- Tokunaga, F., Nakagawa, T., Nakahara, M., Saeki, Y., Taniguchi, M., Sakata, S., et al. (2011). SHARPIN Is a Component of the NF-kappaB-Activating Linear Ubiquitin Chain Assembly Complex. *Nature* 471 (7340), 633–636. doi:10.1038/nature09815
- Torres, M. P., Lee, M. J., Ding, F., Purbeck, C., Kuhlman, B., Dokholyan, N. V., et al. (2009). G Protein Mono-Ubiquitination by the Rsp5 Ubiquitin Ligase. *J. Biol. Chem.* 284 (13), 8940–8950. doi:10.1074/jbc.M809058200
- Tran, H., Hamada, F., Schwarz-Romond, T., and Bienz, M. (2008). Trabd, a New Positive Regulator of Wnt-Induced Transcription with Preference for Binding and Cleaving K63-Linked Ubiquitin Chains. *Genes Dev.* 22 (4), 528–542. doi:10.1101/gad.463208
- Turek, I., Tischer, N., Lassig, R., and Trujillo, M. (2018). Multi-tiered Pairing Selectivity between E2 Ubiquitin-Conjugating Enzymes and E3 Ligases. *J. Biol. Chem.* 293 (42), 16324–16336. doi:10.1074/jbc.RA118.004226
- van der Heden van Noort, G. J., Kooij, R., Elliott, P. R., Komander, D., and Ovaa, H. (2017). Synthesis of Poly-Ubiquitin Chains Using a Bifunctional Ubiquitin Monomer. *Org. Lett.* 19 (24), 6490–6493. doi:10.1021/acs.orglett.7b03085
- van Wijk, S. J. L., Fricke, F., Herhaus, L., Gupta, J., Hotte, K., Pampaloni, F., et al. (2017). Linear Ubiquitination of Cytosolic Salmonella Typhimurium Activates NF-kappaB and Restricts Bacterial Proliferation. *Nat. Microbiol.* 2, 17066. doi:10.1038/nmicrobiol.2017.66
- Varadan, R., Assfalg, M., Haririnia, A., Raasi, S., Pickart, C., and Fushman, D. (2004). Solution Conformation of Lys63-Linked Di-ubiquitin Chain Provides Clues to Functional Diversity of Polyubiquitin Signaling. *J. Biol. Chem.* 279 (8), 7055–7063. doi:10.1074/jbc.M309184200
- Venuto, S., and Merla, G. (2019). E3 Ubiquitin Ligase TRIM Proteins, Cell Cycle and Mitosis. *Cells* 8 (5). doi:10.3390/cells8050510
- Vuorela, M., Pylkas, K., and Winqvist, R. (2011). Mutation Screening of the RNF8, UBC13 and MMS2 Genes in Northern Finnish Breast Cancer Families. *BMC Med. Genet.* 12, 98. doi:10.1186/1471-2350-12-98
- Wagner, S. A., Beli, P., Weinert, B. T., Nielsen, M. L., Cox, J., Mann, M., et al. (2011). A Proteome-wide, Quantitative Survey of *In Vivo* Ubiquitylation Sites Reveals Widespread Regulatory Roles. *Mol. Cell Proteomics* 10 (10), M111–M013284. doi:10.1074/mcp.M111.013284
- Walczak, H., Iwai, K., and Dikic, I. (2012). Generation and Physiological Roles of Linear Ubiquitin Chains. *BMC Biol.* 10, 23. doi:10.1186/1741-7007-10-23
- Wang, B., Jie, Z., Joo, D., Ordureau, A., Liu, P., Gan, W., et al. (2017a). TRAF2 and OTUD7B Govern a Ubiquitin-dependent Switch that Regulates mTORC2 Signalling. *Nature* 545 (7654), 365–369. doi:10.1038/nature22344
- Wang, B., Ma, A., Zhang, L., Jin, W. L., Qian, Y., Xu, G., et al. (2015). POH1 Deubiquitylates and Stabilizes E2F1 to Promote Tumour Formation. *Nat. Commun.* 6, 8704. doi:10.1038/ncomms9704
- Wang, C., Long, W., Peng, C., Hu, L., Zhang, Q., Wu, A., et al. (2016). HTLV-1 Tax Functions as a Ubiquitin E3 Ligase for Direct IKK Activation via Synthesis of Mixed-Linkage Polyubiquitin Chains. *PLoS Pathog.* 12 (4), e1005584. doi:10.1371/journal.ppat.1005584
- Wang, F., Fu, X., Chen, P., Wu, P., Fan, X., Li, N., et al. (2017b). SPSB1-mediated HnRNP A1 Ubiquitylation Regulates Alternative Splicing and Cell Migration in EGF Signaling. *Cell Res.* 27 (4), 540–558. doi:10.1038/cr.2017.7
- Wang, F., Gao, Y., Zhou, L., Chen, J., Xie, Z., Ye, Z., et al. (2022). USP30: Structure, Emerging Physiological Role, and Target Inhibition. *Front. Pharmacol.* 13, 851654. doi:10.3389/fphar.2022.851654
- Wang, G., Long, J., Gao, Y., Zhang, W., Han, F., Xu, C., et al. (2019). SETDB1-mediated Methylation of Akt Promotes its K63-Linked Ubiquitination and Activation Leading to Tumorigenesis. *Nat. Cell Biol.* 21 (2), 214–225. doi:10.1038/s41556-018-0266-1
- Wang, P., Dai, X., Jiang, W., Li, Y., and Wei, W. (2020). RBR E3 Ubiquitin Ligases in Tumorigenesis. *Semin. Cancer Biol.* 67 (Pt 2), 131–144. doi:10.1016/j.semcancer.2020.05.002
- Wang, Q., Huang, L., Hong, Z., Lv, Z., Mao, Z., Tang, Y., et al. (2017c). The E3 Ubiquitin Ligase RNF185 Facilitates the cGAS-Mediated Innate Immune Response. *PLoS Pathog.* 13 (3), e1006264. doi:10.1371/journal.ppat.1006264
- Wang, S., Wang, R., Peralta, C., Yaseen, A., and Pavletich, N. P. (2021a). Structure of the FA Core Ubiquitin Ligase Closing the ID Clamp on DNA. *Nat. Struct. Mol. Biol.* 28 (3), 300–309. doi:10.1038/s41594-021-00568-8
- Wang, X., Zhang, H., Shao, Z., Zhuang, W., Sui, C., Liu, F., et al. (2021b). TRIM31 Facilitates K27-Linked Polyubiquitination of SYK to Regulate Antifungal Immunity. *Signal Transduct. Target Ther.* 6 (1), 298. doi:10.1038/s41392-021-00711-3
- Wang, Y., Shi, M., Feng, H., Zhu, Y., Liu, S., Gao, A., et al. (2018). Structural Insights into Non-canonical Ubiquitination Catalyzed by SidE. *Cell* 173 (5), 1231–1243. e1216. doi:10.1016/j.cell.2018.04.023
- Wenbo, L., and Wang, J. (2017). Uncovering the Underlying Mechanism of Cancer Tumorigenesis and Development under an Immune Microenvironment from Global Quantification of the Landscape. *J. R. Soc. Interface* 14 (131). doi:10.1098/rsif.2017.0105
- Werner, A., Manford, A. G., and Rape, M. (2017). Ubiquitin-Dependent Regulation of Stem Cell Biology. *Trends Cell Biol.* 27 (8), 568–579. doi:10.1016/j.tcb.2017.04.002
- Wertz, I. E., O'Rourke, K. M., Zhou, H., Eby, M., Aravind, L., Seshagiri, S., et al. (2004). De-ubiquitination and Ubiquitin Ligase Domains of A20 Downregulate NF-kappaB Signalling. *Nature* 430 (7000), 694–699. doi:10.1038/nature02794
- Wheaton, K., Sarkari, F., Stanley Johns, B., Davarinejad, H., Egorova, O., Kaustov, L., et al. (2017). Ube2E1/UBCH6 Is a Critical *In Vivo* E2 for the PRC1-Catalyzed Ubiquitination of H2A at Lys-119. *J. Biol. Chem.* 292 (7), 2893–2902. doi:10.1074/jbc.M116.749564
- Whiteaker, J. R., Zhao, L., Ivey, R. G., Sanchez-Bonilla, M., Moore, H. D., Schoenherr, R. M., et al. (2018). Targeted Mass Spectrometry Enables Robust Quantification of FANCD2 Mono-Ubiquitination in Response to DNA Damage. *DNA Repair (Amst)* 65, 47–53. doi:10.1016/j.dnarep.2018.03.003
- Windheim, M., Pegg, M., and Cohen, P. (2008). Two Different Classes of E2 Ubiquitin-Conjugating Enzymes Are Required for the Mono-Ubiquitination of Proteins and Elongation by Polyubiquitin Chains with a Specific Topology. *Biochem. J.* 409 (3), 723–729. doi:10.1042/bj20071338
- Wu, B., Qiang, L., Zhang, Y., Fu, Y., Zhao, M., Lei, Z., et al. (2022). The Deubiquitinase OTUD1 Inhibits Colonic Inflammation by Suppressing RIPK1-Mediated NF-kappaB Signaling. *Cell Mol. Immunol.* 19 (2), 276–289. doi:10.1038/s41423-021-00810-9
- Wu, H. T., Kuo, Y. C., Hung, J. J., Huang, C. H., Chen, W. Y., Chou, T. Y., et al. (2016). K63-polyubiquitinated HAUSP Deubiquitinates HIF-1alpha and Dictates H3K56 Acetylation Promoting Hypoxia-Induced Tumour Progression. *Nat. Commun.* 7, 13644. doi:10.1038/ncomms13644
- Wu, H., Yang, T. Y., Li, Y., Ye, W. L., Liu, F., He, X. S., et al. (2020). Tumor Necrosis Factor Receptor-Associated Factor 6 Promotes Hepatocarcinogenesis by Interacting with Histone Deacetylase 3 to Enhance C-Myc Gene Expression and Protein Stability. *Hepatology* 71 (1), 148–163. doi:10.1002/hep.30801

- Wu, X., and Karin, M. (2015). Emerging Roles of Lys63-Linked Polyubiquitylation in Immune Responses. *Immunol. Rev.* 266 (1), 161–174. doi:10.1111/imr.12310
- Wu, X., Zhang, W., Font-Burgada, J., Palmer, T., Hamil, A. S., Biswas, S. K., et al. (2014a). Ubiquitin-conjugating Enzyme Ubc13 Controls Breast Cancer Metastasis through a TAK1-P38 MAP Kinase Cascade. *Proc. Natl. Acad. Sci. U. S. A.* 111 (38), 13870–13875. doi:10.1073/pnas.1414358111
- Wu, Z., Shen, S., Zhang, Z., Zhang, W., and Xiao, W. (2017). Erratum to: Ubiquitin-Conjugating Enzyme Complex Uev1A-Ubc13 Promotes Breast Cancer Metastasis through Nuclear Factor-K $\kappa$ B Mediated Matrix Metalloproteinase-1 Gene Regulation. *Breast Cancer Res.* 19 (1), 41. doi:10.1186/s13058-017-0833-6
- Wu, Z., Shen, S., Zhang, Z., Zhang, W., and Xiao, W. (2014b). Ubiquitin-conjugating Enzyme Complex Uev1A-Ubc13 Promotes Breast Cancer Metastasis through Nuclear Factor-Small Ka, CyrillicB Mediated Matrix Metalloproteinase-1 Gene Regulation. *Breast Cancer Res.* 16 (4), R75. doi:10.1186/bcr3692
- Wu-Baer, F., Lagrizon, K., Yuan, W., and Baer, R. (2003). The BRCA1/BARD1 Heterodimer Assembles Polyubiquitin Chains through an Unconventional Linkage Involving Lysine Residue K6 of Ubiquitin. *J. Biol. Chem.* 278 (37), 34743–34746. doi:10.1074/jbc.C300249200
- Wu-Baer, F., Ludwig, T., and Baer, R. (2010). The UBXN1 Protein Associates with Autoubiquitinated Forms of the BRCA1 Tumor Suppressor and Inhibits its Enzymatic Function. *Mol. Cell Biol.* 30 (11), 2787–2798. doi:10.1128/MCB.01056-09
- Xiao, N., Li, H., Luo, J., Wang, R., Chen, H., Chen, J., et al. (2012). Ubiquitin-specific Protease 4 (USP4) Targets TRAF2 and TRAF6 for Deubiquitination and Inhibits TNF $\alpha$ -Induced Cancer Cell Migration. *Biochem. J.* 441 (3), 979–986. doi:10.1042/BJ20111358
- Xie, F., Zhang, Z., van Dam, H., Zhang, L., and Zhou, F. (2014). Regulation of TGF- $\beta$  Superfamily Signaling by SMAD Mono-Ubiquitination. *Cells* 3 (4), 981–993. doi:10.3390/cells3040981
- Xing, L., Tang, X., Wu, K., Huang, X., Yi, Y., and Huan, J. (2020). TRIM27 Functions as a Novel Oncogene in Non-triple-negative Breast Cancer by Blocking Cellular Senescence through P21 Ubiquitination. *Mol. Ther. Nucleic Acids* 22, 910–923. doi:10.1016/j.omtn.2020.10.012
- Xu, L., Jia, Y., Yang, X. H., Han, F., Zheng, Y., Ni, Y., et al. (2017). MicroRNA-130b Transcriptionally Regulated by Histone H3 Deacetylation Renders Akt Ubiquitination and Apoptosis Resistance to 6-OHDA. *Biochim. Biophys. Acta Mol. Basis Dis.* 1863 (6), 1678–1689. doi:10.1016/j.bbdis.2017.04.012
- Xu, P., Duong, D. M., Seyfried, N. T., Cheng, D., Xie, Y., Robert, J., et al. (2009). Quantitative Proteomics Reveals the Function of Unconventional Ubiquitin Chains in Proteasomal Degradation. *Cell* 137 (1), 133–145. doi:10.1016/j.cell.2009.01.041
- Xu, X., Kalac, M., Markson, M., Chan, M., Brody, J. D., Bhagat, G., et al. (2020). Reversal of CYLD Phosphorylation as a Novel Therapeutic Approach for Adult T-Cell Leukemia/lymphoma (ATLL). *Cell Death Dis.* 11 (2), 94. doi:10.1038/s41419-020-2294-6
- Xu, Y., Hu, Y., Xu, T., Yan, K., Zhang, T., Li, Q., et al. (2021). RNF8-mediated Regulation of Akt Promotes Lung Cancer Cell Survival and Resistance to DNA Damage. *Cell Rep.* 37 (3), 109854. doi:10.1016/j.celrep.2021.109854
- Xue, B., Li, H., Guo, M., Wang, J., Xu, Y., Zou, X., et al. (2018). TRIM21 Promotes Innate Immune Response to RNA Viral Infection through Lys27-Linked Polyubiquitination of MAVS. *J. Virol.* 92 (14). doi:10.1128/jvi.00321-18
- Yang, J. M., Schiapparelli, P., Nguyen, H. N., Igarashi, A., Zhang, Q., Abbadi, S., et al. (2017). Characterization of PTEN Mutations in Brain Cancer Reveals that Pten Mono-Ubiquitination Promotes Protein Stability and Nuclear Localization. *Oncogene* 36 (26), 3673–3685. doi:10.1038/onc.2016.493
- Yang, Y., Yang, C., Li, T., Yu, S., Gan, T., Hu, J., et al. (2020). The Deubiquitinase USP38 Promotes NHEJ Repair through Regulation of HDAC1 Activity and Regulates Cancer Cell Response to Genotoxic Insults. *Cancer Res.* 80 (4), 719–731. doi:10.1158/0008-5472.CAN-19-2149
- Yao, F., Zhou, Z., Kim, J., Hang, Q., Xiao, Z., Ton, B. N., et al. (2018). SKP2- and OTUD1-Regulated Non-proteolytic Ubiquitination of YAP Promotes YAP Nuclear Localization and Activity. *Nat. Commun.* 9 (1), 2269. doi:10.1038/s41467-018-04620-y
- Yasunaga, J., Lin, F. C., Lu, X., and Jeang, K. T. (2011). Ubiquitin-specific Peptidase 20 Targets TRAF6 and Human T Cell Leukemia Virus Type 1 Tax to Negatively Regulate NF- $\kappa$ B Signaling. *J. Virol.* 85 (13), 6212–6219. doi:10.1128/JVI.00079-11
- Yau, R. G., Doerner, K., Castellanos, E. R., Haakonsen, D. L., Werner, A., Wang, N., et al. (2017). Assembly and Function of Heterotypic Ubiquitin Chains in Cell Cycle and Protein Quality Control. *Cell* 171 (4), 918–933. doi:10.1016/j.cell.2017.09.040
- Yau, R., and Rape, M. (2016). The Increasing Complexity of the Ubiquitin Code. *Nat. Cell Biol.* 18 (6), 579–586. doi:10.1038/ncb3358
- Yeh, C. H., Bellon, M., and Nicot, C. (2018). FBXW7: a Critical Tumor Suppressor of Human Cancers. *Mol. Cancer* 17 (1), 115. doi:10.1186/s12943-018-0857-2
- Yin, H., Gui, Y., Du, G., Frohman, M. A., and Zheng, X. L. (2010). Dependence of Phospholipase D1 Multi-Monoubiquitination on its Enzymatic Activity and Palmitoylation. *J. Biol. Chem.* 285 (18), 13580–13588. doi:10.1074/jbc.M109.046359
- Yin, Q., Han, T., Fang, B., Zhang, G., Zhang, C., Roberts, E. R., et al. (2019). K27-linked Ubiquitination of BRAF by ITCH Engages Cytokine Response to Maintain MEK-ERK Signaling. *Nat. Commun.* 10 (1), 1870. doi:10.1038/s41467-019-09844-0
- Yu, X., Li, W., Deng, Q., Liu, H., Wang, X., Hu, H., et al. (2020a). MYD88 L265P Elicits Mutation-specific Ubiquitination to Drive NF- $\kappa$ B Activation and Lymphomagenesis. *Blood*. doi:10.1182/blood.2020004918
- Yu, X., Li, W., Liu, H., Deng, Q., Wang, X., Hu, H., et al. (2020b). Ubiquitination of the DNA-Damage Checkpoint Kinase CHK1 by TRAF4 Is Required for CHK1 Activation. *J. Hematol. Oncol.* 13 (1), 40. doi:10.1186/s13045-020-00869-3
- Yu, Z., Li, X., Yang, M., Huang, J., Fang, Q., Jia, J., et al. (2021). TRIM41 Is Required to Innate Antiviral Response by Polyubiquitinating BCL10 and Recruiting NEMO. *Signal Transduct. Target Ther.* 6 (1), 90. doi:10.1038/s41392-021-00477-8
- Zhang, H., Hu, H., Greeley, N., Jin, J., Matthews, A. J., Ohashi, E., et al. (2014). STAT3 Restrains RANK- and TLR4-Mediated Signalling by Suppressing Expression of the E2 Ubiquitin-Conjugating Enzyme Ubc13. *Nat. Commun.* 5, 5798. doi:10.1038/ncomms6798
- Zhang, L., Liu, Q., Liu, K. W., Qin, Z. Y., Zhu, G. X., Shen, L. T., et al. (2021). SHARPIN Stabilizes Beta-Catenin through a Linear Ubiquitination-independent Manner to Support Gastric Tumorigenesis. *Gastric Cancer* 24 (2), 402–416. doi:10.1007/s10120-020-01138-5
- Zhang, L., Zhou, F., Garcia de Vinuesa, A., de Kruijff, E. M., Mesker, W. E., Hui, L., et al. (2013a). TRAF4 Promotes TGF- $\beta$  Receptor Signaling and Drives Breast Cancer Metastasis. *Mol. Cell* 51 (5), 559–572. doi:10.1016/j.molcel.2013.07.014
- Zhang, Q., Karnak, D., Tan, M., Lawrence, T. S., Morgan, M. A., and Sun, Y. (2016). FBXW7 Facilitates Nonhomologous End-Joining via K63-Linked Polyubiquitylation of XRCC4. *Mol. Cell* 61 (3), 419–433. doi:10.1016/j.molcel.2015.12.010
- Zhang, X., Meng, T., Cui, S., Liu, D., Pang, Q., and Wang, P. (2022). Roles of Ubiquitination in the Crosstalk between Tumors and the Tumor Microenvironment (Review). *Int. J. Oncol.* 61 (1). doi:10.3892/ijo.2022.5374
- Zhang, X., Zhang, J., Zhang, L., van Dam, H., and ten Dijke, P. (2013b). UBE2O Negatively Regulates TRAF6-Mediated NF- $\kappa$ B Activation by Inhibiting TRAF6 Polyubiquitination. *Cell Res.* 23 (3), 366–377. doi:10.1038/cr.2013.21
- Zhang, Y., Li, Y., Yang, X., Wang, J., Wang, R., Qian, X., et al. (2018). Uev1A-Ubc13 Catalyzes K63-Linked Ubiquitination of RHBDF2 to Promote TACE Maturation. *Cell Signal* 42, 155–164. doi:10.1016/j.cellsig.2017.10.013
- Zhang, Z., Fan, Y., Xie, F., Zhou, H., Jin, K., Shao, L., et al. (2017). Breast Cancer Metastasis Suppressor OTUD1 Deubiquitinates SMAD7. *Nat. Commun.* 8 (1), 2116. doi:10.1038/s41467-017-02029-7
- Zhao, X., Lutz, J., Höllmüller, E., Scheffner, M., Marx, A., and Stengel, F. (2017). Identification of Proteins Interacting with Ubiquitin Chains. *Angew. Chem. Int. Ed. Engl.* 56 (49), 15764–15768. doi:10.1002/anie.201705898
- Zheng, J., Wang, B., Zheng, R., Zhang, J., Huang, C., Zheng, R., et al. (2020). LincRA1 Inhibits Autophagy and Promotes Radioresistance by Preventing H2Bub1/USP44 Combination in Glioma Cells. *Cell Death Dis.* 11 (9), 758. doi:10.1038/s41419-020-02977-x
- Zheng, Y. G., Wu, J., Chen, Z., and Goodman, M. (2008). Chemical Regulation of Epigenetic Modifications: Opportunities for New Cancer Therapy. *Med. Res. Rev.* 28 (5), 645–687. doi:10.1002/med.20120
- Zhi, H., Guo, X., Ho, Y. K., Pasupala, N., Engstrom, H. A. A., Semmes, O. J., et al. (2020). RNF8 Dysregulation and Down-Regulation during HTLV-1 Infection

- Promote Genomic Instability in Adult T-Cell Leukemia. *PLoS Pathog.* 16 (5), e1008618. doi:10.1371/journal.ppat.1008618
- Zhou, H., Liu, Y., Zhu, R., Ding, F., Wan, Y., Li, Y., et al. (2017). FBXO32 Suppresses Breast Cancer Tumorigenesis through Targeting KLF4 to Proteasomal Degradation. *Oncogene* 36 (23), 3312–3321. doi:10.1038/onc.2016.479
- Zhou, H., Wertz, I., O'Rourke, K., Ultsch, M., Seshagiri, S., Eby, M., et al. (2004). Bcl10 Activates the NF-kappaB Pathway through Ubiquitination of NEMO. *Nature* 427 (6970), 167–171. doi:10.1038/nature02273
- Zhou, L., Jiang, H., Du, J., Li, L., Li, R., Lu, J., et al. (2018). USP15 Inhibits Multiple Myeloma Cell Apoptosis through Activating a Feedback Loop with the Transcription Factor NF- $\kappa$ Bp65. *Exp. Mol. Med.* 50 (11), 1–12. doi:10.1038/s12276-018-0180-4
- Zhou, L., and Xu, G. (2022). The Ubiquitination-dependent and -Independent Functions of Cereblon in Cancer and Neurological Diseases. *J. Mol. Biol.* 434 (5), 167457. doi:10.1016/j.jmb.2022.167457
- Zhou, W., Zhong, Z., Lin, D., Liu, Z., Zhang, Q., Xia, H., et al. (2021). Hypothermic Oxygenated Perfusion Inhibits HECTD3-Mediated TRAF3 Polyubiquitination to Alleviate DCD Liver Ischemia-Reperfusion Injury. *Cell Death Dis.* 12 (2), 211. doi:10.1038/s41419-021-03493-2
- Zhu, F., Yi, G., Liu, X., Zhu, F., Zhao, A., Wang, A., et al. (2018). Ring Finger Protein 31-mediated Atypical Ubiquitination Stabilizes Forkhead Box P3 and Thereby Stimulates Regulatory T-Cell Function. *J. Biol. Chem.* 293 (52), 20099–20111. doi:10.1074/jbc.RA118.005802
- Zhu, G., Herlyn, M., and Yang, X. (2021). TRIM15 and CYLD Regulate ERK Activation via Lysine-63-Linked Polyubiquitination. *Nat. Cell Biol.* 23 (9), 978–991. doi:10.1038/s41556-021-00732-8
- Zhu, J., Li, X., Su, P., Xue, M., Zang, Y., and Ding, Y. (2020). The Ubiquitin Ligase RNF181 Stabilizes ERalpha and Modulates Breast Cancer Progression. *Oncogene* 39 (44), 6776–6788. doi:10.1038/s41388-020-01464-z
- Zhu, J., Zhuang, T., Yang, H., Li, X., Liu, H., and Wang, H. (2016). Atypical Ubiquitin Ligase RNF31: the Nuclear Factor Modulator in Breast Cancer Progression. *BMC Cancer* 16, 538. doi:10.1186/s12885-016-2575-8
- Zientara-Rytter, K., and Subramani, S. (2019). The Roles of Ubiquitin-Binding Protein Shuttles in the Degradative Fate of Ubiquitinated Proteins in the Ubiquitin-Proteasome System and Autophagy. *Cells* 8 (1). doi:10.3390/cells8010040
- Ziv, I., Matiuhiu, Y., Kirkpatrick, D. S., Erpapazoglou, Z., Leon, S., Pantazopoulou, M., et al. (2011). A Perturbed Ubiquitin Landscape Distinguishes between Ubiquitin in Trafficking and in Proteolysis. *Mol. Cell Proteomics* 10 (5), M111–M009753. doi:10.1074/mcp.M111.009753

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

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

Copyright © 2022 Yin, Liu, Liu, Tian, Yan, Han and Jiang. 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.



# Epigenetic Regulation in Knee Osteoarthritis

Zhengyu Cai<sup>†</sup>, Teng Long<sup>\*†</sup>, Yaochao Zhao, Ruixin Lin and You Wang<sup>\*</sup>

Department of Bone and Joint Surgery, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

## OPEN ACCESS

### Edited by:

Heng Sun,  
University of Macau, China

### Reviewed by:

Korosh Morshedi,  
Kashan University of Medical  
Sciences, Iran  
Jianjie Li,  
Xuzhou Medical University, China

### \*Correspondence:

Teng Long  
longteng@renji.com  
You Wang  
youwang@renji.com

<sup>†</sup>These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

Received: 13 May 2022

Accepted: 20 June 2022

Published: 08 July 2022

### Citation:

Cai Z, Long T, Zhao Y, Lin R and  
Wang Y (2022) Epigenetic Regulation  
in Knee Osteoarthritis.  
Front. Genet. 13:942982.  
doi: 10.3389/fgene.2022.942982

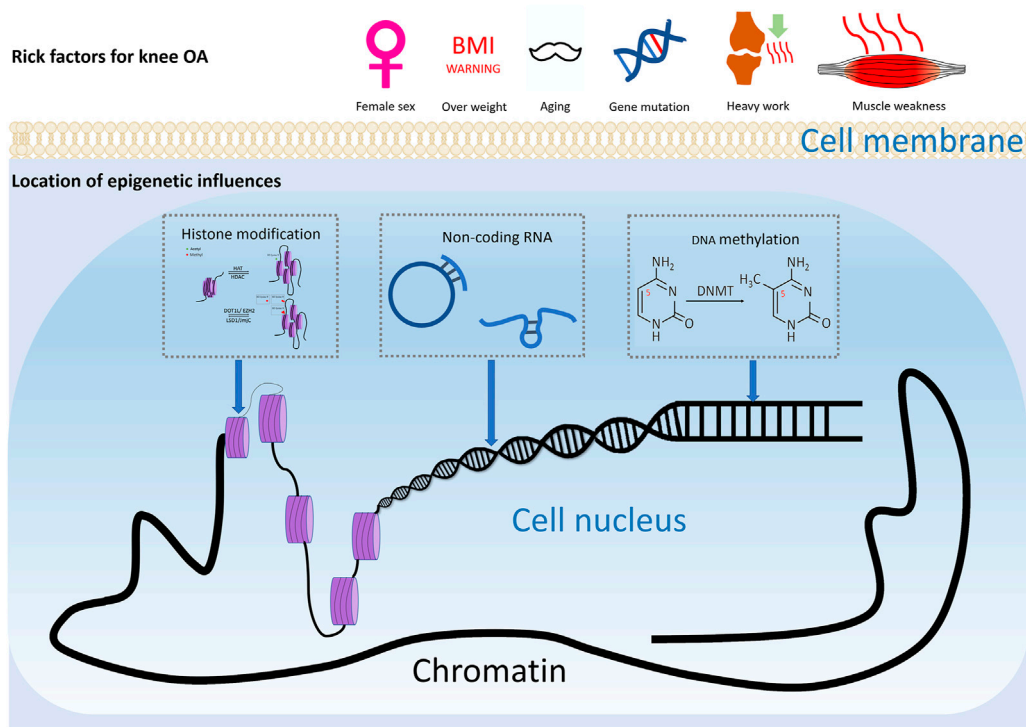
Osteoarthritis (OA) is a complicated disease with both hereditary and environmental causes. Despite an increase in reports of possible OA risk loci, it has become clear that genetics is not the sole cause of osteoarthritis. Epigenetics, which can be triggered by environmental influences and result in transcriptional alterations, may have a role in OA pathogenesis. The majority of recent research on the epigenetics of OA has been focused on DNA methylation, histone modification, and non-coding RNAs. However, this study will explore epigenetic regulation in OA at the present stage. How genetics, environmental variables, and epigenetics interact will be researched, shedding light for future studies. Their possible interaction and control processes open up new avenues for the development of innovative osteoarthritis treatment and diagnostic techniques.

**Keywords:** epigenetics, knee osteoarthritis, DNA methylation, histone modification, noncoding RNA

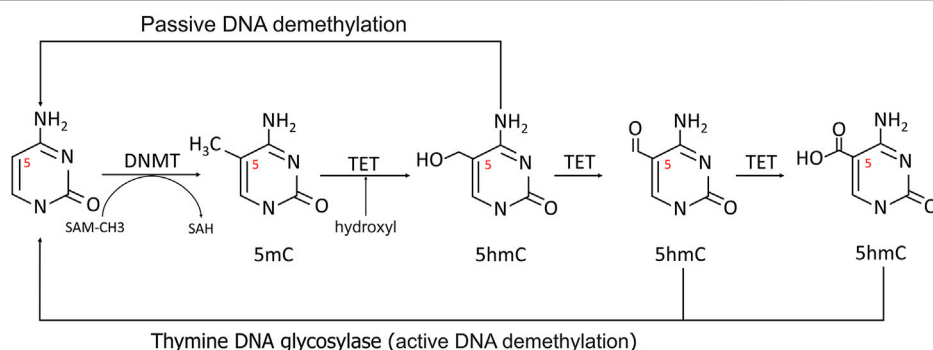
## INTRODUCTION

Osteoarthritis (OA) is a widespread joint disease that affects around 15% of the world's population (Glyn-Jones et al., 2015). OA was once thought to be a degenerative illness arose from chronic wear-and-tear and mechanical stress. However, the emerging paradigm now sees it as a complex process involving interactions between a wide range of internal and environmental factors (Figure 1). The genetic component of the disease is complex. Current evidence supports the theory of a polygenic inheritance, as published studies have reported a variety of OA-related risk loci (Styrkarsdottir et al., 2018; Zengini et al., 2018; Styrkarsdottir et al., 2019; Tachmazidou et al., 2019). However, studies have shown that more than 80% of the disease-related variants are located in non-coding regions of the genome (Maurano et al., 2012). As a result, it has been proposed that changes in gene expression, rather than changes in the genetic code sequence, are more likely to influence OA development. Indeed, epigenetics, a crucial method of gene expression regulation, has been implicated in the start and progression of OA in recent research (Ramos and Meulenbelt, 2017).

Epigenetics is an essential gene-environment interaction process. Epigenetic modification, unlike genomic modifications, is more versatile and reversible. Regulations of epigenetics vary by cell type and gene. Epigenetic phenomena including DNA methylation, genomic imprinting, maternal effects, post-translational modifications of histones, RNA (controlled by non-coding regulatory RNAs), and epigenetic chromatin remodeling (three-dimensional structure of chromatin) have been well studied. (Simon and Jeffries, 2017). With the help of modern testing techniques, epigenetic studies of knee osteoarthritis are now possible to reveal how the external environment affects changes in somatic cells and tissues. Many studies have been conducted on the epigenetics of osteoarthritis, although they have primarily focused on DNA methylation, histone modification, and miRNA (Ramos and Meulenbelt, 2017). This section will review the epigenetic regulation and reciprocity based on previous findings, trying to seek the connection, investigate how genetics and epigenetics interact, and steer future research toward identifying biologically significant alterations and gaining a better understanding of the mechanisms involved.



**FIGURE 1 |** Risk factors of knee OA. Risk factors for knee osteoarthritis includes ageing, gender, injury, and joint overloading, etc. Epigenetics may play a considerable role in how these environmental factors lead to altered gene expression and ultimately pathophysiologic manifestations such as cartilage damage and subchondral osteosclerosis.

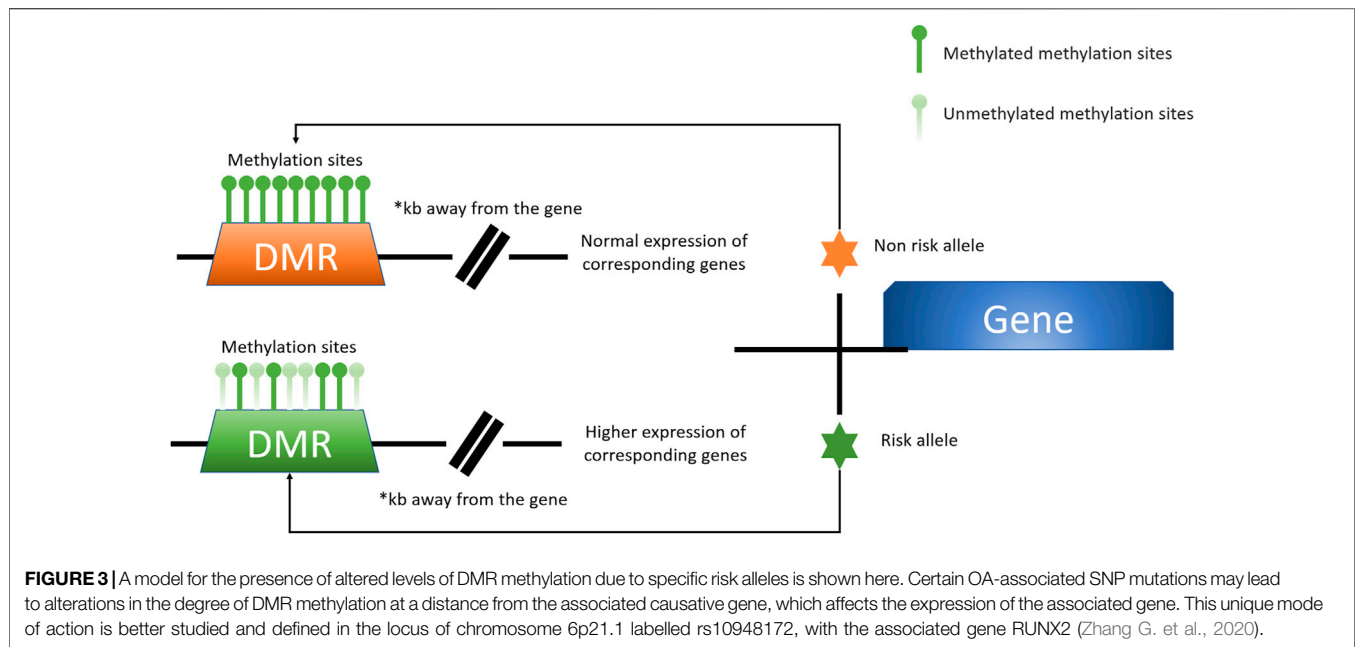


**FIGURE 2 |** Methylation and demethylation of DNA. In methylation process, the 5' carbon of cytosine was methylated to 5-mC with S-adenosine methionine (SAM) as the methyl donor. During active demethylation, the methyl group of 5-mC can be modified by the TETs-mediated addition of hydroxyl groups to generate 5-hydroxymethylcytosine (5hmC), which can be subsequently processed in demethylation process.

## DNA METHYLATION

The most fully studied methylation is 5-methylcytosine (5mC) (Miranda-Duarte, 2018). CpG dinucleotides are distributed unevenly throughout the human genome, with the bulk of CpG islands found in the promoter and exon regions of genes 300 to 3000bp in length. As a result, the methylation level of CpGs can be dynamically controlled by DNMTs and TETs to regulate transcription (**Figure 2**) (Rice et al., 2018).

Recent research into global and gene-specific methylation has proved a clear link between methylation and the development of OA. Helmtrud I. et al. were the first to study individual methylation loci. They investigated the amount of methylation of proximal promoter regions of aggrecanases1 (*ADAMTS 4*) and matrix metalloproteinase (*MMPs*) 2 (gelatinase A), *MMP3* (stromelysin 1), *MMP9* (gelatinase B), and *MMP13* (collagenase 3) (Roach et al., 2005). Despite the fact that the demethylation and sensitivity to demethylation of these four

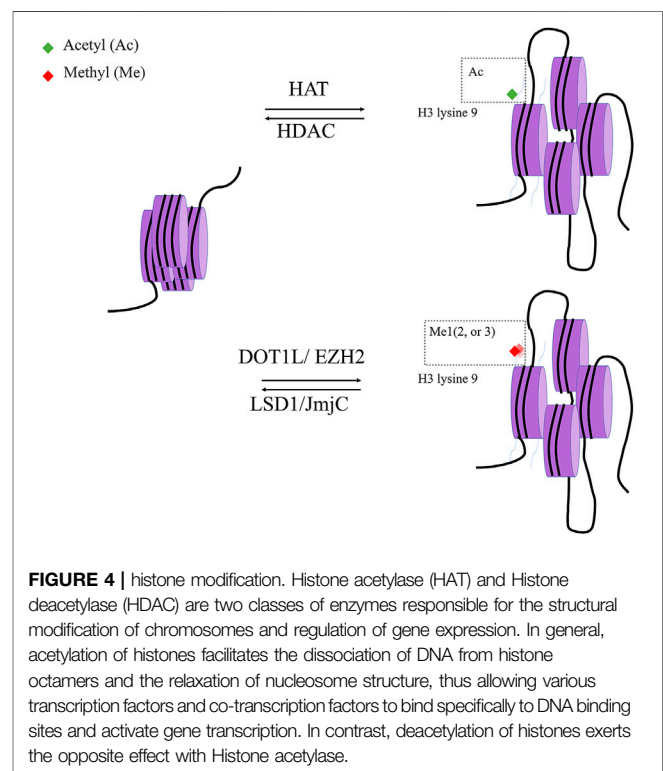


enzymes differed significantly, the degree of demethylation of at least one CpG site in each enzyme in OA was statistically higher than in the control group.

Additional DNA methylation locations encompassing numerous OA-related chemicals or pathways have been discovered in subsequent cartilage tissue studies, and these genes can be classified into four main categories. The first category addresses the extracellular matrix's homeostasis, which includes *COL2A1*, *COL9A1*, *COL10A1*, *ACAN*, *MMP2*, *MMP3*, *MMP9*, *MMP13*, *ADAMTS-4* (Pöschl et al., 2005; Roach et al., 2005; Dickhut et al., 2009; Imagawa et al., 2014). Inflammation-related molecules, such as *IL1b*, *IL8*, *IL32*, *TGF-β2*, *IL1RN*, and *WNT11*, fall under the second category (Hashimoto et al., 2009; Moazedi-Fuerst et al., 2014; Takahashi et al., 2015). *SOX4*, *SOX9*, *RUNX2*, and *SOD2* are members of the third group, which are involved in cartilage maintenance (Ezura et al., 2009; Kim et al., 2013). Growth factors such as *BMP7*, *SOST*, and *GDF5* fall within the fourth category (Loeser et al., 2009; Reynard et al., 2011; Papathanasiou et al., 2015). DNA methylation is known to vary greatly between tissues and stages of illness. Such studies only identify methylation sites that may be related to disease due to sampling issues. Still, these findings add to the current understanding of OA epigenetics.

## CpG Methylation Patterns Correlated With OA-Associated SNPs

Direct SNP correlations or gene-environment interactions may be mediated by these methylation quantitative trait loci (meQTLs). Eilis H. et al. investigated DNA methylation at 850,000 locations across the genome using a MethylationEPIC BeadChip (850K chip) in samples from the Understanding Society UK Household Longitudinal Study (UKHLS) (n = 1,111) (Hannon et al., 2018). They discovered 548 significant



DNA methylation quantitative trait loci correlations between 2,907,234 genetic variations and 93,268 DNA methylation sites, resulting in 548 significant DNA methylation quantitative trait loci. They offered two key concepts: 1) mQTL-associated SNP mutations were more common within 500 kb of the DNA methylation site (designated as cis), and the cis SNP variation had a much more considerable influence on

DNA methylation; 2) proximal methylation sites shared mQTL-associated SNP mutations. Differential methylation region refers to the methylation region regulated by such SNPs (DMR). These results explained the association between SNP changes and surrounding DMR in the proximal regions of several genes correlated with OA risk in osteoarthritis (*RUNX2*, *PLEC*, *ALDH1A2*, *GDF5*, *MGP*, *COLgalt2*, and *COL11A2*) (**Figure 3**) (Reynard et al., 2014; Shepherd et al., 2018; Rice et al., 2019a; Shepherd et al., 2019). Further mQTL research revealed 24 CpGs linked to genes including *COLgalt2*, *COL11A2*, *RAPH1*, *FGFR3* and *WWP2*. Importantly, *WWP2* encloses miRNA 140, which have been shown to be critical in cartilage maintenance, as miRNA 140 have severe cartilage defects (Duan et al., 2020). These findings demonstrate that DNA methylation can influence gene and non-coding RNA expression, as well as gene transcription and translation (**Table 1**).

## Risk Factors and DNA Methylation in OA

Animal models have provided light on the relationship between recognized OA risk factors like injury, aging and altered methylation patterns in the formation of OA. Destabilization of the medial meniscus, for example, can lead to OA development and altered 5mc and 5hmc patterns in 12-week-old male mice (Singh et al., 2019). Furthermore, tamoxifen-induced articular chondrocyte specific deletion of *DNMT3b* (*AGCCRE<sup>ERT2</sup>; DNMT3b<sup>fl/fl</sup>*), the *de novo* methyltransferase, can result in OA-like phenotypes in knee joints; whereas *DNMT3b* overexpression in articular chondrocytes can delay OA development after meniscus ligament injury (Shen et al., 2017).

Other OA risk factors, such as obesity, appear to be mediated by DNA methylation. Obesity has long been assumed to cause OA due to increased mechanical loading and wear and tear. Recent evidence shows that even in non-weight-bearing joints, such as hands, obesity still seems to be a significant risk factor. Leptin, a cytokine-like peptide hormone secreted by white adipose tissue, has long been stipulated to connect obesity with OA. Because obese patients generally have a high level of leptin in their bodies, their blood level is regulated significantly by body fat content. This is most likely owing to the obese patient's altered methylation pattern in the leptin (*lep*) promoter region. According to a recent study, patients after bariatric surgery have lower serum leptin levels and a different methylation profile in the *lep* and *lepr* gene promoter regions. These findings back up previous studies in rat models, which indicated that a high-energy diet could alter *lep* methylation patterns (Milagro et al., 2009). While inhibiting leptin signaling in *ob<sup>-/-</sup>* and *DB<sup>-/-</sup>* mice models resulted in obesity, it did not affect the occurrence of knee OA (Griffin et al., 2014). The suppression of *lep* methylation by 5'-Aza-2-deoxycytidine (AZA) increased MMP13 activity in patient chondrocytes *in vitro*. MMP-13 was elevated after leptin's epigenetic reactivation, and small interfering RNA against *lep* inhibited it directly (Iliopoulos et al., 2007). Overall, existing evidence suggests that leptin methylation could link obesity and OA development.

## The Role of DNA Methyltransferase in OA

Members of the *DNMTs* family include *DNMT1*, *DNMT2*, *DNMT3a*, *DNMT3b*, and *DNMT3l*. The main function of

*DNMT2* is to catalyze the methylation of RNA, while *DNMT3l* does not have catalytic activity. CpG methylation is determined by three DNA methyltransferases: *DNMT1*, *DNMT3A*, and *DNMT3B*. *DNMT1* is responsible for maintaining methylation during DNA replication and damage repair. *DNMT3A* and *DNMT3B* play major role in *de novo* methylation (Gao L. et al., 2020). A previous matched case-control study found a connection between DNMT polymorphisms and primary knee OA. Under a co-dominant and dominant paradigm, the CT haplotype of DNMT1 polymorphisms was related to a lower risk of OA. In contrast, the CC genotype of rs2424913 of DNMT3b was associated with an increased risk (Miranda-Duarte et al., 2020).

Differential expression of *DNMT1* in human chondrocytes collected from different areas in articular chondrocytes can be triggered by IL-1, a pro-inflammatory cytokine. Notably, deep and superficial zone chondrocytes, instead of transition zone chondrocytes, enhanced DNMT1 protein expression and activity in response to IL-1 (Akhtar and Haqqi, 2013). The demethylation of *DNMT1* by AZA can greatly increase the expression of pro-inflammatory or matrix catabolic proteins such as *COX2*, *MMP9*, and *MMP13*. Only *DNMT3b* was expressed in the mature normal cartilage of the knee joint and TMJ, but not *DNMT3a* or *DNMT1*. Additionally, and the occurrence of OA in both joints would lead to the decrease of *DNMT3b* expression (Shen et al., 2017). Overexpression of DNMT3b also appears to slow the onset of osteoarthritis caused by trauma or chemicals (Shen et al., 2017; Zhou et al., 2019).

## HISTONE MODIFICATION

The most in-depth study on OA's overall accessible chromatin landscape change was performed by comparing cartilage from both the outer region of the lateral tibial plateau and the inner region of the medial tibial plateau from the same patient. The accessible chromatin landscapes of injured and undamaged tissues display strikingly distinct chromatin signatures, notwithstanding patient-to-patient differences (Liu et al., 2018). Further analysis revealed that enhancers account for the majority of differentially accessible peaks, including enhancers from *BMP1b*, *WNT5a*, and *FGFR2*, all of which are known to play a role in OA-related cellular activity. Histone modification is shown to regulate some essential chondrocyte regulator genes, such as *SOX9*, as well as functioning in conjunction with epigenetic regulators to affect downstream genes (Kim et al., 2013). *SOX11* and *WNT5b*, two OA-related genes, had differing accessible peaks in their promoter regions.

## Histone Acetylation in OA

Histone acetylation is a dynamic process governed by two distinct enzyme families: HATs and HDACs. HDACs can modify nonhistone proteins to affect a variety of cellular processes in addition to their effects on chromatin structure (**Figure 4**).

Currently, most research on HDACs' functions in OA pathogenesis shows their interactions with non-histone proteins, which change target gene transcription. HDAC1 and

**TABLE 1 |** Genes with a differentially methylated region potentially related to single nucleotide polymorphism

| OA Associated SNP | CpG loci               | Potential gene   |
|-------------------|------------------------|--|
| rs6976            | cg15147215             | cg18099,408 <i>GNL3</i> Rushton et al. (2015)                      |
| rs10948172        | cg18591801             | cg13979708 <i>RUNX2</i> Rushton et al. (2015); Rice et al. (2019a) |
|                   | cg20913747             |  |
|                   | cg18551225             |  |
| rs3204689         | cg12031962             | - <i>ALDH1A2</i> Rushton et al. (2015)                             |
| rs143383          | cg14752227             | - <i>GDF5</i> Rushton et al. (2015)                                |
| rs10471753        | cg25008444             | - <i>PIK3R1</i> Rice et al. (2019a)                                |
| rs11780978        | cg02331830             | cg19405177 <i>PLEC</i>   |
|                   | cg04255391             |  |
|                   | cg14598846             |  |
|                   | cg23299254             |  |
|                   | cg10299941             |  |
| rs4764133         | cg20917083             | - <i>MGP</i> Rice et al. (2019a)                                   |
| rs6516886         | cg20220242             | cg00065302 <i>RWDD2B</i> Rice et al. (2019a)                       |
|                   | cg24751378             |  |
|                   | cg16140273             |  |
|                   | cg18131582             |  |
| rs11583641        | cg10114877             | - <i>COLGALT2</i> Rice et al. (2019b)                              |
| rs62182810        | cg25007799             | - <i>NBEAL1</i> Rice et al. (2019b)                                |
| rs11732213        | cg13921245             | cg20987369 <i>FGFR3</i> Rice et al. (2019b)                        |
| rs9277552         | cg02375585             | cg02197634 <i>COL11A2</i> Rice et al. (2019b)                      |
|                   | cg18170545             |  |
|                   | cg22375663             |  |
| rs60890741        | cg10253484             | cg20040747 <i>ASAP1</i> Rice et al. (2019b)                        |
| rs317630          | cg26736200             | - <i>CPSF1</i> Rice et al. (2019b)                                 |
| rs35206230        | cg26661922             | - <i>SEMA7A</i> Rice et al. (2019b)                                |
| rs6499244         | cg16779580             | - <i>NFAT5</i> Rice et al. (2019b)                                 |
|                   | cg16520312             |  |
| rs2953013         | cg18228076             | - <i>WWP2</i> Rice et al. (2019b)                                  |
| rs62063281        | cg01934064             | - <i>RAB11FIP4</i>   |
|                   | cg15633388             |  |
|                   | cg23616531             |  |
|                   | cg18131582             |  |
|                   | cg20220242             |  |
|                   | Not mentioned in paper |  |
| rs11583641        |                        | cg17117718 <i>LRRc37A</i>  |
| rs6516886         |                        | cg10826688 <i>CRHR1</i>  |
| rs75621460        |                        | cg15295732 <i>MAPT</i>   |
|                   |                        | cg11117266 <i>KANSL1</i> Rice et al. (2019b)                       |
|                   |                        | -  |
|                   |                        | <i>COLGALT2</i> Kehayova et al. (2021)                             |
|                   |                        | <i>RWDD2B</i> Parker et al. (2021)                                 |
|                   |                        | <i>TGFβ1</i> Rice et al. (2021)                                    |

HDAC2 were detected in large amounts in human chondrocytes and synovial cells from OA patients. They may alter the expression of cartilage-specific genes such as *COL9A1*, *COL11A1*, *COMP*, *AGGRECAN*, and *Dermatopontin*, likely via interaction with *SNAIL* protein (Hong et al., 2009). *HDAC4*, which is not indicated in normal knee cartilage, was highly expressed in OA cartilage in other research employing knee cartilage specimens from OA patients and normal donors. modulate *JNK* and *ERK1/2* activation to mediate *IL1β*-induced matrix catabolic protein synthesis (Wang et al., 2018). The *HDAC9-PIASy-RNF4* axis promotes chondrocyte hypertrophy by regulating the sumo and ubiquitination of *Nkx3.2/Bapx1*, which is degraded by the proteasome, according to studies on *HDAC9* in osteoarthritis (Choi et al., 2016).

A few studies on the regulation of mesenchymal stem cell destiny disclose that the epigenetic role of *HDAC* may also be involved in the development of OA. In rat bone marrow stromal cells, *HDAC8* can inhibit osteogenesis via two pathways: it can inhibit histone H3 lysine 9 acetylation, which reduces the osteogenic protein *RUNX2*, *Osterix*, *Osteopontin* and *ALP*. It can also associate with *RUNX2* to repress its transcriptional activity (Lee et al., 2019). Human MSCs increased global histone acetylation following *in situ* stiffening in a hydrogel

environment by reducing *HDAC* activity. Histone acetylation and Lamin A/C, the mechanosensor signaling protein, were considerably low in subchondral bone isolated from OA patients, while *HDAC* activity was significantly higher (Lee et al., 2019).

Sirtuins (*SIRT*s) are histone and protein deacetylases dependent on nicotinamide dinucleotide (NAD<sup>+</sup>). *SIRT*s and *SIRT*-dependent epigenetic control have been widely reported to play vital roles in DNA repair, inflammation, and aging-related illnesses (Lee et al., 2019). Chondrocyte hypertrophy and proteoglycan production were reduced in *SIRT6*-deficient mice from the mesenchyme (Ailixiding et al., 2015). Sirtuin 1 (*SIRT1*), like *SOX9*, is essential for maintaining cartilage homeostasis (Tsuda et al., 2003). Boosted *SIRT1* expression increased the expression of cartilage ECM genes such as *COL2A1*, *COL9A1*, *AGGRECAN* and *COMP*. *SIRT1* has been found to connect with *SOX9* and *p300*, *GN5*, both of which are involved in nucleosome acetylation in the *COL2* promoter region. *SIRT3* is mostly found in mitochondria, and its deacetylation activity regulates mitochondrial function, regeneration, and kinetics (Ansari et al., 2017). These activities are hypothesized to maintain REDOX equilibrium in cell metabolism, preventing oxidative stress (Bause and Haigis, 2013).

**TABLE 2 |** lncRNA which utilize with miRNA in OA progressing.

| lncRNA         | miRNA           | Interrelated target/Regulators  |
|----------------|-----------------|---|
| SNHG14         | miR-124-3-p     | FSTL-1, NLRP3, TLR4/NF- $\kappa$ B pathway Wang B. et al. (2021)  |
| LINC02288      | miR-374a-3p     | RTN3 Fu et al. (2021)   |
| Linc-ROR       | miR-138/miR-145 | SOX9 Feng et al. (2021)   |
| RNA HOTTIP     | miR-663a        | Fyn-related kinase He et al. (2021)   |
| RNA RMRP       | miR-206         | CDK9 Lu et al. (2020)   |
| RNA SNHG16     | miR-373-3p      | PPARGC1B signaling pathway (sponging miR-373-3p) Fan et al. (2021)  |
| RNA SNHG7      | miR-214-5p      | PPARGC1B signaling pathway Xu et al. (2021)   |
| RNA GAS5       | miR-137         | caspase-3, Bax/Bcl-2 Gao S. T. et al. (2020)  |
| PVT1           | miR-93-5p       | HMGB1, TLR4, NF- $\kappa$ B pathway Meng et al. (2020)  |
| RNA XIST       | miR-27b-3p      | ADAMTS-5 AXIS Zhu et al. (2021)   |
| RNA NEAT1      | miR-543         | MMP-3, MMP-9, MMP-13, interleukin (IL)-6 and IL-8 PLA2G4A axis Xiao et al. (2021)   |
| RNA SNHG5      | miR-10a-5p      | IL-1 $\beta$ , H3F3B axis, sponging miR-10a-5P (Jiang et al. (2021)   |
| ARFRP1         | miR-15a-5p      | NF- $\kappa$ b, TLR4 axis Zhang G. et al. (2020)  |
| PCAT1          | miR-27b-3p      | sponging miR-27b-3p (Zhou et al. (2021)   |
| MIR4435-2HG    | miR-510-3p      | MMP1,MMP13, collagen II,IL17-A, <!--Soft-enter Run-on--> p65, phosphorylated (p)-p65, I $\kappa$ B and p-I $\kappa$ B in CHON-001, sponging miR-510-3p Liu et al. (2020a) |
| SNHG9          | miR-34a         | methylation Zhang H. et al. (2020)  |
| H19            | miR-483-5p      | Dusp5 Wang et al. (2020)  |
| LOXL1-AS1      | miR-423-5P      | KDM5C axis, Chen K. et al. (2020)   |
| SNHG15         | miR-7           | KLF4, sponging miR7a Chen et al. (2020a)  |
| Loop LINC00511 | miR-150-5P      | the 3'-UTR of transcription factor (SP1) Zhang Y. et al. (2020)   |
| XIST           | miR-149-5p      | DNMT3A Liu et al. (2020b)   |
| XIST           | miR-653-5p      | DNMT3A Lian and Xi, (2020)  |
| IGHCy1         | miR-6891-3p     | TLR4 Zhang P. et al. (2020)   |
| HOTAIR         | miR-20b         | PTEN Chen et al. (2020b)  |
| H19            | miR-106b-5p     | TIMP2 Tan et al. (2020)   |
| SNHG15         | miR-141-3p      | BCL2L13 Zhang X. et al. (2020)  |

Studies concentrating on deacetylase/acetylases still have significant limitations because *HATs/HDACs* can modify transcription factors as well as signaling molecules, thus affecting the expression of other distant genes (Zhao et al., 2008). *HDAC* inhibitors, for example, can stimulate the acetylation of other proteins in several physiological pathways in addition to increasing histone acetylation. The non-targeted protein deacetylase inhibitors Zolinza (Vorinostat) and Istodax (Romidepsin) have been approved for the treatment of cutaneous T-cell lymphoma. Histone modification medicines for osteoarthritis, on the other hand, have not yet been developed.

## Histone Methylation

Typically, histone methylation is more stable than histone acetylation. Histone methylation is performed by histone methylation transferase (HMT) on lysine and arginine residues. Common locations of methylation include Lys 4, 9, 27, 36, 79 and Arg 2, 17, 26 on H3, Lys 20 and Arg 3 on H4, and Lys 4, 9, 27, 36, 79 and Arg 2, 17, 26 on H3 and H4. Studies have shown that arginine methylation on histone is a relatively dynamic marker. Arginine methylation is associated with gene activation, whereas arginine methylation loss in H3 and H4 is linked to gene silence (Greer and Shi, 2012). In contrast, lysine methylation appears to be a reasonably persistent marker for gene expression regulation.

Similar to histone acetylation, changes in histone methylation are frequently related to altered gene expression and signaling pathways in chondrocytes. For instance, H3k79 methylation was reduced in OA and RA patients (He et al., 2017), while H3K9 methylation was decreased in the temporomandibular joints of elderly mice (Ukita

et al., 2020). DOT1L, an enzyme involved in histone methylation of Lys79 of H3 (H3K79), was a cartilage homeostasis regulator (Monteagudo et al., 2017). *Dot1l* polymorphism was found to be associated with hip joint cartilage thickness and OA risk in research involving 6,532 people (Castaño Betancourt et al., 2012). According to an *in vitro* study, DOT1L can methylate H3K79 of the *LEFL* and *TCF1* genes, hence reducing WNT pathway activation and causing chondrocyte hypertrophy (Monteagudo et al., 2017). *EZH2* is another histone methyltransferase discovered to modulate the WNT pathway. *EZH2*, which is elevated in OA articular chondrocytes, can increase trimethylation of the *SFRP1* promoter and is a WNT inhibitor, leading to hyperactivation of the WNT/ $\beta$ -catenin signaling pathway (Chen L. et al., 2016). Overexpression of *EZH2* increases the expression of *MMP13*, *ADAMTS5*, and *COLX*, via methylation of miR-138 promoter, resulting in cartilage breakdown (Chen L. et al., 2016). In addition to histone methylation, the demethylation process is implicated in OA as well. Jumonji domain-containing 3 (*JMJD3*), HEK27Me3 demethylase, was found at increased level in cartilage from the tibial plateau of the OA knee. Its inhibition *in vitro* by *GSK4*, significantly increases the expression of OA-related genes *MMP13* and *PTGS2*.

## NON-CODING RNA

### microRNA (miRNA) and OA

Mature miRNAs bind to complementary messenger RNA (mRNA) sequences of target genes via RNA-induced silencing complexes (RISCs) or block gene expression directly. Interactions

between miRNAs and highly complementary targets lead to mRNA degradation, while incomplete interactions between miRNAs and target transcripts usually lead to translation suppression. In 2008, Dimitrios et al. found that in the chondrocytes of arthroplasty patients, miR-483, miR-22, miR-377, miR-103, miR-16, miR-223, miR-2b, miR-23b, and miR-509 were elevated, and miR-29a, miR-140, miR-25, miR-337, miR-210, miR-26a, and miR-373 were decreased. The genetic signatures of these 17 miRNAs clearly distinguish OA chondrocytes from normal chondrocytes, and 17 miRNA-protein pairs that may be involved in the progression of osteoarthritis were revealed by matching microRNAs and proteomics. In this study, miR-22 and miR-103 expression were favorably connected with BMI, but miR-25, and miR-337, while miR-29a expression was inversely correlated with BMI (Iliopoulos et al., 2008). Since then, numerous research has investigated the connection between miRNA expression and effector genes in OA, including inflammation, aging, transcription factors, apoptosis, autophagy, and other pathogenic events in the evolution of the illness (Zhang et al., 2015; Wang et al., 2016; D'Adamo et al., 2016; Miyaki et al., 2010). MiR-140 is a well-known miRNA associated with osteoarthritis that can be found in peripheral blood circulation or synovial fluid during the progression of OA (Miyaki et al., 2010). In articular cartilage of OA patients, the extracellular matrix is actively remodeled under inflammatory conditions, altering the local biomechanical properties of chondrocytes and accelerating the course of OA, attesting that MiR-146a has a significant function in OA inflammatory induction (Yamasaki et al., 2009). Other previously unmentioned miRNAs associated with ECM degradation include miR-137, miR-449, miR-221, miR-30a, miR-19b-3p, miR-634, miR-29, miR-107, miR-497-5p, miR-26a, miR-101 (Akbari Dilmaghani et al., 2021). Certain miRNAs, such as miR-378 in synovial fluids and let-7e in blood, might change their expression level as the disease progresses (Beyer et al., 2015; Li et al., 2016). The intra-articular injection of certain microRNAs has the ability to reverse disease progression, which will revolutionize the treatment of osteoarthritis.

## lncRNA and OA

Unlike microRNAs that rely mostly on RNA sequence complementary pairing to suppress target genes, lncRNAs operate in a significantly more sophisticated manner. lncRNAs have a unique role in a variety of gene expression regulation mechanisms, including epigenetic regulation, transcriptional regulation, and post-transcriptional regulation (Hoolwerff et al., 2020; Wang Z. et al., 2021; Statello et al., 2021).

Current research has revealed the way of axial regulation of the highly expressed lncRNA in OA tissue to the appropriate miRNA, which subsequently influences the OA target genes. HOTTIP, for instance, was engaged in the proliferation and death of OA chondrocytes via the miR-663a/FRK axis (He et al., 2021). Other lncRNA and miRNA interaction modes are more unique. Through spatial conformation, the secondary structure created by lncRNA can exert a sponge-like adsorption effect on miRNA, alter the actual contact concentration of miRNA, and the production of inflammatory genes or transcription factors in OA

tissue (Chen et al., 2020a; Liu et al., 2020a; Zhang Y. et al., 2020; Zhou et al., 2021; Fan et al., 2021; Jiang et al., 2021).

## COMBINATION OF EPIGENETIC FACTORS

In the pathological process of osteoarthritis, a multitude of epigenetic variables may influence the expression of OA-related genes. For instance, during articular chondrocyte death, *DNMT3b* facilitated the downregulation of miR-29 by increasing its promoter methylation. This, in turn, led to the overexpression of *PTH1LH*, a process strikingly similar to that observed in tumor disorders (Dou et al., 2020). Numerous studies have identified a connection between HDACs and miRNA cooperation patterns of OA, including *HDAC3*, *HDAC4*, *HDAC7*, and *HDAC8* (Chen et al., 2016b; Chen et al., 2016a; Mao et al., 2018; Meng et al., 2018; Zhang et al., 2019). For instance, miRNA-381 targets the 3'UTR of *HDAC4*, which in turn leads to increase in acetylation of *H3*, *RUNX2* and *MMP13*, and ultimately result in chondrocyte hypertrophy (Chen et al., 2016a). In **Table 2**, a selection of the reasonably well-studied regulatory mechanisms of these models is listed.

## CONCLUSION

As a form of non-coding genetic information, epigenetics plays distinct regulatory roles in disease onset and progression. Epigenetics has made great strides in oncology research, and drugs targeting DNA methylation and histone deacetylases have also emerged. Numerous studies in the field of OA research have demonstrated that epigenetics regulations produce interconnected and dynamic disease progression changes. OA-related signaling pathways, transcription factors, inflammatory factors, extracellular matrix (ECM) proteins, and other variables traditionally associated with OA are all sensitive to epigenetic control to varying degrees. The combination of these genetic and epigenetic variables provides more insight into the pathophysiology of osteoarthritis. Epigenetic control can link a range of genetic and environmental factors, which could provide a holistic understanding of the etiology of osteoarthritis (OA) and guide treatment.

At present, the epigenetics of osteoarthritis focuses primarily on chondrocytes, with only two published works addressing the epigenetic manifestations of subchondral bone (Jeffries et al., 2016; Zhang et al., 2016). Zhang Y et al. discovered that not only did the subchondral bone share 111 differential methylated probes (DMPs) and 41 differential methylated genes (DMGs) with chondrocytes, but they also proposed a novel hypothesis that the subchondral compartment epigenetic changes take precedence over cartilage in the development of osteoarthritis by comparing the tibia zoning analysis (Zhang et al., 2016). Due to the differential expression of epigenetics in various tissues, subchondral bone, synovium, ligaments, and other joint tissues besides chondrocytes may contradict experimental results.

For future approaches to epigenetics in the realm of osteoarthritis, it would be prudent to investigate refined single-cell research. In cartilage tissue isolated from OA samples with variable degrees of injury and degeneration, there exist

chondrocytes, macrophages, fibroblasts, and numerous other types of tissue. This heterogeneity can compromise the accuracy of current knowledge of gene function and epigenetic regulation. In the case of osteoarthritic cartilage, single-cell investigations provide more precise subtyping of diseased tissues. Embryonic cell research has demonstrated the validity of single-cell methylation investigations (Lorthongpanich et al., 2013). Using EpiTOF, histone modification investigations on a single cell are now also feasible. Adopting lanthanide metal isotopes labeled antibodies and mass spectrometry, this method allows epigenetic landscape profiling at single-cell precision (Cheung et al., 2018). In the future, these techniques could be deployed in osteoarthritis research, and this refinement will bring a fresh viewpoint to the field of osteoarthritis research. Large samples and extensive genetic and epigenetic analysis, along with prospective studies of pertinent patient histories and imaging presentations, will eventually provide solid evidence to guide the tertiary prevention and individualized and precise therapy of osteoarthritis.

## REFERENCES

- Ailixiding, M., Aibibula, Z., Iwata, M., Piao, J., Hara, Y., Koga, D., et al. (2015). Pivotal Role of Sirt6 in the Crosstalk Among Ageing, Metabolic Syndrome and Osteoarthritis. *Biochem. Biophysical Res. Commun.* 466 (3), 319–326. doi:10.1016/j.bbrc.2015.09.019
- Akbari Dilmaghani, N., Shoorei, H., Sharifi, G., Mohaqiq, M., Majidpoor, J., Dinger, M. E., et al. (2021). Non-coding RNAs Modulate Function of Extracellular Matrix Proteins. *Biomed. Pharmacother.* 136, 111240. doi:10.1016/j.biopha.2021.111240
- Akhtar, N., and Haqqi, T. M. (2013). In Human Chondrocytes DNMT-1 Plays a Key Role in the Expression of Several Genes Associated with the Pathogenesis of Osteoarthritis (OA). *Osteoarthr. Cartil.* 21 (Suppl. 1), S35–S36. doi:10.1016/j.joca.2013.02.094
- Ansari, A., Rahman, M. S., Saha, S. K., Saikot, F. K., Deep, A., and Kim, K.-H. (2017). Function of the SIRT3 Mitochondrial Deacetylase in Cellular Physiology, Cancer, and Neurodegenerative Disease. *Aging Cell* 16 (1), 4–16. doi:10.1111/acel.12538
- Bause, A. S., and Haigis, M. C. (2013). SIRT3 Regulation of Mitochondrial Oxidative Stress. *Exp. Gerontol.* 48 (7), 634–639. doi:10.1016/j.exger.2012.08.007
- Beyer, C., Zampetaki, A., Lin, N.-Y., Kleyer, A., Perricone, C., Iagnocco, A., et al. (2015). Signature of Circulating microRNAs in Osteoarthritis. *Ann. Rheum. Dis.* 74 (3), e18. doi:10.1136/annrheumdis-2013-204698
- Castaño Betancourt, M. C., Cailotto, F., Kerkhof, H. J., Cornelis, F. M. F., Doherty, S. A., Hart, D. J., et al. (2012). Genome-wide Association and Functional Studies Identify the DOT1L Gene to Be Involved in Cartilage Thickness and Hip Osteoarthritis. *Proc. Natl. Acad. Sci. U.S.A.* 109 (21), 8218–8223. doi:10.1073/pnas.1119899109
- Chen, K., Fang, H., and Xu, N. (2020). LncRNA LOXL1-AS1 Is Transcriptionally Activated by JUND and Contributes to Osteoarthritis Progression via Targeting the miR-423-5p/KDM5C axis. *Life Sci.* 258, 118095. doi:10.1016/j.lfs.2020.118095
- Chen, L., Wu, Y., Wu, Y., Wang, Y., Sun, L., and Li, F. (2016). The Inhibition of EZH2 Ameliorates Osteoarthritis Development through the Wnt/ $\beta$ -Catenin Pathway. *Sci. Rep.* 6, 29176. doi:10.1038/srep29176
- Chen, W., Sheng, P., Huang, Z., Meng, F., Kang, Y., Huang, G., et al. (2016a). MicroRNA-381 Regulates Chondrocyte Hypertrophy by Inhibiting Histone Deacetylase 4 Expression. *Int. J. Mol. Sci.* 17 (9), 1377. doi:10.3390/ijms17091377
- Chen, W., Chen, L., Zhang, Z., Meng, F., Huang, G., Sheng, P., et al. (2016b). MicroRNA-455-3p Modulates Cartilage Development and Degeneration through Modification of Histone H3 Acetylation. *Biochimica Biophysica Acta (BBA) - Mol. Cell Res.* 1863 (12), 2881–2891. doi:10.1016/j.bbamcr.2016.09.010
- Chen, Y., Guo, H., Li, L., Bao, D., Gao, F., Li, Q., et al. (2020a). Long Non-coding RNA (lncRNA) Small Nucleolar RNA Host Gene 15 (SNHG15) Alleviates Osteoarthritis Progression by Regulation of Extracellular Matrix Homeostasis. *Med. Sci. Monit.* 26, e923868. doi:10.12659/MSM.923868
- Chen, Y., Zhang, L., Li, E., Zhang, G., Hou, Y., Yuan, W., et al. (2020b). Long-chain Non-coding RNA HOTAIR Promotes the Progression of Osteoarthritis via Sponging miR-20b/PTEN axis. *Life Sci.* 253, 117685. doi:10.1016/j.lfs.2020.117685
- Cheung, P., Vallania, F., Warsinske, H. C., Donato, M., Schaffert, S., Chang, S. E., et al. (2018). Single-Cell Chromatin Modification Profiling Reveals Increased Epigenetic Variations with Aging. *Cell* 173 (6), 1385–1397. doi:10.1016/j.cell.2018.03.079
- Choi, H.-J., Kwon, S., and Kim, D.-W. (2016). A Post-translational Modification Cascade Employing HDAC9-PIASy-RNF4 axis Regulates Chondrocyte Hypertrophy by Modulating Nkx3.2 Protein Stability. *Cell. Signal.* 28 (9), 1336–1348. doi:10.1016/j.cellsig.2016.06.006
- D'Adamo, S., Alvarez-Garcia, O., Muramatsu, Y., Flamigni, F., and Lotz, M. K. (2016). MicroRNA-155 Suppresses Autophagy in Chondrocytes by Modulating Expression of Autophagy Proteins. *Osteoarthr. Cartil.* 24 (6), 1082–1091. doi:10.1016/j.joca.2016.01.005
- Dickhut, A., Peltari, K., Janicki, P., Wagner, W., Eckstein, V., Egermann, M., et al. (2009). Calcification or Dedifferentiation: Requirement to Lock Mesenchymal Stem Cells in a Desired Differentiation Stage. *J. Cell. Physiol.* 219 (1), 219–226. doi:10.1002/jcp.21673
- Dou, P., He, Y., Yu, B., and Duan, J. (2020). Downregulation of microRNA-29b by DNMT3B Decelerates Chondrocyte Apoptosis and the Progression of Osteoarthritis via PTHLH/CDK4/RUNX2 axis. *Aging (Albany NY)* 13, 7676–7690. doi:10.18632/aging.103778
- Duan, L., Liang, Y., Xu, X., Xiao, Y., and Wang, D. (2020). Recent Progress on the Role of miR-140 in Cartilage Matrix Remodelling and its Implications for Osteoarthritis Treatment. *Arthritis Res. Ther.* 22 (1), 194. doi:10.1186/s13075-020-02290-0
- Ezura, Y., Sekiya, I., Koga, H., Muneta, T., and Noda, M. (2009). Methylation Status of CpG Islands in the Promoter Regions of Signature Genes during Chondrogenesis of Human Synovium-Derived Mesenchymal Stem Cells. *Arthritis Rheum.* 60 (5), 1416–1426. doi:10.1002/art.24472
- Fan, H., Ding, L., and Yang, Y. (2021). LncRNA SNHG16 Promotes the Occurrence of Osteoarthritis by Sponging miR-373-3p. *Mol. Med. Rep.* 23 (2). doi:10.3892/mmr.2020.11756
- Feng, L., Yang, Z. M., Li, Y. C., Wang, H. X., Lo, J. H. T., Zhang, X. T., et al. (2021). Linc-ROR Promotes Mesenchymal Stem Cells Chondrogenesis and Cartilage

## AUTHOR CONTRIBUTIONS

ZC and TL contributed equally to the writing of the article and shared the title of co-first author. TL and YW share the corresponding author.

## FUNDING

This study was funded by Natural Science Foundation of China (no. 81501855).

## ACKNOWLEDGMENTS

Special thanks go to other Department of Joint Surgery colleagues for their insightful comments on this study and Qianxia Huang for English language editing.

- Formation via Regulating SOX9 Expression. *Osteoarthr. Cartil.* 29 (4), 568–578. doi:10.1016/j.joca.2020.12.020
- Fu, Q., Zhu, J., Wang, B., Wu, J., Li, H., Han, Y., et al. (2021). LINC02288 Promotes Chondrocyte Apoptosis and Inflammation through miR-374a-3p Targeting RTN3. *J. Gene Med.* 23, e3314. doi:10.1002/jgm.3314
- Gao, L., Emperle, M., Guo, Y., Grimm, S. A., Ren, W., Adam, S., et al. (2020). Comprehensive Structure-Function Characterization of DNMT3B and DNMT3A Reveals Distinctive De Novo DNA Methylation Mechanisms. *Nat. Commun.* 11 (1), 3355. doi:10.1038/s41467-020-17109-4
- Gao, S. T., Yu, Y. M., Wan, L. P., Liu, Z. M., and Lin, J. X. (2020). LncRNA GAS5 Induces Chondrocyte Apoptosis by Down-Regulating miR-137. *Eur. Rev. Med. Pharmacol. Sci.* 24 (21), 10984–10991. doi:10.26355/eurrev\_202011\_23582
- Glyn-Jones, S., Palmer, A. J. R., Agricola, R., Price, A. J., Vincent, T. L., Weinans, H., et al. (2015). Osteoarthritis. *Lancet* 386 (9991), 376–387. doi:10.1016/s0140-6736(14)60802-3
- Greer, E. L., and Shi, Y. (2012). Histone Methylation: a Dynamic Mark in Health, Disease and Inheritance. *Nat. Rev. Genet.* 13 (5), 343–357. doi:10.1038/nrg3173
- Griffin, T. M., Huebner, J. L., Kraus, V. B., and Guilak, F. (2014). Extreme Obesity Due to Impaired Leptin Signaling in Mice Does Not Cause Knee Osteoarthritis. *Arthritis Rheum.* 60 (10), 2935–2944. doi:10.1002/art.24854
- Hannon, E., Gorrie-Stone, T. J., Smart, M. C., Burrage, J., Hughes, A., Bao, Y., et al. (2018). Leveraging DNA-Methylation Quantitative-Trait Loci to Characterize the Relationship between Methyloic Variation, Gene Expression, and Complex Traits. *Am. J. Hum. Genet.* 103 (5), 654–665. doi:10.1016/j.ajhg.2018.09.007
- Hashimoto, K., Oreffo, R. O. C., Gibson, M. B., Goldring, M. B., and Roach, H. I. (2009). DNA Demethylation at Specific CpG Sites in the IL1Bpromoter in Response to Inflammatory Cytokines in Human Articular Chondrocytes. *Arthritis Rheum.* 60 (11), 3303–3313. doi:10.1002/art.24882
- He, D., Liu, J., Hai, Y., Zhu, Q., Shen, Y., Guo, S., et al. (2017). Increased DOT1L in Synovial Biopsies of Patients with OA and RA. *Clin. Rheumatol.* 37 (5), 1327–1332. doi:10.1007/s10067-017-3941-x
- He, X., Gao, K., Lu, S., and Wu, R. (2021). LncRNA HOTTIP Leads to Osteoarthritis Progression via Regulating miR-663a/Fyn-Related Kinase axis. *BMC Musculoskelet. Disord.* 22 (1), 67. doi:10.1186/s12891-020-03861-7
- Hong, S., Derfoul, A., Pereira-Mouries, L., and Hall, D. J. (2009). A Novel Domain in Histone Deacetylase 1 and 2 Mediates Repression of Cartilage-specific Genes in Human Chondrocytes. *FASEB J.* 23 (10), 3539–3552. doi:10.1096/fj.09-133215
- Hoolwerff, M., Metselaar, P. I., Tuerlings, M., Suchiman, H. E. D., Lakenberg, N., Ramos, Y. F. M., et al. (2020). Elucidating Epigenetic Regulation by Identifying Functional Cis-Acting Long Noncoding RNAs and Their Targets in Osteoarthritic Articular Cartilage. *Arthritis Rheumatol.* 72 (11), 1845–1854. doi:10.1002/art.41396
- Iliopoulos, D., Malizos, K. N., Oikonomou, P., and Tsezou, A. (2008). Integrative microRNA and Proteomic Approaches Identify Novel Osteoarthritis Genes and Their Collaborative Metabolic and Inflammatory Networks. *PLoS One* 3 (11), e3740. doi:10.1371/journal.pone.0003740
- Iliopoulos, D., Malizos, K. N., and Tsezou, A. (2007). Epigenetic Regulation of Leptin Affects MMP-13 Expression in Osteoarthritic Chondrocytes: Possible Molecular Target for Osteoarthritis Therapeutic Intervention. *Ann. Rheumatic Dis.* 66 (12), 1616–1621. doi:10.1136/ard.2007.069377
- Imagawa, K., de Andrés, M. C., Hashimoto, K., Itoi, E., Otero, M., Roach, H. I., et al. (2014). Association of Reduced Type IX Collagen Gene Expression in Human Osteoarthritic Chondrocytes with Epigenetic Silencing by DNA Hypermethylation. *Arthritis & Rheumatology* 66 (11), 3040–3051. doi:10.1002/art.38774
- Jeffries, M. A., Donica, M., Baker, L. W., Stevenson, M. E., Annan, A. C., Beth Humphrey, M., et al. (2016). Genome-Wide DNA Methylation Study Identifies Significant Epigenomic Changes in Osteoarthritic Subchondral Bone and Similarity to Overlying Cartilage. *Arthritis & Rheumatology* 68 (6), 1403–1414. doi:10.1002/art.39555
- Jiang, H., Pang, H., Wu, P., Cao, Z., Li, Z., and Yang, X. (2021). LncRNA SNHG5 Promotes Chondrocyte Proliferation and Inhibits Apoptosis in Osteoarthritis by Regulating miR-10a-5p/H3F3B axis. *Connect. Tissue Res.* 62 (6), 605–614. doi:10.1080/03008207.2020.1825701
- Kehayova, Y. S., Watson, E., Wilkinson, J. M., Loughlin, J., and Rice, S. J. (2021). Genetic and Epigenetic Interplay Regulates COLGALT2, Contributing to Osteoarthritis Genetic Risk. *Arthritis Rheumatol.* doi:10.1002/art.41738
- Kim, K.-I., Park, Y.-S., and Im, G.-I. (2013). Changes in the Epigenetic Status of the SOX-9 promoter in Human Osteoarthritic Cartilage. *J. Bone Min. Res.* 28 (5), 1050–1060. doi:10.1002/jbmr.1843
- Lee, S.-H., Lee, J.-H., Lee, H.-Y., and Min, K.-J. (2019). Sirtuin Signaling in Cellular Senescence and Aging. *BMB Rep.* 52 (1), 24–34. doi:10.5483/bmbrep.2019.52.1.290
- Li, Y.-H., Tavallae, G., Tokar, T., Nakamura, A., Sundararajan, K., Weston, A., et al. (2016). Identification of Synovial Fluid microRNA Signature in Knee Osteoarthritis: Differentiating Early- and Late-Stage Knee Osteoarthritis. *Osteoarthr. Cartil.* 24 (9), 1577–1586. doi:10.1016/j.joca.2016.04.019
- Lian, L. P., and Xi, X. Y. (2020). Long Non-coding RNA XIST Protects Chondrocytes ATDC5 and CHON-001 from IL-1 $\beta$ -induced Injury via Regulating miR-653-5p/SIRT1 axis. *J. Biol. Regul. Homeost. Agents* 34 (2), 379–391. doi:10.23812/19-549-A-65
- Liu, Y., Chang, J.-C., Hon, C.-C., Fukui, N., Tanaka, N., Zhang, Z., et al. (2018). Chromatin Accessibility Landscape of Articular Knee Cartilage Reveals Aberrant Enhancer Regulation in Osteoarthritis. *Sci. Rep.* 8 (1), 15499. doi:10.1038/s41598-018-33779-z
- Liu, Y., Liu, K., Tang, C., Shi, Z., Jing, K., and Zheng, J. (2020b). Long Non-coding RNA XIST Contributes to Osteoarthritis Progression via miR-149-5p/DNMT3A axis. *Biomed. Pharmacother.* 128, 110349. doi:10.1016/j.biopha.2020.110349
- Liu, Y., Yang, Y., Ding, L., Jia, Y., and Ji, Y. (2020a). LncRNA MIR4435-2HG Inhibits the Progression of Osteoarthritis through miR-510-3p Sponging. *Exp. Ther. Med.* 20 (2), 1693–1701. doi:10.3892/etm.2020.8841
- Loeser, R. F., Im, H.-J., Richardson, B., Lu, Q., and Chubinskaya, S. (2009). Methylation of the OP-1 Promoter: Potential Role in the Age-Related Decline in OP-1 Expression in Cartilage. *Osteoarthr. Cartil.* 17 (4), 513–517. doi:10.1016/j.joca.2008.08.003
- Lorthongpanich, C., Cheow, L. F., Balu, S., Quake, S. R., Knowles, B. B., Burkholder, W. F., et al. (2013). Single-cell DNA-Methylation Analysis Reveals Epigenetic Chimerism in Preimplantation Embryos. *Science* 341 (6150), 1110–1112. doi:10.1126/science.1240617
- Lu, J. F., Qi, L. G., Zhu, X. B., and Shen, Y. X. (2020). LncRNA RMRP Knockdown Promotes Proliferation and Inhibits Apoptosis in Osteoarthritis Chondrocytes by miR-206/CDK9 axis. *Pharmazie* 75 (10), 500–504. doi:10.1691/ph.2020.0591
- Mao, G., Hu, S., Zhang, Z., Wu, P., Zhao, X., Lin, R., et al. (2018). Exosomal miR-95-5p Regulates Chondrogenesis and Cartilage Degradation via Histone Deacetylase 2/8. *J. Cell Mol. Med.* 22 (11), 5354–5366. doi:10.1111/jcmm.13808
- Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., et al. (2012). Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science* 337 (6099), 1190–1195. doi:10.1126/science.1222794
- Meng, F., Li, Z., Zhang, Z., Yang, Z., Kang, Y., Zhao, X., et al. (2018). MicroRNA-193b-3p Regulates Chondrogenesis and Chondrocyte Metabolism by Targeting HDAC3. *Theranostics* 8 (10), 2862–2883. doi:10.7150/thno.23547
- Meng, Y., Qiu, S., Sun, L., and Zuo, J. (2020). Knockdown of Exosome-mediated lncPVT1 Alleviates Lipopolysaccharide-induced Osteoarthritis Progression by Mediating the HMGB1/TLR4/NF $\kappa$ B Pathway via miR935p. *Mol. Med. Rep.* 22 (6), 5313–5325. doi:10.3892/mmr.2020.11594
- Milagro, F. I., Campión, J., García-Díaz, D. F., Goyenechea, E., Paternain, L., and Martínez, J. A. (2009). High Fat Diet-Induced Obesity Modifies the Methylation Pattern of Leptin Promoter in Rats. *J. Physiol. Biochem.* 65 (1), 1–9. doi:10.1007/bf03165964
- Miranda-Duarte, A., Borgonio-Cuadra, V. M., González-Huerta, N. C., Rojas-Toledo, E. X., Ahumada-Pérez, J. F., Sosa-Arellano, M., et al. (2020). DNA Methyltransferase Genes Polymorphisms Are Associated with Primary Knee Osteoarthritis: a Matched Case-Control Study. *Rheumatol. Int.* 40 (4), 573–581. doi:10.1007/s00296-019-04474-7
- Miranda-Duarte, A. (2018). DNA Methylation in Osteoarthritis: Current Status and Therapeutic Implications. *Open Rheumatol. J.* 12, 37–49. doi:10.2174/1874312901812010037
- Miyaki, S., Sato, T., Inoue, A., Otsuki, S., Ito, Y., Yokoyama, S., et al. (2010). MicroRNA-140 Plays Dual Roles in Both Cartilage Development and Homeostasis. *Genes Dev.* 24 (11), 1173–1185. doi:10.1101/gad.1915510
- Moazed-Fuerst, F. C., Hofner, M., Gruber, G., Weinhaeusel, A., Stradner, M. H., Angerer, H., et al. (2014). Epigenetic Differences in Human Cartilage between Mild and Severe OA. *J. Orthop. Res.* 32 (12), 1636–1645. doi:10.1002/jor.22722

- Monteagudo, S., Cornelis, F. M. F., Aznar-Lopez, C., Yibmantasiri, P., Guns, L.-A., Carmeliet, P., et al. (2017). DOT1L Safeguards Cartilage Homeostasis and Protects against Osteoarthritis. *Nat. Commun.* 8, 15889. doi:10.1038/ncomms15889
- Papathanasiou, I., Kostopoulou, F., Malizos, K. N., and Tsezou, A. (2015). DNA Methylation Regulates Sclerostin (SOST) Expression in Osteoarthritic Chondrocytes by Bone Morphogenetic Protein 2 (BMP-2) Induced Changes in Smads Binding Affinity to the CpG Region of SOST Promoter. *Arthritis Res. Ther.* 17, 160. doi:10.1186/s13075-015-0674-6
- Parker, E., Hofer, I. M. J., Rice, S. J., Earl, L., Anjum, S. A., Deehan, D. J., et al. (2021). Multi-Tissue Epigenetic and Gene Expression Analysis Combined with Epigenome Modulation Identifies RWDD2B as a Target of Osteoarthritis Susceptibility. *Arthritis Rheumatol.* 73 (1), 100–109. doi:10.1002/art.41473
- Pöschl, E., Fidler, A., Schmidt, B., Kallipolitu, A., Schmid, E., and Aigner, T. (2005). DNA Methylation Is Not Likely to Be Responsible for Aggrecan Down Regulation in Aged or Osteoarthritic Cartilage. *Ann. Rheum. Dis.* 64 (3), 477–480. doi:10.1136/ard.2004.022509
- Ramos, Y. F. M., and Meulenbelt, I. (2017). The Role of Epigenetics in Osteoarthritis: Current Perspective. *Curr. Opin. Rheumatol.* 29 (1), 119–129. doi:10.1097/bor.0000000000000355
- Reynard, L. N., Bui, C., Canty-Laird, E. G., Young, D. A., and Loughlin, J. (2011). Expression of the Osteoarthritis-Associated Gene GDF5 Is Modulated Epigenetically by DNA Methylation. *Hum. Mol. Genet.* 20 (17), 3450–3460. doi:10.1093/hmg/ddr253
- Reynard, L. N., Bui, C., Syddall, C. M., and Loughlin, J. (2014). CpG Methylation Regulates Allelic Expression of GDF5 by Modulating Binding of SP1 and SP3 Repressor Proteins to the Osteoarthritis Susceptibility SNP Rs143383. *Hum. Genet.* 133 (8), 1059–1073. doi:10.1007/s00439-014-1447-z
- Rice, S. J., Roberts, J. B., Tselepi, M., Brumwell, A., Falk, J., Steven, C., et al. (2021). Genetic and Epigenetic Factors Fine-Tune TGFB1 Expression within the Osteoarthritic Articular Joint. *Arthritis Rheumatol.* 73 (10), 1866–1877. doi:10.1002/art.41736
- Rice, S. J., Aubourg, G., Sorial, A. K., Almaraz, D., Tselepi, M., Deehan, D. J., et al. (2018). Identification of a Novel, Methylation-dependent, RUNX2 Regulatory Region Associated with Osteoarthritis Risk. *Hum. Mol. Genet.* 27 (19), 3464–3474. doi:10.1093/hmg/ddy257
- Rice, S. J., Cheung, K., Reynard, L. N., and Loughlin, J. (2019b). Discovery and Analysis of Methylation Quantitative Trait Loci (mQTLs) Mapping to Novel Osteoarthritis Genetic Risk Signals. *Osteoarthr. Cartil.* 27 (10), 1545–1556. doi:10.1016/j.joca.2019.05.017
- Rice, S. J., Tselepi, M., Sorial, A. K., Aubourg, G., Shepherd, C., Almaraz, D., et al. (2019a). Prioritization of PLEC and GRINA as Osteoarthritis Risk Genes through the Identification and Characterization of Novel Methylation Quantitative Trait Loci. *Arthritis Rheumatol.* 71 (8), 1285–1296. doi:10.1002/art.40849
- Roach, H. I., Yamada, N., Cheung, K. S. C., Tilley, S., Clarke, N. M. P., Oreffo, R. O. C., et al. (2005). Association between the Abnormal Expression of Matrix-Degrading Enzymes by Human Osteoarthritic Chondrocytes and Demethylation of Specific CpG Sites in the Promoter Regions. *Arthritis Rheum.* 52 (10), 3110–3124. doi:10.1002/art.21300
- Rushton, M. D., Reynard, L. N., Young, D. A., Shepherd, C., Aubourg, G., Gee, F., et al. (2015). Methylation Quantitative Trait Locus Analysis of Osteoarthritis Links Epigenetics with Genetic Risk. *Hum. Mol. Genet.* 24 (25), 7432–7444. doi:10.1093/hmg/ddv433
- Shen, J., Wang, C., Li, D., Xu, T., Myers, J., Ashton, J. M., et al. (2017). DNA Methyltransferase 3b Regulates Articular Cartilage Homeostasis by Altering Metabolism. *JCI Insight* 2 (12). doi:10.1172/jci.insight.93612
- Shepherd, C., Reese, A. E., Reynard, L. N., and Loughlin, J. (2019). Expression Analysis of the Osteoarthritis Genetic Susceptibility Mapping to the Matrix Gla Protein Gene MGP. *Arthritis Res. Ther.* 21 (1), 149. doi:10.1186/s13075-019-1934-7
- Shepherd, C., Zhu, D., Skelton, A. J., Combe, J., Threadgold, H., Zhu, L., et al. (2018). Functional Characterization of the Osteoarthritis Genetic Risk Residing at ALDH1A2 Identifies Rs12915901 as a Key Target Variant. *Arthritis Rheumatol.* 70 (10), 1577–1587. doi:10.1002/art.40545
- Simon, T. C., and Jeffries, M. A. (2017). The Epigenomic Landscape in Osteoarthritis. *Curr. Rheumatol. Rep.* 19 (6), 30. doi:10.1007/s11926-017-0661-9
- Singh, P., Lessard, S. G., Mukherjee, P., Carballo, C. B., Rodeo, S. A., and Otero, M. (2019). The Progression of Post-traumatic Osteoarthritis in the Murine DMM Model Is Marked by Distinctive Epigenomic Patterns Associated with Transcriptomic Changes in Articular Cartilage. *Osteoarthr. Cartil.* 27, S282. doi:10.1016/j.joca.2019.02.664
- Statello, L., Guo, C.-J., Chen, L.-L., and Huarte, M. (2021). Gene Regulation by Long Non-coding RNAs and its Biological Functions. *Nat. Rev. Mol. Cell Biol.* 22 (2), 96–118. doi:10.1038/s41580-020-00315-9
- Styrkarsdottir, U., Stefansson, O. A., Gunnarsdottir, K., Thorleifsson, G., Lund, S. H., Stefansdottir, L., et al. (2019). GWAS of Bone Size Yields Twelve Loci that Also Affect Height, BMD, Osteoarthritis or Fractures. *Nat. Commun.* 10, 2054. doi:10.1038/s41467-019-09860-0
- Styrkarsdottir, U., Lund, S. H., Thorleifsson, G., Zink, F., Stefansson, O. A., Sigurdsson, J. K., et al. (2018). Meta-analysis of Icelandic and UK Data Sets Identifies Missense Variants in SMO, IL11, COL11A1 and 13 More New Loci Associated with Osteoarthritis. *Nat. Genet.* 50 (12), 1681–1687. doi:10.1038/s41588-018-0247-0
- Tachmazidou, I., Hatzikotoulas, K., Hatzikotoulas, K., Southam, L., Esparza-Gordillo, J., Haberland, V., et al. (2019). Identification of New Therapeutic Targets for Osteoarthritis through Genome-wide Analyses of UK Biobank Data. *Nat. Genet.* 51 (2), 230–236. doi:10.1038/s41588-018-0327-1
- Takahashi, A., de Andrés, M. C., Hashimoto, K., Itoi, E., and Oreffo, R. O. C. (2015). Epigenetic Regulation of Interleukin-8, an Inflammatory Chemokine, in Osteoarthritis. *Osteoarthr. Cartil.* 23 (11), 1946–1954. doi:10.1016/j.joca.2015.02.168
- Tan, F., Wang, D., and Yuan, Z. (2020). The Fibroblast-like Synovocyte Derived Exosomal Long Non-coding RNA H19 Alleviates Osteoarthritis Progression through the miR-106b-5p/TIMP2 Axis. *Inflammation* 43 (4), 1498–1509. doi:10.1007/s10753-020-01227-8
- Tsuda, M., Takahashi, S., Takahashi, Y., and Asahara, H. (2003). Transcriptional Co-activators CREB-Binding Protein and P300 Regulate Chondrocyte-specific Gene Expression via Association with Sox9. *J. Biol. Chem.* 278 (29), 27224–27229. doi:10.1074/jbc.m303471200
- Ukita, M., Matsushita, K., Tamura, M., and Yamaguchi, T. (2020). Histone H3K9 Methylation Is Involved in Temporomandibular Joint Osteoarthritis. *Int. J. Mol. Med.* 45 (2), 607–614. doi:10.3892/ijmm.2019.4446
- Wang, B., Li, J., and Tian, F. (2021). Downregulation of lncRNA SNHG14 Attenuates Osteoarthritis by Inhibiting FSTL-1 Mediated NLRP3 and TLR4/NF-Kb Pathway through miR-124-3p. *Life Sci.* 270, 119143. doi:10.1016/j.lfs.2021.119143
- Wang, C.-L., Zuo, B., Li, D., Zhu, J.-F., Xiao, F., Zhang, X.-L., et al. (2020). The Long Noncoding RNA H19 Attenuates Force-Driven Cartilage Degeneration via miR-483-5p/Dusp5. *Biochem. Biophysical Res. Commun.* 529 (2), 210–217. doi:10.1016/j.bbrc.2020.05.180
- Wang, J., Chen, L., Jin, S., Lin, J., Zheng, H., Zhang, H., et al. (2016). MiR-98 Promotes Chondrocyte Apoptosis by Decreasing Bcl-2 Expression in a Rat Model of Osteoarthritis. *Acta Biochim. Biophys. Sin.* 48 (10), 923–929. doi:10.1093/abbs/gmw084
- Wang, P., Mao, Z., Pan, Q., Lu, R., Huang, X., Shang, X., et al. (2018). Histone Deacetylase-4 and Histone Deacetylase-8 Regulate Interleukin-1 $\beta$ -Induced Cartilage Catabolic Degradation through MAPK/JNK and ERK Pathways. *Int. J. Mol. Med.* 41 (4), 2117–2127. doi:10.3892/ijmm.2018.3410
- Wang, Z., Ni, S., Zhang, H., Fan, Y., Xia, L., and Li, N. (2021). Silencing SGK1 Alleviates Osteoarthritis through Epigenetic Regulation of CREB1 and ABCA1 Expression. *Life Sci.* 268, 118733. doi:10.1016/j.lfs.2020.118733
- Xiao, P., Zhu, X., Sun, J., Zhang, Y., Qiu, W., Li, J., et al. (2021). LncRNA NEAT1 Regulates Chondrocyte Proliferation and Apoptosis via Targeting miR-543/PLA2G4A axis. *Hum. Cell* 34 (1), 60–75. doi:10.1007/s13577-020-00433-8
- Xu, J., Pei, Y., Lu, J., Liang, X., Li, Y., Wang, J., et al. (2021). LncRNA SNHG7 Alleviates IL-1 $\beta$ -induced Osteoarthritis by Inhibiting miR-214-5p-Mediated PPARGC1B Signaling Pathways. *Int. Immunopharmacol.* 90, 107150. doi:10.1016/j.intimp.2020.107150
- Yamasaki, K., Nakasa, T., Miyaki, S., Ishikawa, M., Deie, M., Adachi, N., et al. (2009). Expression of MicroRNA-146a in Osteoarthritis Cartilage. *Arthritis Rheum.* 60 (4), 1035–1041. doi:10.1002/art.24404
- Zengini, E., Hatzikotoulas, K., Tachmazidou, I., Steinberg, J., Hartwig, F. P., Southam, L., et al. (2018). Genome-wide Analyses Using UK Biobank Data Provide Insights into the Genetic Architecture of Osteoarthritis. *Nat. Genet.* 50 (4), 549–558. doi:10.1038/s41588-018-0079-y

- Zhang, C., Zhang, Z., Chang, Z., Mao, G., Hu, S., Zeng, A., et al. (2019). miR-193b-5p Regulates Chondrocytes Metabolism by Directly Targeting Histone Deacetylase 7 in Interleukin-1 $\beta$ -induced Osteoarthritis. *J. Cell. Biochem.* 120 (8), 12775–12784. doi:10.1002/jcb.28545
- Zhang, D., Cao, X., Li, J., and Zhao, G. (2015). MiR-210 Inhibits NF-Kb Signaling Pathway by Targeting DR6 in Osteoarthritis. *Sci. Rep.* 5, 12775. doi:10.1038/srep12775
- Zhang, G., Zhang, Q., Zhu, J., Tang, J., and Nie, M. (2020). LncRNA ARFRP1 Knockdown Inhibits LPS-Induced the Injury of Chondrocytes by Regulation of NF-Kb Pathway through Modulating miR-15a-5p/TLR4 axis. *Life Sci.* 261, 118429. doi:10.1016/j.lfs.2020.118429
- Zhang, H., Li, J., Shao, W., and Shen, N. (2020). LncRNA SNHG9 Is Downregulated in Osteoarthritis and Inhibits Chondrocyte Apoptosis by Downregulating miR-34a through Methylation. *BMC Musculoskelet. Disord.* 21 (1), 511. doi:10.1186/s12891-020-03497-7
- Zhang, P., Sun, J., Liang, C., Gu, B., Xu, Y., Lu, H., et al. (2020). LncRNA IGHG $\gamma$ 1 Acts as a ceRNA to Regulate Macrophage Inflammation via the miR-6891-3p/TLR4 Axis in Osteoarthritis. *Mediat. Inflamm.* 2020, 9743037. doi:10.1155/2020/9743037
- Zhang, X., Huang, C.-r., Pan, S., Pang, Y., Chen, Y.-s., Zha, G.-c., et al. (2020). Long Non-coding RNA SNHG15 Is a Competing Endogenous RNA of miR-141-3p that Prevents Osteoarthritis Progression by Upregulating BCL2L13 Expression. *Int. Immunopharmacol.* 83, 106425. doi:10.1016/j.intimp.2020.106425
- Zhang, Y., Dong, Q., and Sun, X. (2020). Positive Feedback Loop LINC00511/miR-150-5p/SP1 Modulates Chondrocyte Apoptosis and Proliferation in Osteoarthritis. *DNA Cell Biol.* 39 (9), 1506–1512. doi:10.1089/dna.2020.5718
- Zhang, Y., Fukui, N., Yahata, M., Katsuragawa, Y., Tashiro, T., Ikegawa, S., et al. (2016). Identification of DNA Methylation Changes Associated with Disease Progression in Subchondral Bone with Site-Matched Cartilage in Knee Osteoarthritis. *Sci. Rep.* 6, 34460. doi:10.1038/srep34460
- Zhao, W., Kruse, J.-P., Tang, Y., Jung, S. Y., Qin, J., and Gu, W. (2008). Negative Regulation of the Deacetylase SIRT1 by DBC1. *Nature* 451 (7178), 587–590. doi:10.1038/nature06515
- Zhou, L., Gu, M., Ma, X., Wen, L., Zhang, B., Lin, Y., et al. (2021). Long Non-coding RNA PCAT-1 Regulates Apoptosis of Chondrocytes in Osteoarthritis by Sponging miR-27b-3p. *J. Bone Min. Metab.* 39 (2), 139–147. doi:10.1007/s00774-020-01128-8
- Zhou, Y., Chen, M., O'Keefe, R. J., Shen, J., Li, Z., Zhou, J., et al. (2019). Epigenetic and Therapeutic Implications of Dnmt3b in Temporomandibular Joint Osteoarthritis. *Am. J. Transl. Res.* 11 (3), 1736–1747.
- Zhu, Y., Li, R., and Wen, L.-M. (2021). Long Non-coding RNA XIST Regulates Chondrogenic Differentiation of Synovium-Derived Mesenchymal Stem Cells from Temporomandibular Joint via miR-27b-3p/ADAMTS-5 axis. *Cytokine* 137, 155352. doi:10.1016/j.cyto.2020.155352

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

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

Copyright © 2022 Cai, Long, Zhao, Lin and Wang. 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.



## OPEN ACCESS

## EDITED BY

Xiaofan Lu,  
China Pharmaceutical University, China

## REVIEWED BY

Tiantian Zhang,  
Cornell University, United States  
Lin Zhang,  
Clinical Center (NIH), United States  
Jinhui Liu,  
Nanjing Medical University, China  
Jialin Meng,  
First Affiliated Hospital of Anhui Medical  
University, China

## \*CORRESPONDENCE

Juan Du,  
duj656@163.com  
Zhiyun Yang,  
yangzhiyun2016@163.com

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

RECEIVED 10 May 2022

ACCEPTED 29 June 2022

PUBLISHED 22 July 2022

## CITATION

Yu L, Liu X, Wang X, Yan H, Pu Q, Xie Y,  
Du J and Yang Z (2022),  
Glycometabolism-related gene  
signature of hepatocellular carcinoma  
predicts prognosis and  
guides immunotherapy.  
*Front. Cell Dev. Biol.* 10:940551.  
doi: 10.3389/fcell.2022.940551

## COPYRIGHT

© 2022 Yu, Liu, Wang, Yan, Pu, Xie, Du  
and Yang. 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.

# Glycometabolism-related gene signature of hepatocellular carcinoma predicts prognosis and guides immunotherapy

Lihua Yu<sup>1</sup>, Xiaoli Liu<sup>1</sup>, Xinhui Wang<sup>1</sup>, Huiwen Yan<sup>1</sup>, Qing Pu<sup>1</sup>,  
Yuqing Xie<sup>1,2</sup>, Juan Du<sup>3,4,5\*</sup> and Zhiyun Yang<sup>1\*</sup>

<sup>1</sup>Center of Integrative Medicine, Beijing Ditan Hospital, Capital Medical University, Beijing, China, <sup>2</sup>First Clinical Medical College, Beijing University of Chinese Medicine, Beijing, China, <sup>3</sup>Beijing Key Laboratory of Emerging Infectious Diseases, Institute of Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, China, <sup>4</sup>Beijing Institute of Infectious Diseases, Beijing, China, <sup>5</sup>National Center for Infectious Diseases, Beijing Ditan Hospital, Capital Medical University, Beijing, China

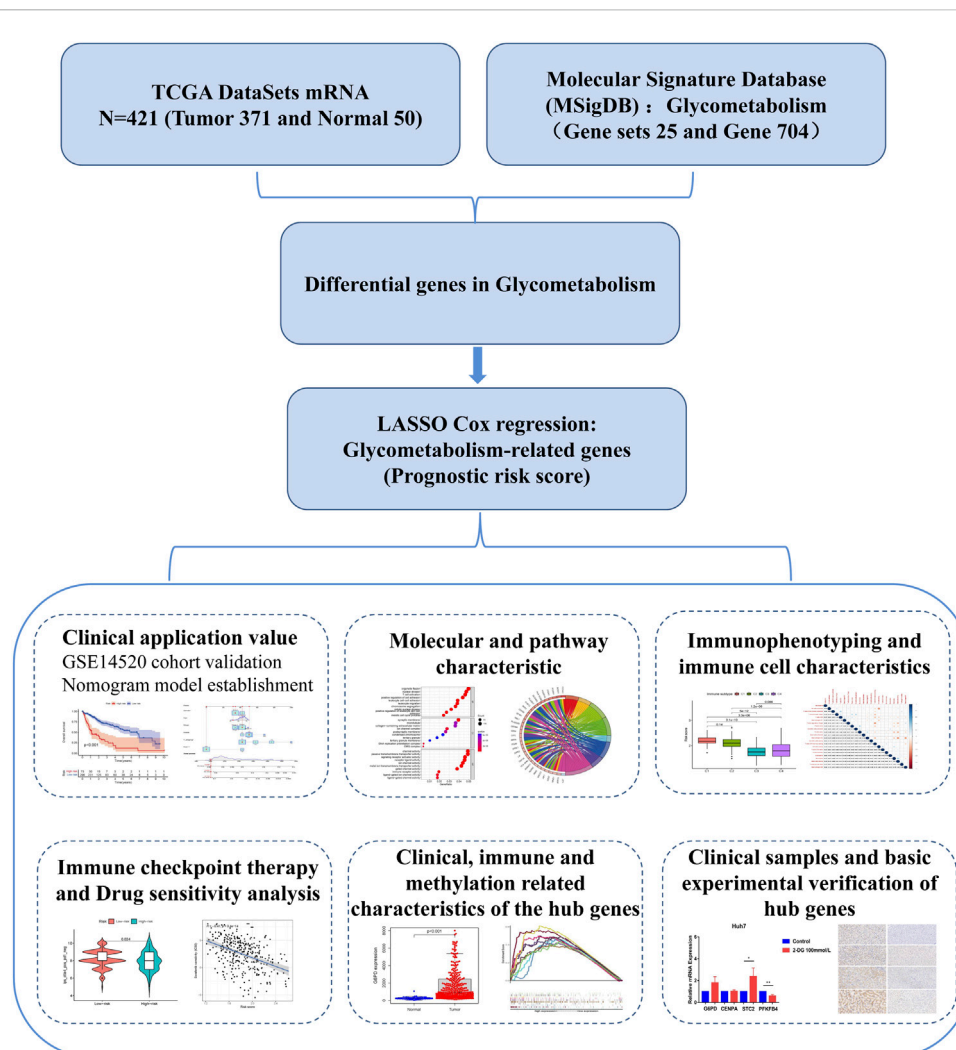
Hepatocellular carcinoma (HCC) is a severe cancer endangering human health. We constructed a novel glycometabolism-related risk score to predict prognosis and immunotherapy strategies in HCC patients. The HCC data sets were obtained from the Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO) database, and the glycometabolism-related gene sets were obtained from the Molecular Signature Database. The least absolute contraction and selection operator (LASSO) regression model was used to construct a risk score based on glycometabolism-related genes. A simple visual nomogram model with clinical indicators was constructed and its effectiveness in calibration, accuracy, and clinical value was evaluated. We also explored the correlation between glycometabolism-related risk scores and molecular pathways, immune cells, and functions. Patients in the low-risk group responded better to anti-CTLA-4 immune checkpoint treatment and benefited from immune checkpoint inhibitor (ICI) therapy. The study found that glycometabolism-related risk score can effectively distinguish the prognosis, molecular and immune-related characteristics of HCC patients, and may provide a new strategy for individualized treatment.

## KEYWORDS

hepatocellular carcinoma, glycometabolism, prognosis, immunotherapy, nomogram

## Introduction

According to the latest global cancer report, hepatocellular carcinoma (HCC) is the sixth most common cancer worldwide and has the third-highest mortality rate (Sung et al., 2021). The World Health Organization (WHO) has projected that the incidence rate of HCC will exceed one million by 2025 (International Agency for Research on Cancer, 2020). China is one of the major countries with a high incidence and mortality of HCC, accounting for approximately half of the total cases worldwide (Feng et al., 2019; Wang et al., 2014). Currently, the effectiveness of various



**FIGURE 1**  
Construction and characterization of glycometabolism-related risk scores.

treatments for HCC is unsatisfactory (Forner et al., 2018). Therefore, there is an urgent need to clarify the specific molecular mechanisms related to HCC for targeted therapy, which is expected to improve the survival rate of patients with HCC, delay tumor progression, and improve the quality of life of patients. Immunotherapy, particularly immune checkpoint inhibitor (ICI) therapy, which has achieved significant clinical breakthroughs, is still faced with challenges, such as low response rate and poor efficacy in some patients (Darvin et al., 2018; Li et al., 2019). The relationship between metabolic restriction and immunity has gradually become a topic of interest and has received considerable attention (DePeaux and Delgoffe, 2021).

In the tumor microenvironment (TME), glycometabolism is the main metabolic pathway of tumor cells and immune cells

(Bose and Le, 2018). Glycometabolism reprogramming is one of the main features of the TME, and tumor cells upregulate glycolytic pathways, undergo tumor escape, and inhibit immune effector cell function until exhaustion occurs (Shevchenko and Bazhin, 2018). In the tumor microenvironment, the IL-10-fc fusion protein has been found to enhance the expansion and effector function of depleted CD8<sup>+</sup> T lymphocytes by promoting the oxidative phosphorylation pathway in glycometabolism (Guo et al., 2021). The NF-E2-related factor 2 (Nrf2) antioxidant pathway was found to restore the metabolism and function of natural killer (NK) cells in human ovarian cancer (Poznanski et al., 2021). Another study found that programmed death-ligand 1 (PD-L1) signaling activates the Akt/mTOR pathway to promote glycolysis in tumor cells. Therapeutic blockade of PD-L1 inhibits tumor

progression by triggering the internalization of PD-L1 and reducing the rate of glycolysis (Chang et al., 2015). Based on this, active efforts are being made to identify therapeutic targets for glycometabolism to enhance the effector function of exhausted immune cells and improve responsiveness to ICI therapy. A few biomarkers of glycometabolism in liver cancer can be used to predict patient prognosis. Identifying potential prognostic markers associated with therapy can enable personalized metabolic immunotherapy in patients with HCC.

This study comprehensively assessed the glycometabolism patterns of patients with HCC and constructed a prognostic risk-score model for glycometabolism. We focused on all glycometabolism-related genes in HCC and constructed a prognostic risk score for these genes. We then characterized the molecular pathways and immune-related features of the prognostic risk score for glycometabolism-related genes, differentiated sorafenib- and 5-fluorouracil-resistant patients, and examined their responsiveness to immunosuppressant therapy. A technical roadmap of this study is shown in Figure 1. This study provides a new perspective for exploring the glycometabolism-immunity mechanism.

## Materials and methods

### Patients and datasets

A total of 421 liver hepatocellular carcinoma (LIHC) samples were downloaded from The Cancer Genome Atlas (TCGA) database, including 371 tumor samples and 50 normal samples with clinical and pathological staging data (<https://portal.gdc.cancer.gov/>). Among these, 369 tumor samples with complete clinical survival information were used as the training set for the model.

Microarray data for 242 HCC samples and their clinical information (GSE14520) were obtained from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>), and the platform was GPL3921. The gene ID of each sample was converted into the corresponding gene symbol using an annotation platform. The average value was calculated if the same gene ID was mapped by multiple probes. Ultimately, 221 HCC samples with complete clinical survival information were included in the validation set. The clinical characteristics of the two cohorts are presented in Supplementary Table S1.

### Acquisition of glycometabolism genes

From the Molecular Signature Database (MSigDB; v7.5.1) (Subramanian et al., 2005), 25 glycometabolism-related gene sets were collected. After removing overlapping genes, 704 glycometabolism-related genes were obtained (Supplementary Table S2).

## Identification of glycometabolism-related hub genes

The differentially expressed genes related to glycometabolism between normal and tumor tissue samples were analyzed using the R package “limma.” Differentially expressed genes related to glycometabolism were obtained and analyzed using Metascape (<http://metascape.org>). Differentially expressed genes with  $|\log_2FC| > 2$  and false discovery rate (FDR) < 0.05 were considered statistically significant.

Using the R package “glmnet,” we sequentially performed univariate Cox and least absolute shrinkage and selection operator (LASSO) Cox regression analyses and identified four glycometabolism-related genes associated with HCC survival.

## Construction and validation of the prognostic risk score

The prognostic risk score was obtained by multiplying the expression values of four glycometabolism-related hub genes by their weights in the LASSO Cox model and then adding them to calculate the risk score of each sample. The Kaplan-Meier (K-M) survival curve was used to evaluate the prognostic ability of the risk score in the TCGA and GEO databases. Univariate and multivariate Cox regression analyses were performed to verify the independent prognostic value of the risk score.

To reflect the clinical application and predictive value of the risk score, a nomogram model was constructed with risk scores and clinical indicators to predict patient prognosis at different times. The model's calibration, accuracy, and clinical value were further evaluated using calibration, receiver operating characteristic (ROC), and decision curve analysis (DCA) curves.

## Comprehensive analysis of molecular and immune characteristics in different prognostic risk score subgroups

The R package “clusterProfiler” was used to analyze the enrichment of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways in different prognostic risk scores to determine the main biological characteristics and the enrichment of cellular functional pathways. A  $p$  value ( $q$  value) < 0.05 was defined as a statistically significant difference. Gene set enrichment analysis (GSEA) was also used to compare the differences in biological processes between the low-risk and high-risk groups, with the “hallmark all. v7.5. symbols” gene set as the internal parameter gene set. Statistical significance was set at  $p < 0.05$  and FDR < 0.05.

According to immune cell composition, 369 patients with HCC in the training set were divided into four immune subsets

(Zhang et al., 2020a). To compare the immune characteristics of the 369 HCC samples with different prognostic risk scores, x-Cell (<https://xcell.ucsf.edu/>) and CIBERSORT (<https://cibersort.stanford.edu/>) were used to evaluate the relative proportions of 22 immune cells and the correlation between immune cells. Subsequently, the immune functions of the different prognostic risk scores were also compared. GSEA was also used to assess changes in immune-related pathways ("c7. immunesigdb. v7.3. symbols") or biological processes with different prognostic risk scores.

To explore the prognostic value of different prognostic risk scores in patients after immunotherapy, the responsiveness to immune checkpoint PD-1 and CTLA-4 treatment was assessed using TCIA (<https://tcia.at/home>). The R package "pRRophetic" was used to analyze the drug sensitivity of patients with different prognostic risk scores.

## Quantitative real-time PCR

RNA was extracted from HepG2 and Huh7 cell lines using the RNeasy Plus Mini Kit (QIAGEN, Germany), and the RNA was reverse transcribed into cDNA using cDNA.

Synthesis kit (Thermo Fisher Scientific, United States), and then qRT-PCR was performed using Master Mix (SYBR Green; Lithuania).  $\beta$ -Actin was used as an internal control, and  $2^{-\Delta\Delta CT}$  methodology was expressed as the relative expression of mRNA. Do at least three independent experiments. All mRNAs were purchased from the Synbio Technologies and the primers were listed in [Supplementary Table S3](#).

## Immunohistochemistry

For immunohistochemical analysis, paraffin sections were performed on tumor tissues and adjacent tissues from patients with HCC. 4- $\mu$ m-thick sections were soaked in xylene for 15 min, then deaffinity and rehydrated with an ethanol gradient, followed by incubation in 3% hydrogen peroxide for 15 min. Blocked with 10% goat serum for 30 min, incubated with primary antibodies to G6PD, CENPA, STC2, and PFKFB4 (abcam, 1:200) overnight at 4°C, and then used secondary antibodies conjugated to horseradish peroxidase (sigma, 1:200) and incubated for 30 min. Sections were scanned with panoramic scanning electron microscope, and positive staining was analyzed by ImageJ software (version 1.8.0).

## Statistical analysis

All statistical analyses were performed using the R software (version 4.1.3) and GraphPad Prism 7 software (version 7.0). Continuous variables were compared between the two groups

using an independent *t*-test, and categorical data were compared using the  $\chi^2$  test. Survival analysis was performed using the K-M survival analysis and log-rank test. Cox regression analysis was used to identify independent prognostic indicators predicting overall survival (OS) in HCC. Statistical significance was set at  $p < 0.05$ .

## Results

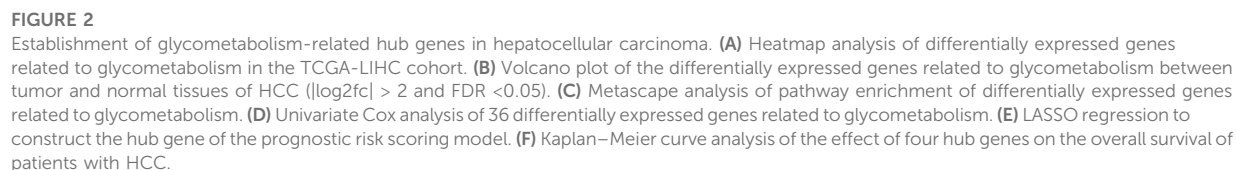
### Acquired glycometabolism-related hub genes in hepatocellular carcinoma patients

In the differential expression analysis (371 tumors vs. 50 normal samples), all genes in the TCGA database intersected with the glycometabolism-related genes obtained from the MSigDB gene set, and 678 glycometabolism-related genes were obtained. Subsequently, 65 differentially expressed genes ( $|\log_2FC| > 2$  and  $FDR < 0.05$ ) were identified between the normal and tumor groups ([Figure 2A](#)). The upregulated and downregulated differential genes with  $|\log_2FC| > 2$  were displayed by volcano plot ([Figure 2B](#)). Furthermore, metascape functional enrichment analysis showed that 65 differentially expressed genes were significantly associated with "carbohydrate metabolic process," "glucose homeostasis," and "HIF1 TFPATHWAY" pathways ([Figure 2C](#)).

In the training set of the TCGA database, 36 differentially expressed genes related to glycometabolism were significantly associated with the prognosis of patients with HCC as revealed by Cox univariate analysis ( $p < 0.05$ ; [Figure 2D](#)). LASSO regression was used to select lambda min as the model with the highest accuracy ([Figure 2E](#)). Four differentially expressed genes related to glycometabolism were more closely related to the OS of patients with HCC: glucose-6-phosphate dehydrogenase (G6PD), centromere protein A (CENPA), stanniocalcin 2 (STC2), and 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase 4 (PFKFB4). The median values of the four hub genes were divided into high and low expression groups. K-M survival analysis revealed that the OS rate of patients with HCC in the high expression group of hub genes was lower than that in the low expression group ( $p < 0.001$ ; [Figure 2F](#)).

### Establishment and validation of the prognostic risk score and nomogram model

Subsequently, a prognostic risk score for all tumor samples was calculated using the formula: prognostic risk score = expression level of G6PD\*(0.15) + expression level of CENPA\*(0.075) + expression level of STC2\*(0.018) + expression level of PFKFB4\*(0.053). From the prognostic risk



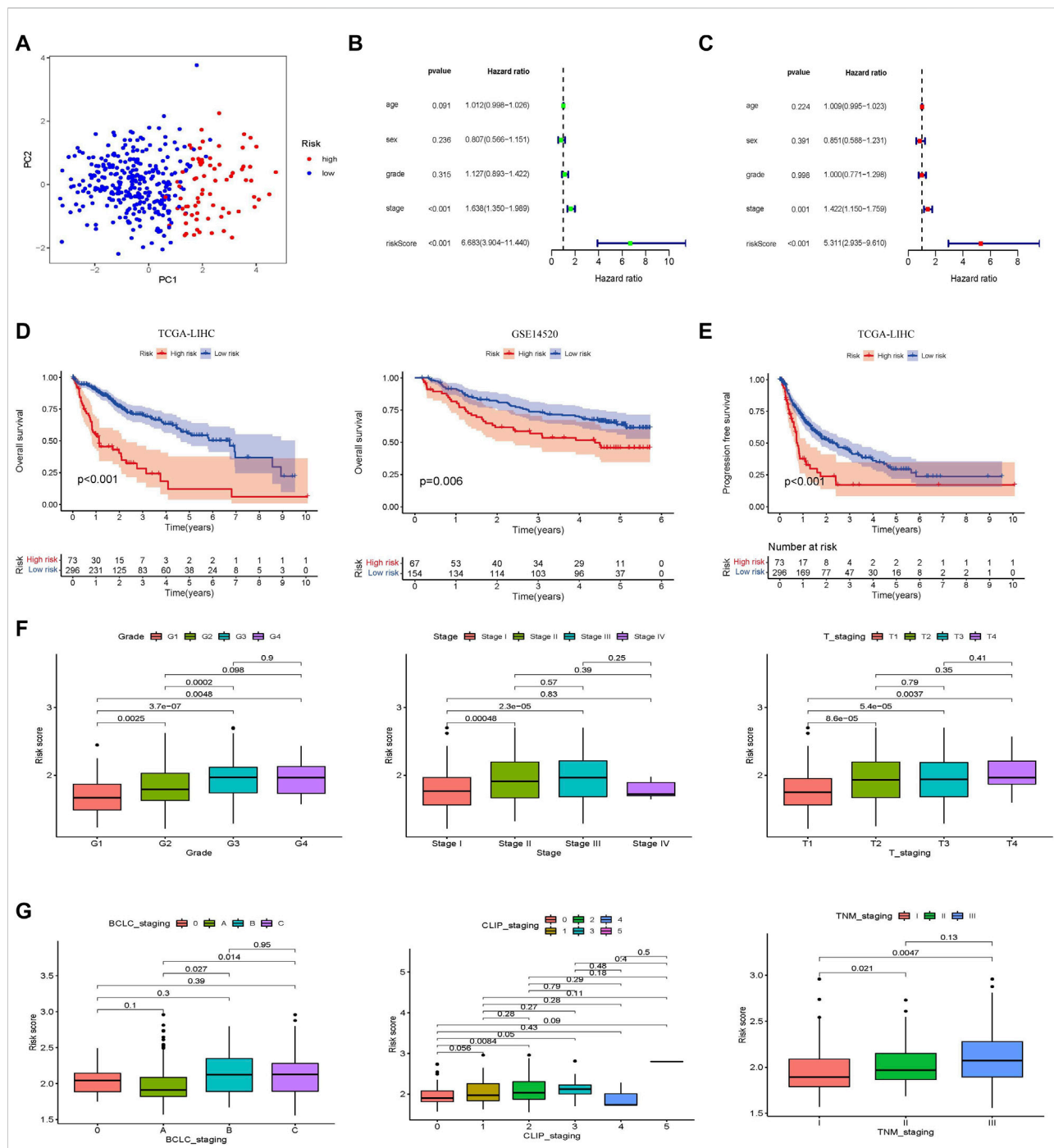


FIGURE 3

Establishment and validation of a prognostic risk score associated with glycometabolism. **(A)** Principal component analysis based on genes related to glycometabolism in patients with HCC. **(B,C)** Forest plot of univariate Cox regression, respectively, in the TCGA-LIHC cohort. **(D)** Kaplan-Meier curve analysis comparing the overall survival between the low-risk and the high-risk groups of patients with HCC in the training and validation cohorts, respectively. **(E)** Kaplan-Meier curve analysis comparing the progression-free survival between the low-risk and high-risk groups of patients with HCC in the TCGA-LIHC cohort. **(F)** Comparison of the relationship between different glycometabolism-related risk scores and clinical stages in the TCGA cohort. **(G)** Comparison of the relationship between different glycometabolism-related risk scores and clinical stages in the GEO cohort.

score, the cut-off value was calculated to be 2.12, and this value was used to divide the risk scores into high- and low-risk groups (Supplementary Figure S1).

Principal component analysis (PCA) was performed to identify significant differences between the low- and high-risk groups. The results showed that the prognostic risk score could distinguish between the high- and low-risk groups (Figure 3A). Clinically relevant indicators, including age, sex, grade, stage, and risk score, were included in the Cox univariate and multivariate analyses; only risk score and stage were independent predictors of OS in patients with HCC ( $p < 0.001$ ; Figures 3B,C). Patients with different prognostic risk scores were divided into high- and low-risk groups according to the cutoff value (2.12). The OS of the patients in the high-risk group was lower than that of the patients in the low-risk group ( $p < 0.001$ ; Figure 3D). The GSE14520 dataset ( $n = 221$ ) was used to verify the effects of the different risk scores on patient survival. Patients in the high-risk group had lower survival rates than those in the low-risk group, which is consistent with the results from the TCGA dataset ( $p = 0.006$ ). The effect of different risk scores on the progression-free survival (PFS) time of patients was also compared, and patients in the high-risk group were found to have shorter PFS times ( $p < 0.001$ ; Figure 3E). Exploring the correlation between prognostic risk scores and clinical indicators furtherly, the training cohort revealed a gradual increase in the prognostic risk score with an increase in different grades, stages, and T (Primary Tumor) stages (Figure 3F). The results of the validation cohort were found to be similar to those of the validation cohort in other staging systems, such as Barcelona Clinic Liver Cancer (BCLC) staging, Cancer of the Liver Italian Program (CLIP) staging, and TNM (Tumor Node Metastasis) staging (Figure 3G). Subsequently, the area under the ROC curve of the prognostic risk score in the training cohort and the validation cohort was observed and compared. In the TCGA database training cohort, the area under curve (AUC) of the prognostic risk score at 1, 3, and 5 years was 0.790, 0.689, and 0.668; in the validation cohort, the AUC of the prognostic risk score at 1, 3, and 5 years was 0.722, 0.635, and 0.587 (Supplementary Figures S2A,B). We also compared the AUC values of the prognostic risk score with other biomarker signatures and found that the prognostic risk score was higher than the other signatures (Supplementary Figures S2C,D) (Zhang et al., 2020a; Zhang et al., 2020b; Pan et al., 2020; Dai et al., 2021). The results showed that compared with other signatures, the prognostic risk score had good predictive performance and the indicators were more concise.

A nomogram model integrating gender, age, stage, grade, T stage, and risk score was constructed and intuitively predicted the 1-, 3-, and 5-years survival rates of patients with HCC (Supplementary Figure S3A). Univariate and multivariate Cox regression analyses showed that the nomogram model was an independent prognostic factor for predicting the outcomes of patients with HCC (Supplementary Figures S3B,C). In addition,

the 1-, 3-, and 5-years calibration curves showed that the nomogram model accurately predicted the prognosis of patients with HCC (Supplementary Figure S3D). The areas under the ROC curve were 0.789, 0.697, and 0.674 for 1, 3, and 5 years, respectively (Supplementary Figure S3E). The DCA curve showed that the nomogram model had a higher rate of clinical benefit than risk score and grade (Supplementary Figure S3F).

## Molecular and pathway characterization of different prognostic risk scores

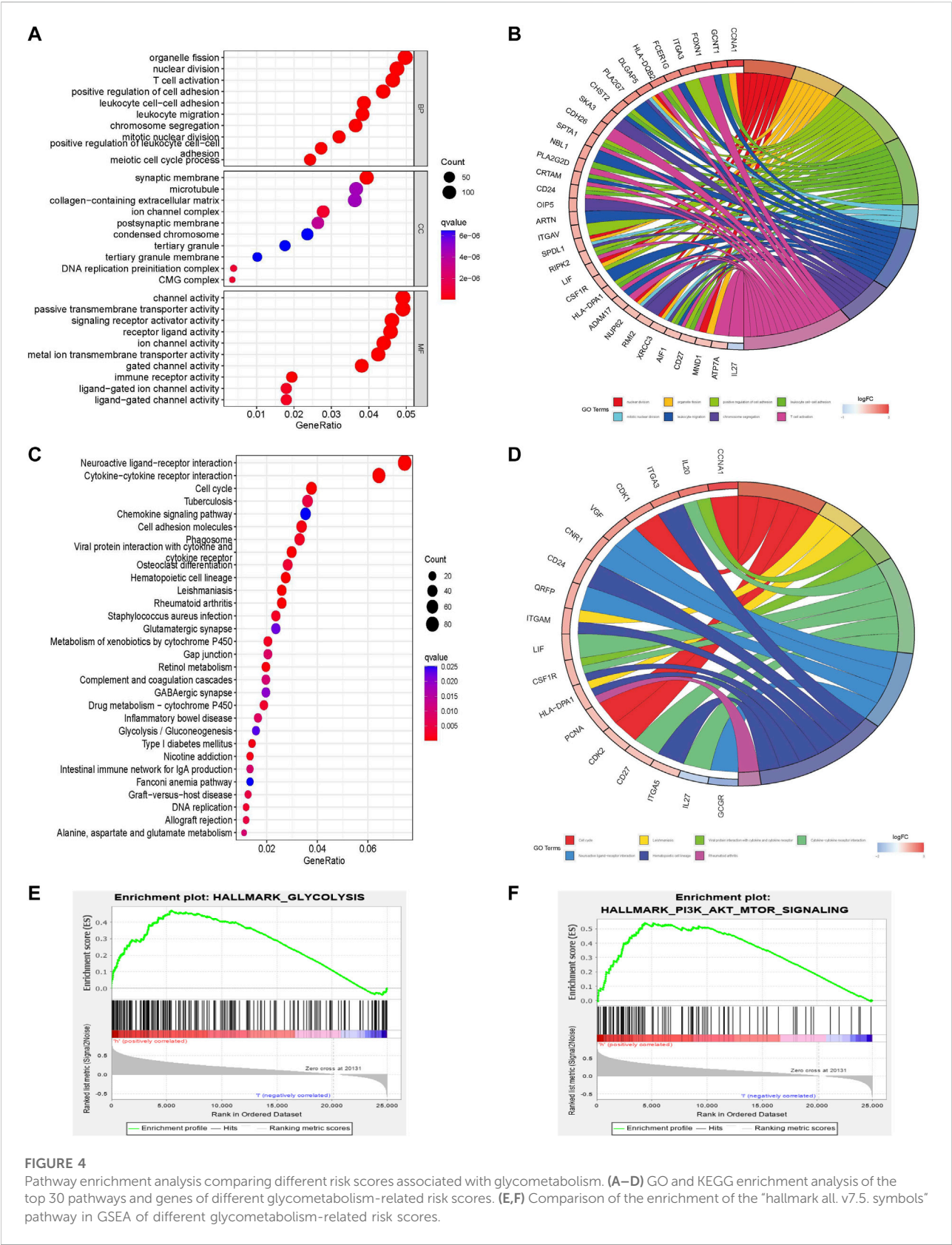
Molecular and pathway characteristics were further compared between the high- and low-risk groups. GO analysis revealed that nuclear division, mitotic nuclear division, and T cell activation pathways were significantly enriched in the high-risk group compared with the low-risk group. And related molecules enriched in high-risk groups include GCNT1, IL-27, and ATP7A, etc. (Figures 4A,B). Meanwhile, KEGG analysis showed that cell cycle, glycolysis and chemokine signaling pathways were significantly enriched in the high-risk group compared with the low-risk group and the related molecules included VGF, CDK2, and IL-27, etc. (Figures 4C,D). GSEA enrichment analysis also revealed that glycolysis (NES = 1.81, Nom  $p = 0.0$ , FDR  $q = 0.033$ ) and PI3K AKT MTOR (NES = 1.73, Nom  $p = 0.002$ , FDR  $q = 0.046$ ) pathways were mainly enriched in the high-risk group (Figures 4E,F; Supplementary Table S4).

## Immune cell characteristics and functions in different prognostic risk scores

To observe the relationship between the prognostic risk score and immune subtype, HCC samples were divided into four immune subtypes: C1 (wound healing), C2 (IFN- $\gamma$  dominant), C3 (inflammatory), and C4 (lymphocyte depleted). The C3 (inflammatory) immune subtype had the lowest risk score and best prognosis, consistent with the results of the original study (Figure 5A) (Vésteinn et al., 2018).

The composition and correlation of immune cells in different risk scores were analyzed, and it was found that CD4<sup>+</sup> memory T-cells, Th1 cells, Th2 cells, basophils, and B cells were more abundant in the high-risk group, whereas macrophages were more abundant in the low-risk group (Figure 5B). Furthermore, the infiltrating immune cells in the low-risk group were mainly correlated with CD4<sup>+</sup> memory T-cells and NK cells, while those in the high-risk group were mainly correlated with macrophages, B cells, and mast cells (Figures 5C,D).

The immune functions between the different risk scores were further explored. More immunosuppressive functions were observed in the high-risk group, such as APC coinhibition, checkpoint, MHC class I, and T-cell coinhibition (Figure 5E).



**FIGURE 4** Pathway enrichment analysis comparing different risk scores associated with glycometabolism. **(A–D)** GO and KEGG enrichment analysis of the top 30 pathways and genes of different glycometabolism-related risk scores. **(E,F)** Comparison of the enrichment of the “hallmark all. v7.5. symbols” pathway in GSEA of different glycometabolism-related risk scores.

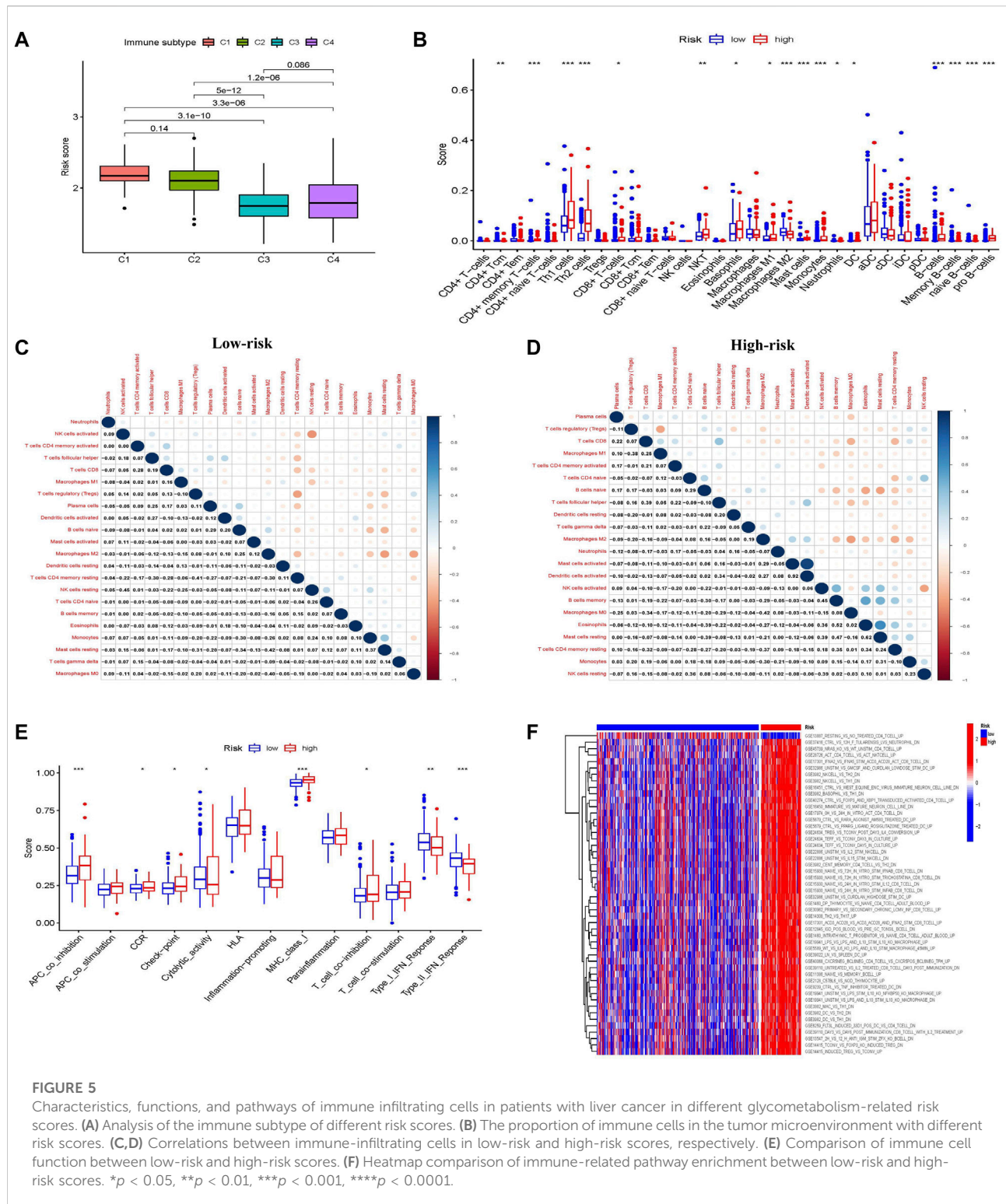
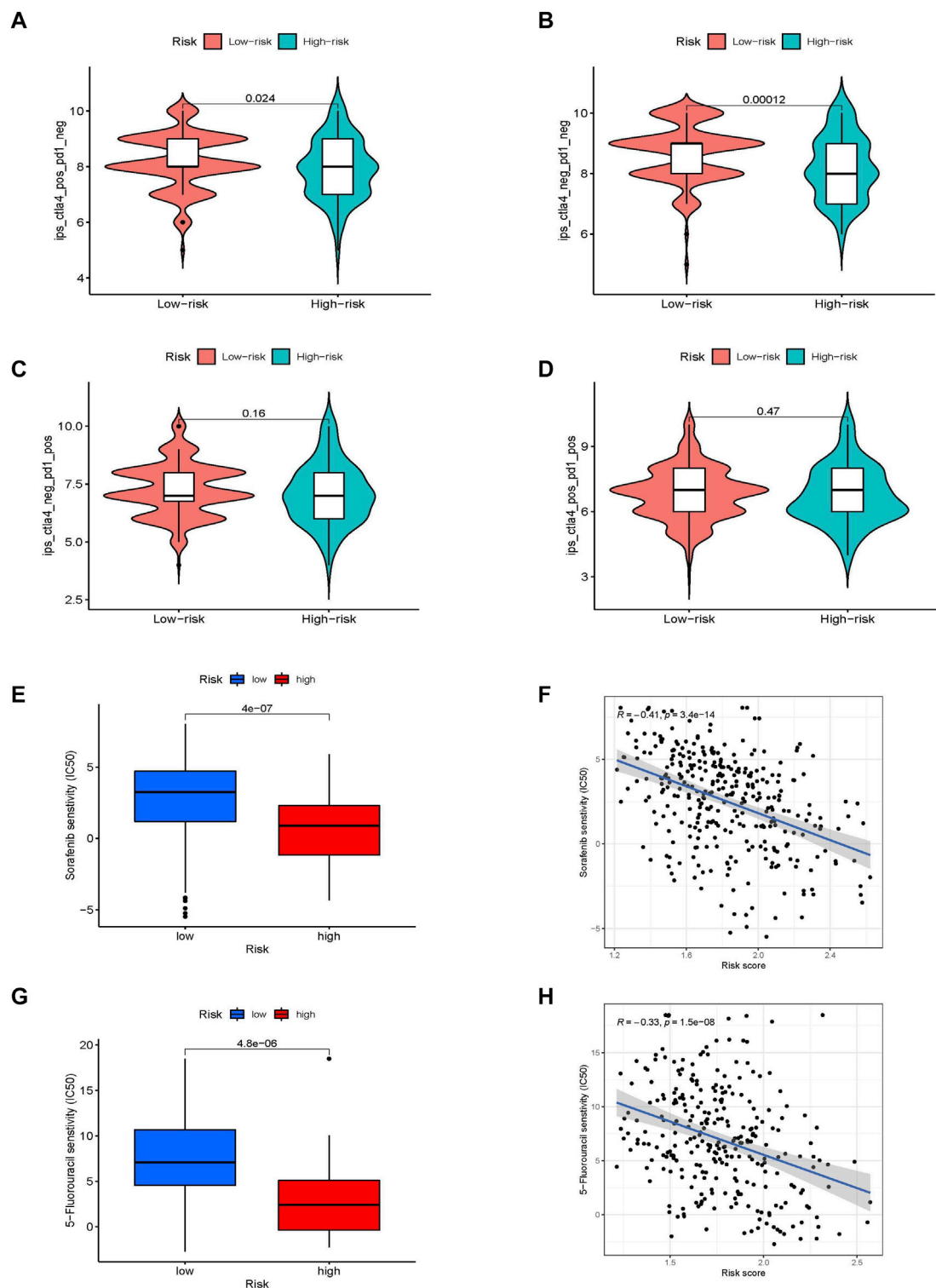


FIGURE 5

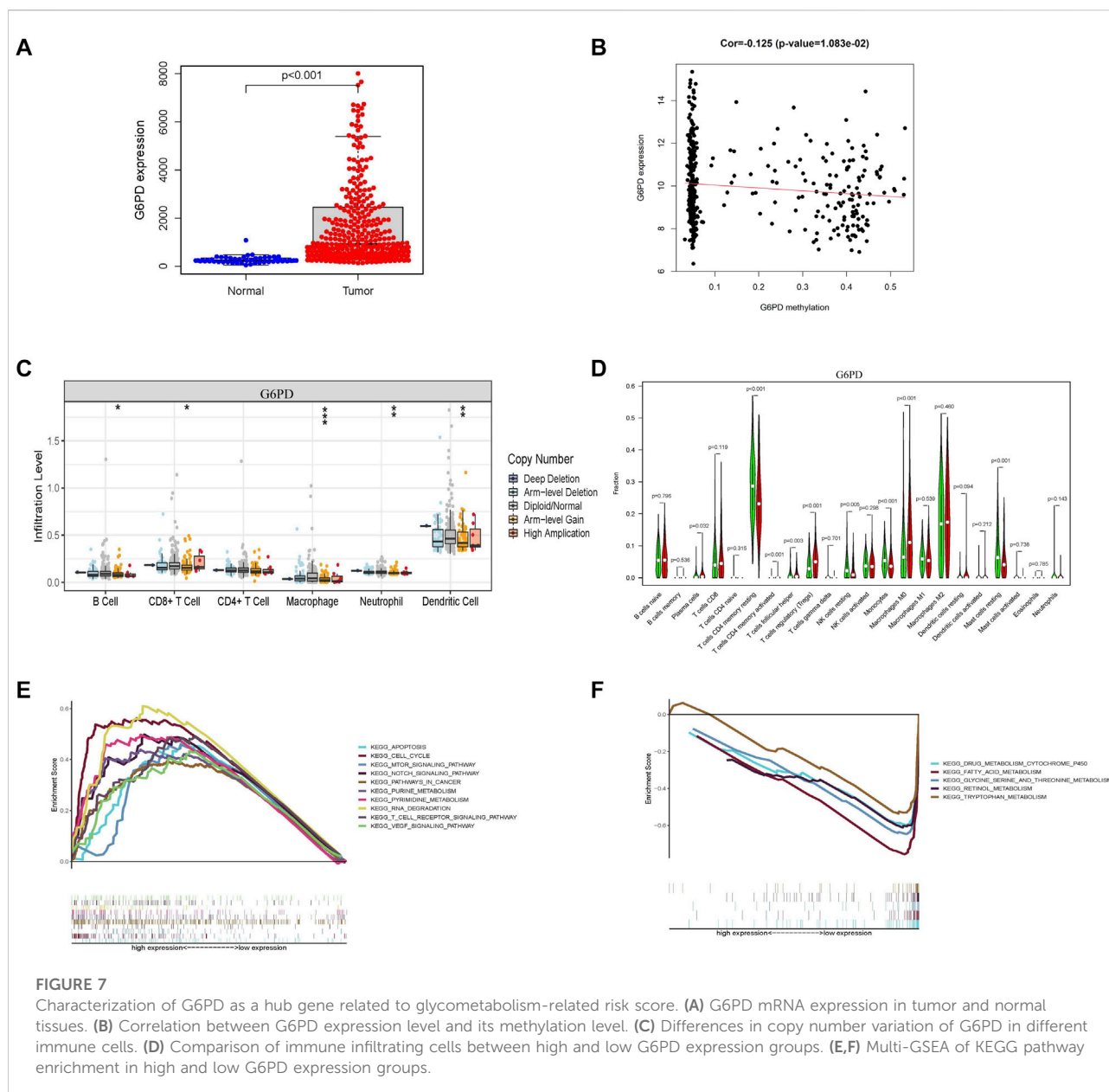
Characteristics, functions, and pathways of immune infiltrating cells in patients with liver cancer in different glycometabolism-related risk scores. **(A)** Analysis of the immune subtype of different risk scores. **(B)** The proportion of immune cells in the tumor microenvironment with different risk scores. **(C,D)** Correlations between immune-infiltrating cells in low-risk and high-risk scores, respectively. **(E)** Comparison of immune cell function between low-risk and high-risk scores. **(F)** Heatmap comparison of immune-related pathway enrichment between low-risk and high-risk scores. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

Subsequently, the “c7.immunesigdb.v7.3. symbols” gene set downloaded from MSigDB was used to enrich gene set variation analysis and compare immune-related biological

functions between the two groups. Interestingly, most immune-related genes were enriched in the high-risk group (Figure 5F).

**FIGURE 6**

The value of different glycometabolism-related risk scores for ICI treatment and drug sensitivity. (A–D) Comparison of the responsiveness to anti-PD-1 and anti-CTLA-4 immunotherapy between samples in high-risk and low-risk groups. (E–H) Comparison of the sensitivities and associations to sorafenib and 5-fluorouracil drugs between high-risk and low-risk groups.



## Immune checkpoint inhibitor treatment benefit and drug sensitivity in different prognostic risk scores

Major breakthroughs have been made in immune checkpoint therapies, particularly in PD-1 and CTLA-4 treatment. We investigated the ability of the prognostic risk score to predict response to immune checkpoint therapy. The results showed that although there was no difference in the clinical response to anti-PD-1 immunotherapy between the two groups, the response to anti-CTLA-4 immunotherapy in the low-risk group was higher than that in the high-risk group. Patients in the low-risk group were considered to be more suitable for CTLA-4 immune

checkpoint therapy (Figures 6A–D). Because of the correlation between drug sensitivity and poor prognosis, we focused on the relationship between risk score and drug sensitivity. Sorafenib and 5-fluorouracil, the most widely used targeted drugs for HCC treatment, were selected to compare the drug sensitivity of the different risk groups. Interestingly, the high-risk group was found to be more sensitive to the targeted drugs sorafenib and 5-fluorouracil ( $p < 0.001$ ), which were significantly negatively correlated with the risk score ( $R = -0.41$ ,  $p < 0.001$  and  $R = -0.33$ ;  $p < 0.001$ , respectively; Figures 6E–H). This indicates that the glycometabolism-related risk score is a novel biomarker for assessing immunotherapy responsiveness and sensitivity to targeted drugs.

## Molecular pathways and immune characteristics of glycometabolism-related hub genes

The hub genes related to glycometabolism in the risk score are also important. The levels of glycometabolism-related hub genes (G6PD, CENPA, STC2, and PFKFB4) in tumor samples were higher than those in normal samples ( $p < 0.001$ ; Figure 7A; Supplementary Figure S4A). The relationships of the four glycometabolism-related hub genes with gender, clinical grade, stage, and T stage were investigated. The expression levels of the four glycometabolism-related hub genes significantly increased with an increase in tumor grade, stage, and T stage (Supplementary Figures S4B–E). Subsequently, we also focused on the methylation expression level of hub genes and found that the expression levels of G6PD ( $p = 1.083^{-02}$ ), CENPA ( $p = 7.08^{-04}$ ), STC2 ( $p = 3.507^{-02}$ ) and PFKFB4 ( $p = 3.541^{-03}$ ) genes were negatively correlated with their methylation levels (Figure 7B; Supplementary Figures S5A–C). The G6PD gene has copy number variations in B cells, CD8<sup>+</sup> T-cells, macrophages, neutrophils, and dendritic cells (Figure 7C, Supplementary Figures S5D–F). At the same time, the high G6PD expression group was mainly composed of regulatory T-cells and macrophage M0 cells, which was significantly different from the composition in the low expression group ( $p < 0.001$ ; Figure 7D; Supplementary Figures S5G–I). Multiple GSEA analysis of G6PD showed that mTOR signaling, Notch signaling, and cancer pathways were mainly enriched in the high expression group, whereas other metabolic pathways, such as drug metabolism, fatty acid metabolism, and tryptophan metabolism were enriched in the low expression group (Figures 7E,F).

## Clinical samples and basic experimental verification of hub genes

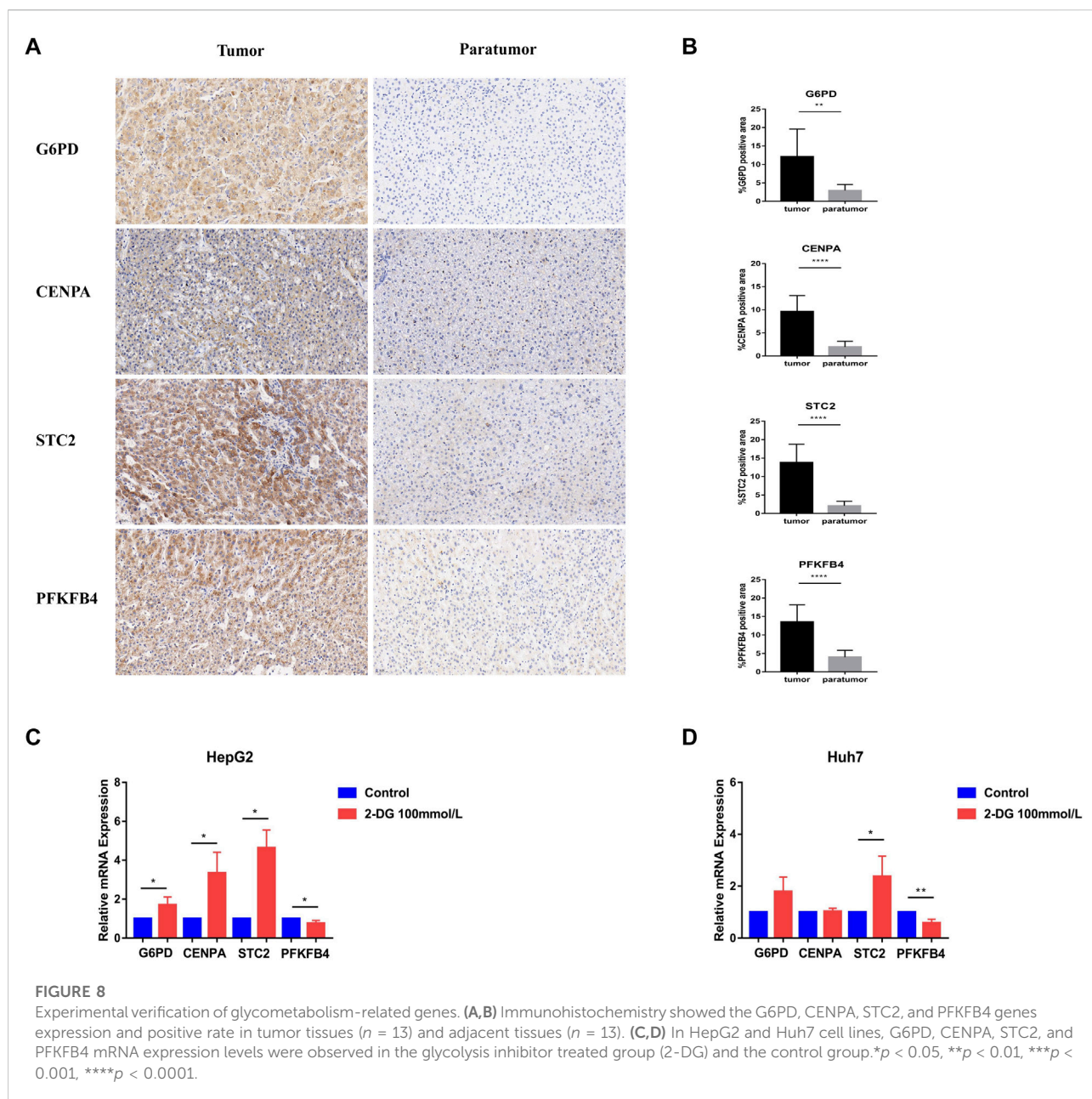
Further, we verified the results in the database through clinical samples and cell line experiments. We selected patients who had undergone hepatectomy for HCC to take tumor tissues and adjacent tissues for immunohistochemical analysis. It was found that the details of four glycometabolism-related genes in tumor tissues were significantly higher than those in adjacent tissues ( $p < 0.01$ ; Figures 8A,B). Subsequently, the expression of four glycometabolism-related genes was detected in HepG2 and Huh7 cell lines after incubation for 48 h and blocking with glycolysis blocker 2-DG. After blocking the glycolytic pathway, only the expression of PFKFB4 decreased in both HepG2 and Huh7 cell lines ( $p < 0.05$ ; Figures 8C,D), while STC2 increased in both HepG2 and Huh7 cell lines ( $p < 0.05$ ). The expression of G6PD and CENPA increased only in HepG2 cell line ( $p < 0.05$ ).

## Discussion

Glucose metabolism is important for all cells to maintain their activity (Bose and Le, 2018). Reprogramming of glucose metabolism can lead to tumor cell progression, metastasis, and recurrence (Peng et al., 2021). However, changes in glucose metabolism in immune cells have attracted increasing attention in immune metabolism. Targeted glucose metabolism may become a promising method for regulating immunity and improving the efficacy of immunotherapy, and may play an important role in the field of tumors and viruses (Hay, 2016; O'Neill et al., 2016). Increasing evidence has revealed the relationship between glucose metabolism reprogramming of immune effector cells, such as T-cells and NK cells, and tumor occurrence and development (Cong et al., 2018; Kang and Tang, 2020). In the past, glycometabolism-related genes have been used to evaluate the prognosis of patients with prostate cancer, gastric cancer, ovarian cancer, and lung adenocarcinoma (Huang et al., 2019; Liu et al., 2020; Zhang et al., 2021a; He et al., 2021). However, there is no prognostic model based on glycometabolism-related genes that evaluates the survival and immune-related characteristics of patients with HCC. Although there are some studies on metabolism-associated molecular classification in HCC (Yang et al., 2020), it is necessary to further focus on the target of glycometabolism, which is beneficial for improving clinical efficacy.

With the rapid development of transcriptomics and bioinformatics, an increasing number of gene markers have become available. The prognostic risk scoring system and nomogram model of glycometabolism-related genes in HCC tumor and normal samples from TCGA and GEO cohorts were established through LASSO Cox regression analysis. The prognostic risk score of glycometabolism-related genes can clarify the roles of hub genes in HCC. The nomogram model intuitively predicted the survival rates of patients at different follow-up times (Balachandran et al., 2015). The OS rate of the patients with HCC in the high-risk group was lower than low-risk group. Both the training and validation groups showed consistent results. The prognostic risk score of glycometabolism-related genes can effectively distinguish high-risk groups and has a particular guiding role in the survival and prognosis of patients with HCC.

The glycometabolism-related prognostic risk scores included G6PD, CENPA, STC2, and PFKFB4. G6PD is the core enzyme of the pentose phosphate pathway (PPP) in glucose metabolism and is involved in tumor cell growth, invasion, and metastasis (Yang et al., 2021). Previous studies have reported that high Nrf2 expression can enhance the expression of G6PD and HIF-1 in breast cancer cells. Nrf-2 activates antioxidant enzymes and upregulates Notch-1 through the G6PD/HIF-1 pathway, thereby affecting the proliferation of breast cancer cells (Zhang et al., 2019a). At the same time, G6PD overexpression was found to be associated with an increase in mTORC1 activity in blood tumors such as acute myeloid leukemia, and also predicts a poor prognosis (Poulain et al., 2017). In this study, we found that the expression of G6PD in tumor tissues was significantly higher than that in normal tissues, which was also



confirmed by immunohistochemical analysis. Meanwhile, it was found that the Notch and mTOR signaling pathways were significantly enriched in the high G6PD expression group, which was consistent with previous reports. CENPA was found to be highly expressed in a variety of tumors and is associated with poor prognosis (Saha et al., 2020; Han et al., 2021). CENPA acts as an upstream transcriptional activator of the karyopherin  $\alpha 2$  subunit gene (KPNA2), indirectly promoting tumor cell growth and glycolysis in patients with colon cancer (Liang et al., 2021). *In vitro* studies have also shown that CENPA can activate the Wnt/ $\beta$ -catenin signaling pathway and promote the proliferation and

metastasis of renal cell carcinoma (Wang et al., 2021a). Glycolysis was the main factor that drives tumor development. STC2 participates in glycolysis-related pathways and phosphorus metabolism (Li et al., 2021). It was highly expressed in pancreatic, lung, colon, and breast cancer and other malignant tumors and is associated with poor prognosis (Na et al., 2015; Zhang et al., 2019b; Jiang et al., 2019; Lin et al., 2019). However, PFKFB4 was a key regulatory enzyme of glycolytic synthesis. In breast, bladder, and pancreatic cancer and other malignant tumors, targeting the glycolysis pathway mediated by PFKFB4 can inhibit the growth and invasion of tumor cells (Zhang et al., 2016; Dasgupta et al., 2018;

Zhang et al., 2021b). Our basic experiments found that glycolysis inhibitors significantly inhibited the level of PFKFB4 mRNA, while other hub genes are not directly controlled by glycolysis pathway, and there may be other metabolic pathways. It was found in the tumor tissue of HCC patients that compared with the adjacent tissue, the tumor tissue significantly overexpressed 4 glycometabolism-related genes. Therefore, glycometabolism-related risk score was a biomarker of tumor promotion and is significantly correlated with poor prognosis.

Understanding the immune cell composition, molecular pathways, and immune functions of different glycometabolism-related risk scores in TME can improve immunotherapy. The relative proportions of 31 immune cells were further assessed using x-Cell, and the correlation between infiltrating immune cells was assessed using CIBERSORT, which were compared in different glycometabolism-related risk scores. The results showed that the main enriched immune cells in the high-risk group were inflammatory cells such as Th1, Th2, basophils, M2 macrophages, and B cells. Through GO and KEGG pathway enrichment analysis, the IL-27 pathway was found to be enriched in the high-risk group. It has been reported that IL-27 can affect multiple effector cells in innate and adaptive immunity (Villarino et al., 2005; Li et al., 2020; Dong et al., 2021). Therefore, enrichment of the IL-27 pathway may be associated with higher scores of CD8<sup>+</sup>T and NKT cells in the high-risk group. It was suggested that glycometabolism-related risk scores and immune inflammatory function may be mediated by IL-27 signaling. The traditional understanding of IL-27 is that its main response cells are immune cells. A recent study in the journal Nature found that IL-27 promotes the browning of adipose tissue by up-regulating the expression of Uncoupling Protein 1 (UCP1) to promote heat production and energy consumption, thereby reducing obesity (Wang et al., 2021b). This was also the latest understanding of the corresponding non-immune cells of IL-27. The results of this study also showed that the high-risk group was mainly involved in immunosuppressive function, including APC coinhibition, checkpoint, and T-cell coinhibition. The data suggested that the high-risk group has characteristics of immune inflammatory cells and immunosuppressive function, and the immune cell infiltration and inflammatory characteristics lead to a poor prognosis in the high-risk group.

Since glycometabolism-related risk scores were associated with poor prognosis in HCC, we explored the relationship between risk scores and resistance to ICI (PD-1 and CTLA-4) and targeted drugs (sorafenib and 5-fluorouracil). PD-1 and CTLA-4 were gradually used for immune checkpoint inhibitors to treat liver cancer (Finkelmeier et al., 2018; Fessas et al., 2020). Sorafenib was a first-line targeted drug used for the treatment of liver cancer. Therefore, it was necessary to analyze their clinical reactivity and drug resistance to better guide clinical medication. We analyzed the relationship between different glycometabolism-related

risk scores and immunophenoscore in patients with HCC. The low-risk group had a higher response to anti-CTLA-4 treatment, indicating that the response to ICI was better in the low-risk group. Subsequently, analysis of sorafenib and 5-Fluorouracil drug resistance revealed that the half maximal inhibitory concentration (IC<sub>50</sub>) of the high-risk group was lower than that of the low-risk group, indicating that drug resistance was less likely to occur in the high-risk group. Glycometabolism-related risk scores provided a possible basis for clinical treatment, and further experimental and clinical verification is needed.

The current study has certain limitations. First, the established prognostic risk score for glycometabolism needs to be validated in a larger multicenter cohort. Second, the downloading of relevant data from public databases is very limited, and it is unknown whether the patient has other metabolic problems. Third, it is necessary to further explore the mechanism of glycometabolism and immunity on the progression and prognosis of HCC. In the future research work, we will pay attention to the clinical application value of this prognostic risk score, and explore the application prospects of glucose metabolism in the field of tumor immunity, which is very instructive and valuable work.

In conclusion, we constructed a prognosis risk score based on glycometabolism-related genes that can predict OS and PFS in patients with HCC and reflect the responsiveness to immune infiltrating cells and immunotherapy in patients. In addition, the prognostic risk score may be a potential biomarker in the field of immunometabolism. Further clinical and experimental studies are required to confirm these findings.

## 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

ZY and JD designed the study. LY and JD wrote the manuscript, collected, and analyzed the data. XL, XW, HY, QP, and YX collected data. ZY approved for final revision and approval.

## Funding

This study was supported by the National Natural Science Foundation of China (Nos. 81874435 and 81902895),

Dengfeng Talent Support Program of Beijing Municipal Administration of Hospitals (No. DFL20191803), the Special Fund of Capital Health Research and Development (No. 2020-2-2173), Beijing Hospitals Authority Clinical Medicine Development of Special Funding Support (No. ZYLX202127), Beijing Natural Science Foundation (No. M21007).

## 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.

## References

- Balachandran, V. P., Gonen, M., Smith, J. J., and DeMatteo, R. P. (2015). Nomograms in oncology: More than meets the eye. *Lancet Oncol.* 16 (4), e173–e180. doi:10.1016/S1470-2045(14)71116-7
- Bose, S., and Le, A. (2018). Glucose metabolism in cancer. *Adv. Exp. Med. Biol.* 1063, 3–12. doi:10.1007/978-3-319-77736-8\_1
- Chang, C. H., Qiu, J., O'Sullivan, D., Buck, M. D., Noguchi, T., Curtis, J. D., et al. (2015). Metabolic competition in the tumor microenvironment is a driver of cancer progression. *Cell* 162 (6), 1229–1241. doi:10.1016/j.cell.2015.08.016
- Cong, J., Wang, X., Zheng, X., Wang, D., Fu, B., Sun, R., et al. (2018). Dysfunction of natural killer cells by FBP1-induced inhibition of glycolysis during lung cancer progression. *Cell Metab.* 28 (2), 243–255. e5. doi:10.1016/j.cmet.2018.06.021
- Dai, Y., Qiang, W., Lin, K., Gui, Y., Lan, X., and Wang, D. (2021). An immune-related gene signature for predicting survival and immunotherapy efficacy in hepatocellular carcinoma. *Cancer Immunol. Immunother.* 70 (4), 967–979. doi:10.1007/s00262-020-02743-0
- Darvin, P., Toor, S. M., Sasidharan Nair, V., and Elkord, E. (2018). Immune checkpoint inhibitors: Recent progress and potential biomarkers. *Exp. Mol. Med.* 50 (12), 1–11. doi:10.1038/s12276-018-0191-1
- Dasgupta, S., Rajapakse, K., Zhu, B., Nikolai, B. C., Yi, P., Putluri, N., et al. (2018). Metabolic enzyme PFKFB4 activates transcriptional coactivator SRC-3 to drive breast cancer. *Nature* 556 (7700), 249–254. doi:10.1038/s41586-018-0018-1
- DePeaux, K., and Delgoffe, G. M. (2021). Metabolic barriers to cancer immunotherapy. *Nat. Rev. Immunol.* 21 (12), 785–797. doi:10.1038/s41577-021-00541-y
- Dong, C., Dang, D., Zhao, X., Wang, Y., Wang, Z., and Zhang, C. (2021). Integrative characterization of the role of IL27 in melanoma using bioinformatics analysis. *Front. Immunol.* 12, 713001. doi:10.3389/fimmu.2021.713001
- Feng, R. M., Zong, Y. N., Cao, S. M., and Xu, R. H. (2019). Current cancer situation in China: Good or bad news from the 2018 global cancer statistics? *Cancer Commun.* 39 (1), 22. doi:10.1186/s40880-019-0368-6
- Fessas, P., Kaseb, A., Wang, Y., Saeed, A., Szafron, D., Jun, T., et al. (2020). Post-registration experience of nivolumab in advanced hepatocellular carcinoma: An international study. *J. Immunother. Cancer* 8 (2), e001033. doi:10.1136/jitc-2020-001033
- Finkelmeier, F., Waidmann, O., and Trojan, J. (2018). Nivolumab for the treatment of hepatocellular carcinoma. *Expert Rev. Anticancer Ther.* 18 (12), 1169–1175. doi:10.1080/14737140.2018.1535315
- Forner, A., Reig, M., and Bruix, J. (2018). Hepatocellular carcinoma. *Lancet* 391 (10127), 1301–1314. doi:10.1016/S0140-6736(18)30010-2
- Guo, Y., Xie, Y. Q., Gao, M., Zhao, Y., Franco, F., Wenes, M., et al. (2021). Metabolic reprogramming of terminally exhausted CD8+ T cells by IL-10 enhances anti-tumor immunity. *Nat. Immunol.* 22 (6), 746–756. doi:10.1038/s41590-021-00940-2
- Han, J., Xie, R., Yang, Y., Chen, D., Liu, L., Wu, J., et al. (2021). CENPA is one of the potential key genes associated with the proliferation and prognosis of ovarian

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2022.940551/full#supplementary-material>

cancer based on integrated bioinformatics analysis and regulated by MYBL2. *Transl. Cancer Res.* TCR 10 (9), 4076–4086. doi:10.21037/tcr-21-175

Hay, N. (2016). Reprogramming glucose metabolism in cancer: Can it be exploited for cancer therapy? *Nat. Rev. Cancer* 16 (10), 635–649. doi:10.1038/nrc.2016.77

He, Z., Chen, D., Wu, J., Sui, C., Deng, X., Zhang, P., et al. (2021). Yes associated protein 1 promotes resistance to 5-fluorouracil in gastric cancer by regulating GLUT3-dependent glycometabolism reprogramming of tumor-associated macrophages. *Archives Biochem. Biophys.* 702, 108838. doi:10.1016/j.abb.2021.108838

Huang, Y., Yang, X., Sun, F., Lu, T., Bi, G., Liang, J., et al. (2019). Prognostic effects of glycometabolism changes in lung adenocarcinoma: A prospective observational study. *Transl. Lung Cancer Res.* 8 (6), 808–819. doi:10.21037/tlcr.2019.10.18

International Agency for Research on Cancer (2020). *Globocan 2018*. IARC. Available at: [https://gco.iarc.fr/today/online-analysis-map?v=2020&mcode=population&mcode\\_population=continents&population=900&populations=900&key=asr&sex=0&cancer=11&type=0&statistic=5&prevalence=0&population\\_group=earth&color\\_palette=default&map\\_scale=quantile&map\\_nb\\_colors=5&continent=0&rotate=%255B10%252C0%255D](https://gco.iarc.fr/today/online-analysis-map?v=2020&mcode=population&mcode_population=continents&population=900&populations=900&key=asr&sex=0&cancer=11&type=0&statistic=5&prevalence=0&population_group=earth&color_palette=default&map_scale=quantile&map_nb_colors=5&continent=0&rotate=%255B10%252C0%255D)

Jiang, S. T., Wang, H. Q., Yang, T. C., Wang, D. W., Yang, L. J., Xi, Y. Q., et al. (2019). Expression of stanniocalcin 2 in breast cancer and its clinical significance. *Curr. Med. Sci.* 39 (6), 978–983. doi:10.1007/s11596-019-2131-2

Kang, S., and Tang, H. (2020). HIV-1 infection and glucose metabolism reprogramming of T cells: Another approach toward functional cure and reservoir eradication. *Front. Immunol.* 11, 572677. doi:10.3389/fimmu.2020.572677

Li, B., Chan, H. L., and Chen, P. (2019). Immune checkpoint inhibitors: Basics and challenges. *Cmc* 26 (17), 3009–3025. doi:10.2174/0929867324666170804143706

Li, S., Huang, Q., Li, D., Lv, L., Li, Y., and Wu, Z. (2021). The significance of Stanniocalcin 2 in malignancies and mechanisms. *Bioengineered* 12 (1), 7276–7285. doi:10.1080/21655979.2021.1977551

Li, W., Xu, M., Li, Y., Huang, Z., Zhou, J., Zhao, Q., et al. (2020). Comprehensive analysis of the association between tumor glycolysis and immune/inflammation function in breast cancer. *J. Transl. Med.* 18 (1), 92. doi:10.1186/s12967-020-02267-2

Liang, Y. C., Su, Q., Liu, Y. J., Xiao, H., and Yin, H. Z. (2021). Centromere protein A (CENPA) regulates metabolic reprogramming in the colon cancer cells by transcriptionally activating karyopherin subunit alpha 2 (KPNA2). *Am. J. Pathology* 191 (12), 2117–2132. doi:10.1016/j.ajpath.2021.08.010

Lin, C., Sun, L., Huang, S., Weng, X., and Wu, Z. (2019). STC2 is a potential prognostic biomarker for pancreatic cancer and promotes migration and invasion by inducing epithelial-mesenchymal transition. *BioMed Res. Int.* 2019, 1–9. doi:10.1155/2019/8042489

Liu, L., Cai, L., Liu, C., Yu, S., Li, B., Pan, L., et al. (2020). Construction and validation of a novel glycometabolism-related gene signature predicting survival in patients with ovarian cancer. *Front. Genet.* 11, 585259. doi:10.3389/fgene.2020.585259

- Na, S. S., Aldonza, M. B., Sung, H. J., Kim, Y. I., Son, Y. S., Cho, S., et al. (2015). Stanniocalcin-2 (STC2): A potential lung cancer biomarker promotes lung cancer metastasis and progression. *Biochimica Biophysica Acta (BBA) - Proteins Proteomics* 1854 (6), 668–676. doi:10.1016/j.bbapap.2014.11.002
- O'Neill, L. A., Kishton, R. J., and Rathmell, J. (2016). A guide to immunometabolism for immunologists. *Nat. Rev. Immunol.* 16 (9), 553–565. doi:10.1038/nri.2016.70
- Pan, L., Fang, J., Chen, M. Y., Zhai, S. T., Zhang, B., Jiang, Z. Y., et al. (2020). Promising key genes associated with tumor microenvironments and prognosis of hepatocellular carcinoma. *Wjg* 26 (8), 789–803. doi:10.3748/wjg.v26.i8.789
- Peng, Y., Yang, H., and Li, S. (2021). The role of glycometabolic plasticity in cancer. *Pathology - Res. Pract.* 226, 153595. doi:10.1016/j.prp.2021.153595
- Poulain, L., Sujobert, P., Zylbersztein, F., Barreau, S., Stuani, L., Lambert, M., et al. (2017). High mTORC1 activity drives glycolysis addiction and sensitivity to G6PD inhibition in acute myeloid leukemia cells. *Leukemia* 31 (11), 2326–2335. doi:10.1038/leu.2017.81
- Poznanski, S. M., Singh, K., Ritchie, T. M., Aguiar, J. A., Fan, I. Y., Portillo, A. L., et al. (2017). Metabolic flexibility determines human NK cell functional fate in the tumor microenvironment. *Cell Metab.* 33 (6), 1205–1220. e5. doi:10.1016/j.cmet.2021.03.023
- Saha, A. K., Contreras-Galindo, R., Niknafs, Y. S., Iyer, M., Qin, T., Padmanabhan, K., et al. (2020). The role of the histone H3 variant CENPA in prostate cancer. *J. Biol. Chem.* 295 (25), 8537–8549. doi:10.1074/jbc.RA119.010080
- Shevchenko, I., and Bazhin, A. V. (2018). Metabolic checkpoints: Novel avenues for immunotherapy of cancer. *Front. Immunol.* 9, 1816. doi:10.3389/fimmu.2018.01816
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U.S.A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA A Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660
- Vesteinn, T., David, L. G., Scott, D. B., Denise, W., and Armaz, M. (2018). The immune landscape of cancer. *Immunity* 48 (4). doi:10.1016/j.immuni.2018.03.023
- Villarino, A. V., Larkin, J., 3rd, Saris, C. J., Caton, A. J., Lucas, S., Wong, T., et al. (2005). Positive and negative regulation of the IL-27 receptor during lymphoid cell activation. *J. Immunol.* 174 (12), 7684–7691. doi:10.4049/jimmunol.174.12.7684
- Wang, F. S., Fan, J. G., Zhang, Z., Gao, B., and Wang, H. Y. (2014). The global burden of liver disease: The major impact of China. *Hepatology* 60 (6), 2099–2108. doi:10.1002/hep.27406
- Wang, Q., Li, D., Cao, G., Shi, Q., Zhu, J., Zhang, M., et al. (2021). IL-27 signalling promotes adipocyte thermogenesis and energy expenditure. *Nature* 600 (7888), 314–318. doi:10.1038/s41586-021-04127-5
- Wang, Q., Xu, J., Xiong, Z., Xu, T., Liu, J., Liu, Y., et al. (2021). CENPA promotes clear cell renal cell carcinoma progression and metastasis via Wnt/ $\beta$ -catenin signaling pathway. *J. Transl. Med.* 19 (1), 417. doi:10.1186/s12967-021-03087-8
- Yang, C., Huang, X., Liu, Z., Qin, W., and Wang, C. (2020). Metabolism-associated molecular classification of hepatocellular carcinoma. *Mol. Oncol.* 14 (4), 896–913. doi:10.1002/1878-0261.12639
- Yang, H. C., Stern, A., and Chiu, D. T. (2021). G6PD: A hub for metabolic reprogramming and redox signaling in cancer. *Biomed. J.* 44 (3), 285–292. doi:10.1016/j.bj.2020.08.001
- Zhang, B. H., Yang, J., Jiang, L., Lyu, T., Kong, L. X., Tan, Y. F., et al. (2020). Development and validation of a 14-gene signature for prognosis prediction in hepatocellular carcinoma. *Genomics* 112 (4), 2763–2771. doi:10.1016/j.ygeno.2020.03.013
- Zhang, C., Chen, S., Ma, X., Yang, Q., Su, F., Shu, X., et al. (2019). Upregulation of STC2 in colorectal cancer and its clinicopathological significance. *Ott Vol.* 12, 1249–1258. doi:10.2147/OTT.S191609
- Zhang, F. P., Huang, Y. P., Luo, W. X., Deng, W. Y., Liu, C. Q., Xu, L. B., et al. (2020). Construction of a risk score prognosis model based on hepatocellular carcinoma microenvironment. *Wjg* 26 (2), 134–153. doi:10.3748/wjg.v26.i2.134
- Zhang, H., Lu, C., Fang, M., Yan, W., Chen, M., Ji, Y., et al. (2016). HIF-1 $\alpha$  activates hypoxia-induced PFKFB4 expression in human bladder cancer cells. *Biochem. Biophysical Res. Commun.* 476 (3), 146–152. doi:10.1016/j.bbrc.2016.05.026
- Zhang, H. S., Zhang, Z. G., Du, G. Y., Sun, H. L., Liu, H. Y., Zhou, Z., et al. (2019). Nrf2 promotes breast cancer cell migration via up-regulation of G6PD/HIF-1 $\alpha$ /Notch1 axis. *J. Cell Mol. Med.* 23 (5), 3451–3463. doi:10.1111/jcmm.14241
- Zhang, P., Tao, W., Lu, C., Fan, L., Jiang, Q., Yang, C., et al. (2021). Bruceine A induces cell growth inhibition and apoptosis through PFKFB4/GSK3 $\beta$  signaling in pancreatic cancer. *Pharmacol. Res.* 169, 105658. doi:10.1016/j.phrs.2021.105658
- Zhang, Y., Mou, Y., Liang, C., Zhu, S., Liu, S., Shao, P., et al. (2021). Promoting cell proliferation, cell cycle progression, and glycolysis: Glycometabolism-related genes act as prognostic signatures for prostate cancer. *Prostate* 81 (3), 157–169. doi:10.1002/pros.24092



## OPEN ACCESS

## EDITED BY

Haitao Wang,  
National Cancer Institute (NIH),  
United States

## REVIEWED BY

Siqi Chen,  
Washington University in St. Louis,  
United States  
Vicky Yau,  
Columbia University Irving Medical  
Center, United States  
Xinwei Wu,  
National Institutes of Health (NIH),  
United States  
Lin Zhang,  
Clinical Center (NIH), United States  
Jiankang Fang,  
University of Pennsylvania, United States

## \*CORRESPONDENCE

Fenghong Cao,  
caofenghong@163.com

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 06 August 2022

ACCEPTED 22 August 2022

PUBLISHED 16 September 2022

## CITATION

Nie S, Huili Y, Yao A, Liu J, Wang Y,  
Wang L, Zhang L, Kang S and Cao F  
(2022), Identification of subtypes of  
clear cell renal cell carcinoma and  
construction of a prognostic model  
based on fatty acid metabolism genes.  
*Front. Genet.* 13:1013178.  
doi: 10.3389/fgene.2022.1013178

## COPYRIGHT

© 2022 Nie, Huili, Yao, Liu, Wang, Wang,  
Zhang, Kang 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.

# Identification of subtypes of clear cell renal cell carcinoma and construction of a prognostic model based on fatty acid metabolism genes

Shiwen Nie<sup>1</sup>, Youlong Huili<sup>1</sup>, Anliang Yao<sup>2</sup>, Jian Liu<sup>2</sup>,  
Yong Wang<sup>2</sup>, Lei Wang<sup>2</sup>, Liguang Zhang<sup>2</sup>, Shaosan Kang<sup>2</sup> and  
Fenghong Cao<sup>2\*</sup>

<sup>1</sup>North China University of Science and Technology, Tangshan, China, <sup>2</sup>Department of Urology, North China University of Science and Technology Affiliated Hospital, Tangshan, China

**Background:** The effects of fatty acid metabolism in many tumors have been widely reported. Due to the diversity of lipid synthesis, uptake, and transformation in clear cell renal cell carcinoma (ccRCC) cells, many studies have shown that ccRCC is associated with fatty acid metabolism. The study aimed to explore the impact of fatty acid metabolism genes on the prognosis and immunotherapy of ccRCC.

**Methods:** Two subtypes were distinguished by unsupervised clustering analysis based on the expression of 309 fatty acid metabolism genes. A prognostic model was constructed by lasso algorithm and multivariate COX regression analysis using fatty acid metabolism genes as the signatures. The tumor microenvironment between subtypes and between risk groups was further analyzed. The International Cancer Genome Consortium cohort was used for external validation of the model.

**Results:** The analysis showed that subtype B had a poorer prognosis and a higher degree of immune infiltration. The high-risk group had a poorer prognosis and higher tumor microenvironment scores. The nomogram could accurately predict patient survival.

**Conclusion:** Fatty acid metabolism may affect the prognosis and immune infiltration of patients with ccRCC. The analysis was performed to understand the potential role of fatty acid metabolism genes in the immune infiltration and prognosis of patients. These findings have implications for individualized treatment, prognosis, and immunization for patients with ccRCC.

**Abbreviations:** FRGs, fatty acid metabolism-related genes; DEGs, Differentially expressed genes; TCGA, The Cancer Genome Atlas; ROC, Receiver operating characteristic; LASSO, least absolute shrinkage and selection operator; PCA, principal component analysis; GSVA, Gene set variation analysis; OS, Overall survival; TME, Tumor microenvironment; ccRCC, clear cell renal cell carcinoma; GEO, Gene Expression Omnibus; ICGC, International Cancer Genome Consortium.

## KEYWORDS

fatty acid metabolic, prognostic model, tumor microenvironment, TCGA, GEO

## Introduction

Renal cell carcinoma (RCC) is one of the top 10 tumors recorded globally (Siegel et al., 2018), and according to the World Health Organization, more than 140,000 renal cell carcinoma patients die each year (Ferlay et al., 2015). The main pathological types of renal cell carcinoma include ccRCC, papillary renal cell carcinoma, and chromophobe cell renal cell carcinoma (Shuch et al., 2015; Hsieh et al., 2017), with ccRCC being the most common pathological subtype worldwide (Shen and Kaelin, 2013). Given the dangers of ccRCC, the identification of effective predictive tools and potential therapeutic targets remains of current interest.

It is well known that metabolic imbalance is one of the main characteristics of tumors, and existing scientific studies confirm that metabolic reprogramming plays a very critical role in the development of tumors (Hanahan and Weinberg, 2011; Rosario et al., 2018; Crunkhorn, 2019). Energy metabolic reprogramming, a new hallmark of tumors, enables accelerated cell growth and proliferation (Veglia et al., 2019; Corn et al., 2020). The first typical metabolic change is the Warburg effect, i.e. aerobic glycolysis. Next, there is glutamine metabolism (Wise and Thompson, 2010). Previously, studies related to abnormal fatty acid metabolism (FA) have not received much attention, but in recent years it has gradually attracted more attention as one of the features of metabolic reprogramming in tumors (Li and Zhang, 2016; Qi et al., 2019; Li et al., 2020). In many cancers, lipid uptake and storage are increased to meet cancer cell energy requirements, and fatty acids provide energy to tumor cells via  $\beta$ -oxidation (Cheng et al., 2018). Renal cell carcinoma has significant changes in cellular metabolism, such as FA metabolism, and RCC characterized by metabolic imbalance is considered to be a metabolic disease (Massari et al., 2015; Wettersten et al., 2017; Akhtar et al., 2018). Fatty acid synthesis is dependent on acetyl-CoA. Mutations in stearoyl-CoA desaturase 1, fatty acid synthase and acetyl-CoA carboxylase in ccRCC can lead to the substantial synthesis of acetyl-CoA, thus causing an abnormal pathway of fatty acid synthesis in ccRCC (Sajjani et al., 2017; Maan et al., 2018). Zhao et al. (2019) confirmed that multiple fatty acid metabolizing enzymes are potential prognostic markers for ccRCC, which suggests that abnormalities in fatty acid metabolism might influence the development of ccRCC. Therefore, further analysis of the impact of fatty acid metabolism-related genes (FRGs) in ccRCC may provide some reference for patient prognosis and individualized treatment.

In our analysis, we aimed to construct a prognostic signature based on TCGA (The Cancer Genome Atlas) database in conjunction with the GEO (Gene Expression Omnibus) database, using FRGs as a predictor. Patients were then distinguished into two different subtypes based on FRGs expression by unsupervised cluster analysis and, finally, the

tumor microenvironment was studied in different risk groups and different subtypes. In this analysis, our results identified FRGs as a potential target for ccRCC as well as a prognostic marker. Furthermore, an attempt was made to explain the alteration of FA metabolism in the prognostic-immune-tumor microenvironment in ccRCC.

## Materials and methods

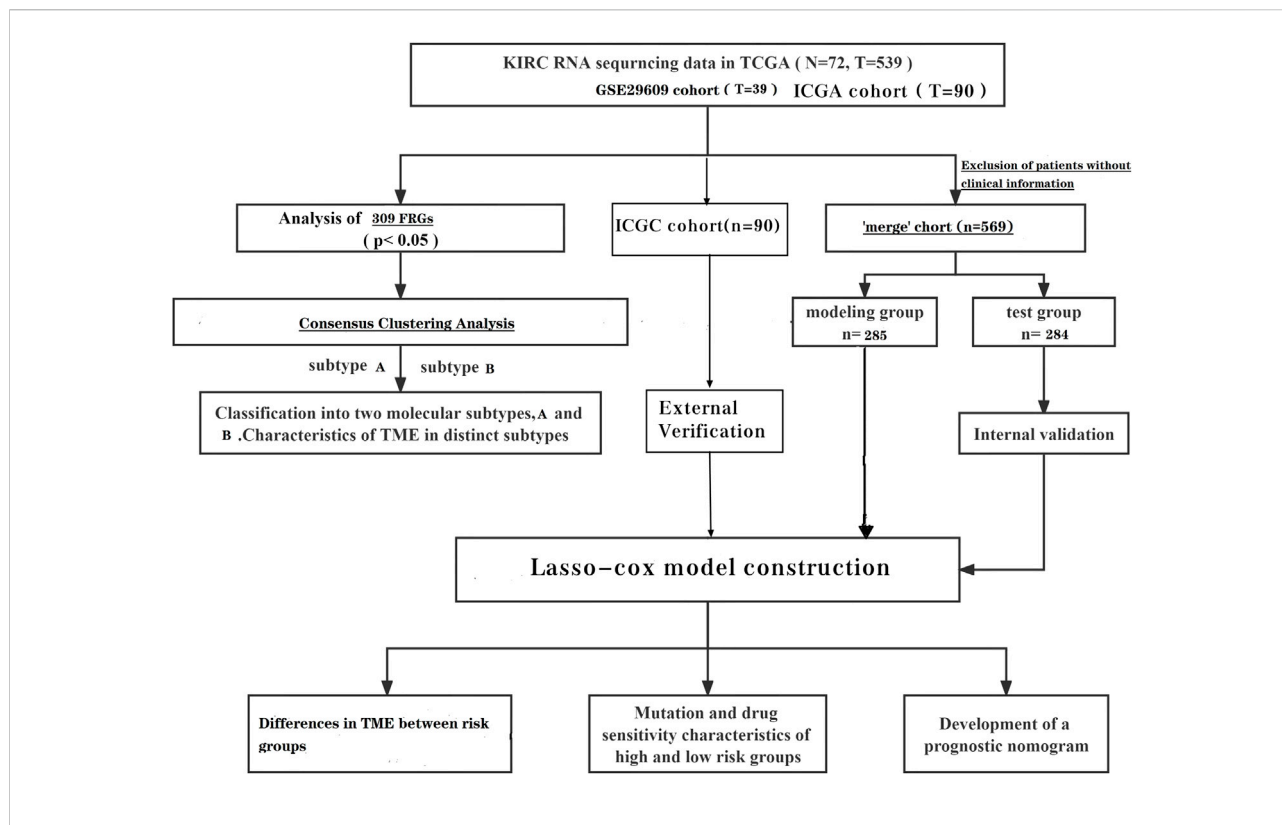
### Data collation and acquisition

To date, many sequencing data are publicly available online; therefore, we acquired the relevant data of ccRCC from TCGA (The Cancer Genome Atlas) database. There were 539 tumor samples and 79 normal samples in this dataset. Similarly, the data set GSE29609, which contains 39 tumor tissues, was obtained from the GEO (Gene Expression Omnibus) database. We also obtained clinically relevant information on the samples, and samples with incomplete clinical information were excluded from further analysis. We converted the downloaded FPKM data to TPM format. We background the adjusted and quantile the normalized TCGA data before performing the analysis. Batch effects were removed by a combat algorithm thus merging the two data sets TCGA-KIRC and GSE29609 (The log2 transform is used to normalize the data, and the combat algorithm is performed using the sva function). After excluding patients with no clinical information, a total of 569 patients were included for analysis. The independent dataset from the International Cancer Genome Consortium (ICGC) database (<https://dcc.icgc.org/>) was used for subsequent external validation. After excluding patients with incomplete clinical information, 90 patients were included in the ICGC cohort.

The gene collections of the KEGG fatty acid metabolic pathway, Hallmark fatty acid metabolic genes, and Reactome fatty acid metabolic genes were obtained from the Molecular Signature Database (MSigDB V7.2), and 309 FRGs were obtained after removing the overlapping parts of the three gene collections, and the specific genes are shown in Supplementary Table S1 (He et al., 2021).

### Prognosis-related differential genes of clear cell renal cell carcinoma

Differential analysis was performed on TCGA dataset. We extracted the DEGs (differential genes) of FRGs using “limma” with a fold change of 1.5;  $p < 0.05$ . Further univariate cox analysis of the “merged” cohort was performed to obtain genes associated



with survival time. The prognosis-related genes intersected with the differential genes of FRGs to obtain the prognosis-related DEGs (Supplementary Table S2).

## Prognostic models associated with fatty acid metabolism-related genes

The “merge” cohort was randomly assigned equally to obtain the train group and the test group. Using the train group as the base sample, the prognosis-related DEGs was analyzed using the lasso analysis and multivariate cox analysis to obtain the prognosis model.

The median FRG score in the train group can distinguish patients in the train group into low risk and high-risk groups. Survival analysis and PCA analysis were performed for the different risk groups. ROC (receiver operating characteristic) curves were plotted to verify the accuracy of the model. Finally, the correlation between clinical indicators and risk scores was assessed.

## Producing a nomogram

The nomogram was created using the “rms” package. Clinical information and risk scores are used as predictors. Each patient

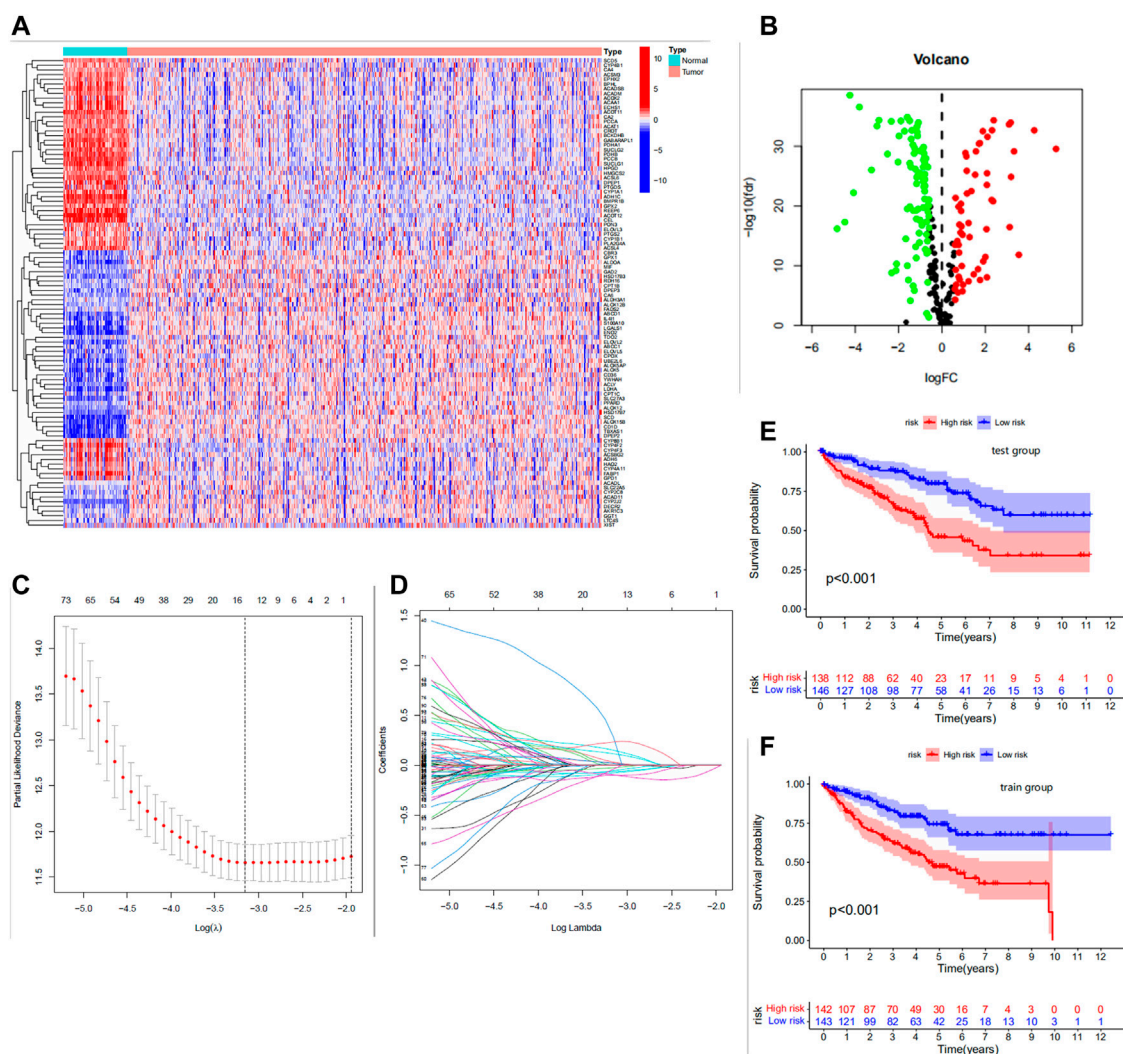
has a different score for different indicators. Each total score has a corresponding 1-year, 3-year and 5-year survival rate.

## Unsupervised consensus cluster analysis for fatty acid metabolism-related genes

We performed unsupervised clustering analysis with the “ConsensusClusterPlus” package. We classified the “merge” samples into different subtypes based on the expression of FRGs. To observe the differences of FRGs in different subtypes, we performed gene set variation analysis on different subtypes to further identify the differences between them. We also investigated the differences in immune infiltration between the different subtypes.

## Immune cell infiltration profile, somatic mutation profile, and drug sensitivity analysis

To further observe potential differences between different risk groups of ccRCC patients, the association between immune cells and modeled genes was studied using the CIBERSORT algorithm. Mutation data from TCGA.KIRC.varscan assessed mutations in different risk



**FIGURE 1**  
**(A)** Heat map of differential expression of the top 100 most significantly differentially expressed fatty acid metabolism genes. Red for high expression, blue for low expression. **(B)** Volcano plot of 151 differentially expressed genes. Red dots indicate upregulated genes, black dots indicate insignificant differences, and green dots indicate downregulated genes. **(C)** Cross-validation was performed to select the minimum lambda value to obtain the modeled genes. **(D)** LASSO coefficient profiles of the 99 FRGS. **(E,F)** The Kaplan-Meier analysis in the test/train group.

groups. The values of semi-inhibitory concentrations (IC50) of ccRCC chemotherapeutic drugs were obtained using the pRRophetic package and further observed for differences in drug sensitivity between the groups.

## Statistical

R software and Perl software were used to perform data analysis (“\*” =  $p < 0.05$ ; “\*\*” =  $p < 0.01$ ; “\*\*\*” =  $p < 0.001$ ). Kaplan-Meier analysis and log-rank test were used to compare OS between subgroups. Immune cell infiltration and TME scores were compared between high- and low-risk groups and between subtypes using the Wilcoxon test. Spearman correlation analysis

was used to compare correlations between the degree of immune cell infiltration and risk scores.

## Results

### Construction and validation of prognostic models

We extracted DEGs of 151 fatty acid metabolism-related genes using “limma,” of which 94 were down-regulated and 57 were up-regulated (Figures 1A,B). Univariate cox analysis was used to obtain genes associated with survival time. The intersection of prognosis-related genes with the differential

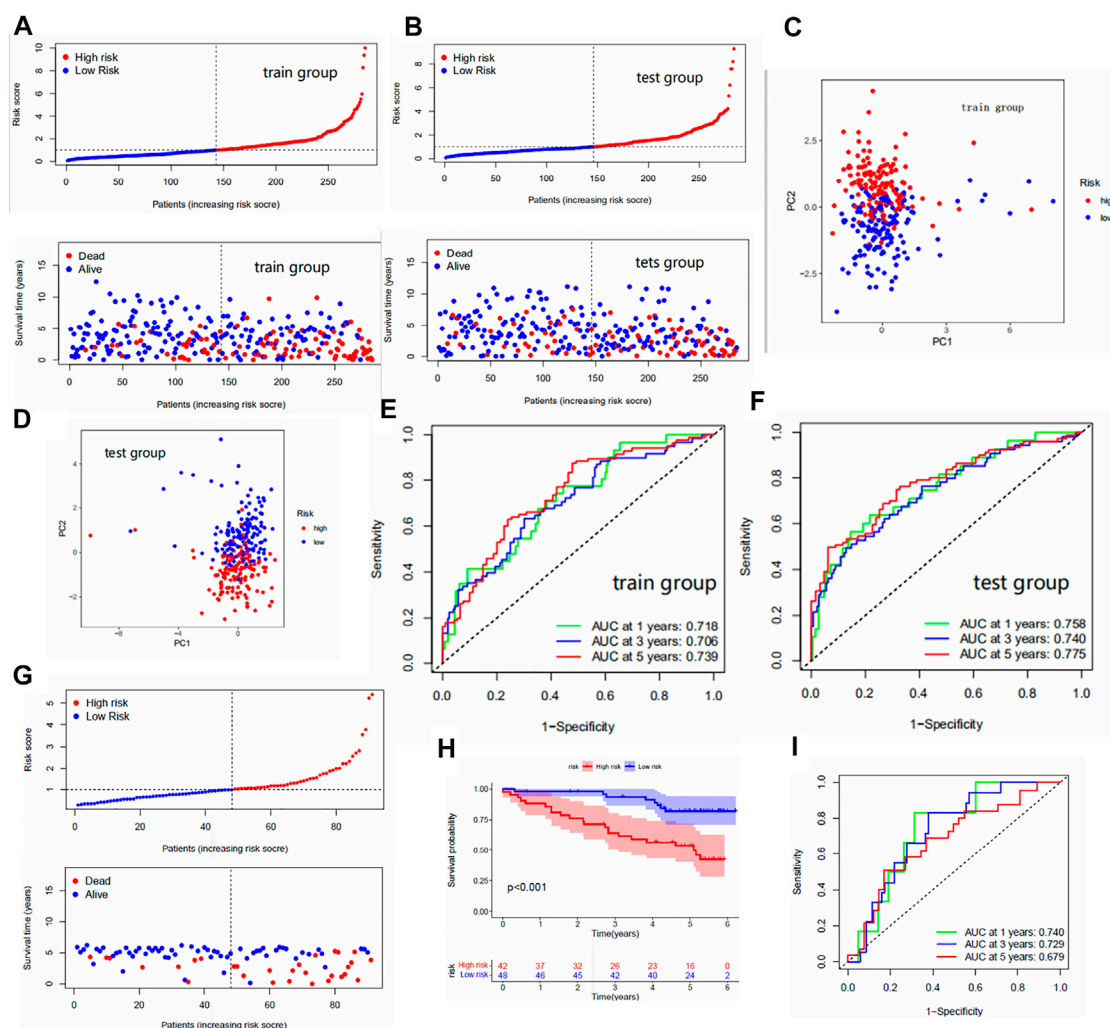


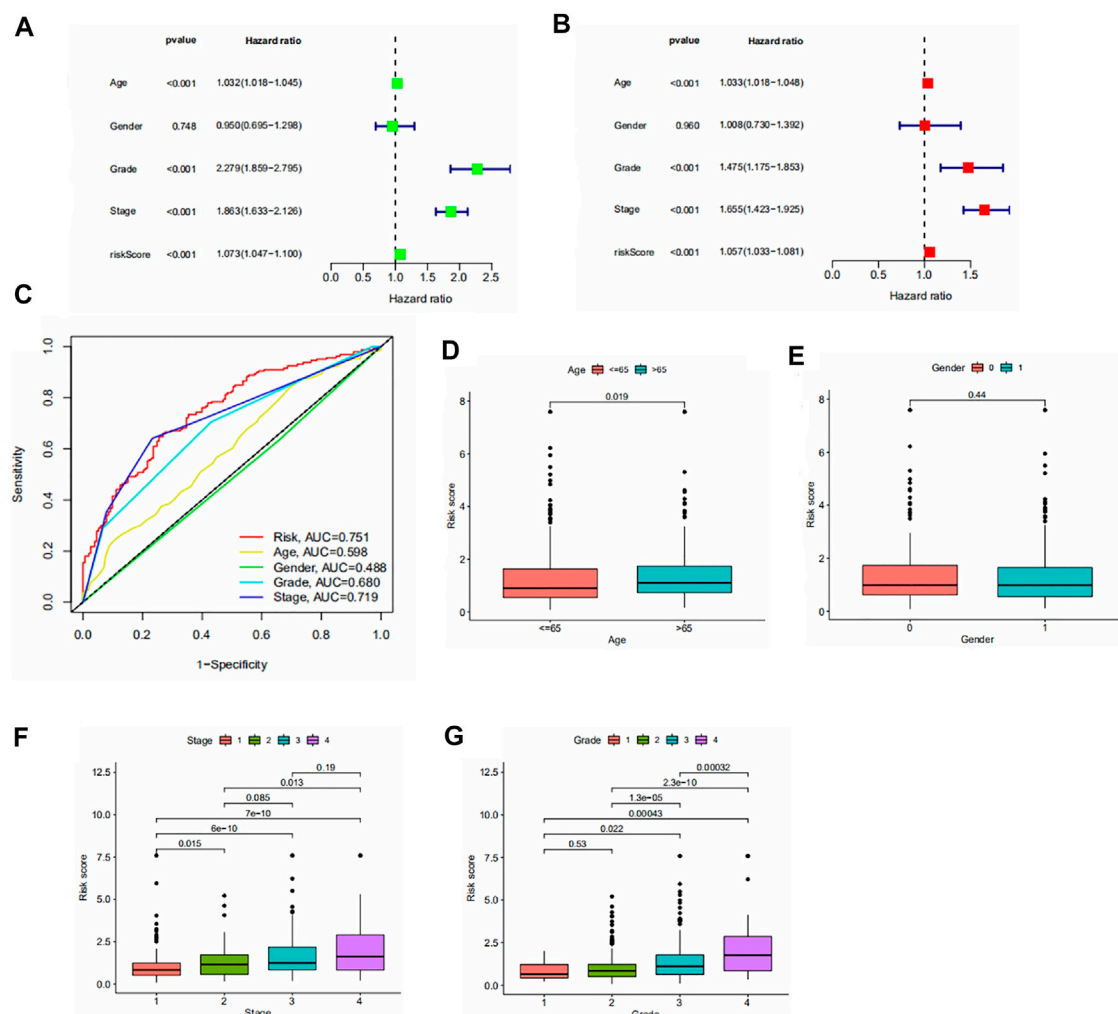
FIGURE 2

(A,B) Risk curves for the train and test groups. These include scatter plots showing risk scores and patient survival status, and ranked plots in order of increasing risk scores. (C,D) PCA results for the train/test groups. Each dot represents a patient. (E,F) The area under the ROC curve for the train/test group. The area under the ROC curve represents the accuracy of the model in predicting patient survival, where the larger the area, the higher the accuracy. (G) Risk curves for the ICGC cohort. (H) The Kaplan-Meier analysis in the ICGC cohort. (I) The area under the ROC curve for the ICGC cohort.

genes of FRGs was taken to obtain 99 prognosis-related DEGs (Supplementary Table S2). The samples were equally divided into the train group ( $n = 285$ ) and the test group ( $n = 284$ ). Based on the 99 prognosis-related DEGs, we performed LASSO regression (Figures 1C,D) and multivariate cox analysis to obtain a prognostic model including 11 predictors (HSD17B3, HPGD, CEL, PTGDS, SCD5, DPEP1, GAD2, ADH6, ALOX12B, IL4I1, and ENO2). This model can be expressed by the formula:

$$\text{Risk score} = (0.2468 * \text{HSD17B3}) + (-0.1651 * \text{HPGD}) + (0.3335 * \text{CEL}) + (0.1054 * \text{PTGDS}) + (-0.1180 * \text{SCD5}) + (-0.1239 * \text{DPEP1}) + (1.2038 * \text{GAD2}) + (-0.1880 * \text{ADH6}) + (-0.5212 * \text{ALOX12B}) + (0.2610 * \text{IL4I1}) + (0.1566 * \text{ENO2})$$

The median FRG score of the train group distinguished patients in the train group into high and low risk groups. Similarly, the test group was also distinguished into two different risk groups. The Kaplan-Meier analysis are shown in Figures 1E,F; we observed significantly lower OS (overall survival) in patients with high risk scores than in patients with low risk scores, which laterally reflects the reliability of the risk scores. The risk curve results show a gradual increase in the number of high-risk patients with increasing risk scores and a higher number of deaths in patients with high risk scores (Figures 2A,B). The PCA analysis showed us the excellent discriminatory ability of the model (Figures 2C,D). The area under the ROC curve was greater than 0.7 for both the train and test groups (Figures 2E,F). This can indicate the high accuracy of the model

**FIGURE 3**

(A,B) Forest plot of univariate/multivariate independent prognostic analysis ( $p < 0.05$  is statistically significant, the larger the hazard ratio, the higher the correlation). (C) Area under the roc curve for risk scores and clinical information. (D–G) Results of the clinical correlation analysis. Differences in risk scores by age, gender, stage, and grade.

prediction ability. To further determine the predictive power of the model, we externally validated the model using the ICGC cohort. The risk curve and survival curve also demonstrated the superior performance of the model (Figures 2G,H). We found that the ROC curves at 1, 3, and 5 years exceeded 0.65 (Figure 2I).

## Independent prognostic analysis

We performed the independent prognostic analysis of clinical indicators and risk scores. We found that age, grade, stage, and risk score could be used as prognostic factors independently of other factors (Figures 3A,B). As shown in Figure 3C, we observed that the risk score had the largest area

under the roc curve, which implies that the risk score was more accurate than age, grade, and stage.

To assess the association between risk scores and clinical parameters, we performed a clinical correlation analysis. The findings are shown in Figures 3D–G. It can be seen that age and gender have no significant effect on the risk score. The risk scores increased with increasing grade and stage levels.

## A new nomogram

To be able to use this model in a clinical setting, we used risk score, stage, and age as predictors thus constructing a Nomogram (Figure 4A). The results of the calibration plots

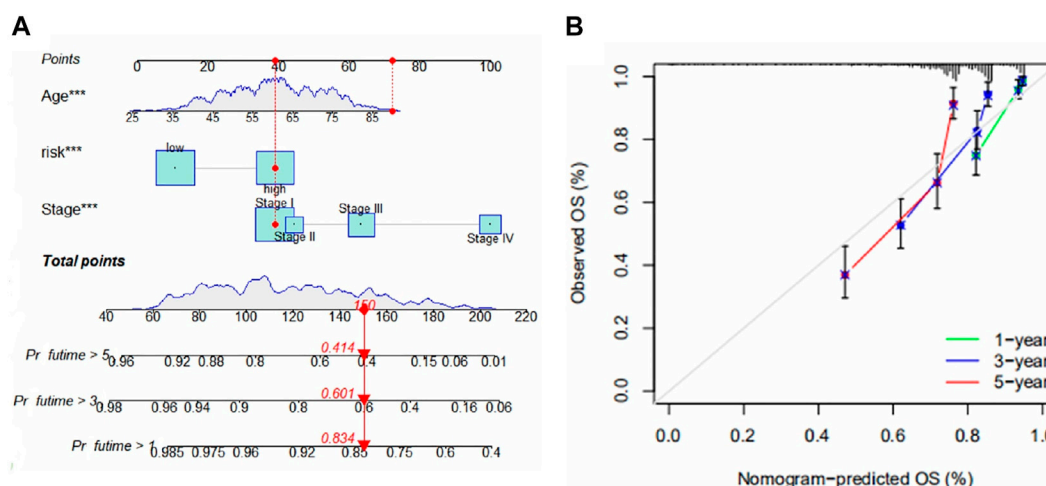


FIGURE 4

(A) Nomogram for predicting 1-, 3-, and 5-year survival for ccRCC. The figure shows the predicted results for the 17th patient in the train group. This patient belongs to the high-risk group of patients, age 90, and at stage I. The scores of each index were 40, 70, and 40, total scores of 150, 1-year, 3-year, and 5-year survival rates were 0.834, 0.601 and 0.414. (B) Calibration chart for the evaluation of nomogram accuracy.

showed a good predictive effect of the nomogram (Figure 4B).

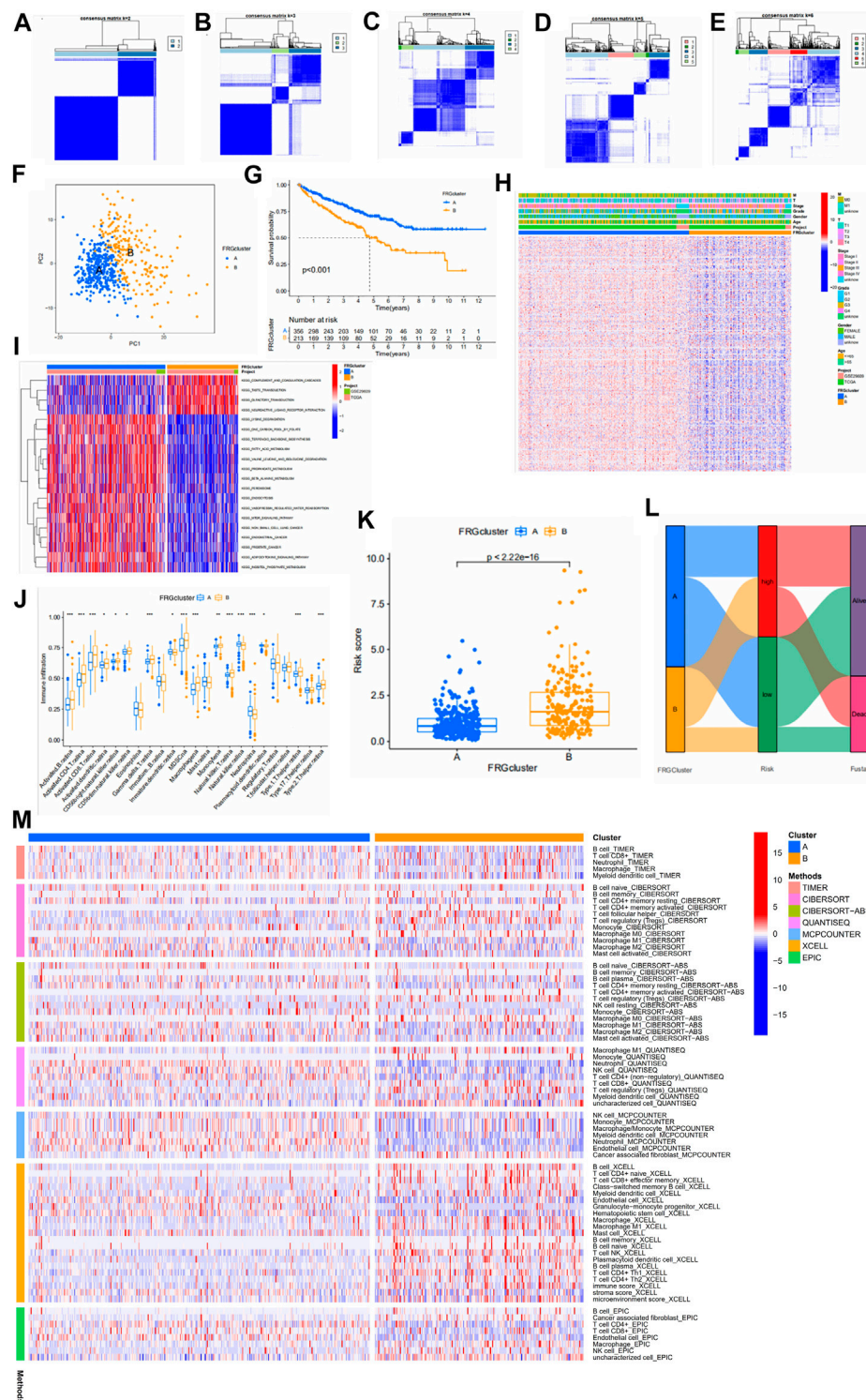
## Identification of fatty acid metabolism subtypes in clear cell renal cell carcinoma

To understand the expression pattern of fatty acid metabolism in ccRCC, all samples were classified into different subtypes using cluster analysis based on the expression of 309 FRGs.  $k = 2$  was appropriate as seen in Figures 5A–E. Therefore, we classified the patients into subtypes A and B. PCA analysis showed a significant difference in FRGs expression between subtypes A and B (Figure 5F). Kaplan-Meier analysis showed higher OS in subtype A than in subtype B (Figure 5G). The comparison of clinical indicators between the two subtypes is shown in Figure 5H, which shows significant differences in FRGs expression and clinical information between the two subtypes. GSEA enrichment analysis showed enrichment in multiple cancer pathways in subtype A, including non-small cell lung cancer, endometrial cancer, and prostate cancer. Tumor signaling pathways were also significantly enriched in subtype A (Figure 5I). The difference in immune cell content between the two subtypes was calculated using the CIBERSORT algorithm. In tcga cohort, we observed that 17 of the 23 immune cells differed significantly between subtypes A and B, and the majority of immune cells were more infiltrated in subtype B than in subtype A (Figure 5J). Some immune cells show a higher degree of infiltration in A than in B. This might be

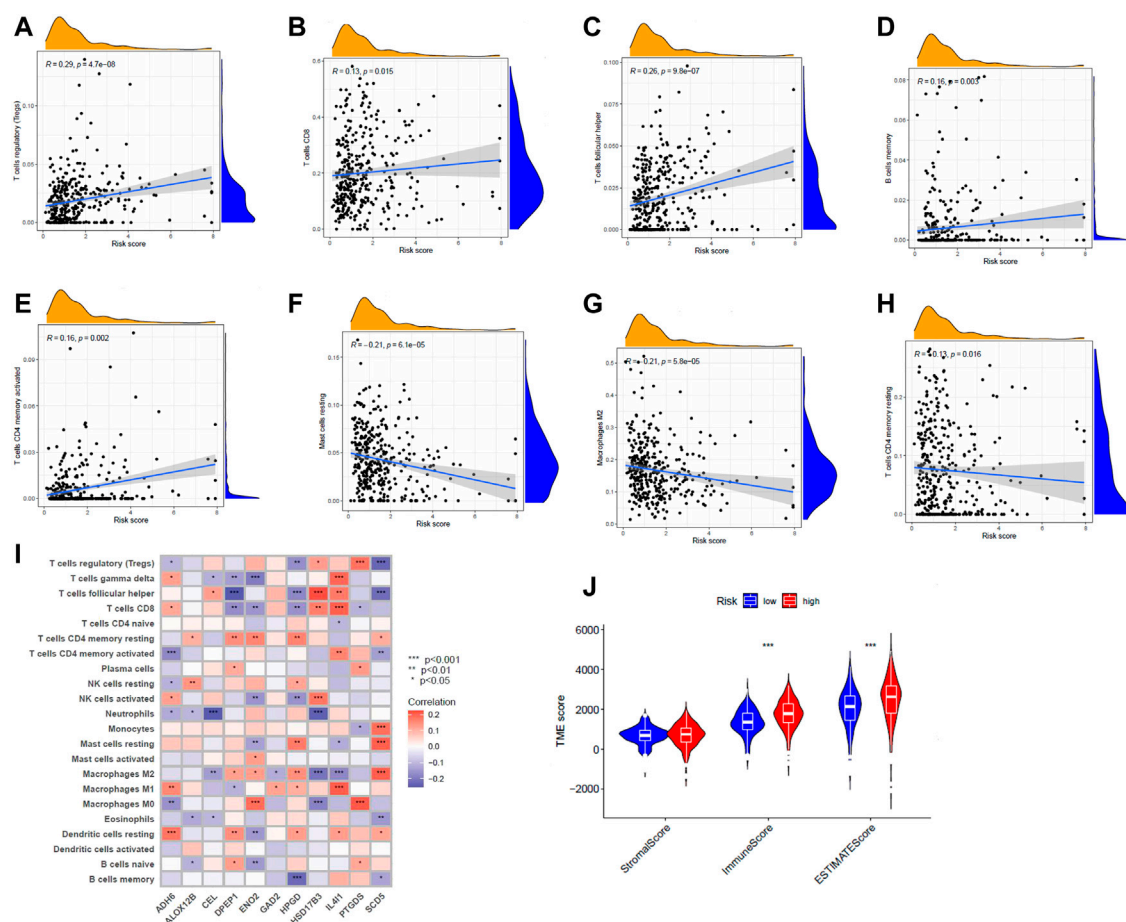
since these immune cells have different roles from other immune cells in both subtypes. In addition, there are significant differences in the risk scores of the different clusters, and it can be seen that the cluster A has a significantly lower risk score than B (Figure 5K). Furthermore, the majority of patients in cluster A belonged to the low risk group and the majority of patients in cluster B belonged to the high-risk group (Figure 5L). Finally, we used multiple platforms to verify the immune infiltration between fatty acid metabolism subtypes. We found an enrichment of immune cells in subtype B, which further suggests a higher degree of immune infiltration in subtype B (Figure 5M).

## Risk scores and tumor microenvironment

The CIBERSORT algorithm was further used to explore the relationship between risk score and immune cells. In tcga cohort, we found that the risk score was positively correlated with Tregs, CD8<sup>+</sup> T cells, T follicular helper cells, memory B cells, and activated memory CD4<sup>+</sup> T cells (Figures 6A–E). Resting mast cells, M2 macrophages, and resting memory CD4<sup>+</sup> T cells were negatively correlated with the risk score (Figures 6F–H). We observed the connection between the 11 genes that constitute the model and immune cell infiltration. These genes could be found to be significantly associated with immune cells (Figure 6I). There were also significant differences in immune scores between risk groups (Figure 6J). In addition, we show the protein expression of model genes in normal and tumor tissues (Supplementary Figure S8). These results were



**FIGURE 5**  
**(A–E)** Results of unsupervised consensus clustering analysis. Defined as two, three, four, five, and six subtypes and their areas. **(F)** The results of PCA analysis between the two subtypes. The figure shows significant differences in fatty acid metabolism-related genes between the two subtypes. **(G)** Survival curves between the two subtypes. **(H)** Comparison between clinical indicators and the two subtypes. **(I)** Differences in biological pathways analyzed by GSVA (gene set variation analysis) in two subtypes. Red for pathways of activation and blue for pathways of inhibition. **(J)** Differences in the level of immune cell infiltration in the two subtypes. **(K)** Association of different subtypes with risk scores. Each point represents a patient. **(L)** Sankey diagram showing the distribution of patients. **(M)** Multiple methods to calculate the immune cell content of two fatty acid metabolic subtypes.



obtained from the HPA (Human Protein Atlas, <https://www.proteinatlas.org/>) database.

## Mutation status and drug sensitivity analysis

The mutation data from TCGA. KIRC.varscan revealed the same top twenty mutated genes between the two groups and a higher number of mutated samples in the low-risk group than in the high-risk group (Figures 7A,B). To observe the differences in drug sensitivity of commonly used chemotherapeutic agents between the different risk groups, using the “pRRophetic” package for drug sensitivity analysis, we observed that patients in the high-risk group were more sensitive to paclitaxel, sunitinib, and rapamycin, while patients in the low-risk group were more sensitive to sorafenib (Figures 7C–F). Subtype B were more sensitive to Paclitaxel, sunitinib, and rapamycin (Supplementary Figure S9).

## Discussion

In the current situation, the mortality rate of ccRCC has still not decreased. As a tumor with metabolic disease nature, the effect of fatty acid metabolism on ccRCC is unknown. To understand the impact of fatty acid metabolic patterns, we distinguished two subtypes with different TME and prognosis by the expression of FRGs and identified the gene signature associated with the prognosis of ccRCC.

Some research tables have demonstrated that the TME has a significant impact on the development of tumors (Hinshaw and Shevde, 2019). The immune cell component of the TME contains lymphocytes, granulocytes, and macrophages. These immune cells play different roles in various immune responses that promote or inhibit tumor survival (Seager et al., 2017). Previous studies have found that CD4<sup>+</sup> T cells have a positive impact on tumor immunity (Saito et al., 2016). Macrophages, on the other hand, have a complex role (Dehne et al., 2017). Among them, M1 macrophages have an

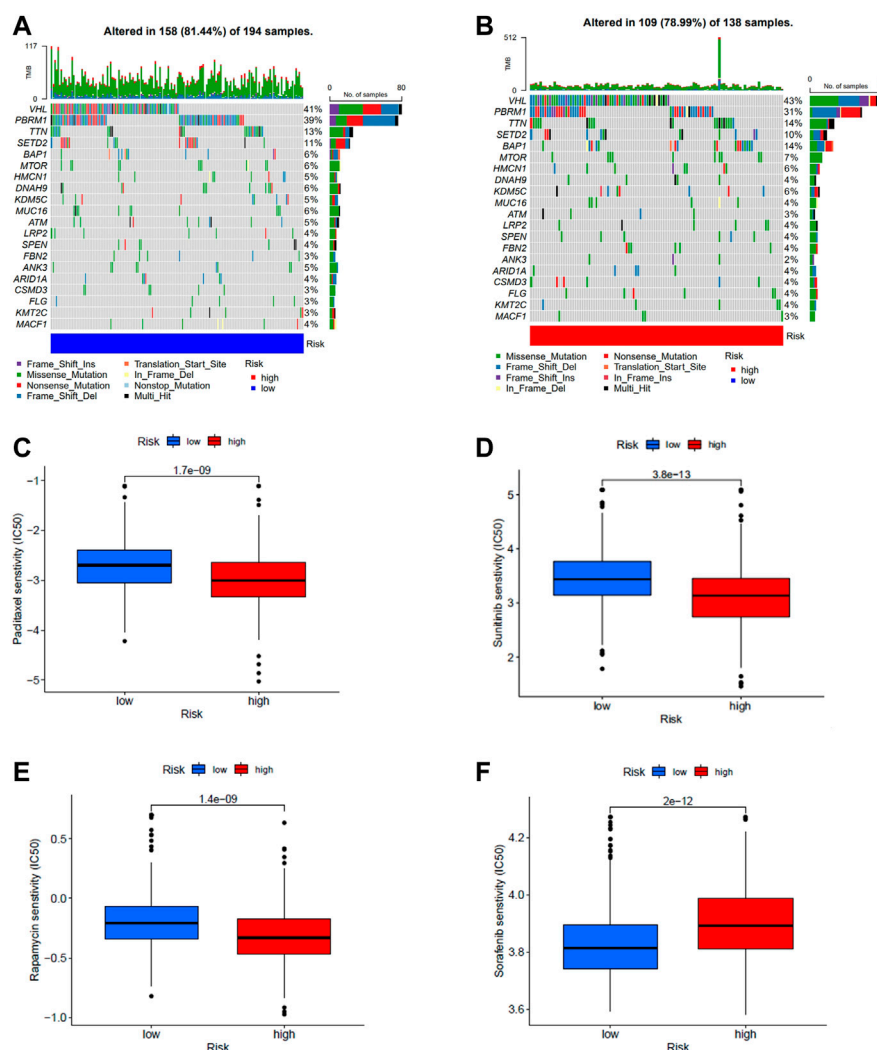


FIGURE 7

(A,B) Waterfall plots of somatic mutation characteristics were created for high and low risk groups. Each column represents a patient. (C–F) Differences in drug sensitivity of paclitaxel, sunitinib, rapamycin, and sorafenib between high- and low-risk groups.

anti-tumor immune role and M2 suppress tumor immunity and promote tumor growth (Sica et al., 2006; Jeannin et al., 2018; Chen et al., 2019). Fatty acids influence the function of immune cells in TME (Tanaka and Sakaguchi, 2017), so it is worth exploring the differences in the degree of immune cell infiltration by different subtypes in our analysis of fatty acid metabolism genes. Some immune cells with antitumor properties can be seen to be enriched in subtype B, yet the Kaplan-Meier analysis of subtype B demonstrated a poor prognosis. We speculate that the effect of fatty acids on immune cell function produced these results. Tanaka and Sakaguchi (2017) found that Tregs infiltration was associated with a poor prognosis, and the amount of Tregs in our analysis was positively correlated with the risk score, which implies that the higher the infiltration of Tregs, the higher the risk, which is consistent with previous reports. We found from the analysis

that different subtypes represent different prognoses. Thus, subtype A can try to predict a good prognosis for ccRCC patients. Likewise, subtype B can predict a poor prognosis in patients. The different levels of immune cell infiltration in the two subtypes may also cause patients to show different outcomes when receiving immunotherapy.

We constructed a prognostic model based on fatty acid metabolism-related genes, which included 11 fatty acid metabolism genes (HSD17B3, HPGD, CEL, PTGDS, SCD5, DPEP1, GAD2, ADH6, ALOX12B, IL4I1, ENO2). Some of these 11 fatty acid metabolism genes have been previously reported, and the results of Song et al.'s analysis of selective splicing signals of more than 10,000 genes in ccRCC demonstrated that SCD5 is one of the potential markers of tumor prognosis (Song et al., 2019). Several reports assessed the expression levels of DPEP1 in various tumor

cells and revealed opposite patterns depending on the tumor type. For example, DPEP1 expression deficiency was associated with breast cancer and Wilms's tumor (Austruy et al., 1993; Green et al., 2009). Antwi et al. (2018) confirmed that mutations in ADH6 are closely associated with the risk of renal cell carcinoma. ALOX12B has been reported to be associated with a variety of cancers, including renal cell carcinoma, lung cancer, breast cancer, and vulvar epidermoid carcinoma (Agarwal et al., 2009; Lee et al., 2009; Shen et al., 2009; Rooney et al., 2015). ALOX12B inhibits immune cytolytic activity in breast and renal cell carcinoma (Rooney et al., 2015), and inhibition of ALOX12B reduces the proliferation of vulvar epidermoid carcinoma cells (Agarwal et al., 2009). IL4I1 is a fatty acid metabolism-related immune checkpoint that activates AHR and accelerates tumor growth (Sadik et al., 2020). The above genes have already been reported and the currently unreported genes could provide leads for further related studies. In addition, compared to previous work, Chen et al. (2021) performed a series of analyses based on metabolic genes in ccRCC, but metabolic alterations encompass multiple pathways. We performed studies in individual pathways (fatty acid metabolic pathway). Wei et al. (2022) constructed a prognostic model based on fatty acid metabolic genes, and the prognostic model we constructed and the identified fatty acid metabolic subtypes can be complementary to it.

The present research has some limitations and a large sample of ccRCC patients is needed to validate the accuracy of the model and the stability of the stratification. Furthermore, *in vivo* or *in vitro* experimental validation would better and more fully confirm the results of our analysis. In addition, treated patients may affect the expression of FRGs, which may have some impact on the results of our analysis.

## Conclusion

In summary, we obtained two subtypes of ccRCC associated with fatty acid metabolism. The two subtypes represent two different prognoses and have different immune infiltration outcomes. We developed 11 FRGs as prognostic models for the signature and established a nomogram to demonstrate the specific prognosis more precisely. These have implications for the

individualized treatment, prognosis, and tumor microenvironment of ccRCC patients.

## Data availability statement

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

## Author contributions

SN conceived and drafted the manuscript. YH, AY, and JL collected data and reviewed relevant information. YW, LW, and LZ performed statistical analysis. SK and FC revised the final manuscript. All authors read and approved the final manuscript.

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1013178/full#supplementary-material>

## References

- Agarwal, S., Achari, C., Praveen, D., Roy, K. R., Reddy, G. V., and Reddanna, P. (2009). Inhibition of 12-LOX and COX-2 reduces the proliferation of human epidermoid carcinoma cells (A431) by modulating the ERK and PI3K-Akt signalling pathways. *Exp. Dermatol.* 18 (11), 939–946. doi:10.1111/j.1600-0625.2009.00874.x
- Akhtar, M., Al-Bozom, I. A., and Al Hussain, T. (2018). Molecular and metabolic basis of clear cell carcinoma of the kidney. *Adv. Anat. Pathol.* 25 (3), 189–196. doi:10.1097/PAP.0000000000000185
- Antwi, S. O., Eckel-Passow, J. E., Diehl, N. D., Serie, D. J., Custer, K. M., Wu, K. J., et al. (2018). Alcohol consumption, variability in alcohol dehydrogenase genes and risk of renal cell carcinoma. *Int. J. Cancer* 142 (4), 747–756. doi:10.1002/ijc.31103
- Austruy, E., Cohen-Salmon, M., Antignac, C., Beroud, C., Henry, I., Nguyen, V. C., et al. (1993). Isolation of kidney complementary DNAs down-expressed in Wilms' tumor by a subtractive hybridization approach. *Cancer Res.* 53 (12), 2888–2894.
- Chen, Y., Liang, Y., Chen, Y., Ouyang, S., Liu, K., and Yin, W. (2021). Identification of prognostic metabolism-related genes in clear cell renal cell carcinoma. *J. Oncol.* 2021, 2042114. doi:10.1155/2021/2042114
- Chen, Y., Song, Y., Du, W., Gong, L., Chang, H., and Zou, Z. (2019). Tumor-associated macrophages: an accomplice in solid tumor progression. *J. Biomed. Sci.* 26 (1), 78. doi:10.1186/s12929-019-0568-z

- Cheng, C., Geng, F., Cheng, X., and Guo, D. (2018). Lipid metabolism reprogramming and its potential targets in cancer. *Cancer Commun.* 38 (1), 27. doi:10.1186/s40880-018-0301-4
- Corn, K. C., Windham, M. A., and Rafat, M. (2020). Lipids in the tumor microenvironment: From cancer progression to treatment. *Prog. Lipid Res.* 80, 101055. doi:10.1016/j.plipres.2020.101055
- Crunkhorn, S. (2019). Targeting cancer cell metabolism in glioblastoma. *Nat. Rev. Cancer* 19 (5), 250. doi:10.1038/s41568-019-0139-3
- Dehne, N., Mora, J., Namgaladze, D., Weigert, A., and Brune, B. (2017). Cancer cell and macrophage cross-talk in the tumor microenvironment. *Curr. Opin. Pharmacol.* 35, 12–19. doi:10.1016/j.coph.2017.04.007
- Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136 (5), E359–E386. doi:10.1002/ijc.29210
- Green, A. R., Krivinskas, S., Young, P., Rakha, E. A., Paish, E. C., Powe, D. G., et al. (2009). Loss of expression of chromosome 16q genes DPEP1 and CTCF in lobular carcinoma *in situ* of the breast. *Breast Cancer Res. Treat.* 113 (1), 59–66. doi:10.1007/s10549-008-9905-8
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013
- He, D., Cai, L., Huang, W., Weng, Q., Lin, X., You, M., et al. (2021). Prognostic value of fatty acid metabolism-related genes in patients with hepatocellular carcinoma. *Aging (Albany NY)* 13 (13), 17847–17863. doi:10.18632/aging.203288
- Hinshaw, D. C., and Shevde, L. A. (2019). The tumor microenvironment innately modulates cancer progression. *Cancer Res.* 79 (18), 4557–4566. doi:10.1158/0008-5472.CAN-18-3962
- Hsieh, J. J., Purdue, M. P., Signoretti, S., Swanton, C., Albiges, L., Schmidinger, M., et al. (2017). Renal cell carcinoma. *Nat. Rev. Dis. Prim.* 3, 17009. doi:10.1038/nrdp.2017.9
- Jeannin, P., Paolini, L., Adam, C., and Delneste, Y. (2018). The roles of CSFs on the functional polarization of tumor-associated macrophages. *FEBS J.* 285 (4), 680–699. doi:10.1111/febs.14343
- Lee, J. Y., Park, A. K., Lee, K. M., Park, S. K., Han, S., Han, W., et al. (2009). Candidate gene approach evaluates association between innate immunity genes and breast cancer risk in Korean women. *Carcinogenesis* 30 (9), 1528–1531. doi:10.1093/carcin/bgp084
- Li, J., Li, Q., Su, Z., Sun, Q., Zhao, Y., Feng, T., et al. (2020). Lipid metabolism gene-wide profile and survival signature of lung adenocarcinoma. *Lipids Health Dis.* 19 (1), 222. doi:10.1186/s12944-020-01390-9
- Li, Z., and Zhang, H. (2016). Reprogramming of glucose, fatty acid and amino acid metabolism for cancer progression. *Cell. Mol. Life Sci.* 73 (2), 377–392. doi:10.1007/s00018-015-2070-4
- Maan, M., Peters, J. M., Dutta, M., and Patterson, A. D. (2018). Lipid metabolism and lipophagy in cancer. *Biochem. Biophys. Res. Commun.* 504 (3), 582–589. doi:10.1016/j.bbrc.2018.02.097
- Massari, F., Ciccarese, C., Santoni, M., Brunelli, M., Piva, F., Modena, A., et al. (2015). Metabolic alterations in renal cell carcinoma. *Cancer Treat. Rev.* 41 (9), 767–776. doi:10.1016/j.ctrv.2015.07.002
- Qi, Y., Chen, D., Lu, Q., Yao, Y., and Ji, C. (2019). Bioinformatic profiling identifies a fatty acid metabolism-related gene risk signature for malignancy, prognosis, and immune phenotype of glioma. *Dis. Markers* 2019, 3917040. doi:10.1155/2019/3917040
- Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160 (1–2), 48–61. doi:10.1016/j.cell.2014.12.033
- Rosario, S. R., Long, M. D., Affronti, H. C., Rowsam, A. M., Eng, K. H., and Smiraglia, D. J. (2018). Pan-cancer analysis of transcriptional metabolic dysregulation using the Cancer Genome Atlas. *Nat. Commun.* 9 (1), 5330. doi:10.1038/s41467-018-07232-8
- Sadik, A., Somarrivas Patterson, L. F., Ozturk, S., Mohapatra, S. R., Panitz, V., Secker, P. F., et al. (2020). IL4I1 is a metabolic immune checkpoint that activates the AHR and promotes tumor progression. *Cell* 182 (5), 1252–1270. doi:10.1016/j.cell.2020.07.038
- Saito, T., Nishikawa, H., Wada, H., Nagano, Y., Sugiyama, D., Atarashi, K., et al. (2016). Two FOXP3(+)CD4(+) T cell subpopulations distinctly control the prognosis of colorectal cancers. *Nat. Med.* 22 (6), 679–684. doi:10.1038/nm.4086
- Sajani, K., Islam, F., Smith, R. A., Gopalan, V., and Lam, A. K. (2017). Genetic alterations in Krebs cycle and its impact on cancer pathogenesis. *Biochimie* 135, 164–172. doi:10.1016/j.biochi.2017.02.008
- Seager, R. J., Hajal, C., Spill, F., Kamm, R. D., and Zaman, M. H. (2017). Dynamic interplay between tumour, stroma and immune system can drive or prevent tumour progression. *Converg. Sci. Phys. Oncol.* 3, 034002. doi:10.1088/2057-1739/aa7e86
- Shen, C., and Kaelin, W. G., Jr. (2013). The VHL/HIF axis in clear cell renal carcinoma. *Semin. Cancer Biol.* 23 (1), 18–25. doi:10.1016/j.semcancer.2012.06.001
- Shen, M., Vermeulen, R., Rajaraman, P., Menashe, I., He, X., Chapman, R. S., et al. (2009). Polymorphisms in innate immunity genes and lung cancer risk in Xuanwei, China. *Environ. Mol. Mutagen.* 50 (4), 285–290. doi:10.1002/em.20452
- Shuch, B., Amin, A., Armstrong, A. J., Eble, J. N., Ficarra, V., Lopez-Beltran, A., et al. (2015). Understanding pathologic variants of renal cell carcinoma: distilling therapeutic opportunities from biologic complexity. *Eur. Urol.* 67 (1), 85–97. doi:10.1016/j.eururo.2014.04.029
- Sica, A., Schioppa, T., Mantovani, A., and Allavena, P. (2006). Tumour-associated macrophages are a distinct M2 polarised population promoting tumour progression: potential targets of anti-cancer therapy. *Eur. J. Cancer* 42 (6), 717–727. doi:10.1016/j.ejca.2006.01.003
- Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer statistics, 2018. *Ca. Cancer J. Clin.* 68 (1), 7–30. doi:10.3322/caac.21442
- Song, J., Liu, Y. D., Su, J., Yuan, D., Sun, F., and Zhu, J. (2019). Systematic analysis of alternative splicing signature unveils prognostic predictor for kidney renal clear cell carcinoma. *J. Cell. Physiol.* 234 (12), 22753–22764. doi:10.1002/jcp.28840
- Tanaka, A., and Sakaguchi, S. (2017). Regulatory T cells in cancer immunotherapy. *Cell Res.* 27 (1), 109–118. doi:10.1038/cr.2016.151
- Veglia, F., Tyurin, V. A., Blasi, M., De Leo, A., Kossenkova, A. V., Donthireddy, L., et al. (2019). Fatty acid transport protein 2 reprograms neutrophils in cancer. *Nature* 569 (7754), 73–78. doi:10.1038/s41586-019-1118-2
- Wei, Z., Cheng, G., Ye, Y., Le, C., Miao, Q., Chen, J., et al. (2022). A fatty acid metabolism signature associated with clinical therapy in clear cell renal cell carcinoma. *Front. Genet.* 13, 894736. doi:10.3389/fgene.2022.894736
- Wettersten, H. I., Aboud, O. A., Lara, P. N., Jr., and Weiss, R. H. (2017). Metabolic reprogramming in clear cell renal cell carcinoma. *Nat. Rev. Nephrol.* 13 (7), 410–419. doi:10.1038/nrneph.2017.59
- Wise, D. R., and Thompson, C. B. (2010). Glutamine addiction: a new therapeutic target in cancer. *Trends biochem. Sci.* 35 (8), 427–433. doi:10.1016/j.tibs.2010.05.003
- Zhao, Z., Liu, Y., Liu, Q., Wu, F., Liu, X., Qu, H., et al. (2019). The mRNA expression signature and prognostic analysis of multiple fatty acid metabolic enzymes in clear cell renal cell carcinoma. *J. Cancer* 10 (26), 6599–6607. doi:10.7150/jca.33024



## OPEN ACCESS

## EDITED BY

Xiaofan Lu,  
China Pharmaceutical University, China

## REVIEWED BY

Jiang Yan,  
First Affiliated Hospital of Shantou  
University Medical College, China  
Shuai Shao,  
The Ohio State University, United States

## \*CORRESPONDENCE

Can Zhao,  
zhaocan921218@163.com  
Li Cheng,  
chengli3@mail.sysu.edu.cn

<sup>†</sup>These authors have contributed equally  
to this work

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 07 August 2022

ACCEPTED 30 August 2022

PUBLISHED 19 September 2022

## CITATION

Zhao T, Liu B, Zhang M, Li S, Zhao C and  
Cheng L (2022), Assessment of  
alterations in histone modification  
function and guidance for death risk  
prediction in cervical cancer patients.  
*Front. Genet.* 13:1013571.  
doi: 10.3389/fgene.2022.1013571

## COPYRIGHT

© 2022 Zhao, Liu, Zhang, Li, Zhao and  
Cheng. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License  
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Assessment of alterations in histone modification function and guidance for death risk prediction in cervical cancer patients

Tingting Zhao<sup>1†</sup>, Bairong Liu<sup>1†</sup>, Mengyuan Zhang<sup>2</sup>, Shiguo Li<sup>3</sup>,  
Can Zhao<sup>1\*</sup> and Li Cheng<sup>1\*</sup>

<sup>1</sup>Department of Obstetrics and Gynecology, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, China, <sup>2</sup>Information Department, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, China, <sup>3</sup>Medical Administration Division, The Seventh Affiliated Hospital, Sun Yat-sen University, Shenzhen, Guangdong, China

**Background:** Cervical cancer is the second most lethal malignancy among women, and histone modification plays a fundamental role in most biological processes, but the prognostic value of histone modification in cervical cancer has not been evaluated.

**Methods:** A total of 594 cervical cancer patients from TCGA-CESC, GSE44001, and GSE52903 cohorts were enrolled in the current study, along with the corresponding clinicopathological features. Patients with a follow-up time less than one month were removed. A total of 122 histone modification-associated signaling pathways were obtained from the MSigDB. The activation scores of these pathways were evaluated using the “GSVA” package, differentially expressed genes were identified by the “limma” package, and pathway enrichment was conducted using the “clusterProfiler 4.0” package. The subsequent least absolute shrinkage and selection operator (LASSO) regression analysis was performed using the “glmnet” package, and a prognostic nomogram was established using the “regplot” package. For the prediction of potential therapeutic drugs, we used the data from GDSC2016 and visualized them via “MOVICS”.

**Results:** Nine of 23 histone modification-associated prognostic genes were identified to construct the prognostic signature by LASSO analysis, named the histone modification-associated gene (HMAG) signature. Cervical patients with HMAG-H in TCGA-CESC cohort showed a 2.68-fold change of death risk, with the 95% CI from 1.533 to 4.671 ( $p < 0.001$ ), as well as the increased death risk of HMAG-H in the GSE44001 cohort (HR: 2.83, 95% CI: 1.370–5.849,  $p = 0.005$ ) and GSE52903 cohort (HR: 4.59, 95% CI: 1.658–12.697,  $p = 0.003$ ). We observed the preferable AUC values of the HMAG signature in TCGA-CESC cohort (1-year: 0.719, 3-year: 0.741, and 5-year: 0.731) and GSE44001 cohort (1-year: 0.850, 3-year: 0.781, and 5-year: 0.755). The C-index of the nomogram showed a prognostic value as high as 0.890, while the C-index for age was only 0.562, and that for grade was only 0.542. Patients with high HMAG scores were more suitable for the treatment of CHIR-99021, embelin, FTI-277, JNK-9L,

JQ12, midostaurin, PF-562271, pyrimethamine, and thapsigargin, and patients with low HMAG scores were more suitable for the treatment of BMS-536924, CP466722, crizotinib, PHA-665752, rapamycin, and TAE684.

**Conclusion:** We comprehensively evaluated the histone modification status in cervical cancer patients and revealed histone modification-associated prognostic genes to construct the HMAG signature, aiming to provide a new insight into prognosis prediction and precise clinical treatment.

#### KEYWORDS

cervical cancer, histone modification, prognosis, signature, prediction

## Introduction

Currently, cervical carcinoma has become the second most lethal malignancy among women worldwide (Bray et al., 2018), with 527,624 new cases and 265,672 tumor-specific deaths annually (Shrestha et al., 2018). For the accurate prediction of prognosis, the International Federation of Gynecology and Obstetrics put forward the cervical cancer staging standard in 2018, according to the depth, greatest dimension of stromal invasion, and the extension of tumor on adjacent regions (such as vagina and pelvis), separating cervical cancer into stages I, II, III, and IV and further substages (Balcacer et al., 2019; Salvo et al., 2020). For different stages of the tumor, the treatments are diverse. Surgical interventions, including trachelectomy and radical hysterectomy, are performed in most early cervical cancers. In addition, radiation and chemoradiation are applied in most advanced or metastatic cervical cancers (Johnson et al., 2019). Unfortunately, these treatments still result in a low response rate and poor prognosis. Therefore, it is essential to develop new predictive prognostic models for optimizing treatment strategies.

Histone proteins, a type of abundant cellular protein, are surrounded by DNA to make nucleosomes (Taylor et al., 2020). The N-terminal tail of each histone protein is the site of posttranslational modifications (PTMs) and can make contact with adjacent nucleosomes (Bannister and Kouzarides, 2011; Bartke et al., 2013; Taylor et al., 2020). There are numerous types of histone PTMs, including acetylation, methylation, ubiquitinylation, and phosphorylation (Bannister and Kouzarides, 2011; Taylor et al., 2020; Zhang et al., 2021). They play fundamental roles in most biological processes (Bannister and Kouzarides, 2011; Taylor et al., 2020). Histone methylation functions in many levels of transcriptional regulation from the chromatin architecture to specific locus regulation and RNA processing (Greer and Shi, 2012). Histone acetylation influences myriad cellular and physiological processes, including transcription, phase separation, autophagy, mitosis, differentiation, and neural function (Shvedunova and Akhtar, 2022). Histone ubiquitinylation works in organization of repair to preserve genomic integrity after the breaking of the DNA double strand

(Uckelmann and Sixma, 2017). Histone phosphorylation provides a rapid and reversible physiological response to DNA damage, nutritional stress, or an altered metabolic state (Uckelmann and Sixma, 2017). More importantly, histone modification affects the accessibility of DNA and recruitment of DNA-binding proteins and thereby regulates gene transcription, which controls transcriptional regulation and corresponding disorders. For example, modification of histones by acetylation/deacetylation influences gene expression and is therefore related to the carcinogenic process (Audia and Campbell, 2016) (Fang et al., 2014).

Many molecular mechanisms of histone modification have also been found. A high level of histone acetylation is associated with the expression of proto-oncogenes; nonetheless, a low level of histone acetylation is linked with the silencing of tumor suppressor genes (Armenta-Castro et al., 2020). Moreover, it has been reported that the repressive expression of osteoprotegerin (OPG) genes is associated with histone modification and further causes chromatin architecture alterations and results in transcriptional repression of OPG genes. OPG plays a tumor suppressive role in tumorigenesis; therefore, histone modification-associated OPG suppression can accelerate the deterioration of cervical cancer (Lu et al., 2009). Some studies also suggest that histone acetylation is linked with the expression of the LGALS9 gene. Histone acetylation is an activator modification for the LGALS9 gene, which can be observed in active promoters. This modification can promote LGALS9 gene transcription. Galectin-9, which is encoded by the LGALS9 gene, can have a positive effect on the apoptosis of tumor cells. Histone acetylation of LGALS9 reduces deterioration in cervical cancer (Armenta-Castro et al., 2020). Relevant molecular mechanisms have been studied to some extent. However, the impact of changes in signaling pathways caused by histone modification on the progression of cervical cancer is not clear. Accordingly, we carried out the current study to illuminate the prognostic value and function of histone modification in cervical cancer patients.

In this study, we gathered information on the signaling pathways related to histone modification. According to the activation status of the pathways, we identified cervical cancer patients with activated and suppressed histone modification.

**TABLE 1** Clinicopathological information of the enrolled cohorts.

|                           | TCGA-CESC (N = 248) | GSE44001 (N = 295) | GSE52903 (N = 51) | Overall (N = 594) |
|---------------------------|---------------------|--------------------|-------------------|-------------------|
| <b>Survival time</b>      |                     |                    |                   |                   |
| Mean (SD)                 | 29.0 (35.5)         | 49.3 (24.5)        | 44.0 (26.5)       | 40.4 (31.2)       |
| Median [Min, Max]         | 16.5 [1.02, 195]    | 48.8 [3.50, 104]   | 58.0 [1.00, 86.0] | 36.4 [1.00, 195]  |
| <b>Events<sup>a</sup></b> |                     |                    |                   |                   |
| No                        | 189 (76.2%)         | 258 (87.5%)        | 31 (60.8%)        | 478 (80.5%)       |
| Yes                       | 59 (23.8%)          | 37 (12.5%)         | 20 (39.2%)        | 116 (19.5%)       |
| <b>Age</b>                |                     |                    |                   |                   |
| Mean (SD)                 | 47.9 (13.9)         | -                  | 50.9 (14.4)       | 48.4 (14.0)       |
| Median [Min, Max]         | 46.0 [20.0, 80.0]   | -                  | 50.0 [24.0, 74.0] | 46.0 [20.0, 80.0] |
| Missing                   | 0 (0%)              | 295 (100%)         | 0 (0%)            | 295 (49.7%)       |
| <b>Stage</b>              |                     |                    |                   |                   |
| Stage I                   | 134 (54.0%)         | 254 (86.1%)        | 24 (47.1%)        | 412 (69.4%)       |
| Stage II                  | 56 (22.6%)          | 41 (13.9%)         | 8 (15.7%)         | 105 (17.7%)       |
| Stage III                 | 34 (13.7%)          | 0 (0%)             | 15 (29.4%)        | 49 (8.2%)         |
| Stage IV                  | 19 (7.7%)           | 0 (0%)             | 4 (7.8%)          | 23 (3.9%)         |
| Unknown                   | 5 (2.0%)            | 0 (0%)             | 0 (0%)            | 5 (0.8%)          |
| <b>Grade</b>              |                     |                    |                   |                   |
| Unknown                   | 26 (10.5%)          | 295 (100%)         | 51 (100%)         | 372 (62.6%)       |
| G1                        | 15 (6.0%)           | 0 (0%)             | 0 (0%)            | 15 (2.5%)         |
| G2                        | 111 (44.8%)         | 0 (0%)             | 0 (0%)            | 111 (18.7%)       |
| G3                        | 95 (38.3%)          | 0 (0%)             | 0 (0%)            | 95 (16.0%)        |
| G4                        | 1 (0.4%)            | 0 (0%)             | 0 (0%)            | 1 (0.2%)          |

<sup>a</sup>The clinical outcome recorded in TCGA-CESC and GSE52903 cohorts is overall survival (OS), and in the GSE44001 cohort, it is disease-free survival (DFS).

Then, we sought differentially expressed genes (DEGs) to reflect the inner alteration of the tumor caused by histone modification. Among these DEGs, we screened for relevant genes that best reflected prognosis and constructed a histone modification-associated gene (HMAG) signature to predict the prognosis of patients with cervical cancer.

## Methods

### Data collection

We obtained the gene expression profile and corresponding clinical information of CESC patients from three independent clinical cohorts: TCGA-CESC, GSE44001, and GSE52903. We downloaded the TCGA-CESC dataset from the GDC TCGA project via the R package “TCGAbiolinks.” The transcripts per million (TPM) format of the gene expression file was chosen and then transformed to the log2 (TPM+1) type to make it comparable with the sequencing results from the microarray. Patient samples in the GSE52903 cohort were collected from Mexico City, which were HPV16-positive and fresh samples with more than 70% tumor tissues. The clinical

outcome recorded in TCGA-CESC and GSE52903 was the overall survival (OS). Samples in the GSE44001 dataset were collected from South Korea, and gene expression values were detected using the GPL14951 Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip platform, along with the prognosis information of disease-free survival (DFS). We further filtered the enrolled patients to exclude those without paired gene expression data and clinical information and those with a follow-up time of less than one month to decrease the potential bias (Table 1).

### Removal of batch effects between cohorts

For the three enrolled cohorts, the potential nonbiological bias was eliminated to make the gene expression profiles of different cohorts more comparable. The “sva” R package was applied with its ComBat algorithms to remove the batch effects, and the gene expression profiles were all adjusted. Then, GSE44001 and GSE52903 were combined as the GEO-combined cohort. Three cohorts were detached before the removal of the batch effect (SupplementaryFigure S1A) and fused together after removal (SupplementaryFigure S1B).

## Collection of the histone modification pathways

To comprehensively reveal the diverse distribution of histone modification in cervical cancer patients, we collected a total of 122 histone modification-associated signaling pathways from the Molecular Signatures Database (MSigDB)-C5 (Liberzon et al., 2011): ontology gene sets, including the process of histone-mediated phosphorylation, methylation, ubiquitination, and acetylation.

## Gene set variation analysis

The activation of 122 histone modification-associated signaling pathways was assessed by the “GSVA, v.3.5” R package, the enrichment score for a specific gene set in each sample was calculated, and GSVA quantified the total gene set activation results (Hanzelmann et al., 2013). Therefore, the gene expression profiles were transferred to gene set activation profiles, including the activated scores of 122 signaling pathways for each cohort.

## Revealing the differentially expressed genes and pathways

To identify the diverse downstream altered biological processes, we carried out the differentially expressed genes (DEGs) among histone modification-activated and -suppressed subgroups via the “limma” package, with the preset threshold value of  $p < 0.05$  and  $|\log_2fc| > 0.4$ . Pathway enrichment was performed by the “clusterProfiler 4.0” R package (Wu et al., 2021), with the employment of 50 HALLMARK pathways and KEGG pathways. In addition, Metascape (<http://metascape.org>) (Zhou et al., 2019) was also used to annotate the DEGs to reveal the activated pathways.

## Construction and calculation of the HMAG signature

The prognostic value of the aforementioned identified DEGs was further assessed by univariate Cox regression analysis. Gene expression was first separated into high and low groups and then entered into the univariate Cox regression analysis. The prognostic genes in both the TCGA-CESC cohort and GEO-combined cohort were selected for a subsequent least absolute shrinkage and selection operator (LASSO) regression analysis, which was conducted using the “glmnet” package. The selected genes were used to calculate the risk score by adding the gene expression multiplied by the corresponding coefficient and

named the HMAG score. The HMAG signature was trained in the TCGA-CESC cohort and validated in the GSE44001 and GSE52903 cohorts. The HMAG score was calculated for each patient and then applied for the subsequent analysis.

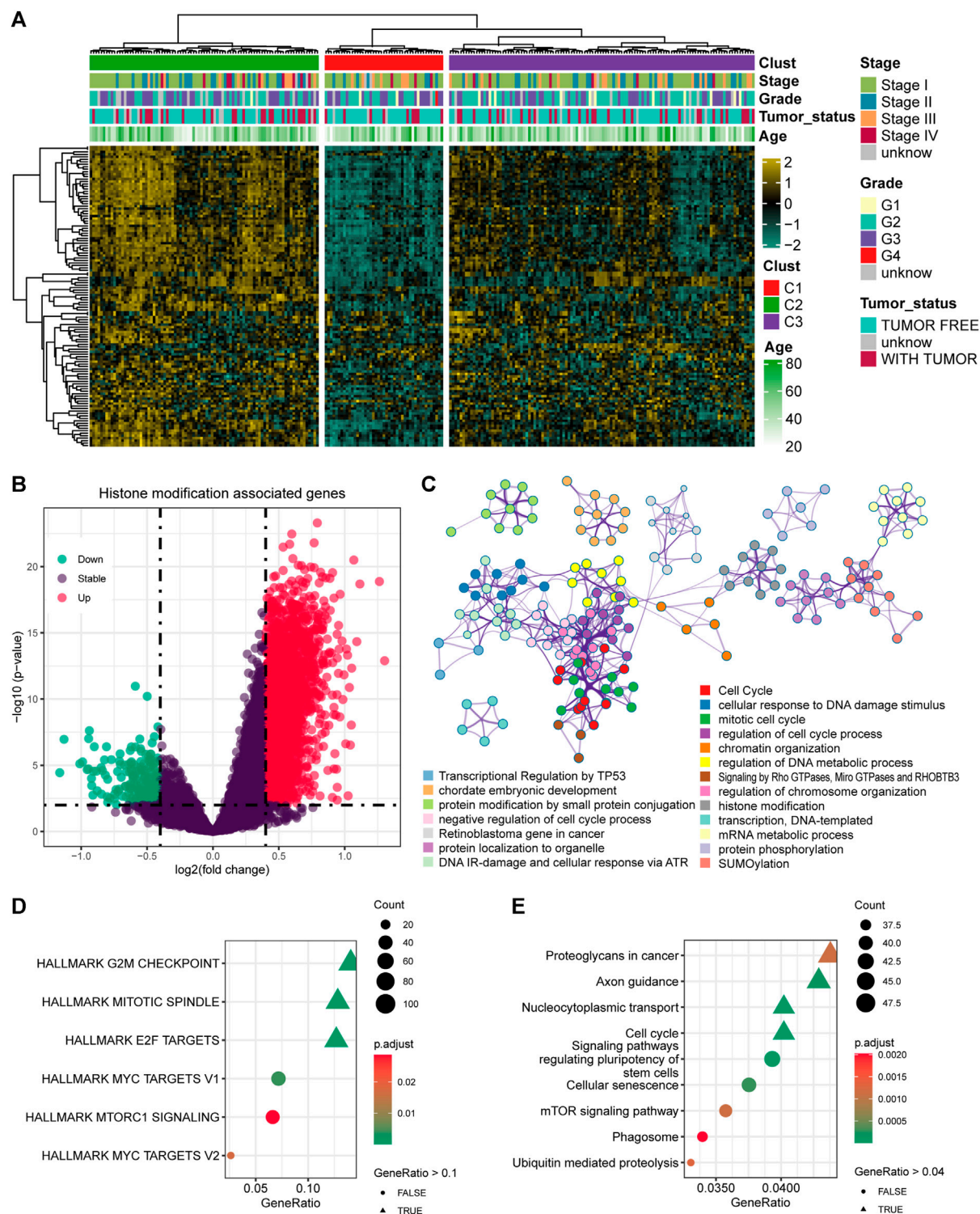
## Multivariate analysis and establishment of the nomogram

The prognostic value of the HMAG score in each cohort was assessed by the K-M plot and receiver operating characteristic (ROC) curve. Multivariate Cox regression analysis was also performed to adjust the potential impact of other clinical features and is presented with a forest plot. Furthermore, we employed the “regplot” package to establish the clinical prognostic nomogram, which could provide a quantitative method for the individualized prediction of cervical cancer. Factors that emerged from the multivariate Cox regression analysis were enrolled for the establishment of the nomogram. C-index curve, calibration curve, and decision curve analyses were all performed to validate the clinical usefulness and accuracy of the nomogram via the “rms” and “rmda” packages.

## Revealing activated pathways and potential therapeutic drugs

We assessed the activated pathways by fast gene set enrichment analysis (fgsea). First, GSEA is performed with the ranking of input molecular readouts, and then the pathway enrichment score is calculated by walking down the list of features, which means that if a feature is felled into the target pathway, the running-sum statistic increases; otherwise, it decreases. The final score is the maximum deviation from zero encountered in the random walk and normalized by computing the z-score of the estimate compared to a null distribution obtained from a random permutation. Moreover, the highly expressed genes in high-HMAG and low-HMAG groups were also annotated to reveal the activated biological process by clusterProfiler 4.0 and visualized via the treeplot model.

For the prediction of potential therapeutic drugs, we used the data from GDSC2016 (<https://www.cancerrxgene.org/>) and visualized them via “MOVICS” (Lu et al., 2020). Ridge regression analysis was performed to predict the potential response result to the chemotherapy drug of each patient and represented by the estimated inhibitory concentration ( $IC_{50}$ ); the lower the  $IC_{50}$  was, the higher the effectiveness of the drug treatment was (Geeleher et al., 2014). Moreover, we also searched for potential therapeutic new drugs through the Gene Set Cancer Analysis (GSCA) online website (<http://bioinfo.life.hust.edu.cn/GSCA>).



**FIGURE 1**  
Identification of the histone modification-altered genes and pathways. **(A)** Heatmap showing the activation status and clinical features in cervical cancer patients. **(B)** Differentially expressed genes (DEGs) between the histone modification-activated and -suppressed subgroups. **(C)** Enriched signaling pathways of the DEGs using Metascape. **(D)** Enriched signaling pathways of the DEGs by HALLMARK pathways. **(E)** Enriched signaling pathways of the DEGs by KEGG pathways.

## Comparison of the HMAG signature with published signatures

To assess the prognostic value of the HMAG signature carried out in the current study, we searched published articles to collect signatures from other studies. Finally, we used four signatures, namely, immune-associated genes (Yu et al., 2021), DNA damage repair-associated genes (Zhou et al., 2022), ferroptosis-related genes (Qi et al., 2021), and autophagy-related genes (Chen et al., 2020). The risk score of each signature was calculated in the TCGA-CESC cohort, and the ROC value and C-index were used to evaluate the prognostic value.

## Statistics

The log-rank test was used to compare the survival outcome between two groups, Student's *t*-test was used to compare the distribution between two groups, and Fisher's exact test was performed to distinguish the difference in categorical data. All statistical analyses were performed using R (Version: 4.1.2). A two-tailed *p*-value < 0.05 was recognized as statistically significant.

## Results

### Characterizing the diversified histone modification-activated status and altered signaling pathways

A total of 122 histone modification-associated pathways were obtained from MSigDB-C5, and the activated score was generated based on the GSVA quantification. Then, the 248 cervical cancer patients from TCGA-CESC cohort were divided into three subtypes based on the distanceMatrix function of the "ClassDiscovery" package, with the preset parameters of "euclidean" and "ward.D." We observed diversified histone modification activation in C1, C2, and C3. The patients in C2 contained the most activated status of histone modification, while the C1 patients contained the suppressed histone modification (Figure 1A). Then, we compared the DEGs between C2 and C1 (Figure 1B), with a preset threshold value of  $p < 0.05$  and  $|\log_2fc| > 0.4$ . A total of 2040 genes were upregulated in C2, and the other 200 genes were upregulated in C1. We combined the DEGs between C1 and C2, annotated the enriched signaling pathways, and observed the altered biological processes of the cell cycle, cellular response to DNA damage stimulus, chromatin organization, and DNA metabolic process (Figure 1C). In addition, we also validated the DEGs in the HALLMARK and KEGG pathways. The DEGs targeted the G2/M checkpoint, mitotic spindle, and E2F targets (Figure 1D) and

were also linked with proteoglycans in cancer, axon guidance, nucleocytoplasmic transport, and cell cycle pathways in KEGG pathways (Figure 1E).

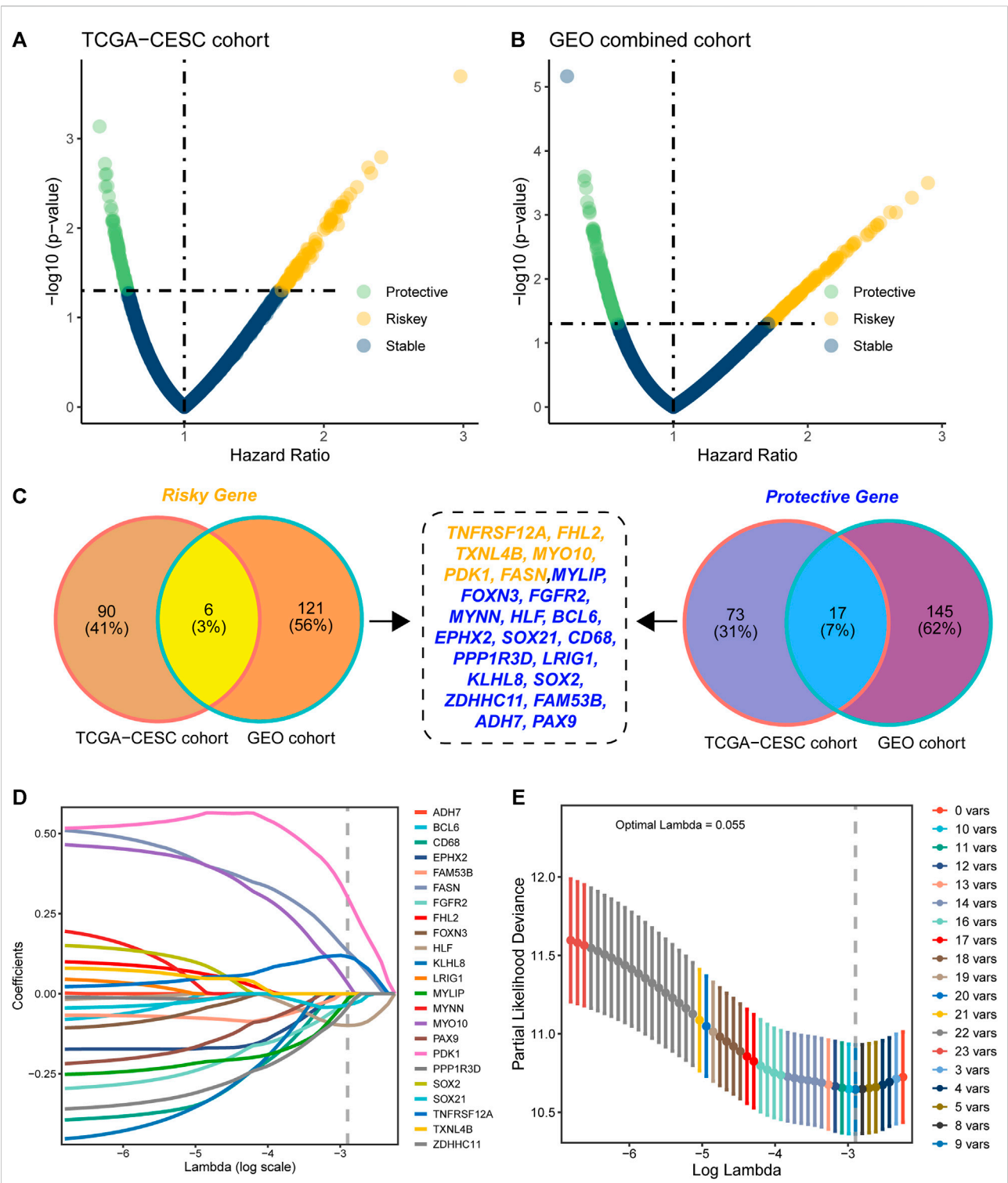
### Distinguishing the prognostic genes and constructing the signature

Based on the 2240 DEGs, we assessed and selected the prognostic DEGs in TCGA-CESC cohort (Figure 2A) and GEO-combined cohort (Figure 2B). Genes with an HR higher than 1 and a *p*-value less than 0.05 were regarded as risky genes, while genes with an HR less than 1 and a *p*-value less than 0.05 were regarded as protective genes. Ninety-six risk genes from TCGA-CESC cohort and 127 risk genes from the GEO-combined cohort merged out six common risk genes; 90 protective genes from TCGA-CESC cohort and 162 protective genes from the GEO-combined cohort merged out 17 common protective genes (Figure 2C). Therefore, 23 genes were finally enrolled for LASSO analysis, and nine genes were identified under the best optimal lambda value of 0.055, leading to the prognostic genes SOC21, HLF, FGFR2, MYLIP, ZDHHC11, FASN, PDK1, MYO10, and TNFRSF12A (Figures 2D,E). The HMAG signature was calculated using the following formula: HMAG score =  $0.115713151 * TNFRSF12A + 0.032286676 * MYO10 + 0.300717617 * PDK1 + 0.131853347 * FASN + -0.015900246 * MYLIP - 0.033923872 * FGFR2 - 0.099329649 * HLF - 0.034241118 * SOX21 - 0.042765486 * ZDHHC11$ .

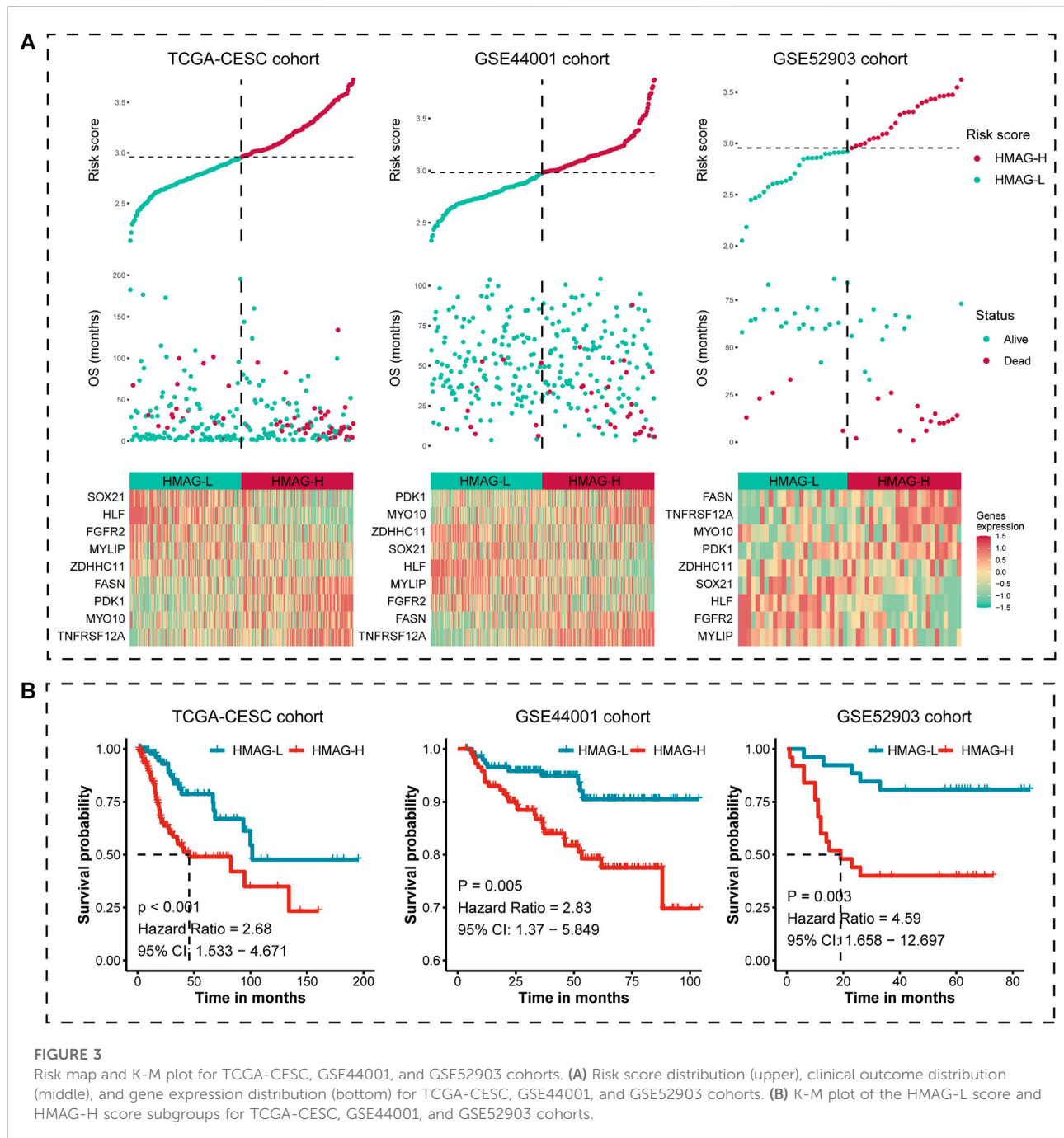
### Prognostic effectiveness of the HMAG signature

We calculated the HMAG score of each patient based on the aforementioned formula. The overall distributions of the risk score, survival status, and gene expression profiles of the nine-gene signature in TCGA-CESC cohort (left), the GSE44001 cohort (middle), and the GSE52903 cohort (right) are shown in Figure 3A. Patients in each cohort were separated into HMAG-L and HMAG-H based on the median HMAG score, and the prognostic value was assessed by Cox regression analysis. Cervical patients with HMAG-H in TCGA-CESC cohort showed a 2.68-fold change of death risk, with the 95% CI from 1.533 to 4.671 ( $p < 0.001$ ), as well as the increased death risk of HMAG-H in the GSE44001 cohort (HR: 2.83, 95% CI: 1.370–5.849,  $p = 0.005$ ) and GSE44001 cohort (HR: 4.59, 95% CI: 1.658–12.697,  $p = 0.003$ ) (Figure 3B).

The prognostic accuracy was evaluated by ROC curves. We observed the preferable AUC values of the HMAG signature in TCGA-CESC cohort (1-year: 0.719, 3-year: 0.741, and 5-year: 0.731) and GSE44001 cohort (1-year: 0.850, 3-year: 0.781, and 5-year: 0.755) (Figure 4A). In addition, we also conducted multivariate Cox regression analysis to adjust for the impact of clinicopathological features. We observed that the HMAG



**FIGURE 2** Construction of the histone modification-associated gene signature. (A) Selection of the prognostic genes in TCGA-CESC cohort; (B) selection of the prognostic genes in the GEO-combined cohort; (C) Venn plot showing the risk genes and protective genes in both TCGA-CESC and GEO-combined cohorts; (D) optimal tuning parameter (lambda) in the LASSO analysis selected with 10-fold cross-validation and one standard error rule; and (E) LASSO coefficient profiles of the 23 candidate genes.

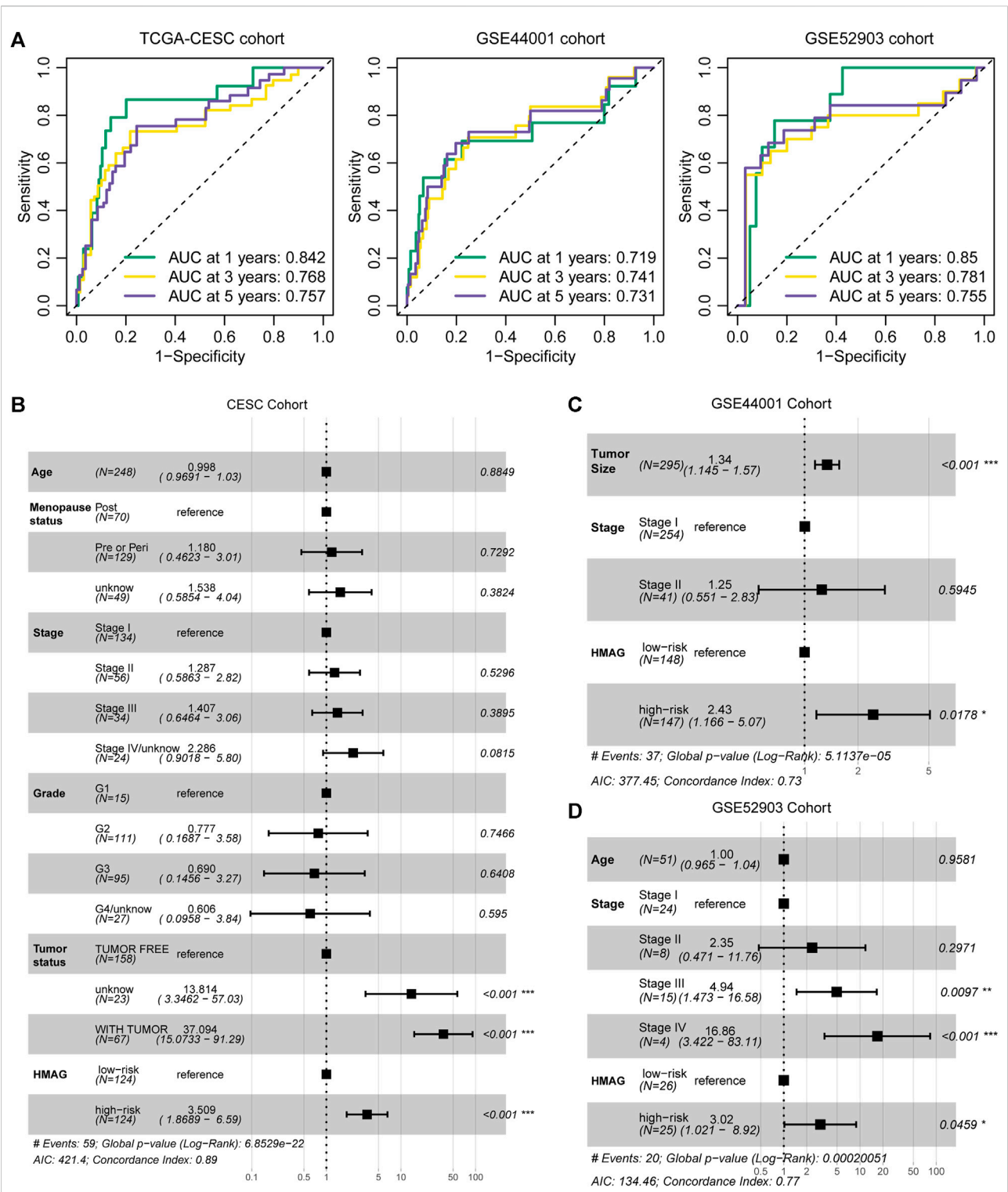


signature (HR: 3.509, 95% CI: 1.869–6.590,  $p < 0.001$ ) and tumor status (tumor vs. tumor free: HR: 37.094, 95% CI: 15.073–91.290,  $p < 0.001$ ) were independent prognostic factors for cervical cancer patients in TCGA-CESC cohort but age, menopausal status, tumor stage, and tumor grade were not (Figure 4B). In the GSE44001 cohort, the HMAG signature also showed an independent risk factor after adjusting for tumor size and tumor stage (HR: 2.43, 95% CI: 1.166–5.070,  $p = 0.0178$ , Figure 4C). Similar results were also shown in the

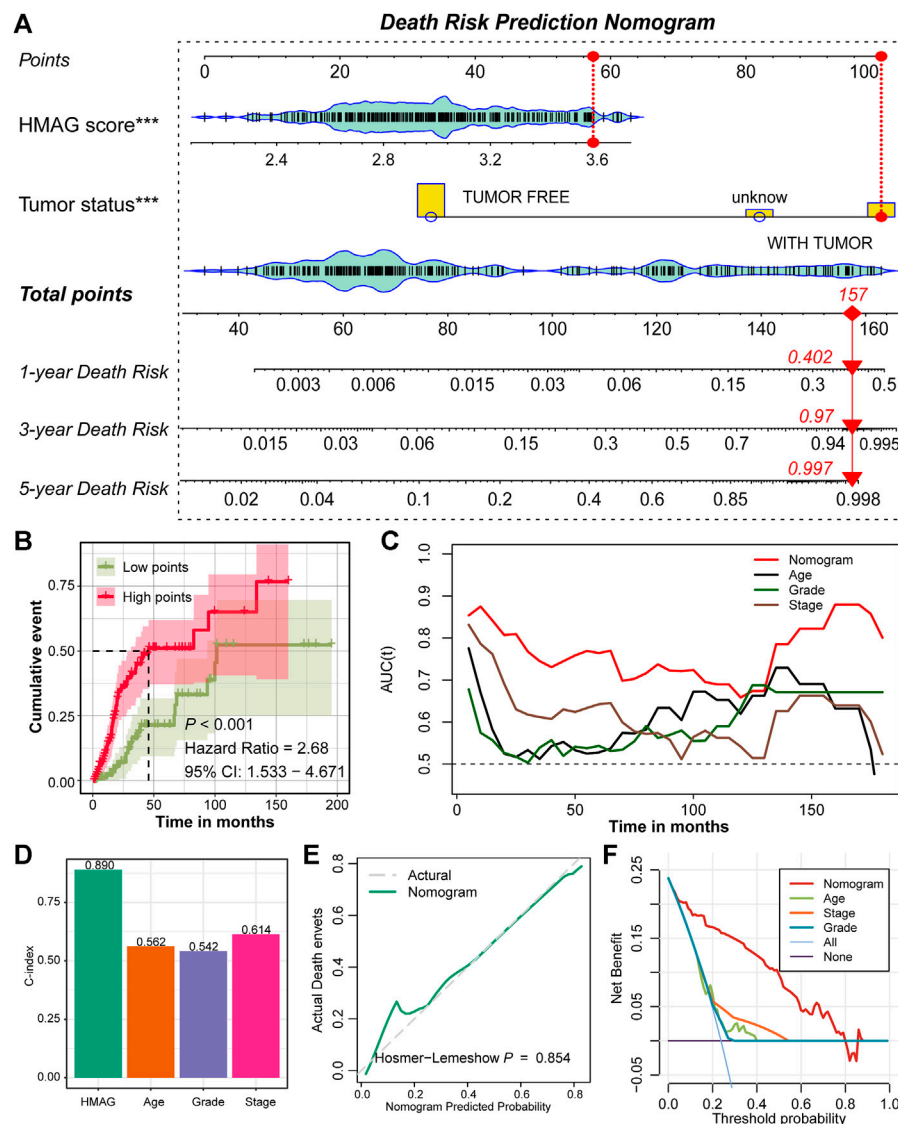
GSE52903 cohort (HR: 3.02, 95% CI: 1.021–8.920,  $p = 0.0459$ , Figure 4D).

## Prognostic nomogram established by the HMAG score and tumor status

With the aforementioned results, we established a prognostic nomogram model containing the factors HMAG



**FIGURE 4** Evaluation of the prognostic value of the HMAG signature. (A) 1-year, 3-year, and 5-year AUC values in TCGA-CESC, GSE44001, and GSE52903 cohorts; (B) forest plot showing the prognostic value of the HMAG signature after adjusting for clinical features in TCGA-CESC cohort; (C) forest plot showing the prognostic value of the HMAG signature after adjusting for clinical features in the GSE44001 cohort; and (D) forest plot showing the prognostic value of the HMAG signature after adjusting for clinical features in the GSE52903 cohort.

**FIGURE 5**

Prognostic nomogram constructed by the HMAG signature and tumor status. (A) Establishment of a nomogram combining the tumor status and HMAG signature; (B) K-M plot showing the diverse cumulative death event in low points and high points subgroups; (C) time ROC lines showing the prognostic value of nomogram, age, grade, and tumor stage; (D) bar plots showing the C-index of nomogram, age, grade, and tumor stage; (E) calibration plot for the nomogram. The dashed line represents the ideal nomogram, the solid line represents our nomogram, and a  $p$ -value of 0.854 indicates that the nomogram-predicted probability is very close to the actual death events; (F) DCA showed that our nomogram had the greatest net benefit compared with the single factor of age, stage, or grade.

score and tumor status (Figure 5A). For a specific patient, the HMAG score and tumor status correspond to the point, and the summary of two points is the total point, with a straight line from the total point site to the bottom line of 1-year, 3-year, and 5-year death risks, indicating the risk of death. With a cumulative incidence plot, we visualized the estimated probability of the death event prior to a specified time. The patients were separated into a low-point subgroup and a high-point subgroup, and we observed significantly diverse

cumulative events in the two groups (HR: 2.68, 95% CI: 1.533–4.671,  $p < 0.001$ , Figure 5B). The AUC values changed over the follow-up time and presented a better prognostic value of the nomogram than regardless of age, grade, or stage (Figure 5C). The overall C-index of the nomogram showed a prognostic value as high as 0.890, while the C-index for age was only 0.562, for grade was only 0.542, and for stage was only 0.614 (Figure 5D). A  $p$ -value of 0.854 calculated by the Hosmer–Lemeshow

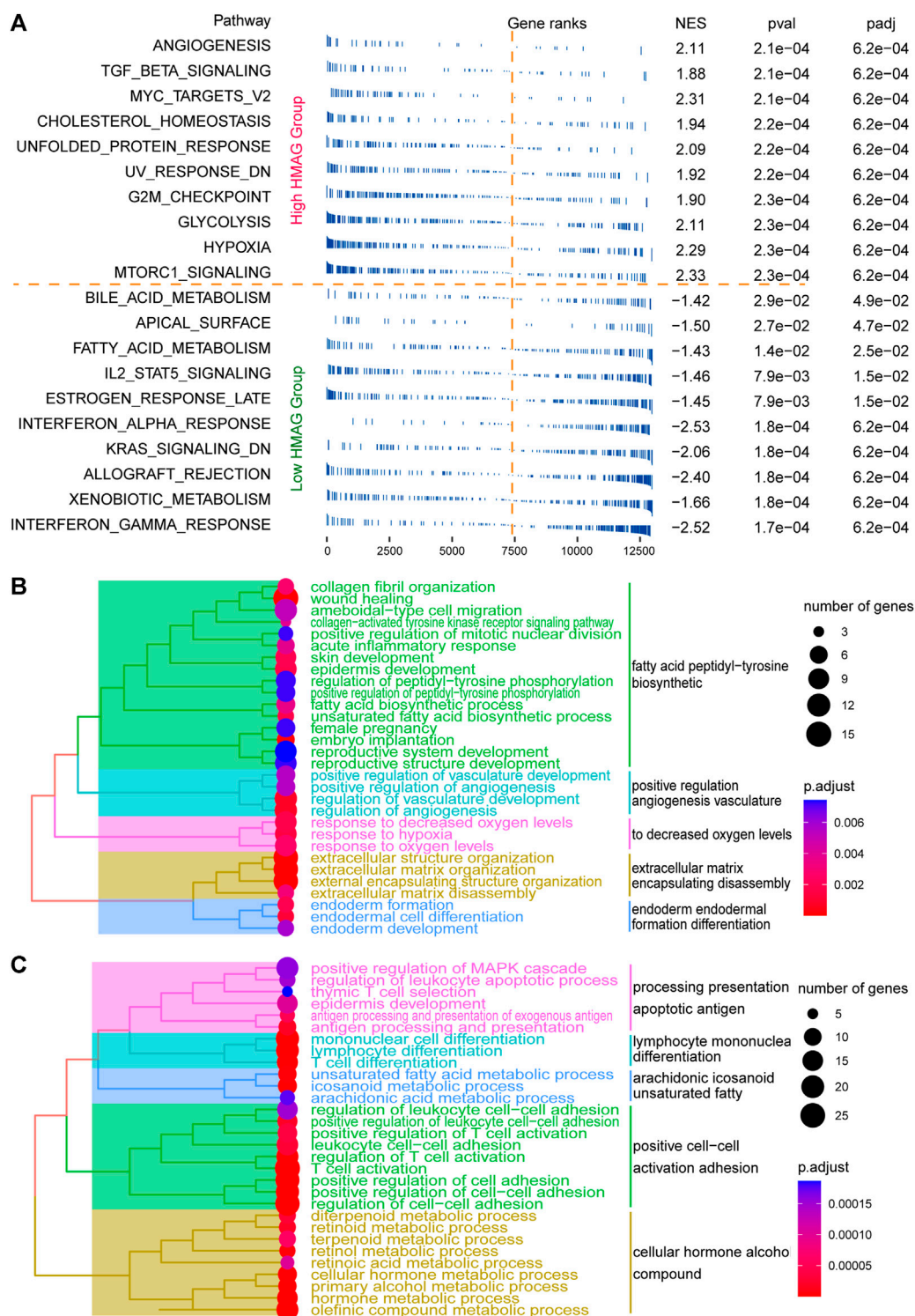
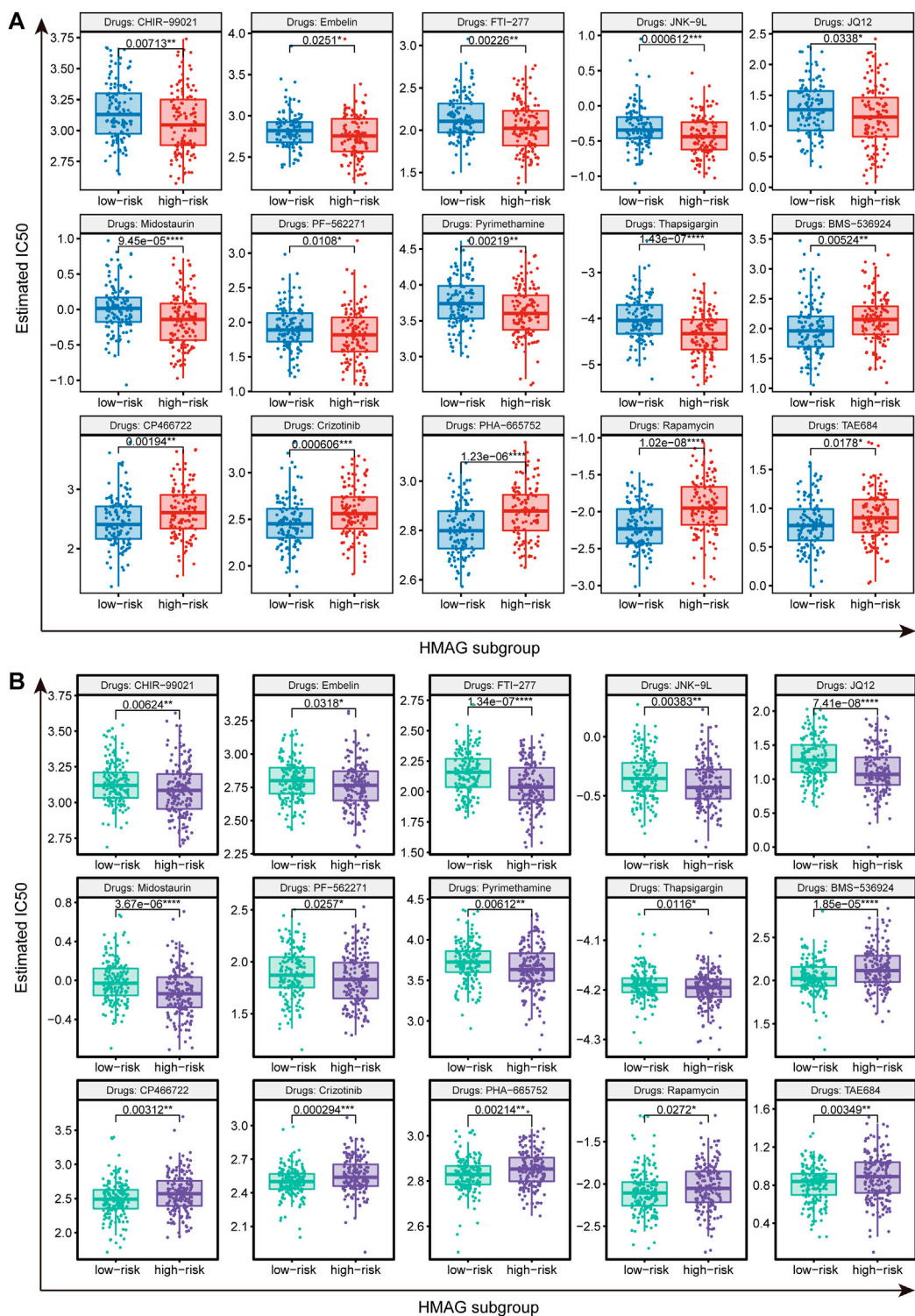


FIGURE 6

Revealing HMAG signature-regulated biological pathways. (A) fgsea showing the diverse activation of tumor biological pathways in the high-HMAG-score group and the low-HMAG-score group; (B) hierarchical clustering of enriched terms in HMAG upregulated genes; and (C) hierarchical clustering of enriched terms in HMAG downregulated genes.



**FIGURE 7**  
Prediction of therapeutic drugs for the HMAG low- and high-risk groups. (A) Potential therapeutic chemo drugs identified in TCGA-CESC cohort and (B) potential therapeutic chemo drugs validated in the GEO cohort.

TABLE 2 Predicting potential chemo drugs via the Cancer Therapeutics Response Portal dataset.

| FASN        |                     |               |             |                    |                    |
|-------------|---------------------|---------------|-------------|--------------------|--------------------|
| Drug        | Dinaciclib          | Alvocidib     | PF-3758309  | Tivantinib         | Fluorouracil       |
| Correlation | −0.20608            | -0.1595       | -0.15792    | -0.14442           | -0.13778           |
| FDR         | 0.000186            | 0.005129      | 0.004505    | 0.010745           | 0.000318           |
| PDK1        |                     |               |             |                    |                    |
| Drug        | Triazolothiadiazine | Vincristine   | Merck60     | Leptomycin B       | CR-1-31B           |
| Correlation | −0.27131            | −0.26454      | −0.26955    | −0.24879           | −0.24931           |
| FDR         | 5E-14               | 1.23E-13      | 1.66E-13    | 5.49E-12           | 7.46E-12           |
| MYO10       |                     |               |             |                    |                    |
| Drug        | Abiraterone         | BRD-K99006945 | VAF-347     | ML334 diastereomer | PD318088           |
| Correlation | -0.29579            | -0.20423      | -0.18466    | -0.14235           | -0.12552           |
| FDR         | 0.01593             | 0.00329       | 0.001264    | 0.025005           | 0.003259           |
| TNFRSF12A   |                     |               |             |                    |                    |
| Drug        | Dasatinib           | VAF-347       | FGIN-1-27   | BRD-K17060750      | ML334 diastereomer |
| Correlation | −0.23802            | −0.2234       | −0.22125    | −0.18099           | −0.1755            |
| FDR         | 1.67E-09            | 9.87E-05      | 0.047048    | 9.99E-05           | 0.006256           |
| SOX21       |                     |               |             |                    |                    |
| Drug        | Cytochalasin B      | Simvastatin   | Fluvastatin |                    |                    |
| Correlation | 0.107668            | 0.106844      | 0.104478    |                    |                    |
| FDR         | 0.022199            | 0.042931      | 0.027822    |                    |                    |
| HLF         |                     |               |             |                    |                    |
| Drug        | Trametinib          | Dasatinib     | Selumetinib | GDC-0879           | PD318088           |
| Correlation | 0.153191            | 0.137379      | 0.135254    | 0.121013           | 0.102907           |
| FDR         | 0.022601            | 0.001002      | 0.001815    | 0.009436           | 0.018969           |
| FGFR2       |                     |               |             |                    |                    |
| Drug        | Tigecycline         | Dabrafenib    | Teniposide  | Isoliquiritigenin  | KW-2449            |
| Correlation | 0.250424            | 0.24845       | 0.2464      | 0.241401           | 0.238217           |
| FDR         | 9.66E-06            | 0.000105      | 2.15E-06    | 0.003093           | 1.02E-10           |
| MYLIP       |                     |               |             |                    |                    |
| Drug        | ML334 diastereomer  | Simvastatin   | Lovastatin  |                    |                    |
| Correlation | 0.155813            | 0.135817      | 0.122877    |                    |                    |
| FDR         | 0.014317            | 0.008074      | 0.005918    |                    |                    |
| ZDHHC11     |                     |               |             |                    |                    |
| Drug        | BRD-K99006945       | AT7867        | ABT-199     | Fluvastatin        | AZD4547            |
| Correlation | 0.148094            | 0.1128        | 0.107107    | 0.102471           | 0.100118           |
| FDR         | 0.032198            | 0.005013      | 0.049336    | 0.031639           | 0.025494           |

analysis in the calibration plot indicated that the prediction performance of this nomogram might be equivalent to an ideal predictive model (Figure 5E). DCA was performed to demonstrate a high clinical net benefit that was almost over the entire threshold probability of the nomogram model compared with other features (Figure 5F).

### Revealing activated pathways and potential therapeutic drugs

The fgsea analysis was performed on TCGA-CESC cohort based on the tumor HALLMARK pathways. We revealed that patients with high HMAG scores contained the activated

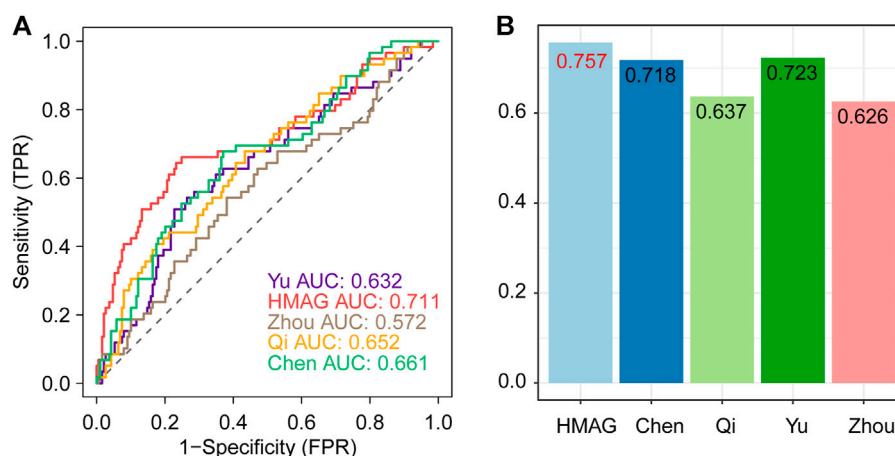


FIGURE 8

Comparison of the prognostic value between HMAG and proposed signatures. (A) ROC curves showing the prognostic value of five signatures and (B) C-index showing the prognostic value of five signatures.

tumor pathways of angiogenesis, G2/M checkpoint, MYC targets, and hypoxia, while patients with low HMAG scores contained the activated tumor pathways of interferon gamma response, interferon alpha response, and IL2/STAT5 signaling (Figure 6A). The DEGs between the high and low HMAG score subgroups also indicated that the pathways of fatty acid peptidyl-tyrosine biosynthetic, positive regulation angiogenesis vasculature, decreased oxygen levels, extracellular matrix encapsulating disassembly, and endoderm formation differentiation terms were activated in high HMAG score patients (Figure 6B), and the pathways of apoptotic antigen presentation processing, lymphocyte mononuclear differentiation, arachidonic eicosanoid unsaturated fatty, positive cell–cell activation adhesion, and cellular hormone alcohol compound terms were activated in low HMAG score patients (Figure 6C).

We identified suitable therapeutic drugs based on the GDSC2016 database and revealed that patients with high HMAG scores were more suitable for the treatment of CHIR-99021, embelin, FTI-277, JNK-9L, JQ12, midostaurin, PF-562271, pyrimethamine, and thapsigargin, and patients with low HMAG scores were more suitable for the treatment of BMS-536924, CP466722, crizotinib, PHA-665752, rapamycin, and TAE684 (all  $p < 0.05$ , Figure 7A). Similar results were also validated in the GEO-combined cohort (all  $p < 0.05$ , Figure 7B). We also searched for new drugs from the GSCA dataset to treat the specific alteration of a single gene, and a total of 41 components were identified that might be functional in the clinical treatment of cervical cancer patients (Table 2).

## The HMAG signature showed better prognostic value than other published signatures

We collected the published signatures of the genes and corresponding indexes from four studies, referring to the genes in immunity, DNA damage repair, ferroptosis, and autophagy. The risk score of each signature was calculated along with the HMAG score using the formula. We observed that the HMAG signature showed more excellent prognostic value than any other signature based on the AUC value (HMAG: 0.711, Yu et al.: 0.632, Zhou et al.: 0.572, Qi et al.: 0.652, and Chen et al.: 0.661, Figure 8A) and C-index (HMAG: 0.757, Yu et al.: 0.723, Zhou et al.: 0.626, Qi et al.: 0.637, and Chen et al.: 0.718, Figure 8B).

## Discussion

Epigenetic modifications are reversible and do not change the DNA sequence of genetic material but can alter how our body reads the DNA sequence, including the basic forms of DNA methylation, posttranslational modifications on histones and noncoding RNAs. However, regulation of the epigenome can distinctly lead to gene malfunction, regulating cell differentiation, proliferation, and even apoptosis and causing disordered cell growth and tumorigenesis. The abnormal control of posttranslational modification-related enzymes, including histone methylase, demethylase, acetylase, and acetyltransferase, acts as a pivotal risk factor for tumors, and these epigenetic alterations may lead to the reprogramming of genomes, activation of oncogenes, or silencing of tumor suppressors (Berry and Janknecht, 2013). Epigenetic regulators

can impact the level of histone modification to the enhancer activity of genes via histone methyltransferases or acetylases (Xiao et al., 2022). Histone deacetylases (HDACs) are essential for maintaining the balance of cell processes by altering histone deacetylation. The abnormal expression of HDACs is tightly linked with several cancers, and their inhibitors have been applied in the clinic to treat several cancers (Patel et al., 2022). Moreover, histone modification impacts the process of epithelial–mesenchymal transition (EMT), and histone acetylation can modulate the acetylation levels of distinct histones at the promoters of EMT-related markers, EMT-inducing transcription factors, and EMT-related long noncoding RNAs to control EMT (Kong et al., 2022).

Histone modification function also emerges in the carcinogenesis of cervical cancer. Higher staining of H3K9ac indicates low grading, negative N-status, and low T-status in cervical cancer. Moreover, the increased expression of H3K4me3 in the cytoplasm was observed to be associated with advanced T stage and unfavorable prognosis in cervical cancer patients (Beyer et al., 2017). Galectin-9, which is encoded by LGALS9, is evidently detected in normal epithelium and endocervical glands but not in cervical intraepithelial neoplasia and cervical squamous cell carcinoma, which indicates that decreased Galectin-9 is a biomarker for the malignant potential of cervical cancer (Liang et al., 2008). Armenta-Castro et al. (2020) reported that histone H3K9 and H3K14 acetylation at the promoter of LGALS9 genes is tightly associated with the protein level of Galectin-9, and the LGALS9 gene presented higher levels of histone acetylation in normal cervical cells than in cancer cells. Sun et al. (2022) also reported that the inhibition of HDACs can activate mitophagy by mediating Parkin acetylation and leading to the inhibition of cervical cancer cell proliferation. Valproic acid/sodium valproate (VPA), a well-known antiepileptic agent, inhibits histone deacetylases, induces histone hyperacetylation, promotes DNA demethylation, and affects the histone methylation status in some cell models. Rocha et al. demonstrated that VPA promotes the abundance of H3K4me2/me3 and increases methyltransferase KMT2D gene expression in HeLa cells. Meanwhile, VPA can also induce the hypomethylation of H3K9me2 and concomitant with the increased gene expression of KDM3A (Rocha et al., 2022). Based on the aforementioned evidence, we sensed the important role of histone modification in cervical cancer tumorigenesis; therefore, we evaluated the histone modification pathway activation status in cervical cancer and generated a prognostic model to predict clinical prognosis.

In the current study, we enrolled a total of 594 cervical cancer patients. We first evaluated the histone modification activation status of 248 patients from TCGA-CESC cohort and identified the activated and suppressed subgroups. Further gene expression variation analysis filtered out 2240 genes and merged them with the results of prognostic analysis. Twenty-three histone modification-associated prognostic genes were enrolled in the LASSO Cox regression analysis. Eventually, we constructed the HMAG signature to predict the prognosis for cervical cancer patients,

with nine genes: TNFRSF12A, MYO10, PDK1, FASN, MYLIP, FGFR2, HLF, SOX21, and ZDHHC11. The HMAG signature showed a preferable prognostic value in TCGA-CESC, GSE44001, and GSE52903 cohorts, and the HMAG signature is an independent prognostic factor after adjusting for the influence of other clinicopathological features. In addition, we established a prognostic nomogram combining the HMAG score and tumor status. The nomogram showed better prognostic value than did age, tumor grade, or tumor stage, with a C-index as high as 0.890. Currently, precise treatment raises concerns in the clinic, and we predicted suitable chemotherapy drugs for patients with high or low HMAG scores. The index of SOX21 in the HMAG signature formula is 0.034241118, which means that the higher expression of SOX21 is associated with a favorable prognosis. The elevated expression of histone H3R26-methylase CARM1 can increase the level of Sox21, and CARM1 inhibition can decrease Sox21 levels, which highlights the importance of epigenetic regulation of Sox21 (Goolam et al., 2016). Caglayan et al. (2013) demonstrated that the exogenous expression of Sox21 in tumor cells resulted in a significant decrease in the tumor size. It seems that Sox21 appears to inhibit stem-like cell properties and initiate the aberrant differentiation of glioma cells. Kurtsdotter et al. also reported that the high levels of SOX5/6/21 in human primary glioblastoma cells enabled the expression of CDK inhibitors and decreased p53 protein turnover, which blocked their tumorigenic capacity through cellular senescence and apoptosis. FASN is a multienzyme protein that serves as the key regulator in lipid metabolism, especially fatty acid synthesis. KDM5C is a histone H3K4-specific demethylase, and the overexpression of KDM5C led to the reduction of H3K4me3 on the promoter and the corresponding downregulation of FASN expression to inhibit FASN-mediated lipid metabolism (Zhang et al., 2020). Du et al. (2022) illustrated that FASN can promote the lymph node metastasis of cervical cancer via cholesterol reprogramming and lymph angiogenesis. The mechanisms of how histone modification influences the tumorigenesis of cervical cancer are complicated and need a further in-depth study.

## Conclusion

In summary, we comprehensively evaluated the histone modification status in cervical cancer patients and revealed histone modification-associated prognostic genes to construct the HMAG signature, aiming to provide new insights into prognosis prediction and precise clinical treatment.

## 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 authors.

## Author contributions

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

## Funding

This work was supported by the 2020–2022 Clinical Research Fund of The Seventh Affiliated Hospital, Sun Yat-sen University.

## 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.

## References

- Armenta-Castro, E., Reyes-Vallejo, T., Maximo-Sanchez, D., Herrera-Camacho, I., Lopez-Lopez, G., and Reyes-Carmona, S. (2020). Histone H3K9 and H3K14 acetylation at the promoter of the LGALS9 gene is associated with mRNA levels in cervical cancer cells. *FEBS Open Bio* 10 (11), 2305–2315. doi:10.1002/2211-5463.12973
- Audia, J. E., and Campbell, R. M. (2016). Histone modifications and cancer. *Cold Spring Harb. Perspect. Biol.* 8 (4), a019521. doi:10.1101/cshperspect.a019521
- Balcacer, P., Shergill, A., and Litkouhi, B. (2019). MRI of cervical cancer with a surgical perspective: Staging, prognostic implications and pitfalls. *Abdom. Radiol.* 44 (7), 2557–2571. doi:10.1007/s00261-019-01984-7
- Bannister, A. J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21 (3), 381–395. doi:10.1038/cr.2011.22
- Bartke, T., Borgel, J., and DiMaggio, P. A. (2013). Proteomics in epigenetics: New perspectives for cancer research. *Brief. Funct. Genomics* 12 (3), 205–218. doi:10.1093/bfgp/elt002
- Berry, W. L., and Janknecht, R. (2013). KDM4/JMJD2 histone demethylases: Epigenetic regulators in cancer cells. *Cancer Res.* 73 (10), 2936–2942. doi:10.1158/0008-5472.CAN-12-4300
- Beyer, S., Zhu, J., Mayr, D., Kuhn, C., Schulze, S., and Hofmann, S. (2017). Histone H3 acetyl K9 and histone H3 tri methyl K4 as prognostic markers for patients with cervical cancer. *Int. J. Mol. Sci.* 18 (3), E477. doi:10.3390/ijms18030477
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 68 (6), 394–424. doi:10.3322/caac.21492
- Caglayan, D., Lundin, E., Kastemar, M., Westermarck, B., and Ferletta, M. (2013). Sox21 inhibits glioma progression *in vivo* by forming complexes with Sox2 and stimulating aberrant differentiation. *Int. J. Cancer* 133 (6), 1345–1356. doi:10.1002/ijc.28147
- Chen, H., Deng, Q., Wang, W., Tao, H., and Gao, Y. (2020). Identification of an autophagy-related gene signature for survival prediction in patients with cervical cancer. *J. Ovarian Res.* 13 (1), 131. doi:10.1186/s13048-020-00730-8
- Du, Q., Liu, P., Zhang, C., Liu, T., Wang, W., Shang, C., et al. (2022). FASN promotes lymph node metastasis in cervical cancer via cholesterol reprogramming and lymphangiogenesis. *Cell Death Dis.* 13 (5), 488. doi:10.1038/s41419-022-04926-2
- Fang, J., Zhang, H., and Jin, S. (2014). Epigenetics and cervical cancer: From pathogenesis to therapy. *Tumour Biol.* 35 (6), 5083–5093. doi:10.1007/s13277-014-1737-z
- Geleher, P., Cox, N. J., and Huang, R. S. (2014). Clinical drug response can be predicted using baseline gene expression levels and *in vitro* drug sensitivity in cell lines. *Genome Biol.* 15 (3), R47. doi:10.1186/gb-2014-15-3-r47

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1013571/full#supplementary-material>

### SUPPLEMENTARY FIGURE S1

Data pretreatment of TCGA-CESC, GSE44001 and GSE52903 cohorts.

- Goolam, M., Scialdone, A., Graham, S. J. L., Macaulay, I. C., Jedrusik, A., Hupalowska, A., et al. (2016). Heterogeneity in Oct4 and Sox2 targets biases cell fate in 4-cell mouse embryos. *Cell* 165 (1), 61–74. doi:10.1016/j.cell.2016.01.047
- Greer, E. L., and Shi, Y. (2012). Histone methylation: A dynamic mark in health, disease and inheritance. *Nat. Rev. Genet.* 13 (5), 343–357. doi:10.1038/nrg3173
- Hanzelmann, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7
- Johnson, C. A., James, D., Marzan, A., and Armaos, M. (2019). Cervical cancer: An overview of pathophysiology and management. *Semin. Oncol. Nurs.* 35 (2), 166–174. doi:10.1016/j.soncn.2019.02.003
- Kong, F., Ma, L., Wang, X., You, H., Zheng, K., and Tang, R. (2022). Regulation of epithelial-mesenchymal transition by protein lysine acetylation. *Cell Commun. Signal.* 20 (1), 57. doi:10.1186/s12964-022-00870-y
- Liang, M., Ueno, M., Oomizu, S., Arikawa, T., Shinonaga, R., and Zhang, S. (2008). Galectin-9 expression links to malignant potential of cervical squamous cell carcinoma. *J. Cancer Res. Clin. Oncol.* 134 (8), 899–907. doi:10.1007/s00432-008-0352-z
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdottir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27 (12), 1739–1740. doi:10.1093/bioinformatics/btr260
- Lu, T. Y., Kao, C. F., Lin, C. T., Huang, D. Y., Chiu, C. Y., and Huang, Y. S. (2009). DNA methylation and histone modification regulate silencing of OPG during tumor progression. *J. Cell. Biochem.* 108 (1), 315–325. doi:10.1002/jcb.22256
- Lu, X., Meng, J., Zhou, Y., Jiang, L., and Yan, F. (2020). Movics: an R package for multi-omics integration and visualization in cancer subtyping. *Bioinformatics* 36, 5539–5541. doi:10.1093/bioinformatics/btaa1018
- Patel, P., Wahan, S. K., Vishakha, S., Kurmi, B. D., Gupta, G. D., and Rajak, H. (2022). Recent progress in histone deacetylase (HDAC) 1 inhibitors as anticancer agent. *Curr. Cancer Drug Targets* 22. doi:10.2174/1568009622666220624090542
- Qi, X., Fu, Y., Sheng, J., Zhang, M., Zhang, M., and Wang, Y. (2021). A novel ferroptosis-related gene signature for predicting outcomes in cervical cancer. *Bioengineered* 12 (1), 1813–1825. doi:10.1080/21655979.2021.1925003
- Rocha, M. A., de Campos Vidal, B., and Mello, M. L. S. (2022). Sodium valproate modulates the methylation status of lysine residues 4, 9 and 27 in histone H3 of HeLa cells. *Curr. Mol. Pharmacol.* 15. doi:10.2174/1874467215666220316110405
- Salvo, G., Odetto, D., Pareja, R., Frumovitz, M., and Ramirez, P. T. (2020). Revised 2018 international federation of Gynecology and Obstetrics (figo) cervical cancer staging: A review of gaps and questions that remain. *Int. J. Gynecol. Cancer* 30 (6), 873–878. doi:10.1136/ijgc-2020-001257
- Shrestha, A. D., Neupane, D., Vedsted, P., and Kallestrup, P. (2018). Cervical cancer prevalence, incidence and mortality in low and middle income countries: A systematic review. *Asian pac. J. Cancer Prev.* 19 (2), 319–324. doi:10.22034/APJCP.2018.19.2.319

- Shvedunova, M., and Akhtar, A. (2022). Modulation of cellular processes by histone and non-histone protein acetylation. *Nat. Rev. Mol. Cell Biol.* 23 (5), 329–349. doi:10.1038/s41580-021-00441-y
- Sun, X., Shu, Y., Ye, G., Wu, C., Xu, M., and Gao, R. (2022). Histone deacetylase inhibitors inhibit cervical cancer growth through Parkin acetylation-mediated mitophagy. *Acta Pharm. Sin. B* 12 (2), 838–852. doi:10.1016/j.apsb.2021.07.003
- Taylor, E. M., Koss, B., Davis, L. E., and Tackett, A. J. (2020). Histone modifications as biomarkers for immunotherapy. *Methods Mol. Biol.* 2055, 213–228. doi:10.1007/978-1-4939-9773-2\_10
- Uckelmann, M., and Sixma, T. K. (2017). Histone ubiquitination in the DNA damage response. *DNA Repair (Amst)* 56, 92–101. doi:10.1016/j.dnarep.2017.06.011
- Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation*. 2 (3), 100141. doi:10.1016/j.xinn.2021.100141
- Xiao, Q., Xiao, Y., Li, L. Y., Chen, M. K., and Wu, M. (2022). Multifaceted regulation of enhancers in cancer. *Biochim. Biophys. Acta. Gene Regul. Mech.* 1865 (6), 194839. doi:10.1016/j.bbagr.2022.194839
- Yu, S., Li, X., Zhang, J., and Wu, S. (2021). Development of a novel immune infiltration-based gene signature to predict prognosis and immunotherapy response of patients with cervical cancer. *Front. Immunol.* 12, 709493. doi:10.3389/fimmu.2021.709493
- Zhang, B., Zhou, B. H., Xiao, M., Li, H., Guo, L., and Wang, M. X. (2020). KDM5C represses FASN-mediated lipid metabolism to exert tumor suppressor activity in intrahepatic cholangiocarcinoma. *Front. Oncol.* 10, 1025. doi:10.3389/fonc.2020.01025
- Zhang, Y., Sun, Z., Jia, J., Du, T., Zhang, N., and Tang, Y. (2021). Overview of histone modification. *Adv. Exp. Med. Biol.* 1283, 1–16. doi:10.1007/978-981-15-8104-5\_1
- Zhou, H., Wu, L., Yu, L., Yang, Y., Kong, L., and Liu, S. (2022). Identify a DNA damage repair gene signature for predicting prognosis and immunotherapy response in cervical squamous cell carcinoma. *J. Oncol.* 2022, 8736575. doi:10.1155/2022/8736575
- Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., and Tanaseichuk, O. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* 10 (1), 1523. doi:10.1038/s41467-019-09234-6



## OPEN ACCESS

## EDITED BY

Xiaofan Lu,  
China Pharmaceutical University, China

## REVIEWED BY

Changjing Wang,  
Hebei Medical University Third Affiliated  
Hospital, China  
Dan Wang,  
Hubei University of Science and  
Technology, China  
Sun Zhaoyang,  
Fudan University, China

## \*CORRESPONDENCE

Huangkai Zhu,  
zhuhuangkai21280@163.com  
Guofang Zhao,  
zhaoguofang@ucas.edu.cn

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 29 July 2022

ACCEPTED 31 August 2022

PUBLISHED 23 September 2022

## CITATION

Sun J, Yang M, Zhao W, Wang F, Yang L,  
Tan C, Hu T, Zhu H and Zhao G (2022),  
Research progress on the relationship  
between the TOR signaling pathway  
regulator, epigenetics, and  
tumor development.  
*Front. Genet.* 13:1006936.  
doi: 10.3389/fgene.2022.1006936

## COPYRIGHT

© 2022 Sun, Yang, Zhao, Wang, Yang,  
Tan, Hu, Zhu and Zhao. This is an open-  
access article distributed under the  
terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Research progress on the relationship between the TOR signaling pathway regulator, epigenetics, and tumor development

Jiaen Sun<sup>1,2</sup>, Minglei Yang<sup>1,2</sup>, Weidi Zhao<sup>1,2</sup>, Fajiu Wang<sup>2</sup>,  
Liangwei Yang<sup>2</sup>, Chuntao Tan<sup>3</sup>, Tianjun Hu<sup>2</sup>, Huangkai Zhu<sup>1,2\*</sup>  
and Guofang Zhao<sup>1,2\*</sup>

<sup>1</sup>School of Medicine, Ningbo University, Ningbo, Zhejiang, China, <sup>2</sup>Department of Thoracic Surgery, Hwa Mei Hospital, University of Chinese Academy of Sciences, Ningbo, Zhejiang, China, <sup>3</sup>Department of Cardiac and Vascular Surgery, Hwa Mei Hospital, University of Chinese Academy of Sciences, Ningbo, Zhejiang, China

Almost all cellular activities depend on protein folding, signaling complex assembly/disassembly, and epigenetic regulation. One of the most important regulatory mechanisms responsible for controlling these cellular processes is dynamic protein phosphorylation/dephosphorylation. Alterations in phosphorylation networks have major consequences in the form of disorders, including cancer. Many signaling cascades, including the target of rapamycin (TOR) signaling, are important participants in the cell cycle, and dysregulation in their phosphorylation/dephosphorylation status has been linked to malignancies. As a TOR signaling regulator, protein phosphatase 2A (PP2A) is responsible for most of the phosphatase activities inside the cells. On the other hand, TOR signaling pathway regulator (TIPRL) is an essential PP2A inhibitory protein. Many other physiological roles have also been suggested for TIPRL, such as modulation of TOR pathways, apoptosis, and cell proliferation. It is also reported that TIPRL was increased in various carcinomas, including non-small-cell lung carcinoma (NSCLC) and hepatocellular carcinomas (HCC). Considering the function of PP2A as a tumor suppressor and also the effect of the TIPRL/PP2A axis on apoptosis and proliferation of cancer cells, this review aims to provide a complete view of the role of TIPRL in cancer development in addition to describing TIPRL/PP2A axis and its epigenetic regulation.

## KEYWORDS

TOR signaling pathway regulator, protein phosphatase, cancer, epigenetic, metabolism

# 1 Introduction

Cancer has overtaken cardiovascular disease as the leading cause of death in many countries. Based on cancer registry data, five frequent malignancies, including breast, prostate, lung, colorectal cancers, and hepatocellular carcinomas (HCC) account for more than half of the yearly observed mortality in the 2016–2020 period (Peto, 2001; Hiatt et al., 2022). Phosphorylation is one of the most well-researched post-translational modifications (PTMs), and it governs a range of biological processes such as cell differentiation, proliferation, death, and cell signaling in healthy conditions. Changes in phosphorylation networks, on the other hand, have major consequences in the form of disorders, including cancer. Many signaling cascades, such as the target of rapamycin (TOR) signaling, are important participants in the cell cycle, and dysregulation in their phosphorylation-dephosphorylation cycle has been linked to many malignancies (Hartl et al., 2011; Brautigan, 2013; Singh et al., 2017). The phosphorylation status of proteins, which controls their function, is determined by the competition between protein kinases and protein phosphatase (PP). Phosphoprotein phosphatases (PPPs) are a class of enzymes (PP1, PP2A, PP2B, and PP2C) necessary for most serine and threonine dephosphorylation in cells.

This is in contrast to the over 400 protein kinases accountable for serine and threonine phosphorylation (Manning et al., 2002). The formation of multimeric holoenzymes is how PPPs attain the specificity and selectivity required of their substrates. The assembly of PPP holoenzymes is subject to stringent regulation, and alterations in the cellular repertoire of PPPs have been connected to cancer development. Combined with the closely related phosphatase PP1, protein phosphatase 2A (PP2A) is responsible for more than 90 percent of the phosphatase activity that occurs within cells. The formation of various heterotrimeric holoenzymes is what controls the functions of PP2A, it is structurally comprised of a scaffold (PP2Aa), a regulatory (PP2Ab), and a catalytic (PP2Ac) subunit. The catalytic c subunit is also called PP2Ac. In addition, a variable B subunit that comes from four regulatory families, including B, B', B'', and B''' (Shi, 2009; Wu et al., 2017). Biogenesis of various PP2A holoenzymes involves PP2A-specific chaperones,  $\alpha 4$  and PP2A phosphatase activator (PTPA), and methylation enzyme leucine carboxyl methyltransferase (LCMT)-1 (Guo et al., 2014).

It is essential for normal physiology to maintain stringent control over the cellular PP2A holoenzymes, and abnormalities in PP2A regulation have been linked to the development of many solid cancers as well as leukaemias (Remmerie and Janssens, 2019). For example, in the chronic myeloid leukaemia, the PP2A phosphatase activity was suppressed. An essential PP2A inhibitory protein is the TOR signaling pathway regulator (TIPRL) (McConnell et al., 2007). The TIPRL structure is a unique butterfly shape consisting of the main core of the

antiparallel  $\beta$ -sheet (Scorsato et al., 2016). TIPRL, as a binding partner for mammalian  $\alpha 4$ /yeast type 2A-associated protein of 42 kDa (Tap42), is the mammalian ortholog of the yeast TAP42 interacting protein of 41 kDa (TIP41) (Jacinto et al., 2001). Ataxia telangiectasia mutated (ATM) and ATM and Rad3-dependent phosphorylation events (ATR) are suppressed by TIPRL interaction with PP2Ac, PP type 4 (PP4), or PP type 6 (PP6), together with  $\alpha 4$ , excluding A and B subunits, finally inhibiting PP function (Nanahoshi et al., 1998; Smetana and Zanchin, 2007a). The TOR signaling pathway is also under the control of the interaction between TIPRL and PP2A (Nakashima et al., 2013). In addition to this pathway, TIPRL has a crucial function in the DNA damage response, apoptosis, and cell proliferation. It is also reported to be increased in various carcinomas, which enables cancer cells to escape apoptotic processes (Song et al., 2012).

Considering the function of PP2A as a tumor suppressor and also the effect of the TIPRL/PP2A axis on apoptosis and proliferation of cancer cells, this review aimed to provide a complete view of the role of TIPRL in tumor development in addition to an overview of the TIPRL/PP2A axis and its epigenetic regulation.

## 2 Overview of TIPRL/PP2A axis

In most of the organisms, the regulation of cell division, cell growth, metabolism, and stress responses depend on the dynamic phosphorylation/dephosphorylation cycle. As one of the most prevalent mechanisms to regulate protein activity, alterations in phosphorylation status affect efficiency, stability, localization, as well as the protein-protein interactions. A phosphate group from ATP is transferred by eukaryotic protein kinases to the hydroxyl groups of serine, threonine, and tyrosine (Ser/Thr/Tyr) residues, while PPs hydrolyze the phosphoester linkage to release phosphate and dephosphorylating protein. In comparison to protein kinases, PPs have received a far smaller amount of research attention for various reasons. At one point, they were viewed as housekeeping (HK) enzymes without significant regulatory activities. Nevertheless, this view has long since altered, and it is now acknowledged that PPs are complexly regulated and extremely selective to various protein substrates (DeLong, 2006; Uhrig et al., 2013a; Brautigan, 2013). The four diverse gene families that make up eukaryotic PPs each have their unique active site signatures. These gene families are as follows: 1) Ser/Thr-specific PPPs, 2)  $Mg^{2+}$ -dependent PPs, 3) Aspartic acid-based PPs, and 4) phospho-Tyr phosphatases (PTPs) (Kerk et al., 2008; Dennis and Bradshaw, 2011).

Approximately 80% of the PP activity in eukaryotic cells is accounted for by the Ser/Thr-specific PPPs, which is one of the most highly conserved proteins among eukaryotic species (Janssens and Goris, 2001). PP1, PP2B, PP4, PP5, PP6, and PP7 are the main subgroups of the PPP family (Uhrig et al.,

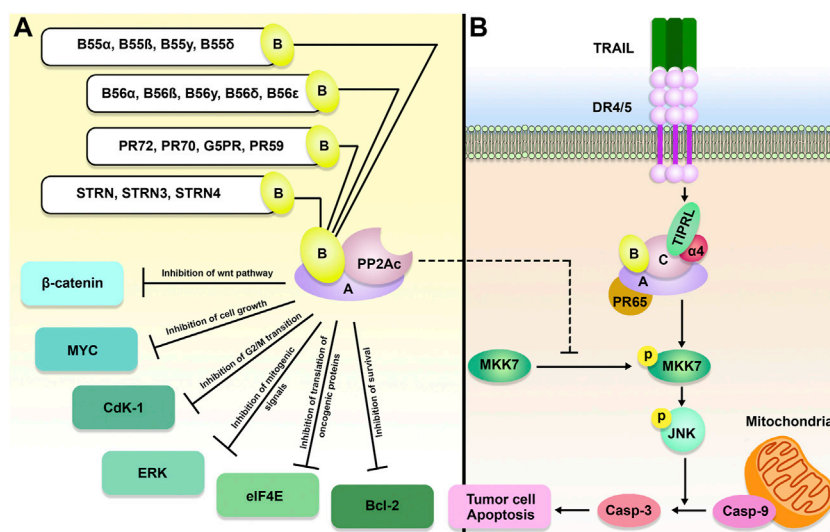


FIGURE 1

TIPRL/PP2A axis and TOR signaling pathway. Active PP2A holoenzyme is an important inhibitor of the TOR signaling pathway. Besides, PP2A inhibits cancer cell growth and survival through dephosphorylation of 4E-binding protein 1 4EBP-1 and S6 kinase (S6K). PP2A holoenzymes can be inhibited as a result of  $\alpha 4$ /TIPRL-mediated recycling and disassembly of scaffold (A) and regulatory (B) subunits.

2013b; Maselli et al., 2014; Rezaeiemanesh et al., 2021). The ten catalytic subunits are responsible for the majority of the dephosphorylation of Ser and Thr in the cell. These subunits include PP1 $\alpha$ , PP1 $\beta$ , PP1 $\gamma$ , PP2Ac $\alpha$ , PP2Ac $\beta$ , PP2Bc, PP4c, PP5c, PP6c, and PP7c. From yeast to humans, the C subunits of PP2A, PP4, and PP6 are closely associated and highly conserved (Chen et al., 2017).

During the process of looking for components of the TOR signaling pathway in yeast, PP2A was shown to be an essential player in the route (Figure 1). As mentioned earlier, subunits A, B, and C make up the PP2A holoenzyme (Xu et al., 2006). The A and C subunits each only have two unique isoforms, while the B subunit is responsible for the majority of PP2A's functional variety. There are currently 16 recognized forms of the B subunit, and several incorporate further splice variants (Vainonen et al., 2021). TIPRL and Tap42/ $\alpha 4$  could engage PP2Ac concurrently, generating a stable ternary complex, according to Smetana et al. (2007b). Tap42/ $\alpha 4$ , as a regulatory component of PP2A, is a downstream effector of the TOR protein kinase. TOR is a protein kinase that controls cell growth in yeast and mammals by coordinating it with the availability of nutrients and environmental factors. PP2A plays a critical part in TOR signaling through its interaction with phosphorylated Tap42/ $\alpha 4$ . Despite being initially discovered as a TAP42 binding protein, *S. cerevisiae* TIP41 operated antagonistically as a suppressor in the TOR axis (Jacinto et al., 2001). Within *S. cerevisiae*, TIP41 was found to engage with the PP2A, PP4, and PP6 C subunits (Gingras et al., 2005). Human TIPRL also binds directly to the PP2A, PP4, and PP6 C subunits. TIPRL, unlike yeast

TIP41, can stimulate the activity of the TOR signaling pathway via its interaction with PP2Ac (Nakashima et al., 2013). Nakashima et al. showed that TIPRL inhibited dephosphorylation of TOR Complex 1 (TORC1) substrates after amino acid deprivation, but TIPRL silencing inhibited phosphorylation of those substrates following amino acid exposure. TIPRL was found to be linked with the PP2Ac, which was essential for TIPRL's impact on TORC1 function (Nakashima et al., 2013).

PP2A, PP4, and PP6 all form distinct clusters within the PPPs, indicating a previously shared ancestor (Uhrig et al., 2013b). A YFL (tyrosine-phenylalanine-leucine) signature, which is conserved among them, can be found near the C-terminal end of the C subunit. PTMs of the catalytic subunits influence holoenzyme assembly by phosphorylating the Tyr residues and methylating the leucine (Leu) residues, respectively (Longin et al., 2007; Janssens et al., 2008). It has been reported that PP2A activity is reduced when phosphorylation occurs at the Tyr (Brautigan, 2013). It has been demonstrated that methylation of the Leu can affect the binding of regulatory subunits, and as a result, methylation is a component of the system that gives protein complexes their substrate specificity (Janssens et al., 2008).

Similar to  $\alpha 4$ , TIPRL is a conserved binding protein to all PP2A-family phosphatases, particularly PP2A, PP4, and PP6 (Gingras et al., 2005; Smetana and Zanchin, 2007a). The fact that all of the residues in PP2A participating in TIPRL binding are identical in PP4 and the fact that there is only one residue that is different in PP6 suggests that TIPRL interacts with PP4/PP6 in

a manner that is analogous to how it interacts with PP2A. The methylation enzyme,  $\alpha 4$ , and TIPRL were among the regulatory proteins that PP4 and PP6 continuously shared with PP2A (McConnell et al., 2007; Jiang et al., 2013). In reaction to demethylation, TIPRL acts as a dynamic PP2A inactivator and inhibits the phosphatase's active site of PP2A, resulting in PP2A disassembly. Besides, Scorsato et al. study suggested that TIPRL could strongly bind to the original PP2A-tail over its tyrosine-phosphorylated counterpart. (Scorsato et al., 2016). The mechanism for dynamic assembly/disassembly of PP2A holoenzymes is established by these findings, which also highlight a dynamic component of PP2A regulation. The main feedback loop for efficient control of PP2A holoenzyme turnover is provided by  $\alpha 4$ /TIPRL, which does not affect the cellular concentration of PP2Ac. These findings explain how PP2Ac concentrations stay stable throughout the cell cycle in mammalian cells (Virshup and Shenolikar, 2009). A key signaling switch governing the stability and disintegration of the holoenzyme is provided by the methylation of the PP2Ac tail in this feedback loop. Demethylation of the PP2Ac tail is essential for the successful attack by TIPRL on the PP2A active site (Rosales et al., 2015). Furthermore, during DNA damage or cellular stress, PP2A holoenzymes may be downregulated as a result of  $\alpha 4$ /TIPRL-mediated recycling. The activity of PP2A in DNA damage-induced ATM/ATR signaling was discovered to be suppressed by TIPRL (Kong et al., 2009).

### 3 Epigenetic regulation of the TIPRL/PP2A axis

To rapidly adjust gene expression in response to alterations in the cellular environment, a broad and complex network of cellular activities balance the reading, deposition, and elimination of epigenetic markers. The three primary types of these epigenetic modifiers are “readers,” “writers,” and “erasers.” On DNA and histones, epigenetic “writers” and “erasers” are responsible for the deposition and removal of epigenetic markers like methylation (Audia and Campbell, 2016; Hyun et al., 2017). The deposited marks are interpreted by the epigenetic readers, which then either engage transcriptional co-factors or further chromatin remodeling complexes to particular places on DNA. To maintain normal gene expression, these modifiers work together in a careful equilibrium. This equilibrium is disturbed by epigenetic dysregulation, which contributes to the aberrant stimulation of oncogenic signaling networks and the beginning stages of the tumorigenesis process (Vicente-Dueñas et al., 2018; Gil and Vagnarelli, 2019). Protein phosphorylation seems to substantially affect the epigenetic state of cancer cells, according to an increasing body of research conducted in recent years (Liu et al., 2016; Yang and Li, 2020). In response to various stimuli inside the cell, kinases and phosphatases make quick and reversible modifications to

proteins. These modifications eventually result in a change in the function, localization, and interaction of partners. Anti-tumor treatments have mostly focused on kinase inhibitors, contributing to customized medicine's growth (Dolgin, 2017). Phosphatases, on the other hand, play an important and sometimes underestimated function in inhibiting neoplastic signaling and are developing as potential targets for medicinal substances (Westermarck, 2018; Vainonen et al., 2021).

The composition of the PP2A holoenzyme is very important to both the function of PP2A and TIPRL and their respective contributions to the process of carcinogenesis. The extraordinary multi-branching mechanism offered by the butterfly-shaped TIPRL may establish some extremely integrative connections with PP2Ac and the A subunit. It is probable that complex interactions are responsible for explaining the multidimensional roles that TIPRL plays in regulating PP2A. These roles include inactivating the phosphatase active site, responding to demethylation, and suppressing some molecular interactions pertinent to holoenzyme formation. In a cellular environment, methylation acts as a protection mark for PP2A holoenzymes, allowing them to avoid being attacked by TIPRL (Vafai and Stock, 2002; Sontag et al., 2004). TIPRL can also be inhibited by single amino acid alterations such as I136T, D71L, D198N, and M196V according to results from reverse two-hybrid experiments (Smetana and Zanchin, 2007b).

Methylation and acetylation of epigenetic targets are facilitated by proteins that are regulated by PP2A. PP2A is responsible for dephosphorylating bromodomain-containing 4 (BRD4), which prevents BRD4 from attaching to acetylated residues and speeds up the transcription process. As a transcriptional and epigenetic regulator that plays an important role in cancer development, BRD4 accumulates at euchromatic sites to stimulate gene expression (Donati et al., 2018). BRD4 is shown to be more likely to be associated with activating acetylation marks after being phosphorylated by casein kinase II (CK2), making BRD4 hyperphosphorylation a predictor of poor prognosis in several malignancies (Wu et al., 2013; Sanz-Álvarez et al., 2021). Additionally, as shown by ERK phosphorylation and increased kinase activity, downregulation of CK2 can result in suppression of the mammalian target of rapamycin (mTOR) cascade, downregulation of Raptor expression levels, and stimulation of the extracellular signaling-regulated protein kinase 1/2 (ERK1/2) signaling mechanism (Olsen et al., 2012).

Histone deacetylase (HDAC) 4/5/7 is dephosphorylated by PP2A, which also prevents HDAC from binding to 14-3-3. This allows HDAC to become nuclear-localized, leading to histone deacetylation. It has been demonstrated that the differentiation status of fibroblasts can be affected by HDAC4 regulation by PP2A (Li et al., 2017). Considering fibroblast infiltration and stimulation contributing to metastasis and treatment resistance in some cancers, fibroblast reprogramming by tumor cells has been established as a crucial aspect of tumor biology (Underwood

TABLE 1 Epigenetic targets of TIPRL/PP2A axis.

| Target | Effect  | Reference   |
|--------|---|---|
| H3     | PP2A reduced the MYC and MYC-related gene transcription   | Huang et al., (2005); Komar and Juszczynski, (2020) |
| BRD4   | PP2A decreased BRD4-associated gene transcription   | Sanz-Álvarez et al. (2021)                          |
| HDAC4  | PP2A inhibited the HDAC binding to 14-3-3   | Paroni et al., (2008); Li et al., (2017)            |
| HDAC7  | PP2A inhibited the HDAC binding to 14-3-3   | Martin et al. (2008)                                |
| HDAC8  | TIPRL is negatively associated with increased apoptosis   | Park and Juhn, (2017)                               |
| PRMT1  | PP2A inhibited PRMT1 activity   | Ichikawa et al. (2020)                              |
| H2AX   | TIPRL upregulation enhanced H2AX phosphorylation in response to DNA damage<br>PP2A promoted DNA damage resolution | Chowdhury et al., (2005); Rosales et al., (2015)    |
| TET2   | PP2A decreased the stability of TET2  | Kundu et al. (2020)                                 |

et al., 2015; Butti et al., 2021; Steele et al., 2021). In addition, HDAC8 overexpression increased the amount of apoptosis generated by cisplatin in H1299 pulmonary cancer cells by suppressing the expression of TIPRL. Additionally, the knockdown of TIPRL increased the amount of apoptosis that was induced in cisplatin-treated cells. In a way dependent on HDAC8, the isoproterenol administration also reduced the amount of transcription of the TIPRL gene caused by cisplatin (Park and Juhn, 2017). TIPRL, as mentioned above, is a negative regulator of PP4. Rosales et al. study showed that the TIPRL depletion boosted PP4 phosphatase activity and the assembly of the functional PP4 complex, which is known to dephosphorylate H2A histone family member X (H2AX). TIPRL upregulation enhanced H2AX phosphorylation in response to DNA damage, whereas TIPRL silencing reduced H2AX phosphorylation. TIPRL upregulation also caused cell death in regard to stress, whereas TIPRL suppression shields cells from genotoxic chemicals, in association with H2AX (Rosales et al., 2015).

A low level of PRMT1/5 methylase activity at the H4R3me2 methylation site is connected with high levels of PP2A activity. Cancer determines the effect of this regulation on gene expression. The most common SDMA methyltransferase is PRMT5, which is linked to gene repression (Yang and Bedford, 2013). There is a mountain of evidence pointing to PRMT5 possessing an oncogenic function in various malignancies, including prostate, pancreatic, and colorectal cancers (Lee et al., 2021; Beketova et al., 2022).

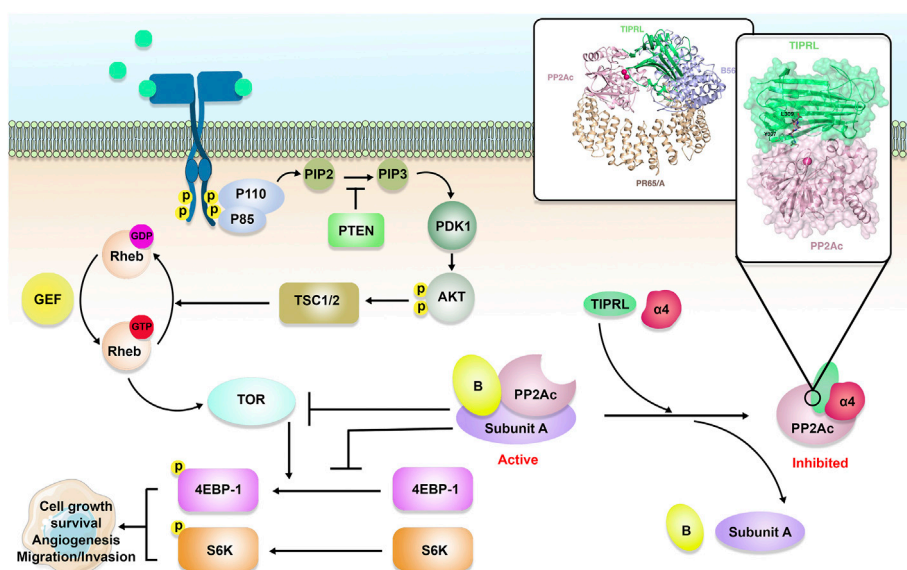
PP2A is responsible for dephosphorylating ten-eleven translocation (TET)-2, which decreases the stability of TET2. This activity inhibits the elimination of methylcytosine in an effective manner. The TET proteins play a key role in the successive oxidation of 5'-methylcytosine to restore the cytosine base's unaltered state and serve primarily as tumor suppressors (Wu and Zhang, 2017). MYC is frequently expressed constitutively in cancerous tissue. There is a correlation between the dysregulation of the MYC family oncogene and a poor prognosis in patients. This dysregulation occurs in more than fifty percent of human malignancies.

Researchers have shown a correlation between MYC upregulation and aggressive human prostate cancer, triple-negative breast cancer, and many different human cancers (Gurel et al., 2008; Palaskas et al., 2011; Dang, 2012). PP2A affects the stability of MYC protein in a direct and immediate post-translational manner, in addition to its potential to regulate MYC transcription (Arnold and Sears, 2006). The activity of PP2A is frequently suppressed during malignancy, which results in an increase in the phosphorylation and the transcriptional activity of MYC (Westermarck and Neel, 2020). The complex processes by which the TIPRL/PP2A axis affects the epigenetic status of cancer cells are summarized in Table 1.

## 4 The role of the TIPRL/PP2A axis in cancer development

During tumorigenesis, PP2A dysregulation can profoundly impact gene expression, genomic stability, and PTM. Because it inhibits the activity of a wide variety of well-known oncogenes, such as MYC, BCL-2, ERK, and AKT, the PP2A holoenzyme is considered a tumor suppressor (Lin et al., 2006; Ruvolo et al., 2011) (Figure 2). In accordance with this function, the PP2A activity is frequently inhibited during cancer (Ruvolo, 2016). PPP2R1A, also known as the PP2A Aα scaffolding subunit, possesses the greatest mutation rate out of all the PP2A subunits in ~1% of all malignancies (Sangodkar et al., 2016). It has been demonstrated that these mutations promote transformation and tumor growth by impairing B or C subunit binding, suppressing the global activity of PP2A (Chen et al., 2005; Haesen et al., 2016).

TIPRL expression pattern in non-small-cell lung carcinoma (NSCLC) was studied by Xu et al. (2020) utilizing The Cancer Genome Atlas (TCGA) dataset. TIPRL was shown to be elevated in NSCLC and was related to an aggressive metastasis stage. TIPRL overexpression was correlated with significantly shorter survival in NSCLC patients. TIPRL silencing inhibited A549, a human Caucasian lung carcinoma cell line, invasion. Besides,



**FIGURE 2**

The role of the TIPRL/PP2A axis in cancer development. PP2A is comprised of a scaffold, a regulatory, and a catalytic (C or PP2Ac) subunit, a variable B subunit that comes from four major families, including B, B', B'', and B'''. PP2A holoenzyme, as a tumor suppressor, can inhibit MYC, Cyclin-dependent kinase 1 (Cdk1), B-cell lymphoma 2 (Bcl-2), Extracellular signal-regulated kinase (ERK), wnt/ $\beta$ -catenin, and eukaryotic translation initiation factor 4E (eIF4E). In addition to inhibition of PP2A, TIPRL can also participate in TRAIL-induced apoptosis and subsequent phosphorylation of MAP kinase 7 (MKK-7) and c-Jun N-terminal kinase (JNK), resulting in cancer cell apoptosis.

TIPRL knockdown dramatically increased cadherin transcription while suppressing vimentin transcription in A549 cells. Another study demonstrated that TIPRL and has-circ-0010235 were upregulated in NSCLC tissues. In NSCLC cells, has-circ-0010235 knockdown or increase of miR-433-3p boosted apoptosis while suppressing proliferation and autophagy. Has-circ-0010235 sponged miR-433-3p to increase TIPRL expression and hence influence NSCLC development. *In vivo*, has-circ-0010235 depletion also inhibited tumor development (Zhang et al., 2021).

According to Jun et al., the expression of TIPRL in liver cancer tissues is highly associated with levels of LC3, a central adaptor in the autophagy cascade, and CD133, also known as prominin-1 and AC133. TIPRL, CD133, and LC3 were shown to be significantly upregulated in hepatocellular carcinomas (HCCs) in comparison to the surrounding normal tissues. TIPRL was found to have a major impact on HCC patient's prognosis. TIPRL knockdown dramatically decreased LC3 and CD133 gene expression, as well as cell survival of HCC cell lines, which were increased by aberrant TIPRL level (Jun et al., 2019). Another research by Jun et al. found that TIPRL and LC3 were significantly upregulated in adult hepatocyte-derived liver disease, but they were downregulated in intrahepatic carcinomas (iCCA). The TIPRL transcription has been found to be the most accurate predictor of survival in patients suffering from liver cancer. This evidence suggests that TIPRL promoted liver cancer cell proliferation by inducing cell survival. TIPRL,

LC3, and CD133 had good efficacy in diagnosing patients with grade 1 iCCA, while TIPRL/LC3/CD133/CD44 axis demonstrated prognostication of both grade 1 HCCs and iCCA (Jun et al., 2021).

Through mRNA microarray analyses, Luan et al. suggested TIPRL as a new metastasis blocker in gastric cancer (GC). TIPRL expression was shown to be lower in clinical GC tissues, and it was linked to a higher metastasis, tumor stage, and a worse clinical prognosis. TIPRL has also been reported to be a direct target of miR-383-5p and miR-216a-5p. TIPRL Upregulation in GC cell lines decreased invasive characteristics, whereas TIPRL-deficient models had the opposite impact. TIPRL also promoted AMP-activated protein kinase (AMPK) activation, which in turn reduced phosphorylation of mTOR, 4EBP-1, and S6K, resulting in mTOR signaling deactivation and consequent inhibition of tumor invasion in GC (Luan et al., 2020).

Tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) is a potential anti-tumor drug (Figure 2), however, some cancerous cells are insensitive to it. Yoon et al. study showed that the co-treatment of Huh7 cells, a permanent HCC cell line, with TRT-0173 or TRT-0029, as TRAIL sensitizers, resulted in TRAIL-induced apoptosis because of suppression of the TIPRL-MKK-7 interaction and subsequent phosphorylation of MAP kinase 7 (MKK-7) and c-Jun N-terminal kinase (JNK). *In vivo*, these chemicals were injected into HCC xenograft tumors, resulting in inhibition of tumor development (Yoon et al., 2017). Yoon et al. also suggested

that combining TRAIL with *Taraxacum officinale* F.H. Wigg (TO) treatment of Huh7 cells resulted in TRAIL-induced apoptosis through a similar pathway (Yoon et al., 2016). *Tussilago farfara* L. (TF) was proposed by Lee et al. (2014) as another TRAIL sensitizer. The combination of TF and TRAIL triggered apoptosis in Huh7 cells by inhibiting the TIPRL-MKK-7 interaction and increasing MKK-7 and JNK phosphorylation. According to Song et al. (2012) study, levels of TIPRL were higher in HCC tissues and cell lines in comparison to the non-tumor tissues. Silencing of TIPRL led to increased apoptosis. Treatment of HCC cells with small interfering RNA (siRNA) against TIPRL (siTIPRL) and TRAIL caused phosphorylation of MKK-7 and JNK and led to apoptosis, indicated by cleavage of caspase-8 and caspase-3. *In vivo*, injection of HCC xenograft tumors with siTIPRL and TRAIL led to the induction of tumor apoptosis.

## 5 Conclusion and future perspective

TIPRL is a novel PP2A inhibitor. It seems that demethylation of the PP2Ac tail is essential for TIPRL-mediated inhibition of PP2A. TIPRL/PP2A also control the methylation, acetylation, and phosphorylation of several epigenetic targets that are important in gene expression and cell cycle. In addition to the effects associated with PP2A inhibition, TIPRL has important regulatory effects in the TOR signaling pathway. Since PP2A holoenzyme is known as a tumor suppressor, its inhibition is expected to contribute to tumor progression. However, studies have shown that this theory is not entirely true. Indeed, there are studies indicating that the high expression of TIPRL is not always associated with tumor development and worse prognosis. Also, there are still contradictions regarding the relationship of TIPRL with the induction of apoptosis in cancer cells. Although TIPRL is expected to prevent cancer cell apoptosis by inhibiting PP2A, which is an inhibitor of BCL-2, a series of studies have shown that this protein can induce TRAIL-related cell death through the phosphorylation of MKK-7. This paradox in the results of studies can be caused by the endless complexities of intracellular signaling pathways and regulatory systems related to each other. However, to understand more about the function of TIPRL in cancer and to use it to enhance the anti-cancer

effects of PP2A, it is necessary to conduct more clinical studies focusing on the epigenetic regulation of the TIPRL/PP2A pathway.

## Author contributions

JS, WZ, and MY is responsible for the collection, collation, and writing of the original manuscript. GZ, HZ, FW, LY, CT, and TH are responsible for the concept development. GZ and HZ are responsible for the revision, and manuscript review. All authors reviewed and accepted the final version.

## Funding

This work was funded by the Ningbo Health Branding Subject Fund (grant no. PPXK 2018-05); the Natural Science Foundation of Ningbo (grant no. 2019A610225); the Hwamei Fund (grant no. 2019HMZDKY04 and 2022HMKY37); the Key discipline of Hwamei Hospital, University of Chinese Academy of Science (grant no. 2020ZDXK03); the Medical Health Science and Technology Project of Zhejiang Provincial Health Commission (Grant No. 2021KY1009 and 2022KY1138).

## 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.

## References

- Arnold, H. K., and Sears, R. C. (2006). Protein phosphatase 2A regulatory subunit B56alpha associates with c-myc and negatively regulates c-myc accumulation. *Mol. Cell. Biol.* 26 (7), 2832–2844. doi:10.1128/MCB.26.7.2832-2844.2006
- Audia, J. E., and Campbell, R. M. (2016). Histone modifications and cancer. *Cold Spring Harb. Perspect. Biol.* 8 (4), a019521. doi:10.1101/cshperspect.a019521
- Beketova, E., Owens, J. L., Asberry, A. M., and Hu, C. D. (2022). PRMT5: A putative oncogene and therapeutic target in prostate cancer. *Cancer Gene Ther.* 29 (3–4), 264–276. doi:10.1038/s41417-021-00327-3
- Brautigan, D. L. (2013). Protein Ser/Thr phosphatases--the ugly ducklings of cell signalling. *Febs J.* 280 (2), 324–345. doi:10.1111/j.1742-4658.2012.08609.x
- Butti, R., Nimma, R., Kundu, G., Bulbule, A., Kumar, T. V. S., Gunasekaran, V. P., et al. (2021). Tumor-derived osteopontin drives the resident fibroblast to myofibroblast differentiation through Twist1 to promote breast cancer progression. *Oncogene* 40 (11), 2002–2017. doi:10.1038/s41388-021-01663-2
- Chen, M. J., Dixon, J. E., and Manning, G. (2017). Genomics and evolution of protein phosphatases. *Sci. Signal.* 10 (474), eaag1796. doi:10.1126/scisignal.aag1796

- Chen, W., Arroyo, J. D., Timmons, J. C., Possemato, R., and Hahn, W. C. (2005). Cancer-associated PP2A Aalpha subunits induce functional haploinsufficiency and tumorigenicity. *Cancer Res.* 65 (18), 8183–8192. doi:10.1158/0008-5472.Can-05-1103
- Chowdhury, D., Keogh, M. C., Ishii, H., Peterson, C. L., Buratowski, S., and Lieberman, J. (2005). gamma-H2AX dephosphorylation by protein phosphatase 2A facilitates DNA double-strand break repair. *Mol. Cell* 20 (5), 801–809. doi:10.1016/j.molcel.2005.10.003
- Dang, C. V. (2012). MYC on the path to cancer. *Cell* 149 (1), 22–35. doi:10.1016/j.cell.2012.03.003
- DeLong, A. (2006). Switching the flip: Protein phosphatase roles in signaling pathways. *Curr. Opin. Plant Biol.* 9 (5), 470–477. doi:10.1016/j.pbi.2006.07.015
- Dennis, E. A., and Bradshaw, R. A. (2011). *Transduction mechanisms in cellular signaling: Cell signaling collection*. Massachusetts, United States: Academic Press.
- Dolgin, E. (2017). The most popular genes in the human genome. *Nature* 551 (7681), 427–431. doi:10.1038/d41586-017-07291-9
- Donati, B., Lorenzini, E., and Ciarrocchi, A. (2018). BRD4 and cancer: Going beyond transcriptional regulation. *Mol. Cancer* 17 (1), 164. doi:10.1186/s12943-018-0915-9
- Gil, R. S., and Vagnarelli, P. (2019). Protein phosphatases in chromatin structure and function. *Biochim. Biophys. Acta. Mol. Cell Res.* 1866 (1), 90–101. doi:10.1016/j.bbamcr.2018.07.016
- Gingras, A. C., Caballero, M., Zarske, M., Sanchez, A., Hazbun, T. R., Fields, S., et al. (2005). A novel, evolutionarily conserved protein phosphatase complex involved in cisplatin sensitivity. *Mol. Cell. Proteomics* 4 (11), 1725–1740. doi:10.1074/mcp.M500231-MCP200
- Guo, F., Stanevich, V., Wlodarchak, N., Sengupta, R., Jiang, L., Satyshur, K. A., et al. (2014). Structural basis of PP2A activation by PTPA, an ATP-dependent activation chaperone. *Cell Res.* 24 (2), 190–203. doi:10.1038/cr.2013.138
- Gurel, B., Iwata, T., Koh, C. M., Jenkins, R. B., Lan, F., Van Dang, C., et al. (2008). Nuclear MYC protein overexpression is an early alteration in human prostate carcinogenesis. *Mod. Pathol.* 21 (9), 1156–1167. doi:10.1038/modpathol.2008.111
- Haesen, D., Abbasi Asbagh, L., Derua, R., Hubert, A., Schrauwen, S., Hoorn, Y., et al. (2016). Recurrent PPP2R1A mutations in uterine cancer act through a dominant-negative mechanism to promote malignant cell growth. *Cancer Res.* 76 (19), 5719–5731. doi:10.1158/0008-5472.Can-15-3342
- Hartl, F. U., Bracher, A., and Hayer-Hartl, M. (2011). Molecular chaperones in protein folding and proteostasis. *Nature* 475 (7356), 324–332. doi:10.1038/nature10317
- Hiatt, R. A., Sibley, A., Venkatesh, B., Cheng, J., Dixit, N., Fox, R., et al. (2022). From cancer epidemiology to policy and practice: The role of a comprehensive cancer center. *Curr. Epidemiol. Rep.* 9 (1), 10–21. doi:10.1007/s40471-021-00280-7
- Huang, W., Batra, S., Atkins, B. A., Mishra, V., and Mehta, K. D. (2005). Increases in intracellular calcium dephosphorylate histone H3 at serine 10 in human hepatoma cells: Potential role of protein phosphatase 2A-protein kinase CbetaII complex. *J. Cell. Physiol.* 205 (1), 37–46. doi:10.1002/jcp.20372
- Hyun, K., Jeon, J., Park, K., and Kim, J. (2017). Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* 49 (4), e324. doi:10.1038/emm.2017.11
- Ichikawa, T., Shanab, O., Nakahata, S., Shimozaki, S., Manachai, N., Ono, M., et al. (2020). Novel PRMT5-mediated arginine methylations of HSP90A are essential for maintenance of HSP90A function in NDRG2(low) ATL and various cancer cells. *Biochim. Biophys. Acta. Mol. Cell Res.* 1867 (2), 118615. doi:10.1016/j.bbamcr.2019.118615
- Jacinto, E., Guo, B., Arndt, K. T., Schmelzle, T., and Hall, M. N. (2001). TIP41 interacts with TAP42 and negatively regulates the TOR signaling pathway. *Mol. Cell* 8 (5), 1017–1026. doi:10.1016/s1097-2765(01)00386-0
- Janssens, V., and Goris, J. (2001). Protein phosphatase 2A: A highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem. J.* 353, 417–439. doi:10.1042/0264-6021:3530417
- Janssens, V., Longin, S., and Goris, J. (2008). PP2A holoenzyme assembly: In cauda venenum (the sting is in the tail). *Trends biochem. Sci.* 33 (3), 113–121. doi:10.1016/j.tibs.2007.12.004
- Jiang, L., Stanevich, V., Satyshur, K. A., Kong, M., Watkins, G. R., Wadzinski, B. E., et al. (2013). Structural basis of protein phosphatase 2A stable latency. *Nat. Commun.* 4, 1699. doi:10.1038/ncomms2663
- Jun, S. Y., Jeon, S. J., Yoon, J. Y., Lee, J. J., Yoon, H. R., Choi, M. H., et al. (2019). The positive correlation of TIPRL with LC3 and CD133 contributes to cancer aggressiveness: Potential biomarkers for early liver cancer. *Sci. Rep.* 9 (1), 16802. doi:10.1038/s41598-019-53191-5
- Jun, S. Y., Yoon, H. R., Yoon, J. Y., Jeon, S. J., Lee, J. J., Halder, D., et al. (2021). The human TOR signaling regulator is the key indicator of liver cancer patients' overall survival: TIPRL/LC3/CD133/CD44 as potential biomarkers for early liver cancers. *Cancers (Basel)* 13 (12), 2925. doi:10.3390/cancers13122925
- Kerk, D., Templeton, G., and Moorhead, G. B. (2008). Evolutionary radiation pattern of novel protein phosphatases revealed by analysis of protein data from the completely sequenced genomes of humans, green algae, and higher plants. *Plant Physiol.* 146 (2), 351–367. doi:10.1104/pp.107.111393
- Komar, D., and Juszczynski, P. (2020). Rebelled epigenome: Histone H3S10 phosphorylation and H3S10 kinases in cancer biology and therapy. *Clin. Epigenetics* 12 (1), 147. doi:10.1186/s13148-020-00941-2
- Kong, M., Ditsworth, D., Lindsten, T., and Thompson, C. B. (2009). Alpha4 is an essential regulator of PP2A phosphatase activity. *Mol. Cell* 36 (1), 51–60. doi:10.1016/j.molcel.2009.09.025
- Kundu, A., Shelar, S., Ghosh, A. P., Ballesta, M., Kirkman, R., Nam, H., et al. (2020). 14-3-3 proteins protect AMPK-phosphorylated ten-eleven translocation-2 (TET2) from PP2A-mediated dephosphorylation. *J. Biol. Chem.* 295 (6), 1754–1766. doi:10.1074/jbc.RA119.011089
- Lee, H. J., Cho, H. S., Jun, S. Y., Lee, J. J., Yoon, J. Y., Lee, J. H., et al. (2014). Tussilago farfara L. augments TRAIL-induced apoptosis through MKK7/JNK activation by inhibition of MKK7-TIPRL in human hepatocellular carcinoma cells. *Oncol. Rep.* 32 (3), 1117–1123. doi:10.3892/or.2014.3279
- Lee, M. K. C., Grimmond, S. M., McArthur, G. A., and Sheppard, K. E. (2021). PRMT5: An emerging target for pancreatic adenocarcinoma. *Cancers (Basel)* 13 (20), 5136. doi:10.3390/cancers13205136
- Li, Y., Tang, C. B., and Kilian, K. A. (2017). Matrix mechanics influence fibroblast-myofibroblast transition by directing the localization of histone deacetylase 4. *Cell. Mol. Bioeng.* 10 (5), 405–415. doi:10.1007/s12195-017-0493-8
- Lin, S. S., Bassik, M. C., Suh, H., Nishino, M., Arroyo, J. D., Hahn, W. C., et al. (2006). PP2A regulates BCL-2 phosphorylation and proteasome-mediated degradation at the endoplasmic reticulum. *J. Biol. Chem.* 281 (32), 23003–23012. doi:10.1074/jbc.M602648200
- Liu, F., Wang, L., Perna, F., and Nimer, S. D. (2016). Beyond transcription factors: How oncogenic signalling reshapes the epigenetic landscape. *Nat. Rev. Cancer* 16 (6), 359–372. doi:10.1038/nrc.2016.41
- Longin, S., Zwaenepoel, K., Louis, J. V., Dilworth, S., Goris, J., and Janssens, V. (2007). Selection of protein phosphatase 2A regulatory subunits is mediated by the C terminus of the catalytic subunit. *J. Biol. Chem.* 282 (37), 26971–26980. doi:10.1074/jbc.M704059200
- Luan, M., Shi, S. S., Shi, D. B., Liu, H. T., Ma, R. R., Xu, X. Q., et al. (2020). TIPRL, a novel tumor suppressor, suppresses cell migration, and invasion through regulating AMPK/mTOR signaling pathway in gastric cancer. *Front. Oncol.* 10, 1062. doi:10.3389/fonc.2020.01062
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* 298 (5600), 1912–1934. doi:10.1126/science.1075762
- Martin, M., Potente, M., Janssens, V., Vertommen, D., Twizere, J. C., Rider, M. H., et al. (2008). Protein phosphatase 2A controls the activity of histone deacetylase 7 during T cell apoptosis and angiogenesis. *Proc. Natl. Acad. Sci. U. S. A.* 105 (12), 4727–4732. doi:10.1073/pnas.0708455105
- Maselli, G. A., Slamovits, C. H., Bianchi, J. I., Vilarrasa-Blasi, J., Caño-Delgado, A. I., and Mora-García, S. (2014). Revisiting the evolutionary history and roles of protein phosphatases with Kelch-like domains in plants. *Plant Physiol.* 164 (3), 1527–1541. doi:10.1104/pp.113.233627
- McConnell, J. L., Gomez, R. J., McCorvey, L. R., Law, B. K., and Wadzinski, B. E. (2007). Identification of a PP2A-interacting protein that functions as a negative regulator of phosphatase activity in the ATM/ATR signaling pathway. *Oncogene* 26 (41), 6021–6030. doi:10.1038/sj.onc.1210406
- Nakashima, A., Tanimura-Ito, K., Oshiro, N., Eguchi, S., Miyamoto, T., Momonami, A., et al. (2013). A positive role of mammalian Tip41-like protein, TIPRL, in the amino-acid dependent mTORC1-signaling pathway through interaction with PP2A. *FEBS Lett.* 587 (18), 2924–2929. doi:10.1016/j.febslet.2013.07.027
- Nanahoshi, M., Nishiuma, T., Tsujishita, Y., Hara, K., Inui, S., Sakaguchi, N., et al. (1998). Regulation of protein phosphatase 2A catalytic activity by alpha4 protein and its yeast homolog Tap42. *Biochem. Biophys. Res. Commun.* 251 (2), 520–526. doi:10.1006/bbrc.1998.9493
- Olsen, B. B., Svenstrup, T. H., and Guerra, B. (2012). Downregulation of protein kinase CK2 induces autophagic cell death through modulation of the mTOR and MAPK signaling pathways in human glioblastoma cells. *Int. J. Oncol.* 41 (6), 1967–1976. doi:10.3892/ijo.2012.1635
- Palaskas, N., Larson, S. M., Schultz, N., Komisopoulou, E., Wong, J., Rohle, D., et al. (2011). 18F-fluorodeoxy-glucose positron emission tomography marks MYC-overexpressing human basal-like breast cancers. *Cancer Res.* 71 (15), 5164–5174. doi:10.1158/0008-5472.Can-10-4633

- Park, J. Y., and Juhn, Y. S. (2017). cAMP signaling increases histone deacetylase 8 expression via the Epac2-Rap1A-Akt pathway in H1299 lung cancer cells. *Exp. Mol. Med.* 49 (2), e297. doi:10.1038/emmm.2016.152
- Paroni, G., Cernotta, N., Dello Russo, C., Gallinari, P., Pallaoro, M., Foti, C., et al. (2008). PP2A regulates HDAC4 nuclear import. *Mol. Biol. Cell* 19 (2), 655–667. doi:10.1091/mbc.e07-06-0623
- Peto, J. (2001). Cancer epidemiology in the last century and the next decade. *Nature* 411 (6835), 390–395. doi:10.1038/35077256
- Remmerie, M., and Janssens, V. (2019). PP2A: A promising biomarker and therapeutic target in endometrial cancer. *Front. Oncol.* 9, 462. doi:10.3389/fonc.2019.00462
- Rezaei-Manesh, A., Mahmoudi, M., Amirzargar, A. A., Vojdani, M., Babaie, F., Mahdavi, J., et al. (2021). Upmodulation of unfolded protein response and ER stress-related IL-23 production in M1 macrophages from ankylosing spondylitis patients. *Inflammation* 45 (2), 665–676. doi:10.1007/s10753-021-01575-z
- Rosales, K. R., Reid, M. A., Yang, Y., Tran, T. Q., Wang, W. I., Lowman, X., et al. (2015). TIPRL inhibits protein phosphatase 4 activity and promotes H2AX phosphorylation in the DNA damage response. *PLoS One* 10 (12), e0145938. doi:10.1371/journal.pone.0145938
- Ruvolo, P. P., Qui, Y. H., Coombes, K. R., Zhang, N., Ruvolo, V. R., Borthakur, G., et al. (2011). Low expression of PP2A regulatory subunit B55α is associated with T308 phosphorylation of AKT and shorter complete remission duration in acute myeloid leukemia patients. *Leukemia* 25 (11), 1711–1717. doi:10.1038/leu.2011.146
- Ruvolo, P. P. (2016). The broken "off" switch in cancer signaling: PP2A as a regulator of tumorigenesis, drug resistance, and immune surveillance. *BBA Clin.* 6, 87–99. doi:10.1016/j.bbaci.2016.08.002
- Sangodkar, J., Farrington, C. C., McClinch, K., Galsky, M. D., Kastrinsky, D. B., and Nara, G. (2016). All roads lead to PP2A: Exploiting the therapeutic potential of this phosphatase. *Febs J.* 283 (6), 1004–1024. doi:10.1111/febs.13573
- Sanz-Álvarez, M., Cristóbal, I., Luque, M., Santos, A., Zazo, S., Madoz-Gürpide, J., et al. (2021). Expression of phosphorylated BRD4 is markedly associated with the activation status of the PP2A pathway and shows a strong prognostic value in triple negative breast cancer patients. *Cancers (Basel)* 13 (6), 1246. doi:10.3390/cancers13061246
- Scorsato, V., Lima, T. B., Righetto, G. L., Zanchin, N. I. T., Brandão-Neto, J., Sandy, J., et al. (2016). Crystal structure of the human Tip41 orthologue, TIPRL, reveals a novel fold and a binding site for the PP2Ac C-terminus. *Sci. Rep.* 6 (1), 30813. doi:10.1038/srep30813
- Shi, Y. (2009). Serine/threonine phosphatases: Mechanism through structure. *Cell* 139 (3), 468–484. doi:10.1016/j.cell.2009.10.006
- Singh, V., Ram, M., Kumar, R., Prasad, R., Roy, B. K., and Singh, K. K. (2017). Phosphorylation: Implications in cancer. *Protein J.* 36 (1), 1–6. doi:10.1007/s10930-017-9696-z
- Smetana, J., and Zanchin, N. (2007). Interaction analysis of the heterotrimer formed by the phosphatase 2A catalytic subunit, alpha4 and the mammalian ortholog of yeast Tip41 (TIPRL). *FEBS J.* 274, 5891–5904. doi:10.1111/j.1742-4658.2007-06112-x
- Smetana, J. H., and Zanchin, N. I. (2007). Interaction analysis of the heterotrimer formed by the phosphatase 2A catalytic subunit, alpha4 and the mammalian ortholog of yeast Tip41 (TIPRL). *Febs J.* 274 (22), 5891–5904. doi:10.1111/j.1742-4658.2007.06112.x
- Song, I. S., Jun, S. Y., Na, H. J., Kim, H. T., Jung, S. Y., Ha, G. H., et al. (2012). Inhibition of MKK7-JNK by the TOR signaling pathway regulator-like protein contributes to resistance of HCC cells to TRAIL-induced apoptosis. *Gastroenterology* 143 (5), 1341–1351. doi:10.1053/j.gastro.2012.07.103
- Sontag, E., Hladik, C., Montgomery, L., Luangpirom, A., Mudrak, I., Ogris, E., et al. (2004). Downregulation of protein phosphatase 2A carboxyl methylation and methyltransferase may contribute to Alzheimer disease pathogenesis. *J. Neuropathol. Exp. Neurol.* 63 (10), 1080–1091. doi:10.1093/jnen/63.10.1080
- Steele, N. G., Biffi, G., Kemp, S. B., Zhang, Y., Drouillard, D., Syu, L., et al. (2021). Inhibition of hedgehog signaling alters fibroblast composition in pancreatic cancer. *Clin. Cancer Res.* 27 (7), 2023–2037. doi:10.1158/1078-0432.Ccr-20-3715
- Uhrig, R. G., Kerk, D., and Moorhead, G. B. (2013). Evolution of bacterial-like phosphoprotein phosphatases in photosynthetic eukaryotes features ancestral mitochondrial or archaeal origin and possible lateral gene transfer. *Plant Physiol.* 163 (4), 1829–1843. doi:10.1104/pp.113.224378
- Uhrig, R. G., Labandera, A. M., and Moorhead, G. B. (2013). Arabidopsis PPP family of serine/threonine protein phosphatases: Many targets but few engines. *Trends Plant Sci.* 18 (9), 505–513. doi:10.1016/j.tplants.2013.05.004
- Underwood, T. J., Hayden, A. L., Derouet, M., Garcia, E., Noble, F., White, M. J., et al. (2015). Cancer-associated fibroblasts predict poor outcome and promote peritumoral-dependent invasion in oesophageal adenocarcinoma. *J. Pathol.* 235 (3), 466–477. doi:10.1002/path.4467
- Vafai, S. B., and Stock, J. B. (2002). Protein phosphatase 2A methylation: A link between elevated plasma homocysteine and Alzheimer's disease. *FEBS Lett.* 518 (1–3), 1–4. doi:10.1016/s0014-5793(02)02702-3
- Vainonen, J. P., Momeny, M., and Westermarck, J. (2021). Druggable cancer phosphatases. *Sci. Transl. Med.* 13 (588), eabe2967. doi:10.1126/scitranslmed.abe2967
- Vicente-Dueñas, C., Hauer, J., Cobaleda, C., Borkhardt, A., and Sánchez-García, I. (2018). Epigenetic priming in cancer initiation. *Trends Cancer* 4 (6), 408–417. doi:10.1016/j.trecan.2018.04.007
- Virshup, D. M., and Shenolikar, S. (2009). From promiscuity to precision: Protein phosphatases get a makeover. *Mol. Cell* 33 (5), 537–545. doi:10.1016/j.molcel.2009.02.015
- Westermarck, J., and Neel, B. G. (2020). Piecing together a broken tumor suppressor phosphatase for cancer therapy. *Cell* 181 (3), 514–517. doi:10.1016/j.cell.2020.04.005
- Westermarck, J. (2018). Targeted therapies don't work for a reason; the neglected tumor suppressor phosphatase PP2A strikes back. *Febs J.* 285 (22), 4139–4145. doi:10.1111/febs.14617
- Wu, C. G., Chen, H., Guo, F., Yadav, V. K., McIlwain, S. J., Rowse, M., et al. (2017). PP2A-B' holoenzyme substrate recognition, regulation and role in cytokinesis. *Cell Discov.* 3, 17027. doi:10.1038/celldisc.2017.27
- Wu, S. Y., Lee, A. Y., Lai, H. T., Zhang, H., and Chiang, C. M. (2013). Phospho switch triggers Brd4 chromatin binding and activator recruitment for gene-specific targeting. *Mol. Cell* 49 (5), 843–857. doi:10.1016/j.molcel.2012.12.006
- Wu, X., and Zhang, Y. (2017). TET-Mediated active DNA demethylation: Mechanism, function and beyond. *Nat. Rev. Genet.* 18 (9), 517–534. doi:10.1038/nrg.2017.33
- Xu, X., Zhu, H., Yang, M., Zheng, E., Zhou, Y., Ni, J., et al. (2020). Knockdown of TOR signaling pathway regulator suppresses cell migration and invasion in non-small cell lung cancer via the regulation of epithelial-to-mesenchymal transition. *Exp. Ther. Med.* 19 (3), 1925–1932. doi:10.3892/etm.2019.8358
- Xu, Y., Xing, Y., Chen, Y., Chao, Y., Lin, Z., Fan, E., et al. (2006). Structure of the protein phosphatase 2A holoenzyme. *Cell* 127 (6), 1239–1251. doi:10.1016/j.cell.2006.11.033
- Yang, Y., and Bedford, M. T. (2013). Protein arginine methyltransferases and cancer. *Nat. Rev. Cancer* 13 (1), 37–50. doi:10.1038/nrc3409
- Yang, Y., and Li, G. (2020). Post-translational modifications of PRC2: Signals directing its activity. *Epigenetics Chromatin* 13 (1), 47. doi:10.1186/s13072-020-00369-1
- Yoon, J. Y., Cho, H. S., Lee, J. J., Lee, H. J., Jun, S. Y., Lee, J. H., et al. (2016). Novel TRAIL sensitizer *Taraxacum officinale* F.H. Wigg enhances TRAIL-induced apoptosis in Huh7 cells. *Mol. Carcinog.* 55 (4), 387–396. doi:10.1002/mc.22288
- Yoon, J. Y., Lee, J. J., Gu, S., Jung, M. E., Cho, H. S., Lim, J. H., et al. (2017). Novel indazole-based small compounds enhance TRAIL-induced apoptosis by inhibiting the MKK7-TIPRL interaction in hepatocellular carcinoma. *Oncotarget* 8 (68), 112610–112622. doi:10.18632/oncotarget.22614
- Zhang, F., Cheng, R., Li, P., Lu, C., and Zhang, G. (2021). Hsa\_circ\_0010235 functions as an oncogenic drive in non-small cell lung cancer by modulating miR-433-3p/TIPRL axis. *Cancer Cell Int.* 21 (1), 73. doi:10.1186/s12935-021-01764-8



## OPEN ACCESS

## EDITED BY

Xiaofan Lu,  
China Pharmaceutical University, China

## REVIEWED BY

Jialin Meng,  
First Affiliated Hospital of Anhui Medical  
University, China  
Lingnan He,  
Tongji University, China  
Xue Fang,  
Fujian Medical University, China

## \*CORRESPONDENCE

Yuan Si,  
siyuan4210@126.com

<sup>†</sup>These authors have contributed equally  
to this work and share first authorship

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 04 August 2022

ACCEPTED 12 September 2022

PUBLISHED 26 September 2022

## CITATION

Zhang X, Wang C, Liu Z and Si Y (2022),  
Identification of m<sup>6</sup>A-related long non-  
coding RNAs for predicting prognosis  
and immune characterizations in  
gastric cancer.  
*Front. Genet.* 13:1011716.  
doi: 10.3389/fgene.2022.1011716

## COPYRIGHT

© 2022 Zhang, Wang, Liu and Si. 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.

# Identification of m<sup>6</sup>A-related long non-coding RNAs for predicting prognosis and immune characterizations in gastric cancer

Xianhui Zhang<sup>1†</sup>, Changjing Wang<sup>2†</sup>, Zhongxin Liu<sup>3</sup> and Yuan Si<sup>4\*</sup>

<sup>1</sup>Department of CT, The XingTai People's Hospital, Xingtai, China, <sup>2</sup>Department of Gastrointestinal Surgery, The Third Hospital of Hebei Medical University, Shijiazhuang, China, <sup>3</sup>Department of Pathology, The XingTai People's Hospital, Xingtai, China, <sup>4</sup>Endoscopic Center, The XingTai People's Hospital, Xingtai, China

**Background:** N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) mRNA modification triggers malignant behavior in tumor cells, which promotes malignant progression and migration of gastric cancer (GC). Nevertheless, studies on the prognostic value of m<sup>6</sup>A-related long non-coding RNA (MRlncRNA) in GC remain quite restricted. The study aimed to develop a reasonable predictive model to explore the prognostic potential of MRlncRNAs in predicting the prognosis of GC patients and monitoring the efficacy of immunotherapy.

**Methods:** Transcriptomic and clinical data for GC were derived from TCGA. Next, univariate Cox, LASSO and multivariate Cox regression analyses were next used to identify prognostic MRlncRNAs, calculate risk scores and build risk assessment models. The predictive power of the risk models was then validated by Kaplan-Meier analysis, ROC curves, DCA, C-index, and nomogram. We attempted to effectively differentiate between groups in terms of immune cell infiltration status, ICI-related genes, immunotherapy responses, and common anti-tumor drug sensitivity.

**Results:** A risk model based on 11 MRlncRNAs was developed with an AUC of 0.850, and the sensitivity and specificity of this model in predicting survival probability is satisfactory. The Kaplan-Meier analysis revealed that the low-risk group in the model had a significantly higher survival rate, and the model was highly associated with survival status, clinical features, and clinical stage. Furthermore, the model was verified to be an independent prognostic risk factor, and the low-risk group in the model had a remarkable positive correlation with a variety of immune cell infiltrates. The expression levels of ICI-related genes differed significantly between the different groups. Lastly, immunotherapy responses and common anti-tumor drug sensitivity also differed significantly between different groups.

**Conclusion:** The risk model on the basis of 11-MRlncRNAs can serve as independent predictors of GC prognosis and may be useful in developing personalized treatment strategies for patients.

## KEYWORDS

gastric cancer, m<sup>6</sup>A, long non-coding RNA, immune, immunotherapy, prognostic model

## 1 Introduction

Gastric cancer (GC) is the third leading cause of cancer-related mortality and the fifth most lethal tumor, with an incidence that widely varies across regions, i.e., >70% in developing countries, mainly in East Asia (Smyth et al., 2020; Sung et al., 2021). Just after lung cancer, GC is the second-largest malignancy in China in terms of morbidity and mortality (Chen et al., 2016). Current treatments that have been shown to be effective in gastric adenocarcinoma include systemic chemotherapy, radiotherapy, surgery, immunotherapy, and targeted therapy. However, significant therapeutic strategies are still needed for the less differentiated histologic subtypes of gastric adenocarcinoma (Sitarz et al., 2018; Joshi and Badgwell, 2021). In addition, because of genetic heterogeneity and the absence of novel treatment methods, the prognosis of patients with GC remains dissatisfactory (Yue et al., 2022). Therefore, appropriate therapies should be developed to forecast the survival probability (SP) of GC patients, better detect tumor growth, and enhance treatment results.

N<sup>6</sup>-methyladenosine (m<sup>6</sup>A), the most frequent modification of mRNA in eukaryotes, regulates practically all RNA cycle phases, including transcription, maturation, translation, degradation, and mRNA stability (An and Duan, 2022). The control of pathological and physiological processes, including cancer, can be influenced by m<sup>6</sup>A RNA methylation. Numerous research conducted in recent years has revealed that m<sup>6</sup>A plays a significant role in the regulation of tumors, which further controls the emergence and growth of tumors through manipulating tumor metabolism. Chen et al. (2020) reported that ALKBH5-mediated m<sup>6</sup>A modification of PVT1 facilitates osteosarcoma tumorigenesis, indicating that ALKBH5 and PVT1 could be potential therapeutic targets for osteosarcoma treatment. In other research, it was discovered that the TME and expressions of crucial immunological checkpoints in hepatocellular carcinoma and lung adenocarcinoma had strong connections with m<sup>6</sup>A-related long non-coding RNAs (MRlncRNA) profiles (Li L. et al., 2021b; Xu et al., 2021). However, further investigation of MRlncRNA signatures in GC patients is still needed.

LncRNAs refer to a class of endogenous cellular RNAs that are longer than 200bp and are encoded by the mammalian genome but are unable to create proteins due to the absence of an open reading frame (Yu et al., 2018). LncRNAs exhibit a variety of powerful capabilities in the tumor microenvironment, including tumorigenesis, tumor metastasis, and the development of associated immune diseases (Wu et al., 2020). Several combination therapies, including immunotherapy in combination with chemotherapy, surgery in combination with

chemotherapy, and even drug combination therapies, have achieved significant clinical efficacy and progress (Marmarelis and Aggarwal, 2018). Therefore, it is compelling to consider combining targeted lncRNA and immunotherapy for cancer treatment.

The goal of this work was to create a MRlncRNA-based GC prognostic risk model that can predict the prognosis of the disease and the effectiveness of immunotherapy. We also hoped to gain new knowledge about the function of MRlncRNAs in GC prognosis and immunotherapy efficacy prediction.

## 2 Materials and methods

### 2.1 Data sources

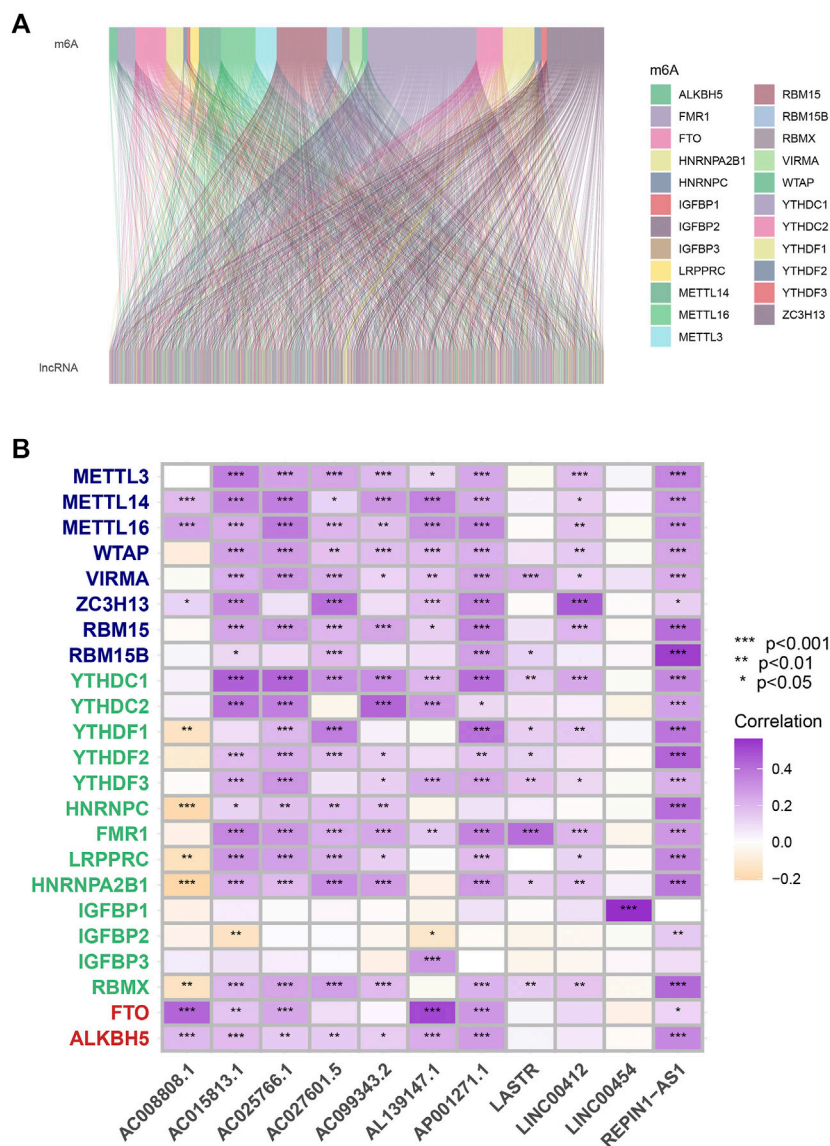
The stomach cancer dataset and the matching clinical information were downloaded from TCGA (<https://tcga-data.nci.nih.gov/tcga>). The expression profiles of mRNAs and lncRNAs were extracted by adding annotations based on the Ensembl database (<http://asia.ensembl.org>). Based on previous literature and databases, we eventually acquired 23 MRGs (MRGs) (Supplementary Table S1) (Wu, 2022; Zheng et al., 2022).

### 2.2 Acquisition of MRlncRNAs and construction of risk model

In this part, MRlncRNAs were recognized applying the Pearson correlation analysis ( $|\text{Pearson } R| > 0.4$  and  $p < 0.001$ ). The entire TCGA set was randomly assigned into a training set and a testing set (ratio, 0.7: 0.3; sample, 224: 94). The specific clinical characteristics of the training and testing sets are shown in Supplementary Table S5. There was no significant difference between the clinical characteristics of the two sets ( $p > 0.05$ ). Next, univariate Cox analysis was employed to recognize prognostic MRlncRNAs ( $p < 0.05$ ), LASSO analysis was applied to distinguish candidate MRlncRNAs, and a risk model was developed by utilizing multivariate Cox analysis.

### 2.3 Verification of prognostic risk model

To verify the prognostic capability of the constructed model, we calculated the risk score for each GC patient using a formula:  $\sum_{i=1}^k \beta_i Si$ , and the samples were categorized as high-or low-risk



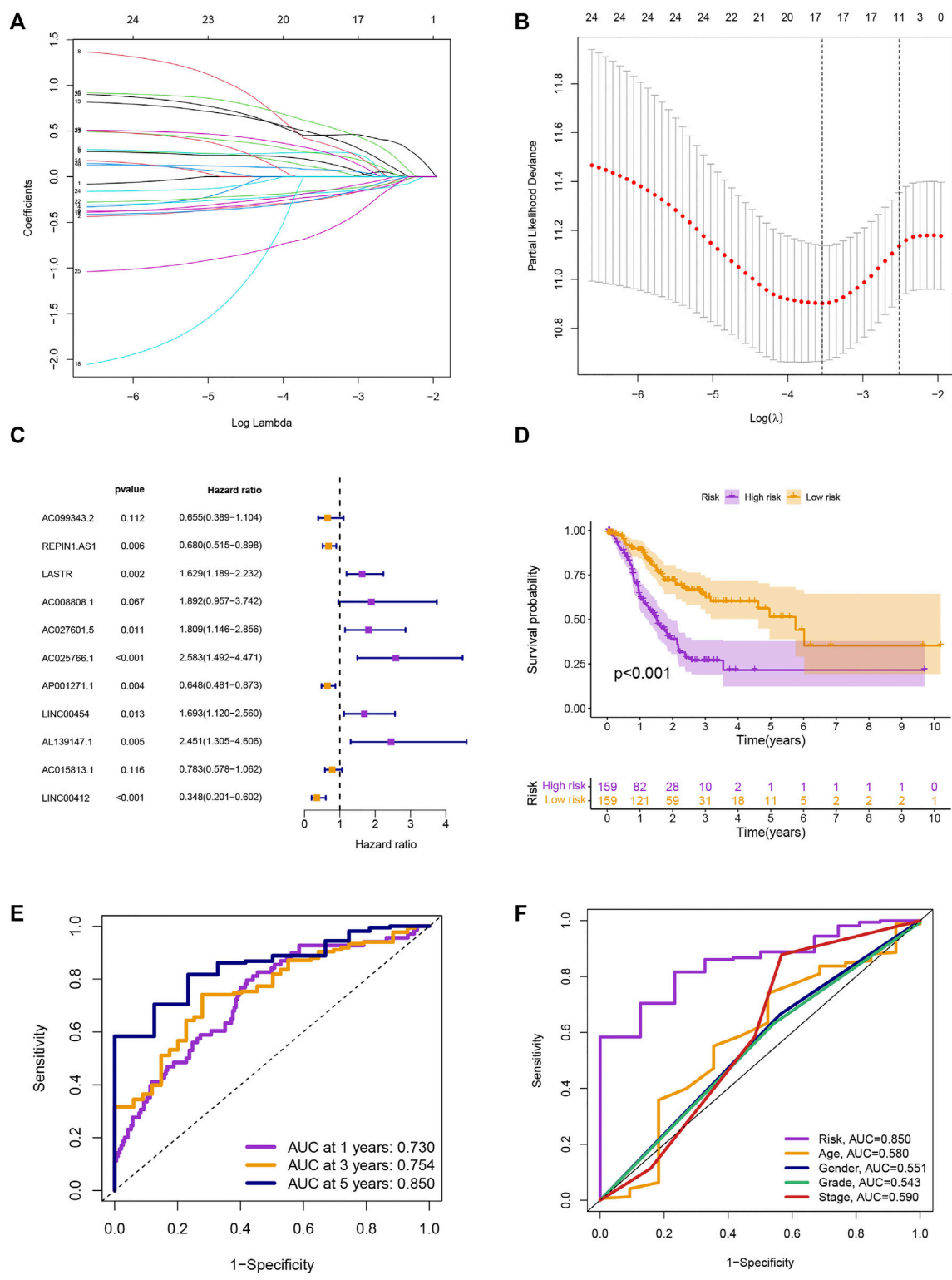
**FIGURE 1**  
Process of MRlncRNAs identification. (A) Sankey relational diagram for all MRlncRNAs and the 23 MRGs. (B) The correlations between 23 MRGs and the 11 prognostic MRlncRNAs.

cohorts based on the median risk score. The Kaplan-Meier technique was utilized to evaluate the ability of this prognostic model to discriminate the survival differences between low-and high-risk groups. The prediction accuracy of the signature for survival in comparison to the conventional clinical characteristics and the known models was estimated by using the time-dependent receiver operating characteristic (ROC) curves and the area under the curve (AUC). On the basis of subgroups divided by clinicopathological traits, we also examined the survival disparities between various groups. Then, univariate and multivariate Cox analyses were conducted in order to verify the model as an independent predictor of prognosis. To evaluate

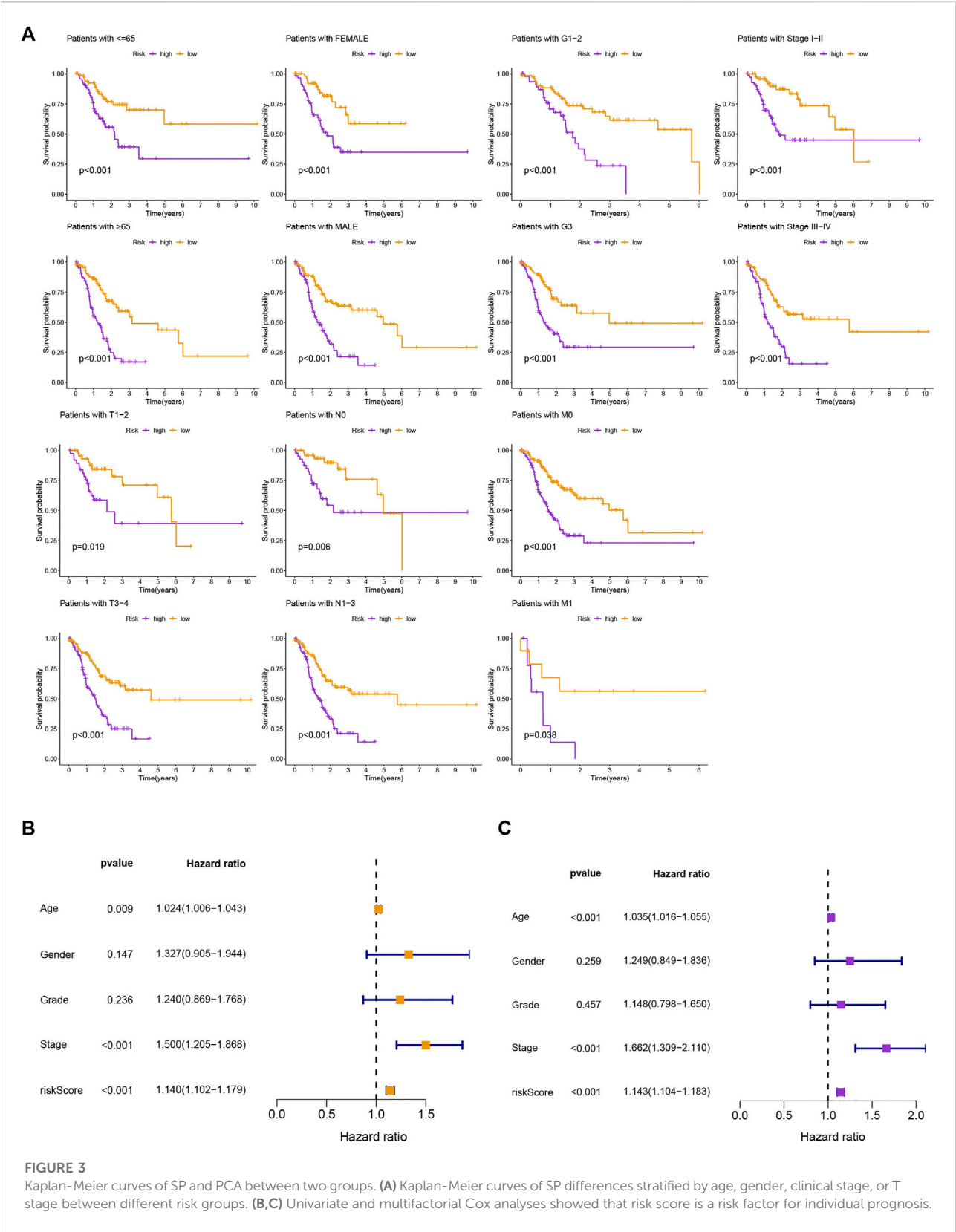
the precision of the signature in comparison to the traditional clinical features, we also adapt decision curve analysis (DCA) and the consistency index (C-index). A nomogram combining the model and clinical features was developed to predict the 1-, 3-, and 5-year SP of patients.

## 2.4 Evaluation of the tumor immune microenvironment landscape

We further investigated the landscape of the tumor immune microenvironment and enrichment level in GC.



**FIGURE 2**  
Development of risk model. (A) The LASSO coefficient profile. (B) Coefficient profile plot. (C) Multivariate Cox regression analysis showed 12 prognostic lncRNAs. (D) Kaplan-Meier curves of the SP in different risk groups. (E) ROC curves to predict the sensitivity of 1-, 3-, and 5-year survival. (F) ROC curves to predict the sensitivity of the risk grade and other clinicopathological characteristics.



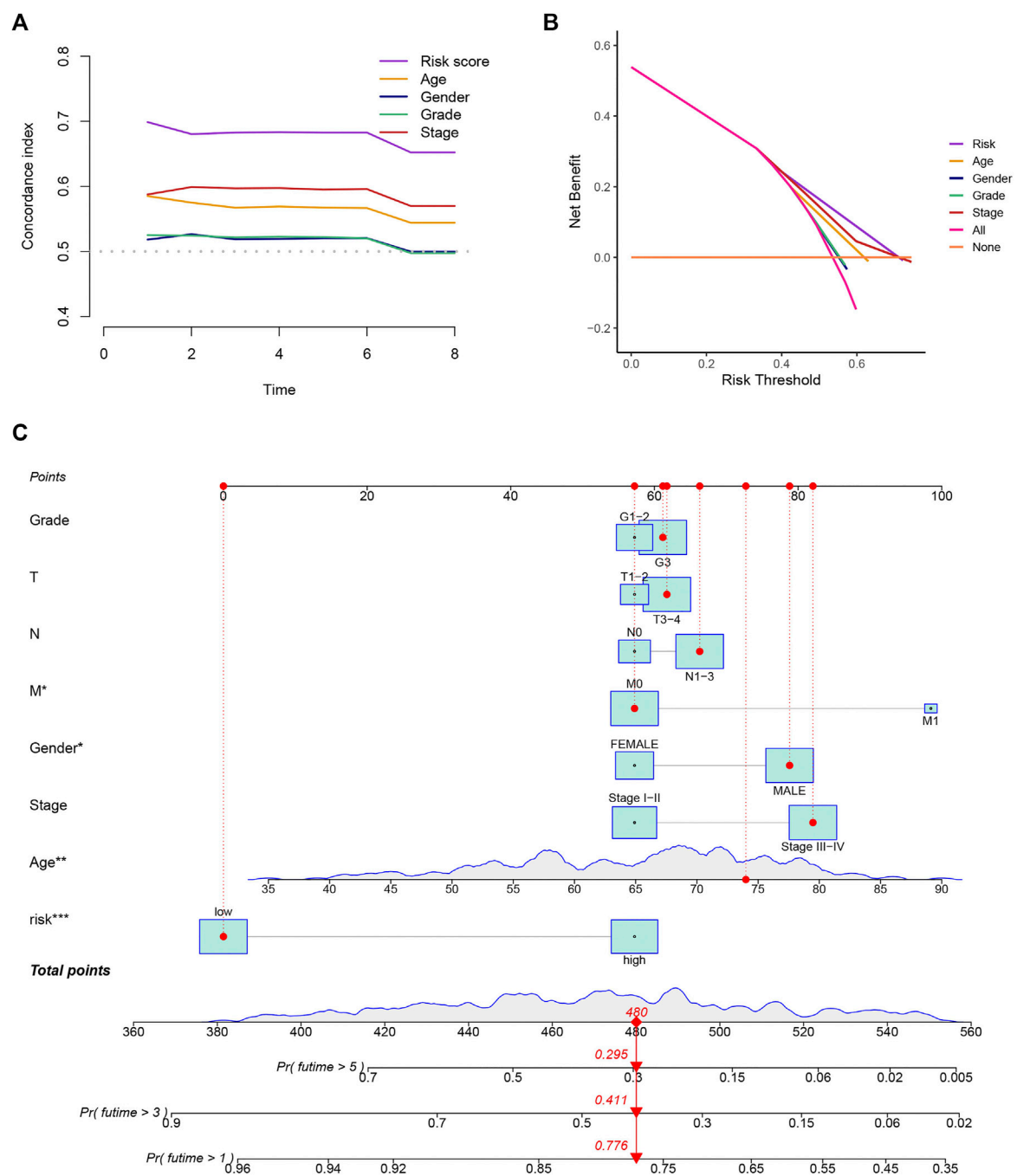
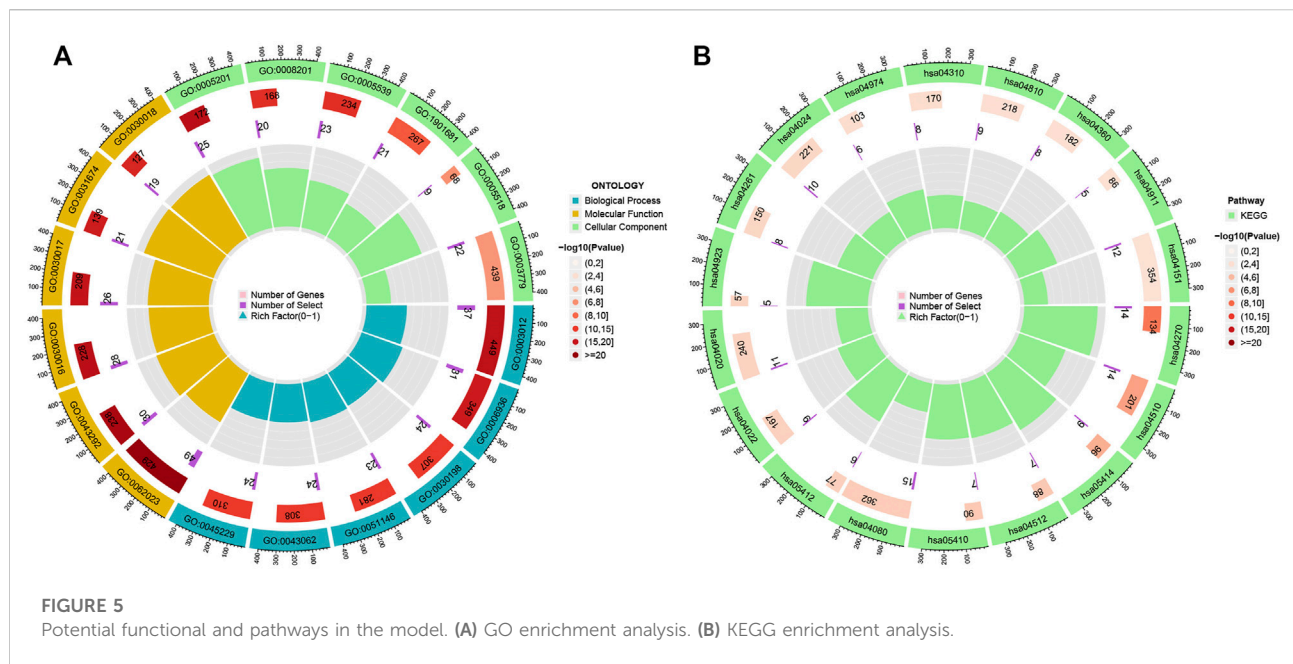


FIGURE 4

Construction of a nomogram. (A,B) The results of the C-index and DCA indicated that the model better predicted the prognosis of GC than other traditional clinical characteristics. (C) The nomogram was constructed according to the risk assessment model and clinical features.

Next, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were carried out to evaluate the potential molecular mechanisms of the risk model. For immune infiltration calculations, we implemented the TIMER, XCELL, QUANTISEQ, MCP-COUNTER, EPIC,

CIBERSORT-ABS, and CIBERSORT algorithms. This allowed us to compare immune cell subpopulations across patients with low-and high-risk. Then, investigating differences in immune function between various groups was performed by using single-sample Gene Set Enrichment Analysis



(ssGSEA). Using the Wilcoxon signed rank test, the expression of immune checkpoint inhibitors (ICIs)-associated molecules in different groups was explored. The amount and quality of gene mutations among various populations were determined using a gene mutation analysis. To predict variations in immunotherapeutic responses among different groups, the tumor mutational burden (TMB) and tumor immune dysfunction and exclusion (TIDE) were calculated as well.

## 2.5 Identification of potential compounds

To evaluate therapy response and explore common anti-tumor drugs for GC treatment in the clinic, we used the R package pRRophetic to calculate the half inhibitory concentration (IC<sub>50</sub>) of drugs and make comparisons in the IC<sub>50</sub> between different groups.

## 2.6 Cell culture

Human gastric mucosal epithelial cells GES-1 was purchased from the Beijing Institute of Cancer Research (Beijing, China) and cultured in RPMI 1640 medium containing 10% fetal bovine serum (FBS, Clark Bioscience, Claymont, United States). Human gastric adenocarcinomic AGS cell line was purchased from the National Collection of Authenticated Cell Cultures (Beijing, China) and cultured in Ham's F-12K (Kaighn's, Thermo Fisher Scientific, Waltham, MA, United States) medium containing 10% FBS. Human

gastric cancer MKN45 cells were purchased from the Beyotime Biotechnology (Nantong, China) and cultured in RPMI 1640 with 10% FBS. All the cells were cultured at 37°C in a 5% CO<sub>2</sub> humidified incubator.

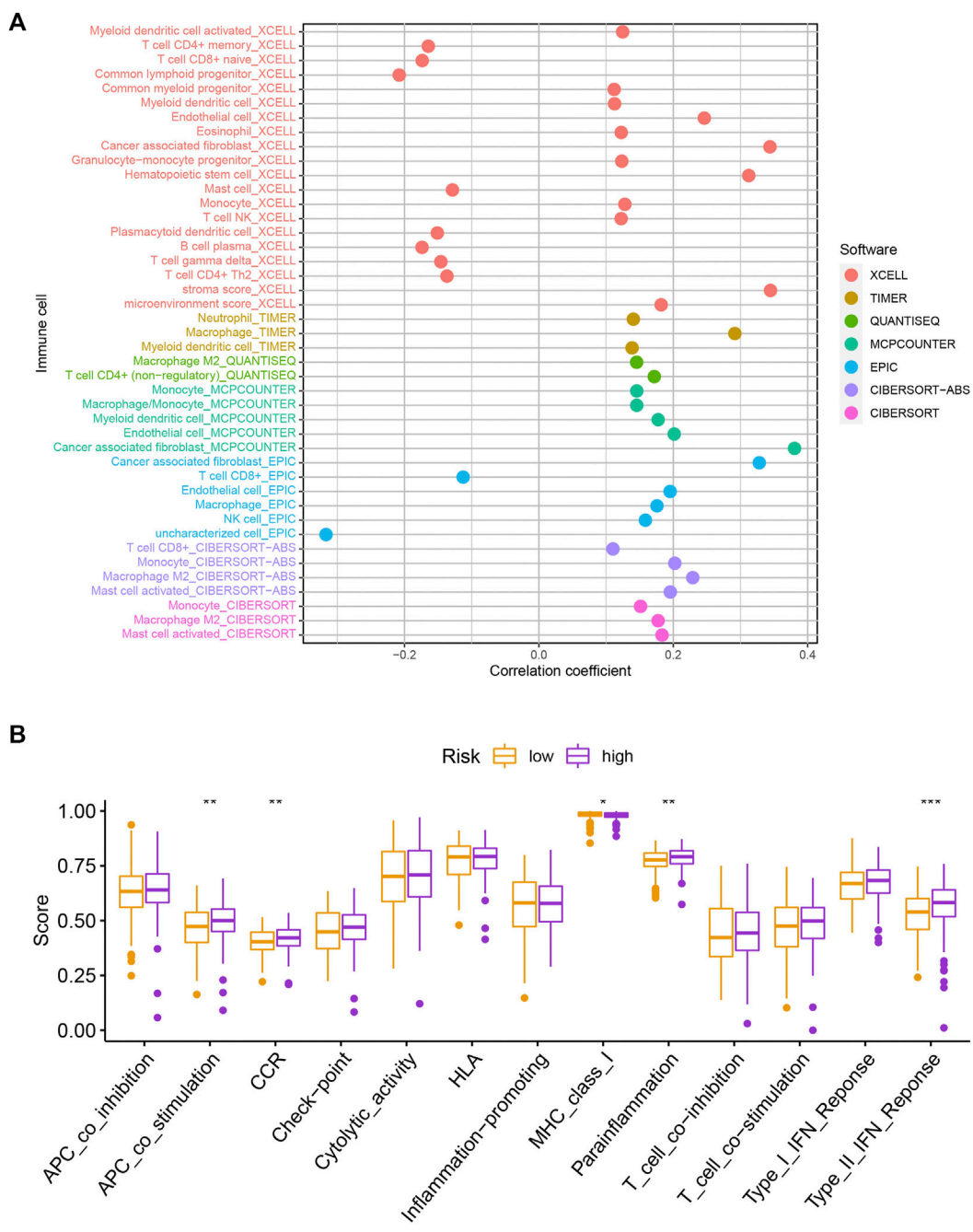
## 2.7 Quantitative real-time PCR analysis

Total RNA was extracted from GES-1, AGS and MKN45 cells using a total RNA extraction kit. Next, 1 µg of total RNA was reverse transcribed into cDNA with the iScript cDNA synthesis kit. A Bio-Rad CFX96 system was used to perform quantitative real-time PCR (qPCR) analysis, and the relative mRNA levels were calculated using the 2<sup>-ΔΔCt</sup> method using GAPDH for normalization. The primer sequences to amplify the genes encoding LASTR, AC008808.1, AC027601.5, AC025766.1, LINC00454 and AL139147.1, and GAPDH are listed in [Supplementary Table S10](#).

## 3 Results

### 3.1 Identification of NRlncRNAs

According to co-expression analysis, 979 MRlncRNAs were recognized (cor > 0.4 and *p* < 0.001) ([Supplementary Table S2](#)). Finally, the m<sup>6</sup>A-lncRNA co-expression network was visualized by using the Sankey diagram in [Figure 1A](#). 11 lncRNAs were selected for the prognostic risk model and, the correlation between MRGs and MRlncRNAs is presented in [Figure 1B](#).

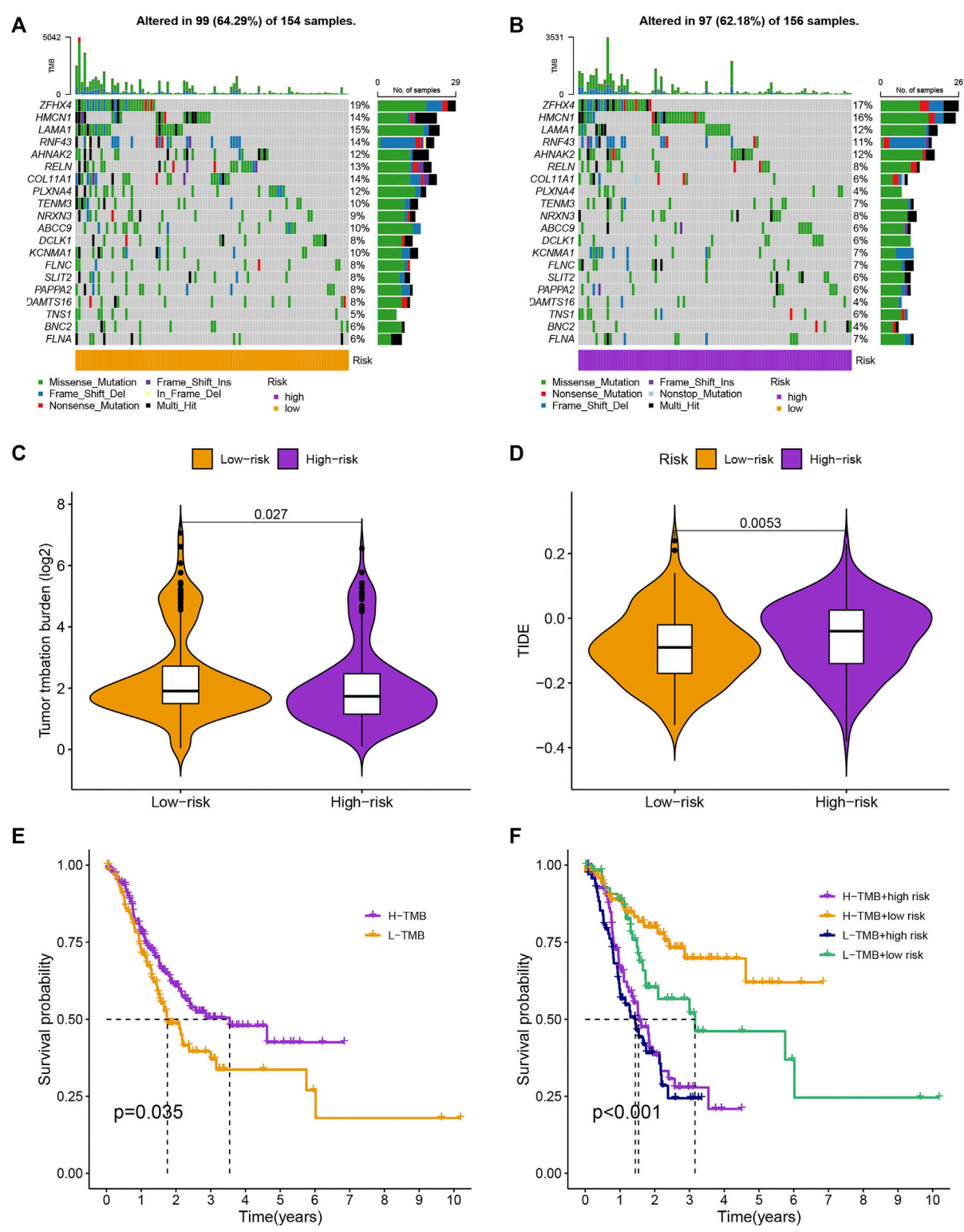


**FIGURE 6**  
Assessment of the immune landscape. **(A)** Changes in immune cell infiltration that are identified by the MRlncRNA-based risk model are linked to an increased chance of developing GC. **(B)** Several immune functions were statistically different between the two groups.

### 3.2 Construction and verification of the risk model

We screened 25 prognostic MRlncRNAs from 979 MRlncRNAs through univariate Cox regression analysis (Supplementary Figure S1A and Supplementary Table S3).

Out of 25 prognostic MRlncRNAs, 17 potential MRlncRNAs were chosen by applying the LASSO analysis (Figures 2A,B). Finally, the multivariate Cox analysis was used to create a prognostic risk model that included 11 MRlncRNAs (Figure 2C and Supplementary Table S4). Based on median risk scores, GC patients were categorized into low-risk and



**FIGURE 7**  
Evaluation of cancer immunotherapy response. (A,B) More genes were mutated in the high-risk group. (C,D) Comparison of the immunotherapy response of the different-risk group to predict TIDE and TMB. (E,F) The SP of patients in different subgroups.

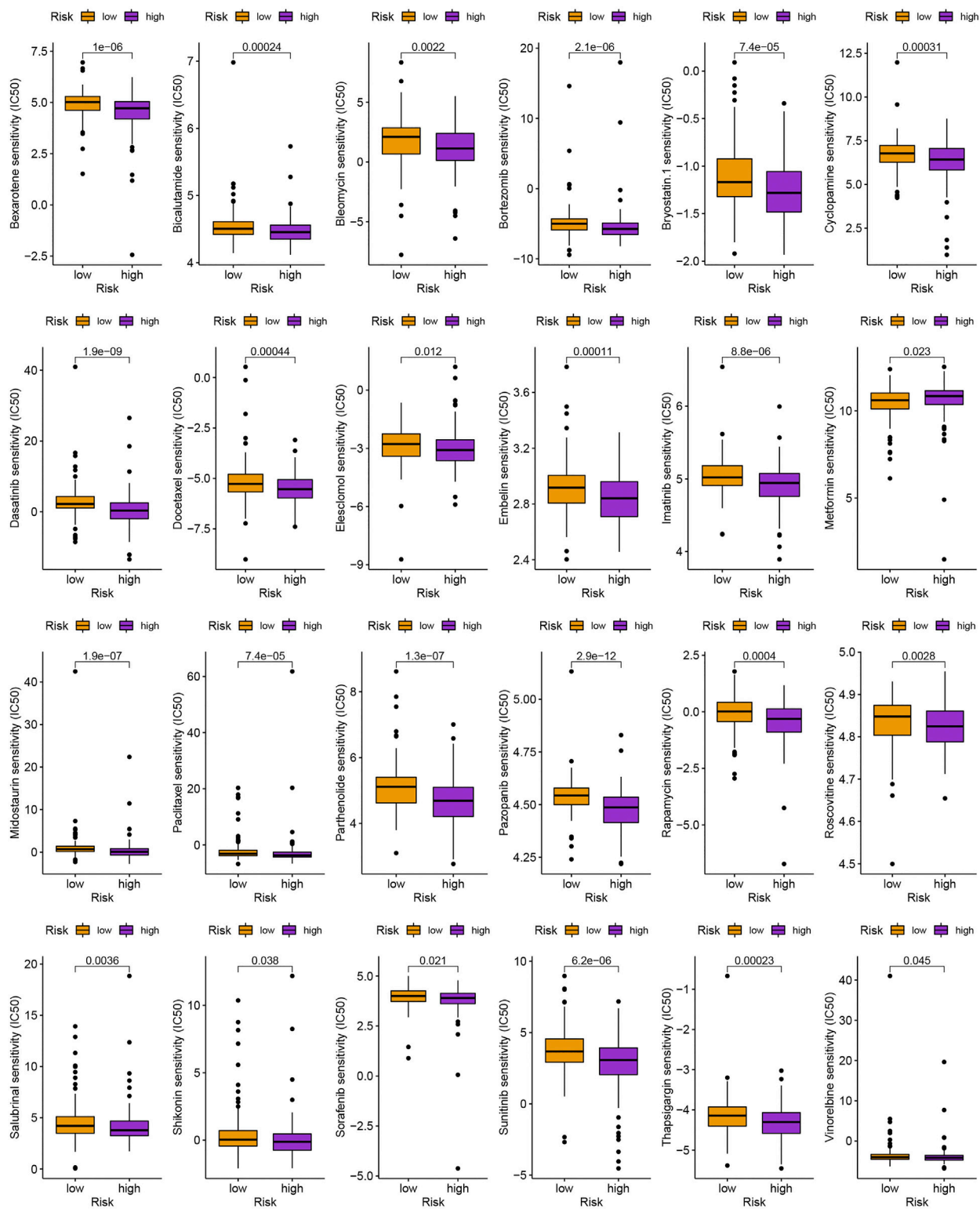
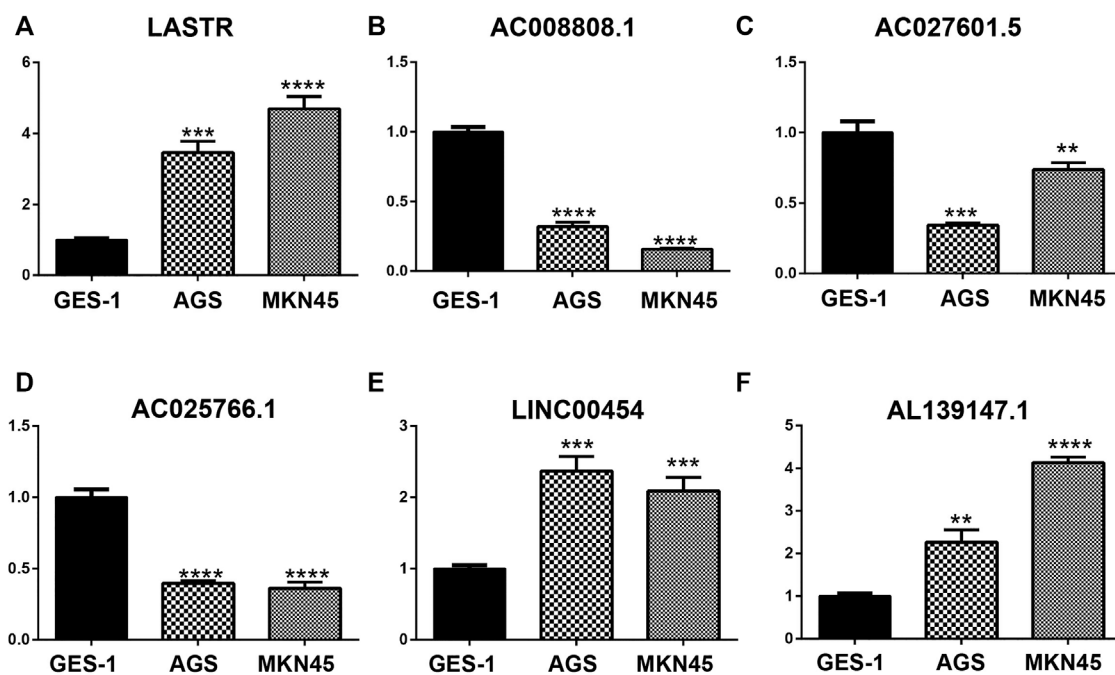


FIGURE 8  
Identification of commonly used anti-tumor drugs targeting the model.

**FIGURE 9**

The mRNA levels of LASTR, LINC00454, and AL139147.1 were increased in AGS and MKN45 cells, whereas the mRNA levels of AC008808.1, AC027601.5, and AC025766.1 were decreased in AGS and MKN45 cells.

high-risk groups. The survival analysis of these two groups suggested that the SP of the low-risk group was higher ( $p < 0.001$ ) (Figure 2D). The 1-year, 3-year, and 5-year survival rates of GC patients are predicted by adopting this risk model, and the prediction accuracy rates are displayed in Figure 2E. The model also demonstrated a higher AUC than other clinicopathological characteristics such as age, gender, grade, and stage, indicating that it was comparably reliable (Figure 2F).

### 3.3 Assessment of the risk assessment model and clinical characteristics

The survival prediction and ROC curves of the testing and entire sets were shown in Supplementary Figures S1B–E, indicating the prediction accuracy of this risk model is satisfactory. In the subgroups separated by age ( $\leq 65$  or  $> 65$ ), gender (female or male), clinical stage (G1-2 or G3, stage I-II or stage III-IV), or TNM stage (T1-2 or T3-4, N0 or N1-3, M0 or M1), the SP was higher in the low-risk group, which indicated that the constructed model was appropriate for various circumstances (Figure 3A). Univariate and multivariate Cox analyses were implemented to evaluate whether this 11 MRlncRNAs risk model had independent prognostic features for GC patients. The hazard ratio of the risk score and the 95% confidence interval were 1.140 and 1.102–1.179

( $p < 0.001$ ) in univariate Cox regression analysis, respectively (Figure 3B). According to multivariate Cox regression analysis, the hazard ratio was 1.143 and the 95% confidence interval was 1.104–1.183 ( $p < 0.001$ ) (Figure 3C), indicating that the risk model of the 11 MRlncRNAs was independent and had nothing to do with clinicopathological features including age, gender, clinical stage, T stage, and risk scores (Supplementary Table S6).

To evaluate the precision of the signature in comparison to the traditional clinical features, we also adopted the C-index and DCA, demonstrating that the signature has a greater ability to forecast the prognosis of GC than other clinical features (Figures 4A,B). We combined this risk model with clinicopathologic characteristics and evaluated patients' total risk scores according to them to build a nomogram to predict 1-year, 3-year, and 5-year survival rates of GC patients (Figure 4C).

### 3.4 Estimation of the immune landscape

To explore the potential functional and pathway differences in different risk groups, we identified 347 differentially expressed genes in high- and low-risk groups for GO and KEGG enrichment analysis. The results of GO and KEGG enrichment analysis are presented in Figures 5A,B, and their details are shown in the Supplementary Tables S7, S8. By examining potential connections between immune cell sub-

populations and GC risk, we explored whether the prognostic MRlncRNA pairings included in the risk model were connected to activities in the tumor immune milieu. We found a statistically significant association between changes in the immune cell landscape and elevated GC risk (Figure 6A and Supplementary Table S9). Statistical differences were found in some immune functions, including parainflammation, response to type II interferon, CCR, and APC stimulation (Figure 6B).

### 3.5 Evaluation of cancer immunotherapy response and drug sensitivity

As is shown in Figures 7A,B, ZFHx4, HMCN1, LAMA1, RNF43, AHNK2, RELN, COL11A1, PLXNA4, TENM3, and NRXN3 were the top 10 most mutated genes for differentially expressed genes between high- and low-risk groups. ZFHx4, LAMA1, and RNF43 mutations were much more high in patients in the low-risk group than in those in the high-risk group, but the exact reverse was seen for the HMCN1 mutation levels. The relationship between the MRlncRNAs model and immunotherapeutic biomarkers was then examined. As expected, we observed a significantly greater response to immunotherapy in the high-risk group than in the low-risk group, indicating that our m<sup>6</sup>A-based classifier score can be applied for TIDE and TMB prediction (Figures 7C,D). Next, based on TMB, we categorized all samples in the high- and low-risk groups into two subgroups: high-TMB and low-TMB, respectively. The patients with low TMB after immunotherapy had a higher SP (Figure 7E). Then, we predicted the SP of patients with different TMB subgroups in the two risk groups separately and found that patients with low TMB in the low-risk group had the highest probability of survival (Figure 7F). This implies that we can select appropriate immunotherapeutic agents for GC patients according to their risk patterns. In drug susceptibility analysis, statistical differences were found in the IC<sub>50</sub> among the 24 chemical or targeted drugs used for GC treatment in the different groups (Figure 8).

### 3.6 Quantitative real-time PCR analysis

The qPCR results showed that the mRNA levels of LASTR, LINC00454 and AL139147.1 were increased in AGS and MKN45 cells, whereas the mRNA levels of AC008808.1, AC027601.5 and AC025766.1 were decreased in AGS and MKN45 cells, which further validates the prognostic risk model we have constructed (Figure 9).

## 4 Discussion

GC is the third most common cause of cancer-related death and the fifth most deadly malignancy, and the prognosis of GC

patients remains unsatisfactory (Sexton et al., 2020). It is crucial to create innovative methods to raise the survival rate of this illness due to the dismal prognosis of patients with late-stage GC (Johnston and Beckman, 2019). Immunotherapy is currently regarded as a cutting-edge treatment option for diseases like breast, stomach, and lung cancer. In addition, lncRNAs have been demonstrated to have a significant role in the regulation of gene expression in various malignancies, including GC (Denaro et al., 2019). Therefore, we created a prognostic model based on MRlncRNAs in GC to predict the prognosis of GC patients, explore TME and cancer immunotherapy responses, and add fresh ideas to the clinical treatment of GC.

In our study, we identified 979 MRlncRNAs from the TCGA to discuss the prognostic function of MRlncRNAs. We identified 000 prognostic MRlncRNAs and a risk model to predict SP in GC patients was built based on 11 MRlncRNAs: AC099343.2, REPIN1, AS1, LASTR, AC008808.1, AC027601.5, AC025766.1, AP001271.1, LINC00454, AL139147.1, AC015813.1, and LINC00412. Of these, AC099343.2 was identified as an autophagy-related lncRNA signature for potential prognostic biomarkers of patients with cervical cancer (Feng et al., 2021); REPIN1, AS1 and LASTR were developed to improve the prognosis prediction of stomach adenocarcinoma patients (Luo et al., 2022). In another study, AP001271.1 and three other lncRNAs were selected to construct a risk model for predicting prognosis for GC patients (Wei et al., 2021). Furthermore, other lncRNAs were found for the first time in this study.

Then, patients were grouped into low- and high-risk groups according to median risk score, and some analyses were made, including Kaplan-Meier analysis, univariate and multivariate Cox analyses. We found that the MRlncRNAs risk model was an independent risk factor of SP. Through ROC analysis, we also discovered that the model was more accurate than conventional clinical features in predicting GC survival. Finally, according to a nomogram developed to make predictions in SP of GC patients, we found that the predicted and measured values for the SPs are highly consistent. On the basis of the above analysis, this risk model based on 11 MRlncRNAs that were independently related to SP was pretty accurate.

TMB, which stands for total number of somatic coding mutations, has gained much attention as a new predictive biomarker that is closely related to the development of neoantigens that trigger anti-tumor response (Allgauer et al., 2018; Addeo et al., 2021). The context of TMB identified at diagnosis represents the immune response and chemotherapy benefit, and variations in the numbers of CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, macrophages, and cancer-associated fibroblasts infiltrating in the TME correlate with clinical outcomes in a variety of malignancies, including GC, melanoma, urothelial cancer, lung cancer, and breast cancer (Zeng et al., 2019; DeBerardinis, 2020). In our study, we discovered that the TMB of the low-risk group was higher

than that of the high-risk group. Furthermore, the TIDE prediction score has also been deployed in numerous investigations, and its ability to predict prognosis has been successfully established (Jiang P. et al., 2018a; Chen et al., 2022). In our study, the TIDE algorithm predicted a more favourable response to immunotherapy in individuals with high-risk subtypes. Based on the aforementioned results, we draw the conclusion that our prediction model could deliver precise immunological biomarkers for oncology. What's more, the conclusions of our study provide information on the molecular biology of lncRNAs that are connected to m<sup>6</sup>A in GC. In this study, the TIDE algorithm predicted that immunotherapy is more effective for patients with the high-risk subtype. We come to the conclusion that our prediction model could provide accurate immunological indicators for oncotherapy based on the aforementioned findings. Additionally, the molecular biological mechanisms of MRlncRNAs in GC are also newly revealed by this work.

In general, TNM stage is the most important determinant of GC prognosis in clinical practice (Jiang Y. et al., 2018b). However, due to tumor heterogeneity, even patients with similar TNM staging exhibit widely varying prognoses (Li K. et al., 2021a). It suggests that extant periodization algorithms are deficient in capturing GC heterogeneity and accurately predicting prognosis in GC patients. Therefore, more research remains to be carried out in investigating potential predictive and therapeutic biomarkers. The study suggests that MRlncRNA models may provide a new tool for GC prognosis prediction. This study used several analytical methods to validate this new model so that we could select the optimal model and apply it in a rational way. We hypothesized that the predictive model would still be feasible without external data validation. However, we are aware that there are flaws and limitations in this study. As the molecular mechanism of MRlncRNAs is not fully understood, it would be sensible to validate this with more convincing basic experiments. And the sample size needs to be expanded in future studies to increase the confidence. Furthermore, we will investigate the role of MRlncRNAs and their interactions with MRGs through *in vitro* experiments, seeking to assess the accuracy of the model in future studies and provide new ideas for clinical treatment.

## 5 Conclusion

In summary, this study reveals that the processes and mechanisms of MRlncRNAs are based on a novel prognostic model that provides new insights into GC prognosis prediction and clinical treatment. Furthermore, the model we developed was accurate and effective in predicting GC prognosis and showed sensitivity in identifying GC patients who responded well to immunotherapy.

## Data availability statement

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

## Author contributions

XZ and CW contributed equally to this work and wrote the paper. YS conceived and designed the manuscript. ZL collected and analyzed the data. All authors agree to publish the article.

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1011716/full#supplementary-material>

### SUPPLEMENTARY FIGURE S1

(A) 25 prognostic MRlncRNAs recognized by univariate Cox analysis. (B–E) Kaplan-Meier survival curves, and 1, 3 and 5-year ROC curves in the testing set and total set.

### SUPPLEMENTARY TABLE S1

23 MRGs.

### SUPPLEMENTARY TABLE S2

The results of MRGs-MRlncRNAs.

### SUPPLEMENTARY TABLE S3

25 prognostic MRlncRNAs were selected through univariate Cox regression analysis.

### SUPPLEMENTARY TABLE S4

11 MRlncRNAs were selected by the multivariate Cox analysis.

### SUPPLEMENTARY TABLE S5

There was no significant difference between the clinical characteristics of the two sets.

**SUPPLEMENTARY TABLE S6**

The details of univariate and multivariate Cox analyses.

**SUPPLEMENTARY TABLE S7**

The details of GO enrichment analysis.

**SUPPLEMENTARY TABLE S8**

The details of KEGG enrichment analysis.

**SUPPLEMENTARY TABLE S9**

The details of immune cell landscape.

**SUPPLEMENTARY TABLE S10**

Primer sequences for quantitative real-time PCR analysis.

## References

- Addeo, A., Friedlaender, A., Banna, G. L., and Weiss, G. J. (2021). TMB or not TMB as a biomarker: That is the question. *Crit. Rev. Oncol. Hematol.* 163, 103374. doi:10.1016/j.critrevonc.2021.103374
- Allgäuer, M., Budczies, J., Christopoulos, P., Endris, V., Lier, A., Rempel, E., et al. (2018). Implementing tumor mutational burden (TMB) analysis in routine diagnostics—A primer for molecular pathologists and clinicians. *Transl. Lung Cancer Res.* 7, 703–715. doi:10.21037/tlcr.2018.08.14
- An, Y., and Duan, H. (2022). The role of m6A RNA methylation in cancer metabolism. *Mol. Cancer* 21, 14. doi:10.1186/s12943-022-01500-4
- Chen, F., Yang, J., Fang, M., Wu, Y., Su, D., and Sheng, Y. (2022). Necroptosis-related lncRNA to establish novel prognostic signature and predict the immunotherapy response in breast cancer. *J. Clin. Lab. Anal.* 36, e24302. doi:10.1002/jcla.24302
- Chen, S., Zhou, L., and Wang, Y. (2020). ALKBH5-mediated m6A demethylation of lncRNA PVT1 plays an oncogenic role in osteosarcoma. *Cancer Cell Int.* 20, 34. doi:10.1186/s12935-020-1105-6
- 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
- DeBerardinis, R. J. (2020). Tumor microenvironment, metabolism, and immunotherapy. *N. Engl. J. Med.* 382, 869–871. doi:10.1056/NEJMcibr1914890
- Denaro, N., Merlano, M. C., and Lo Nigro, C. (2019). Long noncoding RNA s as regulators of cancer immunity. *Mol. Oncol.* 13, 61–73. doi:10.1002/1878-0261.12413
- Feng, Q., Wang, J., Cui, N., Liu, X., and Wang, H. (2021). Autophagy-related long non-coding RNA signature for potential prognostic biomarkers of patients with cervical cancer: A study based on public databases. *Ann. Transl. Med.* 9, 1668. doi:10.21037/atm-21-5156
- Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018a). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24, 1550–1558. doi:10.1038/s41591-018-0136-1
- Jiang, Y., Zhang, Q., Hu, Y., Li, T., Yu, J., Zhao, L., et al. (2018b). ImmunoScore signature: A prognostic and predictive tool in gastric cancer. *Ann. Surg.* 267, 504–513. doi:10.1097/SLA.0000000000002116
- Johnston, F. M., and Beckman, M. (2019). Updates on management of gastric cancer. *Curr. Oncol. Rep.* 21, 67. doi:10.1007/s11912-019-0820-4
- Joshi, S. S., and Badgwell, B. D. (2021). Current treatment and recent progress in gastric cancer. *Ca. Cancer J. Clin.* 71, 264–279. doi:10.3322/caac.21657
- Li, K., Zhang, A., Li, X., Zhang, H., and Zhao, L. (2021a). Advances in clinical immunotherapy for gastric cancer. *Biochim. Biophys. Acta. Rev. Cancer* 1876, 188615. doi:10.1016/j.bbcan.2021.188615
- Li, L., Xie, R., and Lu, G. (2021b). Identification of m6A methyltransferase-related lncRNA signature for predicting immunotherapy and prognosis in patients with hepatocellular carcinoma. *Biosci. Rep.* 41, BSR20210760. doi:10.1042/BSR20210760
- Luo, L., Li, L., Liu, L., Feng, Z., Zeng, Q., Shu, X., et al. (2022). A necroptosis-related lncRNA-based signature to predict prognosis and probe molecular characteristics of stomach adenocarcinoma. *Front. Genet.* 13. doi:10.3389/fgene.2022.833928
- Marmarelis, M. E., and Aggarwal, C. (2018). Combination immunotherapy in non-small cell lung cancer. *Curr. Oncol. Rep.* 20, 55. doi:10.1007/s11912-018-0697-7
- Sexton, R. E., Al Hallak, M. N., Diab, M., and Azmi, A. S. (2020). Gastric cancer: A comprehensive review of current and future treatment strategies. *Cancer Metastasis Rev.* 39, 1179–1203. doi:10.1007/s10555-020-09925-3
- Sitarz, R., Skierucha, M., Mielko, J., Offerhaus, J., Maciejewski, R., and Polkowski, W. (2018). Gastric cancer: Epidemiology, prevention, classification, and treatment. *Cancer Manag. Res.* 10, 239–248. doi:10.2147/CMAR.S149619
- Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C., and Lordick, F. (2020). Gastric cancer. *Lancet* 396, 635–648. doi:10.1016/S0140-6736(20)31288-5
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660
- Wei, J., Zeng, Y., Gao, X., and Liu, T. (2021). A novel ferroptosis-related lncRNA signature for prognosis prediction in gastric cancer. *BMC Cancer* 21, 1221. doi:10.1186/s12885-021-08975-2
- Wu, M., Fu, P., Qu, L., Liu, J., and Lin, A. (2020). Long noncoding RNAs, new critical regulators in cancer immunity. *Front. Oncol.* 10, 550987. doi:10.3389/fonc.2020.550987
- Wu, Z., Zhang, X., Chen, D., Li, Z., Wu, X., Wang, J., et al. (2022). N6-Methyladenosine-Related lncRNAs are potential remodeling indicators in the tumor microenvironment and prognostic markers in osteosarcoma. *Front. Immunol.* 12, 806189. doi:10.3389/fimmu.2021.806189
- Xu, F., Huang, X., Li, Y., Chen, Y., and Lin, L. (2021). m6A-related lncRNAs are potential biomarkers for predicting prognoses and immune responses in patients with LUAD. *Mol. Ther. Nucleic Acids* 24, 780–791. doi:10.1016/j.omtn.2021.04.003
- Yu, W., Wang, H., He, Q., Xu, Y., and Wang, X. (2018). Long noncoding RNAs in cancer-immunity cycle. *J. Cell. Physiol.* 233, 6518–6523. doi:10.1002/jcp.26568
- Yue, Q., Zhang, Y., Bai, J., Duan, X., and Wang, H. (2022). Identification of five N6-methyladenosine-related ncRNA signatures to predict the overall survival of patients with gastric cancer. *Dis. Markers* 2022, 7765900. doi:10.1155/2022/7765900
- Zeng, D., Li, M., Zhou, R., Zhang, J., Sun, H., Shi, M., et al. (2019). Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol. Res.* 7, 737–750. doi:10.1158/2326-6066.CIR-18-0436
- Zheng, D., Yu, L., Wei, Z., Xia, K., and Guo, W. (2022). N6-Methyladenosine-Related lncRNAs are potential prognostic biomarkers and correlated with tumor immune microenvironment in osteosarcoma. *Front. Genet.* 12, 805607. doi:10.3389/fgene.2021.805607



## OPEN ACCESS

EDITED BY  
Hehuang Xie,  
Virginia Tech, United States

REVIEWED BY  
Valentin Vetter,  
Charité Universitätsmedizin Berlin,  
Germany  
Yaping Liu,  
Cincinnati Children's Hospital Medical  
Center, United States

\*CORRESPONDENCE  
Shuxia Wang,  
singlesail@126.com

SPECIALTY SECTION  
This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 29 May 2022  
ACCEPTED 19 August 2022  
PUBLISHED 26 September 2022

CITATION  
Li M, Bao L, Zhu P and Wang S (2022),  
Effect of metformin on the epigenetic  
age of peripheral blood in patients with  
diabetes mellitus.  
*Front. Genet.* 13:955835.  
doi: 10.3389/fgene.2022.955835

COPYRIGHT  
© 2022 Li, Bao, Zhu and Wang. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Effect of metformin on the epigenetic age of peripheral blood in patients with diabetes mellitus

Man Li<sup>1</sup>, Litao Bao<sup>2</sup>, Ping Zhu<sup>1</sup> and Shuxia Wang<sup>1\*</sup>

<sup>1</sup>Department of Geriatrics, The Second Medical Center and National Clinical Research Center for Geriatric Diseases, Chinese PLA General Hospital, Beijing, China, <sup>2</sup>Institute of Gerontology, Second Medical Center, PLA General Hospital, Beijing, China

**Background:** Metformin has been proven to have an antiaging effect. However, studies on how metformin affects global epigenetic regulation and its effect on the epigenetic clock in diabetes mellitus (DM) patients are limited. This study aims to investigate the impact of metformin on the epigenetic age in subjects with type 2 DM.

**Results:** We collected the peripheral blood of the metformin group and the no-metformin group of the 32 DM patients. Three previously established epigenetic clocks (Hannum, Horvath, and DNAmPhenoAge) were used to estimate the epigenetic age acceleration of the two groups. We defined biological age acceleration for each group by comparing the estimated biological age with the chronological age. Results were presented as follows: 1) all three epigenetic clocks were strongly correlated with chronological age. 2) We found a strong association between metformin intake and slower epigenetic aging by Horvath's clock and Hannum's clock.

**Conclusions:** Here, we found an association between metformin intake and slower epigenetic aging.

## KEYWORDS

aging, metformin, epigenetics, DNA methylation, biomarker

## Background

DNA methylation is the most widely studied epigenetic mark, and it is the process by which methyl groups are added or removed from the DNA sequence, usually at cytosine-guanine dinucleotides (CpGs). DNA methylation patterns change over the life in response to environmental factors such as diet, smoking,

**Abbreviations:** AMPK, adenosine activated protein kinase; DM, diabetes mellitus; DMPs, differential methylation positions; mTOR, mammalian target of rapamycin; BMI, body mass index; BCC, blood cell composition; EAA, epigenetic age acceleration; CpGs, cytosine-guanine dinucleotides; NAD, nicotinamide adenine dinucleotide; NADH, reduced nicotinamide adenine dinucleotide.

and stress and also with age (Rakyan et al., 2010; Bell et al., 2012; Maegawa et al., 2010). For example, while hypomethylation is common with aging, some CpG islands and gene-rich regions become hypermethylated with age (Benayoun et al., 2015). Based on age-related changes in DNA methylation, several research groups have identified what are known as DNA methylation clocks (Horvath, 2013; Nwanaji-Enwerem et al., 2018; Wagner, 2017). The epigenetic clocks have been established to be predictive of all-cause mortality (Marioni et al., 2015a; Chen et al., 2016), cancer (Kresovich et al., 2019; Yang et al., 2016), frailty, and cognitive and physical functioning (Marioni et al., 2015b). These DNA methylation clocks have been identified to predict chronological age with high accuracy and are considered as the most promising marker of aging (Lu et al., 2019). Age acceleration, a discrepancy between DNAm age and chronological age, tells us whether the person is biologically younger or older compared to his/her chronological age (Gibson et al., 2019; El Khoury et al., 2019). Accelerated epigenetic aging has been proven to be associated with many aging-related and other diseases, including cancer, Down's syndrome, physical and cognitive decline, and all-cause mortality (Marioni et al., 2015b; Horvath et al., 2015; Marioni et al., 2015a; Daniel et al., 2018).

Metformin is a widely used medication that has been used as the first-line oral treatment for type 2 diabetes (Inzucchi, 2002). Recent advances revealed that this drug, in addition to its glucose-lowering action, might be a promising target for aging (Barzilai et al., 2016; Pernicova and Korbonits, 2014; Martin-Montalvo et al., 2013; Johnson et al., 2013; Cabreiro et al., 2013). It appears to target a number of age-related mechanisms. At a molecular level, metformin leads to the activation of adenosine-activated protein kinase (AMPK) and increases antioxidant protection (Pernicova and Korbonits, 2014; Martin-Montalvo et al., 2013). Metformin could exert the inhibition of the mammalian target of rapamycin (mTOR) signaling (Nair et al., 2014). Inhibition of this pathway extends the lifespan in model organisms and confers protection against a growing list of age-related pathologies (Johnson et al., 2013). Preclinical studies of metformin suggest that metformin robustly increases the lifespan in *C. elegans* by up to 36% (Cabreiro et al., 2013). There is a lot of evidence that metformin is a promising antiaging drug. However, studies on how metformin affects global epigenetic regulation and its effect on the epigenetic clock are limited. The aim of the study was to investigate the metformin-induced antiaging effect and its effects on the genome-wide DNA methylation in human peripheral blood. We conducted this study to investigate the pathways of metformin in real-life physiological conditions in humans. This is important given the polypotent effects of metformin, and such research could lead to new and important targets not only for the treatment of DM but also for other diseases.

## Methods

### Study population

Diabetes mellitus (DM) was defined as the presence of diabetes symptoms and a resting plasma glucose concentration  $\geq 200$  mg/dl, a fasting plasma concentration  $\geq 126$  mg/dl, a 2 h plasma glucose concentration  $\geq 200$  mg/dl in a 75 g oral glucose tolerance test or use of a hypoglycemic agent or other medications for DM. The candidates were excluded if they had severe heart failure, active infectious disease, history of malignancy, or end-stage renal disease, or were in a deep coma. Patients were also excluded if they have any other chronic disease other than diabetes, such as cardiovascular disease, respiratory disease, tumor, and rheumatic immune disease, confirm that the participants did not use antibiotics, immunosuppressive medications, corticosteroids, or pharmaceutical-grade probiotics during at least 2 months leading up to blood collection. Also, confirm that no subject had diarrhea within 7 days leading into the study. All the participants reported that they were on a normal diabetic diet. These patients were divided into two groups based on whether they were on metformin medication with a stable dosage of 0.5 g/d for at least 5 years. The characteristics of the study population were summarized in Table 1. None of the two groups included in the study had any disease other than diabetes and neither group had received any other medications in the past 5 years. All participants provided written informed consent and the study was approved by the ethics board of the Chinese PLA general hospital.

### Genomic DNA extraction and quality control

Fasting peripheral blood samples were collected in the morning and then stored at  $-80^{\circ}\text{C}$  until use. Approximately, 500 ng of genomic DNA from each sample was used for sodium bisulfite conversion using the EZ DNA Methylation-Gold Kit (Zymo Research, United States) in accordance with the manufacturer's instructions. DNA from a total of 32 participants was assayed using Illumina's Infinium human methylation 850 Beadchip as previously described. Genome-wide DNA methylation was assessed using the Illumina Infinium human methylation 850K BeadChip (Illumina Inc, United States) which covers 99% of all RefSeq genes and contains 853,307 sites. The process was performed according to the manufacturer's protocols. Quality control of methylation data including intensity readouts, filtering, and cell composition was adapted from Hannon et al. (2016) and was done using the R package (version 4.0.0). Load EPIC's original IDAT files and filter out probe sites according to the following principles: 1) filter out probes used to filter  $p$  values ( $\geq 0.01$ ) 2). In more than 5% of samples, the probes in beads smaller than three were filtered out. 3) Filter out all non-CpG probes

TABLE 1 Characteristics of the study group.

**Characteristics of the study group**

| Characteristic                                | Total              | No-metformin group | Metformin group    | <i>p</i> value |
|---|--------------------|--------------------|--------------------|----------------|
| Male  | 100%               | 100%               | 100%               | -              |
| Age, years, mean $\pm$ SD                     | 73.3 $\pm$ 5.60    | 73.6 $\pm$ 5.90    | 72.9 $\pm$ 6.30    | 0.56           |
| Smoking, n (%)                                | 12(37.50)          | 5(31.25)           | 7(43.75)           | 0.09           |
| BMI and mean $\pm$ SD                         | 25.3 $\pm$ 3.40    | 25.2 $\pm$ 2.90    | 25.4 $\pm$ 3.70    | 0.98           |
| Fasting plasma glucose, mmol/l, mean $\pm$ SD | 6.27 $\pm$ 2.70    | 6.19 $\pm$ 2.08    | 6.31 $\pm$ 3.12    | 0.12           |
| HbA1c, %                                      | (5.35 $\pm$ 1.07)% | (5.65 $\pm$ 0.98)% | (5.24 $\pm$ 1.20)% | 0.65           |

contained in the dataset. 4) Filter all SNP-related probes. 5) Filter all probes in chromosomes X and Y.

## Determinations of epigenetic age and epigenetic age acceleration

The array data (IDAT files) was analyzed using the ChAMP package in R for deriving the methylation level. The methylation status of all the probes was denoted as  $\beta$  value, which is the ratio of the methylated probe intensity to the overall probe intensity (sum of methylated and unmethylated probe intensities plus constant  $\alpha$ , where  $\alpha = 100$ ). CpG site (probe) intensities were transformed to  $\beta$  values with a standard equation in which beta is the ratio of the methylated probe (m) intensities to the overall intensities ( $m + u + \alpha$ , where  $\alpha$  is the constant offset, 100, and  $u$  is the unmethylated probe intensity). The resulting  $\beta$  values ranged from 0 (completely unmethylated) to 1 (fully methylated). It is generally believed that a  $\beta$  value greater than 0.8 is hypermethylated, while a  $\beta$  value less than 0.2 is hypomethylated, and a  $\beta$  value in the range of 0.2–0.8 is partially methylated. The DNA methylation ages were measured using the epigenetic clock (Horvath, 2013). The epigenetic clock was defined as an age prediction method based on DNA methylation levels at 353 CpG sites. The epigenetic age acceleration value was calculated by subtracting the actual chronological age from the DNAm age (Horvath, 2013). Another measure of acceleration (acceleration residual) equaled the residual resulting from linear regressing the DNAm age on the chronological age (Horvath, 2013). Using the processed DNAm data, DNAm age was calculated using the R code provided by the clock developers (Hannum et al., 2013; Horvath, 2013; Levine et al., 2018).

## Statistical analysis

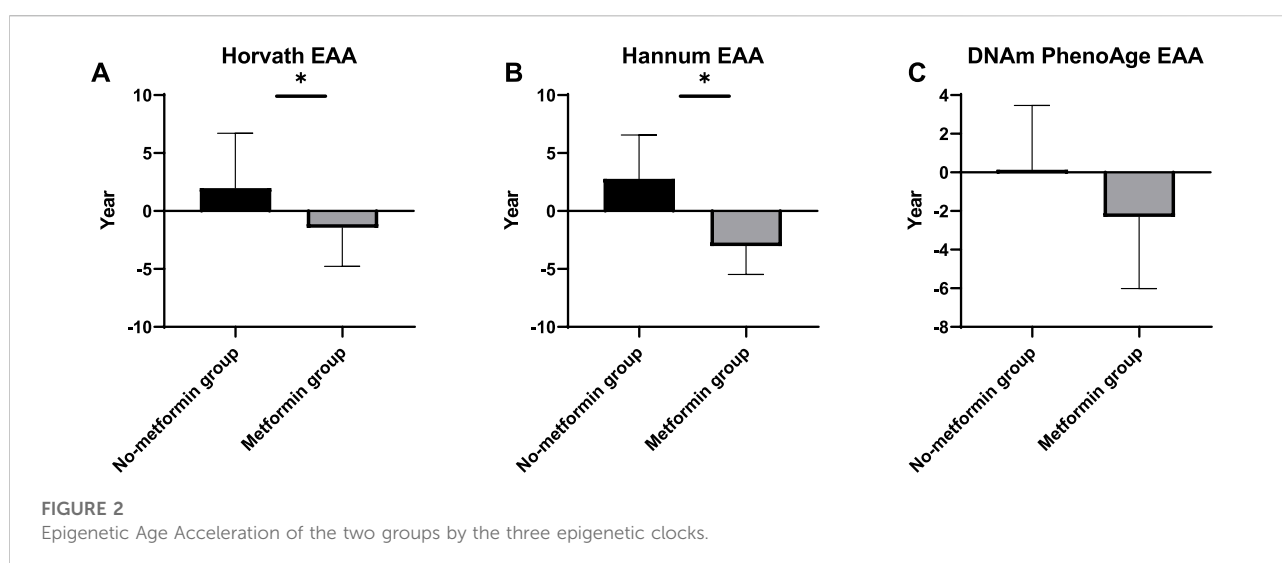
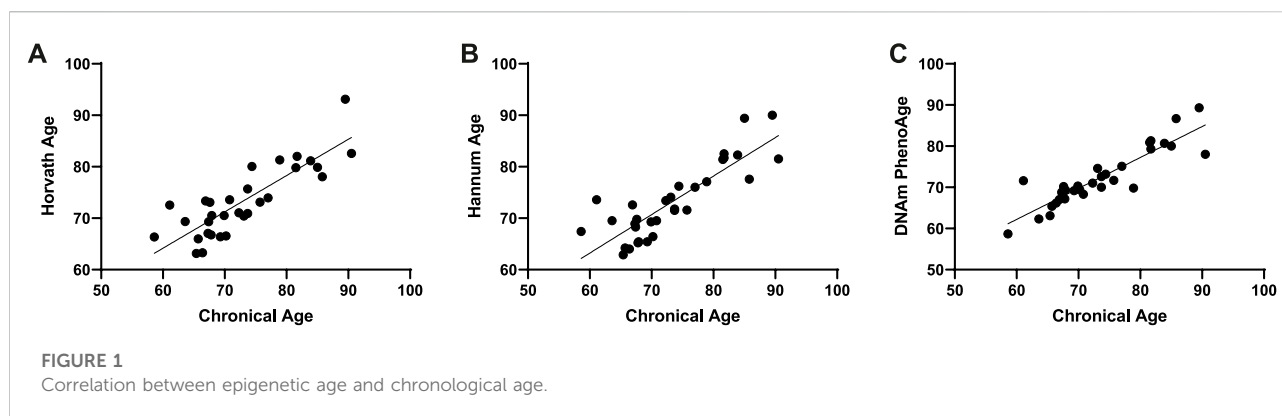
Baseline characteristics were conducted using Stata 14 software (StataCorp, Inc., College Station, TX,

United States). The data are presented as averages and standard deviations unless otherwise stated. Associations between chronological age and DNAm age were analyzed using standardized regression coefficients. The Kruskal–Wallis test was performed to determine the significant difference between the metformin group and the control group. For each of the three epigenetic clocks, we defined epigenetic age acceleration by regressing the DNAm age to chronological age and calculating the difference between observed and fitted DNAm age as epigenetic age acceleration (EAA). We calibrated the fitted DNAm age for the entire analyzed sample and restricted it to a random subcohort; we did not observe meaningful differences using either method; therefore, we used the DNAm age fit to the full dataset. We estimated age acceleration with and without adjustment for blood cell composition (BCC), as it varies with age (Chen et al., 2016). Pearson correlations were used to examine DNAm age and epigenetic age acceleration metric correlations with chronological age. We adjusted for confounding by age, body mass index (BMI; kg/m<sup>2</sup>, continuous), and smoking because they may influence DNAm age. R Studio was used to perform the statistical analysis.

## Results

### Characteristics of the study participants

All subjects were male, with an average of 73.3 years. Based on the medications of metformin, the patients were divided into the no-metformin group and the metformin group. A total of 16 DM patients with at least 5-year of medications with a stable dosage of 0.5 g/d were included in the metformin group and the other 16 DM patients with no metformin medications in the past 5 years were included in the no-metformin medication group. A total of 32 patients were included in the study. The characteristics of the study population are summarized in Table 1.



## Prediction of epigenetic age using the epigenetic clock

We used an Illumina Infinium 850k array to evaluate the effect of metformin on DNA methylation in patients. To verify the epigenetic clock predictors in our cohort, we correlated epigenetic age with chronological age, as described earlier. As expected, all three epigenetic clocks were strongly correlated with chronological age (Hannum: Pearson  $r = 0.88$ ,  $p < 0.001$ ; Horvath: Pearson  $r = 0.85$ ,  $p < 0.001$ ; DNAmPhenoAge: Pearson  $r = 0.83$ ,  $p < 0.001$ ) (Figures 1A–C).

## The association between metformin intake and slower epigenetic aging in patients with DM

We evaluate age acceleration developed by the three clocks. *t*-tests were conducted to verify whether the metformin intake was associated

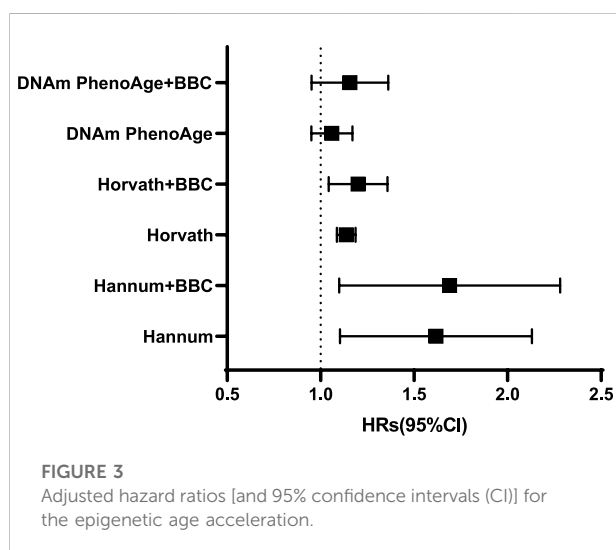
with slower epigenetic aging in the peripheral blood of the patients with DM. The meta-analyzed estimates for EAA, adjusted for chronological age, sex, BMI, and smoking variables, exhibited statistically significant associations for metformin intake and EAA measures with the exception of the DNAmPhenoAge clock. A positive (negative) value of accelerated age indicates that the DNAm age is older (younger) than its actual age. The Horvath EAA of the no-metformin group exceeded that of the metformin group by 2.77 years ( $p = 0.04$ ) (Figure 2A). The Hannum EAA of the no-metformin group exceeded that of the metformin group by 3.43 years ( $p = 0.04$ ) (Figure 2B). The DNAmPhenoAge EAA values did not show significant differences between the two groups, but these differences show a similar trend to accelerating differences (Figure 2C).

## Multivariable adjustment

The meta-analyzed estimates for EAA, adjusted for chronological age, sex, race, and smoking variables, exhibited

**TABLE 2** Epigenetic age acceleration metrics and metformin intake using the full sample of DM patients.

| Epigenetic clocks  | HR (95%CI)      | P      |
|--------------------|-----------------|--------|
| Hannum             | 1.75(1.05–2.05) | < 0.01 |
| Hannum + BBC       | 1.87(1.03–2.17) | < 0.01 |
| Horvath            | 1.19(1.09–1.13) | 0.04   |
| Horvath + BBC      | 1.23(1.03–1.34) | 0.02   |
| DNAmPhenoAge       | 1.06(0.95–1.17) | 0.14   |
| DNAmPhenoAge + BBC | 1.21(0.93–1.22) | 0.1    |



statistically significant associations for metformin intake. After adjusting for other risk factors in the multivariable model, statistical significance persisted for EAA measures for two clocks: the Horvath clock (HR, 1.19 [95% CI, 1.09–1.31];  $P < 0.05$ ) and the Hannum clock (HR, 1.75 [95% CI, 1.05–2.05];  $P < 0.05$ ; Table 2). The meta-analyzed point estimate for Hannum EAA had the largest magnitude association (hazard ratio [HR], 1.75 [95% CI, 1.05–2.05]) (Figure 3). Although the meta-analyzed measures of association for EAA measures for the DNAmPhenoAge EAA were not statistically significant after multivariable adjustment, their directionality remained consistent relative to the associations observed on unadjusted analysis.

## Discussion

We performed a genome-wide methylation study to investigate the effect of metformin on DNA methylation using baseline blood samples. We demonstrated that metformin intake was associated with slower epigenetic aging. In the present study,

three clocks were used to determine the effect of metformin on DM patients. Our research showed that Hannum EAA had the strongest association with metformin intake. Therefore, our study strengthens the view reported earlier that the epigenetic clock is a relatively accurate reflection of a person's biological age.

## Metformin and aging

Metformin is an approved drug to treat diabetes, but it also seems to target some of the mechanisms associated with aging. Particularly for aging, metformin can lead to decreased insulin levels, a decreased IGF-1 signaling pathway (Yang et al., 2020; Admasu et al., 2018), and mTOR inhibition (Vazirpanah et al., 2019; Howell et al., 2017), inhibition of mitochondrial complex 1 in the electron transport chain and reduction of endogenous reactive oxygen species (ROS) production (Nguyen et al., 2019; Bridges et al., 2014), activation of AMPK (Howell et al., 2017; Duca et al., 2015), and a decrease of DNA damage (Zhang et al., 2016; Na et al., 2013). It has been reported that metformin increases the lifespan of *C. elegans* by altering microbial folate and methionine metabolism (Cabreiro et al., 2013). It also improves the health span and lifespan of mice and humans (Martin-Montalvo et al., 2013; Forslund et al., 2015). Moreover, metformin has been proven to interact with several known longevity pathways such as dietary restriction (DR) (Dhahbi et al., 2005; Madeo et al., 2019).

## The epigenetic clock

Recent evidence suggests that the epigenetic clock is the most promising marker of aging. The epigenetic clock has been reported to track biological aging associated with morbidity and mortality. The result of a meta-analysis including 13,089 participants showed that epigenetic clocks can predict all-cause mortality (Chen et al., 2016). Epigenetics clock is considered to be the most promising biomarker of biological age when compared with telomere length and other biomarkers (Jylhävä et al., 2017). Similarly, by comparing different estimation methods, the researchers identified DNA methylation as the most promising biomarker for predicting age (Lee et al., 2016). Therefore, many antiaging measures use epigenetic age to evaluate the effectiveness of interventions (Fahy et al., 2019; Sae-Lee et al., 2018). Metformin has long been considered as an “anti-aging” drug, based on preclinical experiments with lower-order organisms and numerous retrospective data (Barzilai et al., 2016). However, the molecular mechanisms remained unclear and the underlying mechanisms need to be better understood. Previous studies have reported the effect of metformin on epigenetics (Banerjee et al., 2016; Zhong et al., 2017; García-Calzón et al., 2017). In a small sample size study ( $n = 12$ ), Elbere et al. (2018) showed an altered blood DNA methylation profile following the use of metformin in nondiabetic participants. However, studies focused

on metformin's effect on the epigenetic clock are limited. The analysis of epigenetic differences between elderly diabetic patients with and without metformin is helpful to find possible intervention targets.

## The underlying mechanisms: metformin and AMPK

Our study confirms that epigenetic ages are younger in DM patients with medication of metformin. AMPK plays a major regulatory role in cell energy homeostasis by directly phosphorylating metabolic enzymes and nutrient transporters, and indirectly promoting mitochondrial biogenesis and the deactivation of nuclear genes in functional mitochondrial biogenesis. AMPK as a target for promoting healthy aging is associated with its role in multiple signaling pathways. 1) TOR pathway: downregulation of the TOR pathway extends the lifespan in *C. elegans*, fruit flies, and mice (Johnson et al., 2013). The prevailing view of AMPK/TOR interaction is that AMPK is primarily an upstream inhibitor of TOR (Inoki et al., 2003; Gwinn et al., 2008). 2) FOXOs pathway: rIIS is the most powerful and least controversial candidate to promote healthy aging, rIIS dramatically increases life expectancy and extends healthy aging in a variety of organisms, including mammals. The only member of the FOXO transcription factor family that is activated by rIIS completely requires rIIS-mediated longevity. In *C. elegans*, AMPK might activate FOXO, thus performing the antiaging effect. 3) Sirtuins pathway: SIRT1 gene plays an antiaging role by improving efficiency in inducing and maintaining pluripotent states. AMPK can activate SIRT1 by changing the nicotinamide adenine dinucleotide: reduced nicotinamide adenine dinucleotide (NAD: NADH) ratio. To sum up, AMPK influences the aging process through a variety of pathways (Ong and Ramasamy, 2018).

It has been reported that AMPK activator metformin leads to increased trimethylation of H3K79 and regulates mitochondrial biogenesis and senescence through H3K79 methylation (Karnewar et al., 2018). AMPK-mediated phosphorylation resulted in the activation of histone acetyltransferase 1 (HAT1) (Marin et al., 2017). H3K4me3 antagonizes the HIR/Asf1/Rtt106 repressor complex to promote histone gene expression and extend chronological lifespan. New research indicated that changes in AMPK phosphorylation following the changes in histone 3 (H3K9) acetylation and methylation status (Dziewulska et al., 2020).

## Limitations

Here, we found an association between metformin intake and slower epigenetic aging of the DM patients, and metformin is a potentially promising antiaging drug; there are still some limitations to our study. The sample size is relatively small and limited to a single center. There are still some lifestyle factors that can influence DNA methylation, such as any potential difference between diet and

physical activity levels among the groups. Although all of the participants reported that they were on a normal diabetic diet, we know that the Chinese diet is complex and these factors cannot be compared in most cases because we try to avoid the effects of diet on DNA methylation levels.

## Conclusion

We found an association between metformin intake and slower epigenetic aging in DM 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 in the article/Supplementary Material.

## Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Chinese People's Liberation Army General Hospital. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

SW and PZ contributed substantially to the conception or design of the work. ML contributed to data collection, data interpretation, critical review of the manuscript, and drafting of the manuscript. LB contributed to data collection. All authors read and approved the final manuscript.

## Funding

This work was supported by the National Key Research and Development Program of China (2020YFC2008900), National Defense Science and Technology Innovation Special Zone Project (19-163-15-ZD-009-001-10), and the Key Projects of Logistics Scientific Research Project of the Chinese PLA (19BJZ30).

## 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

## References

- Admasu, T. D., Chaithanya Batchu, K., Barardo, D., Ng, L. F., Lam, V. Y. M., Xiao, L., et al. (2018). Drug synergy slows aging and improves healthspan through IGF and SREBP lipid signaling. *Dev. Cell.* 47, 67–79. doi:10.1016/j.devcel.2018.09.001
- Banerjee, P., Surendran, H., Chowdhury, D. R., Prabhakar, K., and Pal, R. (2016). Metformin mediated reversal of epithelial to mesenchymal transition is triggered by epigenetic changes in E-cadherin promoter. *J. Mol. Med.* 94, 1397–1409. doi:10.1007/s00109-016-1455-7
- Barzilai, N., Crandall, J. P., Kritchevsky, S. B., and Espeland, M. A. (2016). Metformin as a tool to target aging. *Cell. Metab.* 23, 1060–1065. doi:10.1016/j.cmet.2016.05.011
- Bell, J. T., Tsai, P.-C., Yang, T.-P., Pidsley, R., Nisbet, J., Glass, D., et al. (2012). Epigenome-wide scans identify differentially methylated regions for age and age-related phenotypes in a healthy ageing population. *PLoS Genet.* 8, e1002629. doi:10.1371/journal.pgen.1002629
- Benayoun, B. A., Pollina, E. A., and Brunet, A. (2015). Epigenetic regulation of ageing: Linking environmental inputs to genomic stability. *Nat. Rev. Mol. Cell. Biol.* 16, 593–610. doi:10.1038/nrm4048
- Bridges, H. R., Jones, A. J. Y., Pollak, M. N., and Hirst, J. (2014). Effects of metformin and other biguanides on oxidative phosphorylation in mitochondria. *Biochem. J.* 462, 475–487. doi:10.1042/BJ20140620
- Cabreiro, F., Au, C., Leung, K.-Y., Vergara-Irigaray, N., Cochemé, H. M., Noori, T., et al. (2013). Metformin retards aging in *C. elegans* by altering microbial folate and methionine metabolism. *Cell.* 153, 228–239. doi:10.1016/j.cell.2013.02.035
- Chen, B. H., Marioni, R. E., Colicino, E., Peters, M. J., Ward-Caviness, C. K., Tsai, P.-C., et al. (2016). DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging* 8, 1844–1865. doi:10.18632/aging.101020
- Daniel, S., Nylander, V., Ingerslev, L. R., Zhong, L., Fabre, O., Clifford, B., et al. (2018). T cell epigenetic remodeling and accelerated epigenetic aging are linked to long-term immune alterations in childhood cancer survivors. *Clin. Epigenetics* 10, 138. doi:10.1186/s13148-018-0561-5
- Dhahbi, J. M., Mote, P. L., Fahy, G. M., and Spindler, S. R. (2005). Identification of potential caloric restriction mimetics by microarray profiling. *Physiol. Genomics* 23, 343–350. doi:10.1152/physiolgenomics.00069.2005
- Duca, F. A., Côté, C. D., Rasmussen, B. A., Zadeh-Tahmasebi, M., Rutter, G. A., Filippi, B. M., et al. (2015). Metformin activates a duodenal Ampk-dependent pathway to lower hepatic glucose production in rats. *Nat. Med.* 21, 506–511. doi:10.1038/nm.3787
- Dziewulska, A., Dobosz, A. M., Dobrzyn, A., Smolinska, A., Kolczynska, K., Ntambi, J. M., et al. (2020). SCD1 regulates the AMPK/SIRT1 pathway and histone acetylation through changes in adenine nucleotide metabolism in skeletal muscle. *J. Cell. Physiol.* 235, 1129–1140. doi:10.1002/jcp.29026
- El Khoury, L. Y., Gorrie-Stone, T., Smart, M., Hughes, A., Bao, Y., Andrayas, A., et al. (2019). Systematic underestimation of the epigenetic clock and age acceleration in older subjects. *Genome Biol.* 20, 283. doi:10.1186/s13059-019-1810-4
- Elbere, I., Silamikelis, I., Ustinova, M., Kalnina, I., Zaharenko, L., Peculis, R., et al. (2018). Significantly altered peripheral blood cell DNA methylation profile as a result of immediate effect of metformin use in healthy individuals. *Clin. Epigenetics* 10, 156. doi:10.1186/s13148-018-0593-x
- Fahy, G. M., Brooke, R. T., Watson, J. P., Good, Z., Vasanawala, S. S., Maecker, H., et al. (2019). Reversal of epigenetic aging and immunosenescent trends in humans. *Aging Cell.* 18, e13028. doi:10.1111/accel.13028
- Forslund, K., Hildebrand, F., Nielsen, T., Falony, G., Le Chatelier, E., Sunagawa, S., et al. (2015). Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. *Nature* 528, 262–266. doi:10.1038/nature15766
- García-Calzón, S., Perflyev, A., Männistö, V., de Mello, V. D., Nilsson, E., Pihlajamäki, J., et al. (2017). Diabetes medication associates with DNA methylation of metformin transporter genes in the human liver. *Clin. Epigenetics* 9, 102. doi:10.1186/s13148-017-0400-0
- Gibson, J., Russ, T. C., Clarke, T.-K., Howard, D. M., Hillary, R. F., Evans, K. L., et al. (2019). A meta-analysis of genome-wide association studies of epigenetic age acceleration. *PLoS Genet.* 15, e1008104. doi:10.1371/journal.pgen.1008104
- Gwinn, D. M., Shackelford, D. B., Egan, D. F., Mihaylova, M. M., Mery, A., Vazquez, D. S., et al. (2008). AMPK phosphorylation of raptor mediates a metabolic checkpoint. *Mol. Cell.* 30, 214–226. doi:10.1016/j.molcel.2008.03.003
- Hannon, E., Dempster, E., Viana, J., Burrage, J., Smith, A. R., Macdonald, R., et al. (2016). An integrated genetic-epigenetic analysis of schizophrenia: Evidence for colocalization of genetic associations and differential DNA methylation. *Genome Biol.* 17, 176. doi:10.1186/s13059-016-1041-x
- Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sada, S., et al. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell.* 49, 359–367. doi:10.1016/j.molcel.2012.10.016
- Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biol.* 14, R115. doi:10.1186/gb-2013-14-10-r115
- Horvath, S., Garagnani, P., Bacalini, M. G., Pirazzini, C., Salvioli, S., Gentilini, D., et al. (2015). Accelerated epigenetic aging in Down syndrome. *Aging Cell.* 14, 491–495. doi:10.1111/accel.12325
- Howell, J. J., Hellberg, K., Turner, M., Talbott, G., Kolar, M. J., Ross, D. S., et al. (2017). Metformin inhibits hepatic mTORC1 signaling via dose-dependent mechanisms involving AMPK and the TSC complex. *Cell. Metab.* 25, 463–471. doi:10.1016/j.cmet.2016.12.009
- Inoki, K., Zhu, T., and Guan, K.-L. (2003). TSC2 mediates cellular energy response to control cell growth and survival. *Cell.* 115, 577–590. doi:10.1016/s0092-8674(03)00929-2
- Inzucchi, S. E. (2002). Oral antihyperglycemic therapy for type 2 diabetes: Scientific review. *JAMA* 287, 360–372. doi:10.1001/jama.287.3.360
- Johnson, S. C., Rabinovitch, P. S., and Kaeblerlein, M. (2013). mTOR is a key modulator of ageing and age-related disease. *Nature* 493, 338–345. doi:10.1038/nature11861
- Jylhävä, J., Pedersen, N. L., and Hägg, S. (2017). Biological age predictors. *EBioMedicine* 21, 29–36. doi:10.1016/j.ebiom.2017.03.046
- Karnewar, S., Neeli, P. K., Panuganti, D., Kotagiri, S., Mallappa, S., Jain, N., et al. (2018). Metformin regulates mitochondrial biogenesis and senescence through AMPK mediated H3K79 methylation: Relevance in age-associated vascular dysfunction. *Biochim. Biophys. Acta. Mol. Basis Dis.* 1864, 1115–1128. doi:10.1016/j.bbdis.2018.01.018
- Kresovich, J. K., Xu, Z., O'Brien, K. M., Weinberg, C. R., Sandler, D. P., and Taylor, J. A. (2019). Methylation-based biological age and breast cancer risk. *J. Natl. Cancer Inst.* 111, 1051–1058. doi:10.1093/jnci/djz020
- Lee, H. Y., Lee, S. D., and Shin, K.-J. (2016). Forensic DNA methylation profiling from evidence material for investigative leads. *BMB Rep.* 49, 359–369. doi:10.5483/bmbrep.2016.49.7.070
- Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., et al. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging* 10, 573–591. doi:10.18632/aging.101414
- Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., et al. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* 11, 303–327. doi:10.18632/aging.101684
- Madeo, F., Carmona-Gutierrez, D., Hofer, S. J., and Kroemer, G. (2019). Caloric restriction mimetics against age-associated disease: Targets, mechanisms, and therapeutic potential. *Cell. Metab.* 29, 592–610. doi:10.1016/j.cmet.2019.01.018
- Maegawa, S., Hinkal, G., Kim, H. S., Shen, L., Zhang, L., Zhang, J., et al. (2010). Widespread and tissue specific age-related DNA methylation changes in mice. *Genome Res.* 20, 332–340. doi:10.1101/gr.096826.109
- Marin, T. L., Gongol, B., Zhang, F., Martin, M., Johnson, D. A., Xiao, H., et al. (2017). AMPK promotes mitochondrial biogenesis and function by phosphorylating the epigenetic factors DNMT1, RBBP7, and HAT1. *Sci. Signal.* 10, eaaf7478. doi:10.1126/scisignal.aaf7478

- Marioni, R. E., Shah, S., McRae, A. F., Chen, B. H., Colicino, E., Harris, S. E., et al. (2015). DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol.* 16, 25. doi:10.1186/s13059-015-0584-6
- Marioni, R. E., Shah, S., McRae, A. F., Ritchie, S. J., Muniz-Terrera, G., Harris, S. E., et al. (2015). The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. *Int. J. Epidemiol.* 44, 1388–1396. doi:10.1093/ije/dyu277
- Martin-Montalvo, A., Mercken, E. M., Mitchell, S. J., Palacios, H. H., Mote, P. L., Scheibye-Knudsen, M., et al. (2013). Metformin improves healthspan and lifespan in mice. *Nat. Commun.* 4, 2192. doi:10.1038/ncomms3192
- Na, H.-J., Park, J.-S., Pyo, J.-H., Lee, S.-H., Jeon, H.-J., Kim, Y.-S., et al. (2013). Mechanism of metformin: Inhibition of DNA damage and proliferative activity in *Drosophila* midgut stem cell. *Mech. Ageing Dev.* 134, 381–390. doi:10.1016/j.mad.2013.07.003
- Nair, V., Sreevalsan, S., Basha, R., Abdelrahim, M., Abudayyeh, A., Rodrigues Hoffman, A., et al. (2014). Mechanism of metformin-dependent inhibition of mammalian target of rapamycin (mTOR) and ras activity in pancreatic cancer: Role of specificity protein (sp) transcription factors. *J. Biol. Chem.* 289, 27692–27701. doi:10.1074/jbc.M114.592576
- Nguyen, T. T., Ung, T. T., Li, S., Lian, S., Xia, Y., Park, S. Y., et al. (2019). Metformin inhibits lithocholic acid-induced interleukin 8 upregulation in colorectal cancer cells by suppressing ROS production and NF- $\kappa$ B activity. *Sci. Rep.* 9, 2003. doi:10.1038/s41598-019-38778-2
- Nwanaji-Enwerem, J. C., Weisskopf, M. G., and Baccarelli, A. A. (2018). Multi-tissue DNA methylation age: Molecular relationships and perspectives for advancing biomarker utility. *Ageing Res. Rev.* 45, 15–23. doi:10.1016/j.arr.2018.04.005
- Ong, A. L. C., and Ramasamy, T. S. (2018). Role of Sirtuin1-p53 regulatory axis in aging, cancer and cellular reprogramming. *Ageing Res. Rev.* 43, 64–80. doi:10.1016/j.arr.2018.02.004
- Pernicova, I., and Korbonits, M. (2014). Metformin-mode of action and clinical implications for diabetes and cancer. *Nat. Rev. Endocrinol.* 10, 143–156. doi:10.1038/nrendo.2013.256
- Rakyan, V. K., Down, T. A., Maslau, S., Andrew, T., Yang, T.-P., Beyan, H., et al. (2010). Human aging-associated DNA hypermethylation occurs preferentially at bivalent chromatin domains. *Genome Res.* 20, 434–439. doi:10.1101/gr.103101.109
- Sae-Lee, C., Corsi, S., Barrow, T. M., Kuhnle, G. G. C., Bollati, V., Mathers, J. C., et al. (2018). Dietary intervention modifies DNA methylation age assessed by the epigenetic clock. *Mol. Nutr. Food Res.* 62, e1800092. doi:10.1002/mnfr.201800092
- Vazirpanah, N., Ottria, A., van der Linden, M., Wichers, C. G. K., Schuiveling, M., van Lochem, E., et al. (2019). mTOR inhibition by metformin impacts monosodium urate crystal-induced inflammation and cell death in gout: a prelude to a new add-on therapy? *Ann. Rheum. Dis.* 78, 663–671. doi:10.1136/annrheumdis-2018-214656
- Wagner, W. (2017). Epigenetic aging clocks in mice and men. *Genome Biol.* 18, 107. doi:10.1186/s13059-017-1245-8
- Yang, X., Kord-Varkaneh, H., Talaei, S., Clark, C. C. T., Zanghelini, F., Tan, S. C., et al. (2020). The influence of metformin on IGF-1 levels in humans: A systematic review and meta-analysis. *Pharmacol. Res.* 151, 104588. doi:10.1016/j.phrs.2019.104588
- Yang, Z., Wong, A., Kuh, D., Paul, D. S., Rakyan, V. K., Leslie, R. D., et al. (2016). Correlation of an epigenetic mitotic clock with cancer risk. *Genome Biol.* 17, 205. doi:10.1186/s13059-016-1064-3
- Zhang, Q.-S., Tang, W., Deater, M., Phan, N., Marcogliese, A. N., Li, H., et al. (2016). Metformin improves defective hematopoiesis and delays tumor formation in Fanconi anemia mice. *Blood* 128, 2774–2784. doi:10.1182/blood-2015-11-683490
- Zhong, T., Men, Y., Lu, L., Geng, T., Zhou, J., Mitsuhashi, A., et al. (2017). Metformin alters DNA methylation genome-wide via the H19/SAHH axis. *Oncogene* 36, 2345–2354. doi:10.1038/onc.2016.391



## OPEN ACCESS

## EDITED BY

Haitao Wang,  
National Cancer Institute, United States

## REVIEWED BY

Qin Xu,  
National Institute of Allergy and  
Infectious Diseases (NIH), United States  
Yaqiang Cao,  
National Institutes of Health,  
United States  
Dawei Zhou,  
Wexner Medical Center, The Ohio State  
University, United States

## \*CORRESPONDENCE

Yan V. Sun,  
yan.v.sun@emory.edu

<sup>†</sup>These authors contributed equally to  
this work and share first authorship

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 16 August 2022

ACCEPTED 22 September 2022

PUBLISHED 11 October 2022

## CITATION

Titanji BK, Lee M, Wang Z, Chen J, Hui Q,  
Lo Re III V, So-Armah K, Justice AC, Xu K,  
Freiberg M, Gwinn M, Marconi VC and  
Sun YV (2022), Epigenome-wide  
association study of biomarkers of liver  
function identifies albumin-associated  
DNA methylation sites among male  
veterans with HIV.  
*Front. Genet.* 13:1020871.  
doi: 10.3389/fgene.2022.1020871

## COPYRIGHT

© 2022 Titanji, Lee, Wang, Chen, Hui, Lo  
Re III, So-Armah, Justice, Xu, Freiberg,  
Gwinn, Marconi and Sun. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Epigenome-wide association study of biomarkers of liver function identifies albumin-associated DNA methylation sites among male veterans with HIV

Boghuma K. Titanji<sup>1†</sup>, Mitch Lee<sup>2†</sup>, Zeyuan Wang<sup>2</sup>, Junyu Chen<sup>2</sup>,  
Qin Hui<sup>2</sup>, Vincent Lo Re III<sup>3</sup>, Kaku So-Armah<sup>4</sup>, Amy C. Justice<sup>5,6</sup>,  
Ke Xu<sup>5,6</sup>, Matthew Freiberg<sup>7</sup>, Marta Gwinn<sup>8</sup>,  
Vincent C. Marconi<sup>1,9,10,11</sup> and Yan V. Sun<sup>2,9\*</sup>

<sup>1</sup>Division of Infectious Disease, Emory School of Medicine, Atlanta, GA, United States, <sup>2</sup>Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, United States, <sup>3</sup>Division of Infectious Diseases Department of Medicine and Center for Clinical Epidemiology and Biostatistics Perelman School of Medicine University of Pennsylvania, Philadelphia, PA, United States, <sup>4</sup>Boston University Medical School, Boston, MA, United States, <sup>5</sup>Connecticut Veteran Health System, West Haven, CT, United States, <sup>6</sup>Yale University School of Medicine, New Haven, CT, United States, <sup>7</sup>Cardiovascular Medicine Division and Tennessee Valley Healthcare System, Vanderbilt University Medical Center, Nashville, TN, United States, <sup>8</sup>Centers for Disease Control and Prevention, Atlanta, GA, United States, <sup>9</sup>Atlanta Veterans Affairs Health Care System, Decatur, GA, United States, <sup>10</sup>Hubert Department of Global Health, Rollins School of Public Health, Atlanta, GA, United States, <sup>11</sup>Emory Vaccine Center, Yerkes National Primate Research Center, Emory University, Atlanta, GA, United States

**Background:** Liver disease (LD) is an important cause of morbidity and mortality for people with HIV (PWH). The molecular factors linked with LD in PWH are varied and incompletely characterized. We performed an epigenome-wide association study (EWAS) to identify associations between DNA methylation (DNAm) and biomarkers of liver function—aspartate transaminase, alanine transaminase, albumin, total bilirubin, platelet count, FIB-4 score, and APRI score—in male United States veterans with HIV.

**Methods:** Blood samples and clinical data were obtained from 960 HIV-infected male PWH from the Veterans Aging Cohort Study. DNAm was assessed using the Illumina 450K or the EPIC 850K array in two mutually exclusive subsets. We performed a meta-analysis for each DNAm site

**Abbreviations:** AA, Age acceleration; ALT, Alanine amino transferase; APRI, AST to platelet ratio index; ART, Antiretroviral therapy; AST, Aspartate amino transferase; BMI, Body mass index; CpG, 5'—Cytosine—phosphate—Guanine—3'; DNAm, Deoxyribonucleic acid methylation; EAAA, Extrinsic epigenetic age acceleration; EWAS, Epigenome-wide association study; FIB, Fibrosis; FKBP, FK506 binding protein; HBV, Hepatitis B virus; HCC, Hepatocellular carcinoma; HCV, Hepatitis C virus; HIV, Human immunodeficiency virus; IEAA, Intrinsic epigenetic age acceleration; LD, Liver Disease; NACD, Non-AIDS related chronic disease; NAFLD, Non-Alcohol fatty liver disease; PWH, People with HIV; SOCS, suppressor of cytokine signaling; STAT, Signal transducer and activator of transcription; TGF, Tissue growth factor; TMEM, transmembrane; TNF $\alpha$ , tumor necrosis factor-  $\alpha$ ; VACS, Veterans aging cohort; ZEB, Zinc Finger E-Box Binding Homeobox 1.

measured by either platform. We also examined the associations between four measures of DNAm age acceleration (AA) and liver biomarkers.

**Results:** Nine DNAm sites were positively associated with serum albumin in the meta-analysis of the EPIC and 450K EWAS after correcting for multiple testing. Four DNAm sites (cg16936953, cg18942579, cg01409343, and cg12054453), annotated within the *TMEM49* and four of the remaining five sites (cg18181703, cg03546163, cg20995564, and cg23966214) annotated to *SOCS3*, *FKBP5*, *ZEB2*, and *SAMD14* genes, respectively. The DNAm site, cg12992827, was not annotated to any known coding sequence. No significant associations were detected for the other six liver biomarkers. Higher PhenoAA was significantly associated with lower level of serum albumin ( $\beta = -0.007$ ,  $p$ -value =  $8.6 \times 10^{-4}$ , CI: -0.011116, -0.002884).

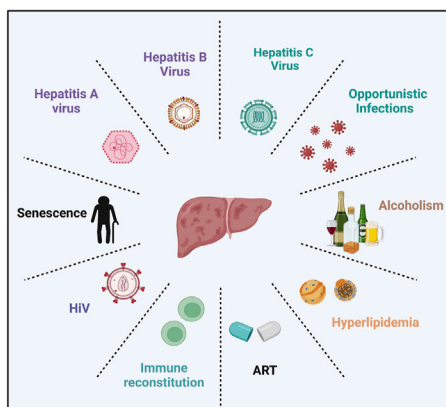
**Conclusion:** We identified epigenetic associations of both individual DNAm sites and DNAm AA with liver function through serum albumin in men with HIV. Further replication analyses in independent cohorts are warranted to confirm the epigenetic mechanisms underlying liver function and LD in PWH.

#### KEYWORDS

EWAS, liver disease, HIV, biomarkers, albumin (ALB)

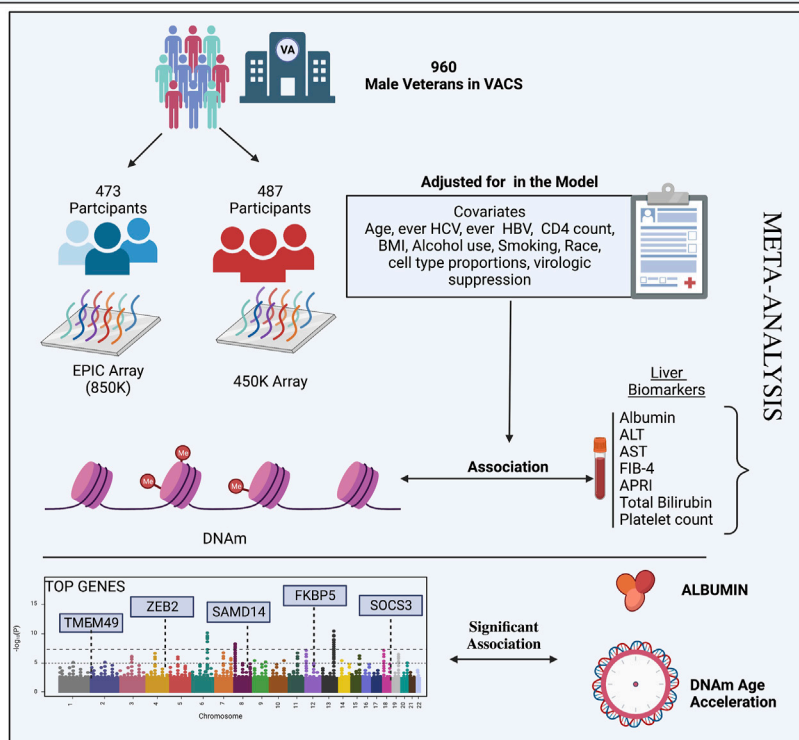
### EPIGENOME-WIDE ASSOCIATION STUDY OF BIOMARKERS OF LIVER FUNCTION IDENTIFIES ALBUMIN ASSOCIATED DNA METHYLATION SITES AMONG VETERANS WITH HIV

#### BACKGROUND



People with HIV have increased morbidity and mortality from liver disease. The drivers of which are incompletely characterized

#### GRAPHICAL ABSTRACT



## Introduction

The widespread availability and use of effective combination antiretroviral therapy to treat HIV has transformed the outlook for people with HIV (PWH), many of whom now live longer and are facing the challenges of aging with HIV. The burden of non-AIDS related chronic disease (NACD) is increasingly recognized in this patient population as an urgent concern that needs to be addressed to improve care for PWH.

Liver disease (LD) accounts for an estimated 13%–18% all-cause mortality in PWH, and PWH have a high prevalence of traditional causes of LD (Palella et al., 2006; Smith et al., 2010; Smith et al., 2014; Kaspar and Sterling, 2017). Co-infection with hepatitis B and C viruses is estimated between 5%–30% based on cohort studies (Konopnicki et al., 2005; Shepard et al., 2005), and up to 30%–40% of PWH have signs of non-alcoholic fatty liver disease (NAFLD) (Lemoine et al., 2012). Poorly controlled HIV-infection itself is an independent risk factor for liver fibrosis (Kim et al., 2016a), which is the most common outcome of chronic liver injury. Even in individuals with well-controlled HIV (undetectable HIV RNA) and immunologic recovery ( $CD4^+$  cell counts  $> 500$  cells/mm<sup>3</sup>), oxidative stress, mitochondrial injury, toxic metabolite accumulation, gut microbial translocation, systemic inflammation, senescence, and nodular regenerative hyperplasia [reviewed in (Kim et al., 2016a)] contribute to liver injury and fibrosis. Given all of the factors contributing to LD among PWH, there is a need to improve prevention, risk prediction, and treatment of LD in PWH.

Progress in elucidating the biological mechanisms underlying LD has been challenging due to the complexity of the multiple roles played by the liver in homeostasis and metabolic functions. Epigenome-wide association study (EWAS) have the potential to identify novel mechanisms of LD and generate new hypotheses that may provide insights into prevention and treatment of LD. Previous studies on the role of DNA methylation (DNAm) in LD have focused on non-alcoholic fatty LD (NAFLD), since determinants of NAFLD are not as easily identifiable as LD resulting from e.g., alcoholism or viral hepatitis [reviewed in (Zhang et al., 2021)]. Studies on the association between alcoholic liver disease and DNAm have identified genes which may be associated with the incidence and progression of hepatocellular carcinoma (HCC) (French, 2013; Varela-Rey et al., 2013; Zakhari, 2013; Rosen et al., 2018). Other EWAS have uncovered DNAm sites and genes associated with the development of HCC in patients with chronic hepatitis B infections (Su et al., 2007; Nishida et al., 2008; Zhao et al., 2014). PWH are significantly underrepresented in these studies. Moreover, most studies have focused on LD as an overall clinical diagnosis and have not assessed potential associations between DNAm and specific metabolic functions of the liver. Biomarkers are useful surrogates of

the liver's metabolic functions and may reflect its capacity to perform specific functions. The absence of data on the relationship between DNAm and specific markers of liver function, particularly among PWH, represents an important gap in the literature which we hope the present study will begin to fill.

DNAm age acceleration (AA) is a novel biomarker of biological aging, and predicts age-related disease outcomes and mortality (Bell et al., 2019). Several studies have also shown that PWH exhibit elevated DNAm AA above the average level exhibited by non-infected people of equivalent chronological age. (Horvath and Levine, 2015; Rickabaugh et al., 2015). These observations suggest that DNAm AA may partially explain the early onset of chronic comorbid conditions among PWH (Levine et al., 2015a). Relative acceleration of DNAm age from blood-based assays is often measured by four metrics: intrinsic epigenetic age acceleration (IEAA) (Levine et al., 2015a; Horvath and Levine, 2015), extrinsic epigenetic age acceleration (EEAA) (Hannum et al., 2013; Levine et al., 2015b), phenotypic age acceleration (PhenoAA) (Levine et al., 2018), and Grim age acceleration (GrimAA) (Lu et al., 2019).

The present study included available markers of liver function from the Veterans Aging Cohort Study (VACS): aspartate transaminase (AST), alanine transaminase (ALT), serum albumin, total bilirubin, platelet count, fibrosis-4 (FIB-4) score, and the AST-to-platelet-ratio index (APRI) score. We conducted epigenetic association analyses between individual DNAm sites, measures of DNAm AA (IEAA, EEAA, PhenoAA and GrimAA) and these seven biomarkers of liver function in a cohort of male veterans with HIV infection in the United States.

## Methods

### Study population

A total of 960 male veterans with HIV at the time of blood collection were included from the VACS, having both phenotypic and epigenetic data available from previous studies (Justice et al., 2006; Chen et al., 2020). Participants with cancer at the time of blood collection were excluded because the cancer treatment and disease can influence liver function and DNAm which is beyond the scope of the present study. The VACS consists of electronic medical record data from patients with HIV receiving care at Veterans Health Administration (VA) medical facilities across the United States. Each PWH is matched on age, sex, race/ethnicity, and site to two persons without HIV. Data include hospital and outpatient diagnoses (recorded using International Classification of Diseases, Ninth Revision [ICD-9] and Tenth [ICD-10] codes), procedures (recorded

using Current Procedural Terminology [CPT] codes), laboratory results, and pharmacy data (Justice et al., 2006). Information on age, race, smoking status, body mass index (BMI), diabetes status (glucose level  $\geq 200$  mg/dl on 2 separate occasions or glucose level  $\geq 200$  mg/dl on 1 occasion plus treatment with an oral hypoglycemic or insulin for  $\geq 30$  days) (Mathur et al., 2019), ever infection with HBV (defined as HBV surface antigen positive, acute resolved HBV or ICD-9 code for HBV diagnosis), ever infection with HCV (defined as HCV antibody positive regardless of HCV RNA, ICD-9 code for HCV diagnosis, or HCV genotype), and antiretroviral (ART) use at time of blood draw were obtained from VA's clinical data warehouse. Alcohol use status was obtained from the VACS survey (Justice et al., 2006). At the visit of blood draw, AST, ALT, serum albumin, total bilirubin, platelet count, CD4<sup>+</sup> T-cell count, and plasma HIV-1 RNA viral load were measured for each patient using approved clinical assays from certified clinical laboratories. Participants were categorized as HIV suppressed if their HIV RNA was  $< 200$  copies/mm<sup>3</sup> or as unsuppressed VL otherwise. A fibrosis-4 (FIB-4) score and an AST-to-platelet ratio index (APRI) score were calculated for each participant following the standard equations (Wai et al., 2003; Sterling et al., 2006). All participants provided written consent for the use of their data.

## DNAm data generation, processing, and quality control

The genome-wide DNAm profiles were measured using genome DNA samples extract from whole blood. Genomic DNA extraction, epityping, data processing, and quality control were performed as previously described. (Chen et al., 2020). Genome-wide DNAm levels for 473 of the participants were assessed using the Infinium HumanMethylation450 (450K) array platform (Illumina), while those for the remaining 487 participants were assessed using the Infinium HumanMethylationEPIC (EPIC) array platform (Illumina) at the Yale Center for Genomic Analysis. Quality control procedures were followed as previously described in published studies (Zhang et al., 2017; Solomon et al., 2018; Chen et al., 2020). We compared the characteristics between the 450K and EPIC subsets using t-tests for continuous variables and chi-square tests for categorical variables.

DNAm sites for 412,583 and 846,604 DNAm sites were included in the analyses for the dataset acquired with the 450K and EPIC arrays, respectively. Differences in the proportions of the six main leukocyte cell types present in whole blood (CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, monocytes, B cells, granulocytes, and natural killer cells) across samples are well described confounders of associations between DNAm in the blood and many phenotypes. To account for this, the proportions

of these six cell types for each participant were determined based on the top cell-type-specific DNAm sites in a reference panel of known proportions following the standard algorithm through the *minfi* package in R (Houseman et al., 2012; Moran et al., 2016). The estimated cell type proportions were then controlled for in all EWAS analyses.

The association between each covariate included in the EWAS model—race, smoking status, BMI, diabetes status, hazardous alcohol use (Freiberg et al., 1999), ever HCV infection, ever HBV infection, ART use, CD4 count, viral suppression, and leukocyte cell-type proportions, top ten principal components—and each liver marker was assessed using a linear model controlling for chronological age as in the following model.

### Individual liver marker ~ covariate + age

Associations between each liver marker and chronological age were also assessed using a univariate model.

### Individual liver marker ~ age

Regardless of statistical significance, all covariates listed above were included in the final EWAS model to account for the reasonable possibility that they might confound the relationship between DNAm and the liver markers as in the following model.

Individual liver marker ~ DNAm + age + race + current smoking + BMI + diabetes + hazardous alcohol use + ever infection with HCV + ever infection with HBV + ART use + CD4<sup>+</sup> T-cell count + VL + leukocyte cell type proportions + top ten principal components.

## Correlations among liver biomarkers

To identify correlations among the biomarkers of liver function included in this study, a Spearman correlation coefficient and corresponding *p*-value were calculated for every pair of selected biomarkers using the R statistical software package. To normalize the distributions of strongly right-skewed biomarkers for this and all subsequent analyses, a natural log transformation was performed on the measured ALT, AST, and calculated FIB-4 and APRI scores.

## Associations of DNAm age acceleration with liver biomarkers

DNAm AA for each participant was measured using the IEAA, EEAA, PhenoAA, and GrimAA metrics as specified by

TABLE 1 Characteristics of participants grouped by platform used for genome-wide DNA methylation.

| Characteristic   | 450K (N = 473) | EPIC (N = 488) | Overall (N = 961) | p-value              |
|--|----------------|----------------|-------------------|----------------------|
| ALT (units/L)  | 38.8 (24.9)    | 34.1 (21.2)    | 36.4 (23.2)       | $1.9 \times 10^{-3}$ |
| AST (units/L)  | 43.4 (28.1)    | 37.0 (20.1)    | 40.2 (24.6)       | $7.3 \times 10^{-5}$ |
| Serum Albumin (mg/dl)                                  | 3.84 (0.489)   | 3.95 (0.457)   | 3.90 (0.476)      | $2.9 \times 10^{-4}$ |
| Total Serum Bilirubin (mg/dl)                          | 0.78 (0.485)   | 0.76 (0.513)   | 0.77 (0.500)      | 0.61                 |
| Platelet Count [ $\times 10^3$ (Zhang et al., 2021)/L] | 224 (71.3)     | 232 (69.4)     | 228 (70.4)        | $8.0 \times 10^{-2}$ |
| FIB-4 Score  | 1.81 (1.22)    | 1.57 (0.943)   | 1.69 (1.09)       | $8.5 \times 10^{-4}$ |
| APRI Score   | 0.58 (0.602)   | 0.47 (0.393)   | 0.52 (0.509)      | $8.4 \times 10^{-4}$ |
| Age (years)  | 51.8 (7.53)    | 50.7 (7.48)    | 51.2 (7.52)       | $3.1 \times 10^{-2}$ |
| Race   |                |                |                   | 1.00                 |
| Black  | 402 (85.0%)    | 391 (80.1%)    | 793 (82.5%)       |                      |
| Non-Black  | 71 (15.0%)     | 97 (19.9%)     | 168 (17.5%)       |                      |
| Smoking Status   |                |                |                   | 1.00                 |
| Not current  | 203 (42.9%)    | 210 (43.0%)    | 413 (43.0%)       |                      |
| Current  | 270 (57.1%)    | 278 (57.0%)    | 548 (57.0%)       |                      |
| Diabetes Status  |                |                |                   | 1.00                 |
| No   | 390 (82.5%)    | 406 (83.2%)    | 796 (82.8%)       |                      |
| Yes  | 83 (17.5%)     | 82 (16.8%)     | 165 (17.2%)       |                      |
| BMI  | 25.4 (4.44)    | 25.9 (4.43)    | 25.6 (4.44)       | 0.14                 |
| Missing  | 8 (1.7%)       | 9 (1.8%)       | 17 (1.8%)         |                      |
| Alcohol Use  |                |                |                   | 1.00                 |
| Non-Hazardous  | 195 (41.2%)    | 191 (39.1%)    | 386 (40.2%)       |                      |
| Hazardous  | 276 (58.4%)    | 297 (60.9%)    | 573 (59.6%)       |                      |
| Ever Hepatitis C Infection                             |                |                |                   | 1.00                 |
| No   | 241 (51.0%)    | 312 (63.9%)    | 553 (57.5%)       |                      |
| Yes  | 232 (49.0%)    | 176 (36.1%)    | 408 (42.5%)       |                      |
| Ever Hepatitis B Infection                             |                |                |                   | 1.00                 |
| No   | 399 (84.4%)    | 430 (88.1%)    | 829 (86.3%)       |                      |
| Yes  | 52 (11.0%)     | 40 (8.2%)      | 92 (9.6%)         |                      |
| CD4 <sup>+</sup> Cell Count                            | 411 (252)      | 442 (263)      | 427 (258)         | $6.4 \times 10^{-2}$ |
| HIV viral Load   |                |                |                   | 1.00                 |
| Suppressed   | 196 (41.4%)    | 204 (41.8%)    | 400 (41.6%)       |                      |
| Unsuppressed   | 274 (57.9%)    | 284 (58.2%)    | 558 (58.1%)       |                      |

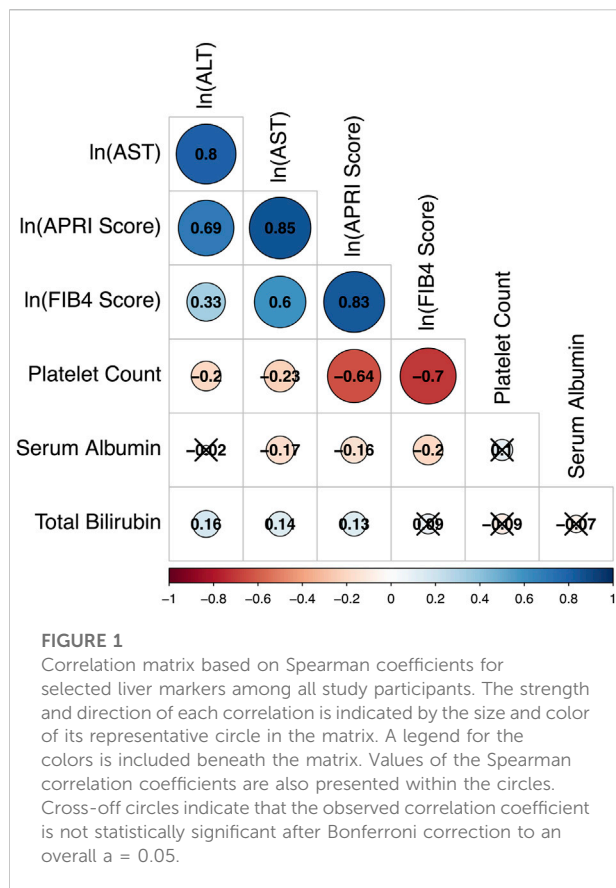
Statistics for numeric variables are presented as mean (SD), while those for categorical variables are presented as count (%). Counts across levels of some categorical variables may not sum to the corresponding total due to missing values. N= number; P-value, level of statistical significance.

original reporting articles (Chen et al., 2016; Levine et al., 2018; Lu et al., 2019). IEAA is calculated as the residual of a person's DNAm level in blood cells across a set of DNAm sites proposed by Horvath after regressing DNAm level on chronological age and controlling for proportions of different leukocytes within the sample (Levine et al., 2015a; Horvath and Levine, 2015). EEAA is calculated similarly, except that a different set of DNAm sites proposed by Hannum are used while controlling for proportions of a different set of cell types (Hannum et al., 2013; Levine et al., 2018). PhenoAA, in contrast, is calculated by first estimating phenotypic age (in years) with a linear regression model that uses clinical variables as inputs, then estimating DNAm age (in years) based on methylation levels at a set of DNAm sites proposed by Levine et al. (2018) and finally calculating the residual that results

from regressing calculated DNAm age on calculated phenotypic age. GrimAA is calculated in a similar way as PhenoAA but utilizes levels of plasma biomarkers indicative of physiological stress to estimate phenotypic age and a different set of DNAm sites to assess DNAm age (Lu et al., 2019). Each of these measures of DNAm AA perform differently depending on the outcome being evaluated and can complement each other to assess the biological aging process.

The association between DNAm AA, as measured by each of these four metrics, and each of the selected biomarkers of liver function was then assessed by linear regression while controlling for all covariates included in the final EWAS model.

Individual liver marker ~ DNAm AA + age + race + current smoking + BMI + diabetes + hazardous alcohol use + ever



infection with HCV + ever infection with HBV + ART use +  $CD4^+$  T-cell count + VL + leukocyte cell type proportions + top ten principal components.

## Meta-analysis of EWAS for liver biomarkers

To investigate the association between each biomarker of liver function and individual DNAm sites across the autosomal chromosomes, data from the 450K dataset were analyzed in parallel with data from the EPIC dataset using the same model, and a fixed effect model meta-analysis was conducted using inverse variance weighted effect size method. The final model was adjusted for age, race, current smoking, BMI, diabetes, hazardous alcohol use, ever infection with HCV, ever infection with HBV, ART use,  $CD4^+$  T-cell count, VL, leukocyte cell type proportions, and the top ten principal components within the group being analyzed. Where necessary for any given combination of DNAm site and biomarker, participants missing values required for the model were excluded from the analysis for that combination.

A total of 385,062 DNAm sites measured by both the 450K, and EPIC arrays were included in the EWAS meta-analysis. For

DNAm sites not covered by both platforms, results were obtained only from the cohort profiled by the covering platform and no meta-analysis was conducted. Separate false-discovery rate adjustments ( $q < 0.05$ ) were conducted for each liver biomarker. All analyses were performed in R (v4.0.3).

## Results

### Participant characteristics

After data processing and quality control, the analysis dataset for those profiled using the 450K and EPIC platforms included observations from 473 to 487 individuals, respectively. Characteristics of the two EWAS subsets are summarized in Table 1. All participants were male veterans with HIV who were never diagnosed with cancer and had an average age of  $51.2 \pm 7.5$  years. The chronological age between the 450K and EPIC sub-cohorts was moderately different ( $\Delta = 1.1$  years,  $p$ -value = 0.031). The distribution of two sub-cohorts did not differ significantly ( $p$ -value > 0.05) in race (Black vs. non-Black) or in prevalence of current smoking, diabetes, hazardous alcohol use, ever infection with HBV, ever infection with HCV, or viral load suppression. They also did not differ significantly in their average BMI or  $CD4^+$  T-cell count. Average biomarker values from the 450K cohort differed significantly from those from the EPIC cohort for ALT ( $\Delta = 4.7$  units/L,  $p$ -value =  $1.93 \times 10^{-3}$ ), AST ( $\Delta = 6.4$  units/L,  $p$ -value =  $7.34 \times 10^{-5}$ ), serum albumin ( $\Delta = -0.11$  mg/dl,  $p$ -value =  $2.89 \times 10^{-4}$ ), FIB-4 score ( $\Delta = 0.24$ ,  $p$ -value =  $8.49 \times 10^{-4}$ ), and APRI score ( $\Delta = 0.112$ ,  $p$ -value =  $8.35 \times 10^{-4}$ ). Total bilirubin and platelet count did not differ significantly between the two sub-cohorts.

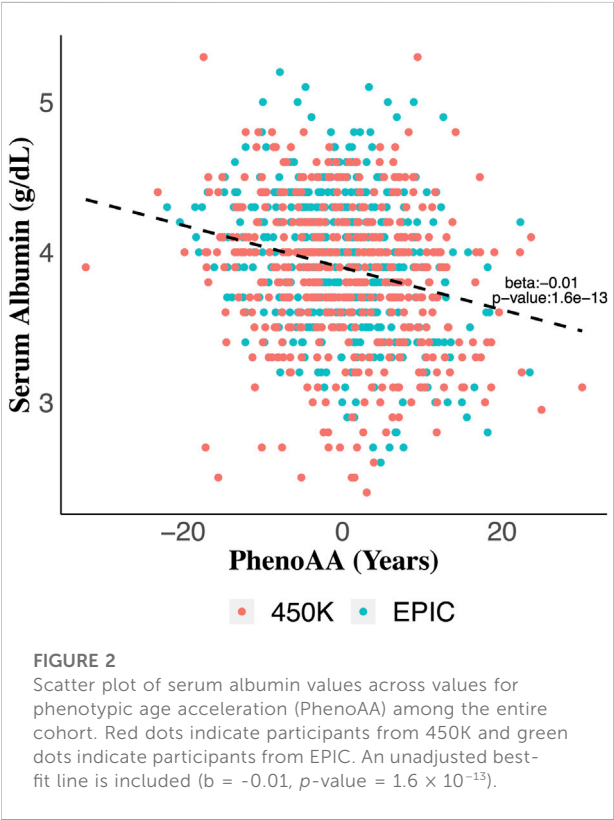
### Correlations among liver biomarkers

To illustrate the overall correlation between seven liver function markers, we summarized the pair-wise correlation in Figure 1. Observed values for ALT, AST, FIB-4 score, and APRI score all correlated positively and strongly with each other (Spearman correlation coefficient  $\rho > 0.6$ ), except for ALT and FIB-4 score with moderate correlation ( $\rho = 0.36$ ) (Figure 1). Platelet count correlated negatively with AST, ALT, FIB-4 score, and APRI score to varying degrees and weakly with albumin ( $\rho = 0.13$ ), but not with total bilirubin ( $\rho = -0.06$ ). Serum albumin correlated negatively with AST, APRI score, and FIB-4 score ( $-0.25 \leq \rho \leq -0.18$ ) but not with ALT ( $\rho = -0.01$ ). Total bilirubin did not correlate significantly with serum albumin, platelet count, FIB-4 score, or APRI score ( $p > 0.05$ ) and correlated only weakly with AST and ALT ( $\rho < 0.14$ ). Distributions of the biomarkers before and after transformation, are presented in Supplementary Figure S1.

TABLE 2 Associations between DNA methylation age acceleration and selected biomarkers of liver function.

|              |           | Model   |   |   |   |
|--------------|-----------|---|---|---|---|
| Liver Marker |           | PhenoAA   | GrimAA  | IEAA  | EEAA  |
| ln (AST)     | Beta (SE) | 2.5×10 <sup>-3</sup> (2.1 × 10 <sup>-3</sup> )    | 7.9×10 <sup>-3</sup> (4.0 × 10 <sup>-3</sup> )    | 2.6 × 10 <sup>-3</sup> (2.8 × 10 <sup>-3</sup> )  | 1.9×10 <sup>-3</sup> (2.8 × 10 <sup>-3</sup> )    |
|              | p         | 0.23  | 0.05  | 0.36  | 0.51  |
| ln (ALT)     | Beta (SE) | 1.9 × 10 <sup>-4</sup> (2.5 × 10 <sup>-3</sup> )  | 3.4 × 10 <sup>-3</sup> (4.7 × 10 <sup>-3</sup> )  | 1.3 × 10 <sup>-3</sup> (3.3 × 10 <sup>-3</sup> )  | 5.5 × 10 <sup>-4</sup> (3.3 × 10 <sup>-3</sup> )  |
|              | p         | 0.94  | 0.47  | 0.69  | 0.87  |
| ALB          | Beta (SE) | -7.0 × 10 <sup>-3</sup> (2.1 × 10 <sup>-3</sup> ) | -7.2 × 10 <sup>-3</sup> (4.0 × 10 <sup>-3</sup> ) | -4.1 × 10 <sup>-3</sup> (2.8 × 10 <sup>-3</sup> ) | -4.3 × 10 <sup>-3</sup> (2.8 × 10 <sup>-3</sup> ) |
|              | p         | 8.6 × 10 <sup>-4</sup>                            | 0.07  | 0.15  | 0.12  |
| TBILI        | Beta (SE) | -4.9 × 10 <sup>-3</sup> (2.3 × 10 <sup>-3</sup> ) | -1.8 × 10 <sup>-3</sup> (4.4 × 10 <sup>-3</sup> ) | 2.1 × 10 <sup>-3</sup> (3.2 × 10 <sup>-3</sup> )  | 2.0 × 10 <sup>-3</sup> (3.1 × 10 <sup>-3</sup> )  |
|              | p         | 0.04  | 0.69  | 0.50  | 0.52  |
| PLT          | Beta (SE) | -0.48 (0.32)                                      | -0.12 (0.61)                                      | -0.2 (0.44)                                       | -0.72 (0.43)                                      |
|              | p         | 0.14  | 0.84  | 0.64  | 0.09  |
| ln (FIB-4)   | Beta (SE) | 1.4 × 10 <sup>-3</sup> (2.2 × 10 <sup>-3</sup> )  | -5.5 × 10 <sup>-4</sup> (4.1 × 10 <sup>-3</sup> ) | -1.4 × 10 <sup>-3</sup> (2.9 × 10 <sup>-3</sup> ) | 1.8 × 10 <sup>-3</sup> (2.9 × 10 <sup>-3</sup> )  |
|              | p         | 0.53  | 0.89  | 0.63  | 0.54  |
| ln (APRI)    | Beta (SE) | 1.3 × 10 <sup>-3</sup> (2.9 × 10 <sup>-3</sup> )  | 3.2 × 10 <sup>-3</sup> (5.5 × 10 <sup>-3</sup> )  | 2.6 × 10 <sup>-4</sup> (3.9 × 10 <sup>-3</sup> )  | 3.7 × 10 <sup>-3</sup> (3.9 × 10 <sup>-3</sup> )  |
|              | p         | 0.64  | 0.56  | 0.95  | 0.34  |

Biomarker values were ln-transformed when needed to achieve a more normal distribution as indicated on the left and then regressed on each DNA methylation age acceleration metric in linear models that control for the covariates included in the EWAS model. Associations that remained significant after FDR adjustment at  $Q < 0.05$  are highlighted in yellow. Abbreviations: AST, aspartate aminotransferase (units/L); ALT, alanine aminotransferase (units/L); ALB, serum albumin (mg/dl); TBILI, total bilirubin level (mg/dl); PLT, platelet count (cells/ $\mu$ L); FIB-4, FIB-4 score; APRI, APRI score; PhenoAA, phenotype age acceleration; GrimAA, Grim age acceleration; IEAA, intrinsic epigenetic age acceleration; EEAA, extrinsic epigenetic age acceleration.



### Covariate associations with liver biomarkers

Associations between covariates included in the final EWAS model and the selected liver biomarkers are presented in Table 2. Only “ever infection with HCV” was associated with all seven selected liver biomarkers when controlling for age, while diabetes status was not associated with any of the selected markers. All liver markers associated with at least four of the covariates included in the model excluding cell type proportions. Regardless of the  $p$ -value for the observed association between each covariate and each liver biomarker, all covariates were consistently included in the final epigenetic analysis model for each liver biomarker. Participants with complete data of all covariates were analyzed.

### Associations of DNAm age acceleration with liver biomarkers

Serum albumin was significantly associated with PhenoAA ( $\text{beta} = -0.014$ ,  $p\text{-value} = 1.6 \times 10^{-13}$ ) in an unadjusted model (Figure 2). That association remained significant after adjusting for all covariates in the final regression model ( $\text{beta} = -0.007$ ,  $p\text{-value} = 8.6 \times 10^{-4}$ ) (Table 3). No other metric of DNAm AA (i.e., IEAA, EEAA, and GrimAA) was associated with any other

TABLE 3 DNAm sites are significantly associated with selected liver markers after meta-analysis and FDR correction.

Serum albumin

| DNAm       | Chr | Pos       | Gene          | Meta-analysis |                      |       | EPIC       |                      | 450K        |                      |
|------------|-----|-----------|---------------|---------------|----------------------|-------|------------|----------------------|-------------|----------------------|
|            |     |           |               | Beta (SE)     | p                    | FDR   | Beta (SE)  | p                    | Beta (SE)   | p                    |
| cg16936953 | 17  | 57915665  | <i>TMEM49</i> | 1.2 (0.21)    | $4.2 \times 10^{-9}$ | 0.001 | 1.0 (0.29) | $6.6 \times 10^{-4}$ | 1.4 (0.29)  | $1.7 \times 10^{-6}$ |
| cg18181703 | 17  | 76354621  | <i>SOCS3</i>  | 1.9 (0.35)    | $9.5 \times 10^{-8}$ | 0.009 | 2.3 (0.45) | $9.9 \times 10^{-7}$ | 1.3 (0.55)  | $2.1 \times 10^{-2}$ |
| cg03546163 | 6   | 35654363  | <i>FKBP5</i>  | 1.1 (0.20)    | $1.1 \times 10^{-7}$ | 0.009 | 1.2 (0.26) | $9.2 \times 10^{-6}$ | 0.94 (0.33) | $4.5 \times 10^{-3}$ |
| cg18942579 | 17  | 57915773  | <i>TMEM49</i> | 1.4 (0.26)    | $1.1 \times 10^{-7}$ | 0.009 | 1.2 (0.34) | $7.3 \times 10^{-4}$ | 1.7 (0.41)  | $3.2 \times 10^{-5}$ |
| cg01409343 | 17  | 57915740  | <i>TMEM49</i> | 1.7 (0.32)    | $1.2 \times 10^{-7}$ | 0.009 | 1.7 (0.43) | $1.4 \times 10^{-4}$ | 1.8 (0.48)  | $3.1 \times 10^{-4}$ |
| cg20995564 | 2   | 145172035 | <i>ZEB2</i>   | 1.2 (0.23)    | $1.9 \times 10^{-7}$ | 0.012 | 1.5 (0.29) | $3.0 \times 10^{-7}$ | 0.66 (0.39) | $9.1 \times 10^{-2}$ |
| cg23966214 | 17  | 48203188  | <i>SAMD14</i> | 3.5 (0.70)    | $4.7 \times 10^{-7}$ | 0.024 | 3.2 (0.81) | $8.7 \times 10^{-5}$ | 4.4 (1.4)   | $1.6 \times 10^{-3}$ |
| cg12054453 | 17  | 57915717  | <i>TMEM49</i> | 1.0 (0.20)    | $5.0 \times 10^{-7}$ | 0.024 | 0.7 (0.29) | $1.1 \times 10^{-2}$ | 1.3 (0.28)  | $8.4 \times 10^{-6}$ |
| cg12992827 | 3   | 101901234 | None Mapped   | 1.9 (0.38)    | $8.0 \times 10^{-7}$ | 0.034 | 2.2 (0.47) | $4.0 \times 10^{-6}$ | 1.3 (0.66)  | $5.5 \times 10^{-2}$ |

Chr: chromosome number; Pos: basepair position (human genome build hg19); beta: beta coefficient; SE: standard error; p: p-value, i.e. level of statistical significance. Results for DNAm sites where DNA methylation remained significantly associated with serum albumin after meta-analysis and FDR correction to  $Q < 0.05$  are presented, including statistics from the meta-analysis and from the separate EWAS of the two cohorts. No significant associations were identified after meta-analysis and FDR correction for AST, ALT, total serum albumin, platelet count, FIB-4 score, or APRI score, so those markers are excluded from the table. coefficients represent the average change in serum albumin (mg/dl) expected for an increase in DNA methylation from 0% to 100%.

selected biomarker of liver function in the adjusted model (Table 3).

EWAS and meta-analysis of liver biomarkers

Meta-analysis of the EWAS results from the EPIC and 450K datasets for each selected liver biomarker and each DNAm site included in both platforms identified that hypermethylation at nine DNAm sites were significantly associated with increased serum albumin among male veterans with HIV after adjusting for all covariates (Table 3; Figure 3A). A regional plot for the section of chromosome 17 that contains five significant DNAm sites is presented in Figure 3B. Quantile-quantile analysis of the expected and observed p-values from the meta-analysis for serum albumin showed very moderate global inflation (inflation factor of 1.02) of the nominal p-values (Figure 3C), so no further corrections were applied. Notably, among the top one hundred DNAm sites most strongly associated with (i.e., with the lowest p-values) serum albumin in the meta-analysis, the beta coefficients obtained from the two EWAS sub-cohorts (the EPIC and 450K) show a strong, positive correlation (beta = 0.89, p-value < 0.001) (Figure 3D).

Four of the nine DNAm sites was positively associated with serum albumin—cg16936953, cg18942579, cg01409343, and cg12054453—annotated to be located within the *TMEM49* gene. A 10% increase in methylation at each of those sites corresponded to average increases in serum albumin of 0.12 g/dl (95% CI: 0.08, 0.16 g/dl), 0.14 g/dl (95% CI: 0.09, 0.19 g/dl), 0.17 g/dl (95% CI: 0.11, 0.23 g/dl), and 0.10 g/dl (95% CI: 0.06, 0.14 g/dl), respectively. Four of the

remaining five DNAm sites—cg18181703, cg03546163, cg20995564, and cg23966214—annotated to correspond to the *SOCS3*, *FKBP5*, *ZEB2*, and *SAMD14* genes, respectively. A 10% increase in methylation at each of those DNAm sites corresponded to average increases in serum albumin of 0.19 g/dl (95% CI: 0.12, 0.26 g/dl), 0.11 g/dl (95% CI: 0.07, 0.15 g/dl), 0.12 g/dl (95% CI: 0.07, 0.17 g/dl), and 0.35 g/dl (95% CI: 0.21, 0.49 g/dl). The remaining DNAm site, cg12992827, was not annotated to correspond to any known coding sequence. A 10% increase in methylation at that site corresponded to an average increase in serum albumin of 0.19 g/dl (95% CI: 0.11, 0.26 g/dl). A quantile-quantile plot and Manhattan plot of the nominal p-values obtained from meta-analysis of the EPIC and 450K sub-cohorts are presented in Supplementary Figures S2, S3 for the other six liver biomarkers. Additionally, statistics and annotations for the ten DNAm sites with the lowest p-values after meta-analysis are presented in Supplementary Table S1 for each tested liver marker. Among the DNAm sites that were included in only one of the two methylation assay platforms used for this study, and that were not included in the meta-analysis, none was significantly associated with any of the liver biomarkers after correcting for multiple testing. Statistics and annotations for the ten DNAm sites with the lowest p-values among those that could not be meta-analyzed are presented for each liver biomarker in Supplementary Table S2.

Discussion

In this study we evaluated the epigenetic associations with seven clinically relevant biomarkers of liver function—aspartate

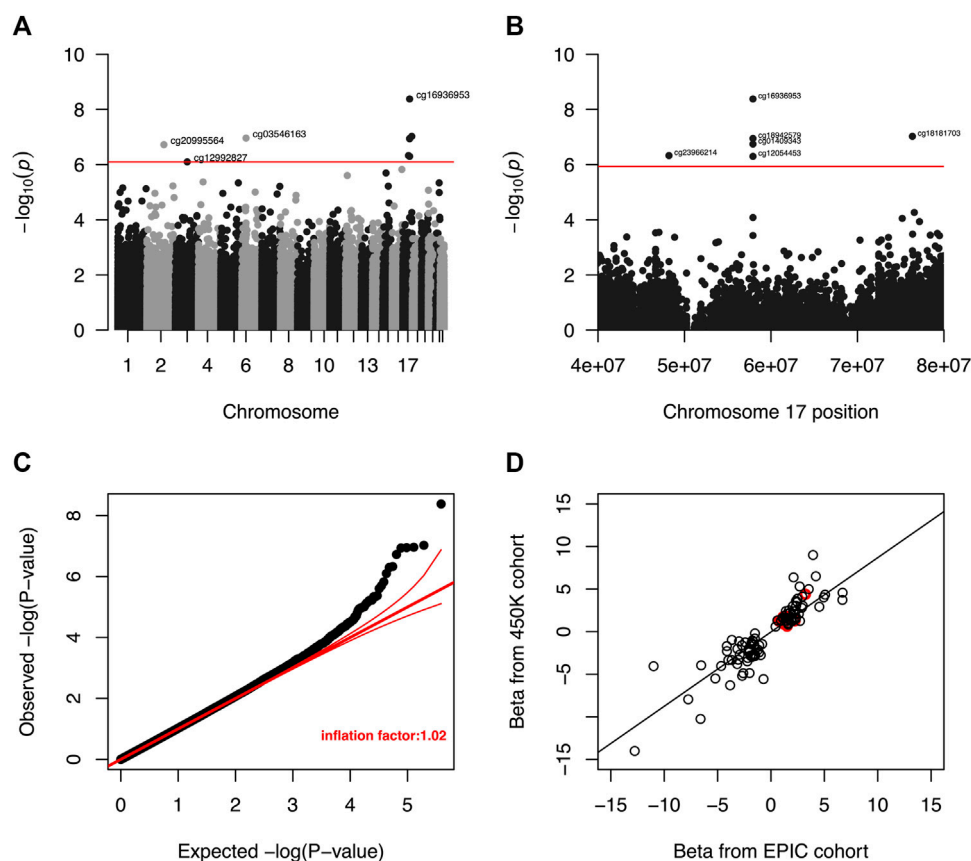


FIGURE 3

(A) Manhattan plot of unadjusted  $p$ -values from meta-analysis of EWAS results for serum albumin. The red line indicates the unadjusted  $p$ -value that corresponds to a threshold for FDR significance at  $Q < 0.05$ . DNA methylation was not significantly associated with serum albumin at any DNAm site after a more restrictive Bonferroni correction to an overall  $\alpha = 0.05$ . (B) Regional Manhattan plot of unadjusted  $p$ -values from meta-analysis of EWAS results for CpG sites in chromosome 17 and serum albumin. The red line indicates the unadjusted  $p$ -value that corresponds to a threshold for FDR significance (among all CpG sites from across entire genome) at  $Q < 0.05$ . DNA methylation was not significantly associated with serum for any CpG site after a more restrictive Bonferroni correction to an overall  $\alpha = 0.05$ . (C) Quantile-quantile plot of unadjusted  $p$ -values from meta-analysis of EWAS results for serum albumin. The global inflation factor was 1.02. Red lines represent a distribution of null associations with a 95% confidence interval. (D) Correlation of beta coefficients obtained from separate EWAS of serum albumin among the EPIC and 450K cohorts for the 100 DNAm sites with the smallest unadjusted  $p$ -values after meta-analysis. Beta coefficients from the separate analyses are positively correlated ( $r = 0.89$ ,  $p$ -value  $< 0.001$ ), indicating that the signs of observed beta coefficient are consistent between the two datasets. Circles representing  $p$ -value pairs that were significant after meta-analysis and FDR correction to  $Q < 0.05$  are highlighted in red.

transaminase, alanine transaminase, serum albumin, total bilirubin, platelet count, FIB-4 score, and APRI score—among male United States veterans with HIV. We identified that nine DNAm sites mapped to the genes *TMEM49*, *SOCS3*, *FKBP5*, *ZEB2*, and *SAMD14* were significantly associated with serum albumin levels. We also demonstrate an association between higher PhenoAA (an aging marker for DNAm AA) and low serum levels of albumin.

Serum albumin is a highly abundant protein primarily synthesized by the liver. Synthesis of albumin occurs in polysomes of hepatocytes at a rate of 10 g/day–15 g/day and accounts of 10% of liver protein synthesis (Garcovich et al., 2009). Albumin is not stored in the liver and therefore is not released on demand. Instead in situations of increased need in healthy adults, hepatocytes can increase albumin synthesis by up

to 300% (Garcovich et al., 2009). Impaired albumin synthesis and function have been reported in many liver diseases (Carvalho and Machado, 2018), and low serum levels of albumin may be a useful indicator of LD. Several studies have shown that serum albumin is an important prognostic factor and a significant predictor of death in patients with cirrhosis (Carvalho and Machado, 2018) and also a predictor of serious non-AIDS events in PWH (Ronit et al., 2018a; Ronit et al., 2018b).

Four of the nine serum albumin associated DNAm sites are located near the *TMEM49* gene (also known as *VMP1*). DNAm site *TMEM49*-cg16936953 was also associated with HIV infection, C-reactive protein and inflammatory bowel disease (Gross et al., 2016; Ligthart et al., 2016; Ventham et al., 2016). *TMEM49* encodes a transmembrane protein that drives cellular

autophagy (Molejon et al., 2013) and is overexpressed in pancreatitis-affected acinar cells (Vaccaro et al., 2003). Cellular autophagy, therefore, may be linked to regulation of serum albumin. Previous studies have shown that serum albumin suppresses autophagy *via* mTOR activation and that depriving cultured cells of serum albumin induces autophagy which is protective against the accumulation of harmful reactive oxygen species (ROS).

SOCS3 encodes a protein that inhibits cytokine signaling in the STAT pathway in response to elevated cytokine levels (Jo et al., 2005; Carow and Rottenberg, 2014). Suppression of SOCS3 contributes to liver fibrosis by increasing fibrogenic signaling *via* STAT3-mediated upregulation of tissue growth factor (TGF)- $\beta$  (Ogata et al., 2006; Jadid et al., 2018; Dees et al., 2020). In addition, obesity has been associated with downregulation of SOCS3, which raises the possibility that an association between DNAm near SOCS3 and liver dysfunction could result from the well described relationship between obesity and liver dysfunction. DNAm site SOCS3-cg18181703 was associated with BMI in several recent EWAS (Ali et al., 2016; Mendelson et al., 2017; Wahl et al., 2017; Xu et al., 2018; Sun et al., 2019). In this cohort of PWH, SOCS3-cg18181703 was associated with BMI ( $p$ -value:0.044) further supporting this hypothesis. Epigenetic associations between SOCS3-cg18181703 and C-reactive protein, inflammatory bowel disease, type-2 diabetes and cognitive abilities (Chambers et al., 2015; Ligthart et al., 2016; Marioni et al., 2018; Juvinao-Quintero et al., 2021) also suggest potential pleiotropic effects of SOCS3 gene.

FKBP5 encodes an immunoregulatory protein that contributes to basic protein folding and trafficking (Zannas et al., 2016). FKBP5 has been demonstrated to positively regulate the stress response and drive acquisition of metabolic disorders including obesity, insulin resistance, and diabetes (Sidibeh et al., 2018). Moreover, FKBP5 is thought to contribute to liver dysfunction (Kusumanchi et al., 2021) and in one study, deletion of FKBP5 protected knock-out mice from fatty liver disease despite high fat diets (Stechschulte et al., 2016). ZEB2 encodes a zinc-finger DNA-binding protein expressed in hepatocytes (Cai et al., 2012). Studies have demonstrated that repression of ZEB2 expression by microRNAs induces apoptosis in hepatocytes (Zhao et al., 2018), suggesting that ZEB2 might influence preservation of liver integrity in the face of hepatocellular damage. The remaining DNAm site associated positively with serum albumin levels was located near the SMAD14 gene. The function of the SMAD14 gene is unknown, but it has been linked to gastric cancer and hypothesized to act as a tumor suppressor gene (Xu et al., 2020). No studies have associated SMAD14 or its tumor suppressor properties to either general liver health and function or specifically to serum albumin regulation.

Despite the association between PhenoAA and serum albumin, no other significant associations were noted

between the four assessed metrics of DNAm AA—PhenoAA, GrimAA, IEAA, and EEAA—and other biomarkers. Compared to the other metrics of DNAm aging, PhenoAA is unique in that it is highly predictive of morbidity and mortality even after adjusting for chronological age (Liu et al., 2019). This may explain its significant association with albumin, which is also a prognostic marker of survival in LD (Carvalho and Verdelho Machado, 2018). It is plausible that specific DNAm sites may regulate the expression of genes in a manner that influences behavior of the selected biomarkers of liver health, either by impacting the progression of conditions that damage the liver and inhibit function—like fibrosis or hepatocellular carcinoma—or by affecting pathways that are specifically involved in regulation of a specific liver biomarkers.

Low serum level of albumin, though an important prognostic biomarker for LD (Carvalho and Verdelho Machado, 2018), is not specific. Thus, the associations we have uncovered in this study may be mediated by other non-hepatic factors in PWH. For example, hypoalbuminemia is associated with inflammation and synthesis of albumin is suppressed by pro-inflammatory cytokines TNF- $\alpha$  and Interleukin-1 (Dinarello, 1984; Perlmutter et al., 1986), both of which are important mediators of chronic inflammation in PWH (Deeks et al., 2013; De Pablo-Bernal et al., 2014). It is thus presumptuous to infer that the association between DNAm and albumin reflects underlying LD alone and is not combination of multiple interacting factors.

## Strengths and limitations

This study combines data from two different microarray platforms to assess consistent epigenetic associations of overlapping DNAm sites. This enabled a robust meta-analysis of epigenetic associations with a large sample size. The significant associations are likely to be true positive findings, and should be replicated in future studies. Secondly, we examine multiple biomarkers which when combined reflect overall liver function and individually inform on the liver's capacity to perform specific metabolic functions. This allowed us to examine DNAm sites across a wider range of genes that may link to specific aspects of liver function. Thirdly, unlike other EWAS of liver function or LD, our cohort included exclusively people with HIV, a group which has a higher burden of liver dysfunction and related morbidity. Our findings contribute to the identification of genes that may be relevant to liver health and function in the context of HIV infection.

One limitation of our study is that EWAS discovery was sub-optimally powered for a subset of DNAm sites included in only one of the platforms. Meta-analysis results from DNAm

sites measured on both platforms were highly consistent, suggesting that combining data from the two platforms successfully augments the power to detect significant associations. Future replication and meta-analysis with extended epigenomic coverage would improve the discovery of liver function-associated DNAm sites. Secondly, our study population was restricted to male veterans with HIV, limiting the generalizability of our findings to women with HIV. Thirdly, we used DNAm levels measured in blood cells to assess associations between DNAm and biomarkers of liver function, which may not represent DNAm patterns in the liver cells that are responsible for regulating these biomarkers. The lack of significant associations between DNAm and AST, ALT, total serum bilirubin, platelet count, FIB-4 score and, APRI score might result from an absence of epigenetic associations with liver functions of large effects, suggesting better powered future studies are needed to identify epigenetic associations with liver functions with smaller effects. Also, the potential for misclassifying liver fibrosis with the use of clinical scores like FIB-4 and APRI as has been shown in hepatitis B infection (Kim et al., 2016b) and may have limited our ability to detect an association with these markers.

## Conclusion

We identified epigenetic associations of both individual DNAm sites and DNAm AA with liver function through serum albumin in men with HIV. EWAS may inform on disease pathogenesis and generate new hypotheses for predicting disease progression and treatment of LD in HIV. Identifying specific genes linked with liver function among PWH can inform disease progression to improve the health of this patient population. Further replication analyses in independent cohorts are warranted to confirm and improve the discovery of the epigenetic markers of liver function and LD in PWH.

## Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: Due to US Department of Veterans Affairs (VA) regulations and our ethics agreements, the analytic data sets used for this study are not permitted to leave the VA firewall without a Data Use Agreement. This limitation is consistent with other studies based on VA data. However, VA data are made freely available to researchers with an approved VA study protocol. For more information, please visit <https://www.virec.research.va.gov> or contact the VA Information Resource Center at [VIREC@va.gov](mailto:VIREC@va.gov). Requests to access these datasets should be directed to [yan.v.sun@emory.edu](mailto:yan.v.sun@emory.edu).

## Ethics statement

The datasets for this article are not publicly available due to concerns of ethics agreement with the US Department of Veterans Affairs (VA). These data are part of the VACS database. Requests to access these datasets should be directed to the corresponding author, YS, [yan.v.sun@emory.edu](mailto:yan.v.sun@emory.edu). The institutional review board approved the original VACS cohort, and all participants provided written informed consent.

## Author contributions

BT and ML—Original manuscript draft and data analysis ZW, JC, QH—Data analysis VR, KA, KX, MF, MG, VM, AJ—Manuscript revision YS—Study design, data and, analysis and, manuscript revision.

## Funding

This work was supported by the following grants: BT, ML, ZW, JC, QH, VM and YS received support from the National Institute of Health (R01DK125187), VM and YS received support from Emory CFAR (P30AI050409), KX received support from the National Institute on Drug Abuse (R01DA042691, R01DA047820, R01DA047063), KS received support from NIH K01HL134147, VACS - NIH NIAAA U24 AA020794, U01 AA020790, U10 AA013566 completed.

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1020871/full#supplementary-material>

## References

- Ali, O., Cerjak, D., Kent, J. W., Jr., James, R., Blangero, J., Carless, M. A., et al. (2016). Methylation of SOCS3 is inversely associated with metabolic syndrome in an epigenome-wide association study of obesity. *Epigenetics* 11, 699–707. doi:10.1080/15592294.2016.1216284
- Bell, C. G., Lowe, R., Adams, P. D., Baccarelli, A. A., Beck, S., Bell, J. T., et al. (2019). DNA methylation aging clocks: Challenges and recommendations. *Genome Biol.* 20, 249. doi:10.1186/s13059-019-1824-y
- Cai, M. Y., Luo, R. Z., Chen, J. W., Pei, X. Q., Lu, J. B., Hou, J. H., et al. (2012). Overexpression of ZEB2 in peritumoral liver tissue correlates with favorable survival after curative resection of hepatocellular carcinoma. *PLoS One* 7, e32838. doi:10.1371/journal.pone.0032838
- Carow, B., and Rottenberg, M. E. (2014). SOCS3, a major regulator of infection and inflammation. *Front. Immunol.* 5, 58. doi:10.3389/fimmu.2014.00058
- Carvalho, J. R., and Machado, M. V. (2018). New insights about albumin and liver disease. *Ann. Hepatol.* 17, 547–560. doi:10.5604/01.3001.0012.0916
- Carvalho, J. R., and Verdelho Machado, M. (2018). New insights about albumin and liver disease. *Ann. Hepatol.* 17, 547–560. doi:10.5604/01.3001.0012.0916
- Chambers, J. C., Loh, M., Lehne, B., Drong, A., Kriebel, J., Motta, V., et al. (2015). Epigenome-wide association of DNA methylation markers in peripheral blood from Indian asians and Europeans with incident type 2 diabetes: A nested case-control study. *Lancet. Diabetes Endocrinol.* 3, 526–534. doi:10.1016/S2213-8587(15)00127-8
- Chen, B. H., Marioni, R. E., Colicino, E., Peters, M. J., Ward-Caviness, C. K., Tsai, P. C., et al. (2016). DNA methylation-based measures of biological age: meta-analysis predicting time to death. *Aging (Albany NY)* 8, 1844–1865. doi:10.18632/aging.101020
- Chen, J., Huang, Y., Hui, Q., Mathur, R., Gwinn, M., So-Armah, K., et al. (2020). Epigenetic associations with estimated glomerular filtration rate among men with human immunodeficiency virus infection. *Clin. Infect. Dis.* 70, 667–673. doi:10.1093/cid/ciz240
- De Pablo-Bernal, R. S., Ruiz-Mateos, E., Rosado, I., Dominguez-Molina, B., Alvarez-Rios, A. I., Carrillo-Vico, A., et al. (2014). TNF- $\alpha$  levels in HIV-infected patients after long-term suppressive cART persist as high as in elderly, HIV-uninfected subjects. *J. Antimicrob. Chemother.* 69, 3041–3046. doi:10.1093/jac/dku263
- Deeks, S. G., Tracy, R., and Douek, D. C. (2013). Systemic effects of inflammation on health during chronic HIV infection. *Immunity* 39, 633–645. doi:10.1016/j.immuni.2013.10.001
- Dees, C., Pötter, S., Zhang, Y., Bergmann, C., Zhou, X., Lubner, M., et al. (2020). TGF- $\beta$ -induced epigenetic deregulation of SOCS3 facilitates STAT3 signaling to promote fibrosis. *J. Clin. Invest.* 130, 2347–2363. doi:10.1172/JCI122462
- Dinarello, C. A. (1984). Interleukin-1 and the pathogenesis of the acute-phase response. *N. Engl. J. Med.* 311, 1413–1418. doi:10.1056/NEJM198411293112205
- Freiberg, M. S., McGinnis, K. A., Kraemer, K., Samet, J. H., Conigliaro, J., CuRtis Ellison, R., et al. (1999). The association between alcohol consumption and prevalent cardiovascular diseases among HIV-infected and HIV-uninfected men. *J. Acquir. Immune Defic. Syndr.* 53, 247–253. doi:10.1097/QAI.0b013e3181c6c4b7
- French, S. W. (2013). Epigenetic events in liver cancer resulting from alcoholic liver disease. *Alcohol Res.* 35, 57–67.
- Garcovich, M., Zocco, M. A., and Gasbarrini, A. (2009). Clinical use of albumin in hepatology. *Blood Transfus.* 7, 268–277. doi:10.2450/2008.0080-08
- Gross, A. M., Jaeger, P. A., Kreisberg, J. F., Licon, K., Jepsen, K. L., Khosroheidari, M., et al. (2016). Methylome-wide analysis of chronic HIV infection reveals five-year increase in biological age and epigenetic targeting of HLA. *Mol. Cell* 62, 157–168. doi:10.1016/j.molcel.2016.03.019
- Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sada, S., et al. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* 49, 359–367. doi:10.1016/j.molcel.2012.10.016
- Horvath, S., and Levine, A. J. (2015). HIV-1 infection accelerates age according to the epigenetic clock. *J. Infect. Dis.* 212, 1563–1573. doi:10.1093/infdis/jiv277
- Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., et al. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinforma.* 13, 86. doi:10.1186/1471-2105-13-86
- Jadid, F. Z., Chihab, H., Alj, H. S., Elfahry, R., Zaidane, I., Tazi, S., et al. (2018). Control of progression towards liver fibrosis and hepatocellular carcinoma by SOCS3 polymorphisms in chronic HCV-infected patients. *Infect. Genet. Evol.* 66, 1–8. doi:10.1016/j.meegid.2018.08.027
- Jo, D., Liu, D., Yao, S., Collins, R. D., and Hawiger, J. (2005). Intracellular protein therapy with SOCS3 inhibits inflammation and apoptosis. *Nat. Med.* 11, 892–898. doi:10.1038/nm1269
- Justice, A. C., Dombrowski, E., Conigliaro, J., Fultz, S. L., Gibson, D., Madenwald, T., et al. (2006). Veterans aging cohort study (VACS): Overview and description. *Med. Care* 44, S13–S24. doi:10.1097/01.mlr.0000223741.02074.66
- Juvinao-Quintero, D. L., Marioni, R. E., Ochoa-Rosales, C., Russ, T. C., Deary, I. J., van Meurs, J. B. J., et al. (2021). DNA methylation of blood cells is associated with prevalent type 2 diabetes in a meta-analysis of four European cohorts. *Clin. Epigenetics* 13, 40. doi:10.1186/s13148-021-01027-3
- Kaspar, M. B., and Sterling, R. K. (2017). Mechanisms of liver disease in patients infected with HIV. *BMJ Open Gastroenterol.* 4, e000166. doi:10.1136/bmjgast-2017-000166
- Kim, H. N., Nance, R., Van Rompaey, S., Delaney, J. C., Crane, H. M., Cachay, E. R., et al. (2016). Poorly controlled HIV infection: An independent risk factor for liver fibrosis. *J. Acquir. Immune Defic. Syndr.* 72, 437–443. doi:10.1097/QAI.0000000000000992
- Kim, W. R., Berg, T., Asselah, T., Flisiak, R., Fung, S., Gordon, S. C., et al. (2016). Evaluation of APRI and FIB-4 scoring systems for non-invasive assessment of hepatic fibrosis in chronic Hepatitis B patients. *J. Hepatol.* 64, 773–780. doi:10.1016/j.jhep.2015.11.012
- Konopnicki, D., Mocroft, A., de Wit, S., Antunes, F., Ledergerber, B., Katlama, C., et al. (2005). Hepatitis B and HIV: Prevalence, AIDS progression, response to highly active antiretroviral therapy and increased mortality in the EuroSIDA cohort. *AIDS* 19, 593–601. doi:10.1097/01.aids.0000163936.99401.f6
- Kusumanchi, P., Liang, T., Zhang, T., Ross, R. A., Han, S., Chandler, K., et al. (2021). Stress-responsive gene FK506-binding protein 51 mediates alcohol-induced liver injury through the hippo pathway and chemokine (C-X-C motif) ligand 1 signaling. *Hepatology* 74, 1234–1250. doi:10.1002/hep.31800
- Lemoine, M., Serfaty, L., and Capeau, J. (2012). From nonalcoholic fatty liver to nonalcoholic steatohepatitis and cirrhosis in HIV-infected patients: Diagnosis and management. *Curr. Opin. Infect. Dis.* 25, 10–16. doi:10.1097/QCO.0b013e32834ef599
- Levine, M. E., Hosgood, H. D., Chen, B., Absher, D., Assimes, T., and Horvath, S. (2015). DNA methylation age of blood predicts future onset of lung cancer in the women's health initiative. *Aging (Albany NY)* 7, 690–700. doi:10.18632/aging.100809
- Levine, M. E., Lu, A. T., Bennett, D. A., and Horvath, S. (2015). Epigenetic age of the pre-frontal cortex is associated with neuritic plaques, amyloid load, and Alzheimer's disease related cognitive functioning. *Aging (Albany NY)* 7, 1198–1211. doi:10.18632/aging.100864
- Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., et al. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)* 10, 573–591. doi:10.18632/aging.101414
- Ligthart, S., Marzi, C., Aslibekyan, S., Mendelson, M. M., Conneely, K. N., Tanaka, T., et al. (2016). DNA methylation signatures of chronic low-grade inflammation are associated with complex diseases. *Genome Biol.* 17, 255. doi:10.1186/s13059-016-1119-5
- Liu, Z., Kuo, P.-L., Horvath, S., Crimmins, E., Ferrucci, L., and Levine, M. (2019). A new aging measure captures morbidity and mortality risk across diverse subpopulations from nhanes IV: A cohort study. *PLoS Med.* 15, e1002718. doi:10.1371/journal.pmed.1002718
- Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., et al. (2019). DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging (Albany NY)* 11, 303–327. doi:10.18632/aging.101684
- Marioni, R. E., McRae, A. F., Bressler, J., Colicino, E., Hannon, E., Li, S., et al. (2018). Meta-analysis of epigenome-wide association studies of cognitive abilities. *Mol. Psychiatry* 23, 2133–2144. doi:10.1038/s41380-017-0008-y
- Mathur, R., Hui, Q., Huang, Y., Gwinn, M., So-Armah, K., Freiberg, M. S., et al. (2019). DNA methylation markers of type 2 diabetes mellitus among male veterans with or without human immunodeficiency virus infection. *J. Infect. Dis.* 219, 1959–1962. doi:10.1093/infdis/jiz273
- Mendelson, M. M., Marioni, R. E., Joeannes, R., Liu, C., Hedman, A. K., Aslibekyan, S., et al. (2017). Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: A mendelian randomization approach. *PLoS Med.* 14, e1002215. doi:10.1371/journal.pmed.1002215
- Molejon, M. I., Ropolo, A., Re, A. L., Boggio, V., and Vaccaro, M. I. (2013). The VMP1-Becn1 interaction regulates autophagy induction. *Sci. Rep.* 3, 1055. doi:10.1038/srep01055

- Moran, S., Arribas, C., and Esteller, M. (2016). Validation of a DNA methylation microarray for 850,000 CpG sites of the human genome enriched in enhancer sequences. *Epigenomics* 8, 389–399. doi:10.2217/epi.15.114
- Nishida, N., Nagasaka, T., Nishimura, T., Ikai, I., Boland, C. R., and Goel, A. (2008). Aberrant methylation of multiple tumor suppressor genes in aging liver, chronic hepatitis, and hepatocellular carcinoma. *Hepatology* 47, 908–918. doi:10.1002/hep.22110
- Ogata, H., Chinen, T., Yoshida, T., Kinjyo, I., Takaesu, G., Shirai, H., et al. (2006). Loss of SOCS3 in the liver promotes fibrosis by enhancing STAT3-mediated TGF-beta1 production. *Oncogene* 25, 2520–2530. doi:10.1038/sj.onc.1209281
- Palella, F. J., Baker, R. K., Moorman, A. C., Chmiel, J. S., Wood, K. C., Brooks, J. T., et al. (2006). Mortality in the highly active antiretroviral therapy era: Changing causes of death and disease in the HIV outpatient study. *J. Acquir. Immune Defic. Syndr.* 43, 27–34. doi:10.1097/01.qai.0000233310.90484.16
- Perlmuter, D. H., Dinarello, C. A., Punsal, P. I., and Colten, H. R. (1986). Cachectin/tumor necrosis factor regulates hepatic acute-phase gene expression. *J. Clin. Invest.* 78, 1349–1354. doi:10.1172/JCI112721
- Rickabaugh, T. M., Baxter, R. M., Sehl, M., Sinsheimer, J. S., Hultin, P. M., Hultin, L. E., et al. (2015). Acceleration of age-associated methylation patterns in HIV-1-infected adults. *PLoS One* 10, e0119201. doi:10.1371/journal.pone.0119201
- Ronit, A., Hatleberg, C. I., Ryom, L., Bonnet, F., El-Sadr, W., Reiss, P., et al. (2018). Associations between serum albumin and serious non-AIDS events among people living with HIV. *AIDS* 32, 1837–1848. doi:10.1097/QAD.0000000000001900
- Ronit, A., Sharma, S., Baker, J. V., Mngqibisa, R., Delory, T., Caldeira, L., et al. (2018). Serum albumin as a prognostic marker for serious non-AIDS endpoints in the strategic timing of antiretroviral treatment (START) study. *J. Infect. Dis.* 217, 405–412. doi:10.1093/infdis/jix350
- Rosen, A. D., Robertson, K. D., Hlady, R. A., Muench, C., Lee, J., Philibert, R., et al. (2018). DNA methylation age is accelerated in alcohol dependence. *Transl. Psychiatry* 8, 182. doi:10.1038/s41398-018-0233-4
- Shepard, C. W., Finelli, L., and Alter, M. J. (2005). Global epidemiology of hepatitis C virus infection. *Lancet. Infect. Dis.* 5, 558–567. doi:10.1016/S1473-3099(05)70216-4
- Sidibeh, C. O., Pereira, M. J., Abalo, X. M., J Boersma, G., Skrtic, S., Lundkvist, P., et al. (2018). FKBP5 expression in human adipose tissue: Potential role in glucose and lipid metabolism, adipogenesis and type 2 diabetes. *Endocrine* 62, 116–128. doi:10.1007/s12020-018-1674-5
- Smith, C., Sabin, C. A., Lundgren, J. D., Thiebaut, R., Weber, R., et al. (2010). Factors associated with specific causes of death amongst HIV-positive individuals in the D:A:D Study. *AIDS* 24, 1537–1548. doi:10.1097/QAD.0b013e32833a0918
- Smith, C. J., Ryom, L., Weber, R., Morlat, P., Pradier, C., Reiss, P., et al. (2014). Trends in underlying causes of death in people with HIV from 1999 to 2011 (D:A:D): A multicohort collaboration. *Lancet* 384, 241–248. doi:10.1016/S0140-6736(14)60604-8
- Solomon, O., MacIsaac, J., Quach, H., Tindula, G., Kobor, M. S., Huen, K., et al. (2018). Comparison of DNA methylation measured by Illumina 450K and EPIC BeadChips in blood of newborns and 14-year-old children. *Epigenetics* 13, 655–664. doi:10.1080/15592294.2018.1497386
- Stechschulte, L. A., Qiu, B., Warrier, M., Hinds, T. D., Zhang, M., Gu, H., et al. (2016). FKBP51 null mice are resistant to diet-induced obesity and the PPAR $\gamma$  agonist rosiglitazone. *Endocrinology* 157, 3888–3900. doi:10.1210/en.2015-1996
- Sterling, R. K., Lissen, E., Clumeck, N., Sola, R., Correa, M. C., Montaner, J., et al. (2006). Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology* 43, 1317–1325. doi:10.1002/hep.21178
- Su, P. F., Lee, T. C., Lin, P. J., Lee, P. H., Jeng, Y. M., Chen, C. H., et al. (2007). Differential DNA methylation associated with Hepatitis B virus infection in hepatocellular carcinoma. *Int. J. Cancer* 121, 1257–1264. doi:10.1002/ijc.22849
- Sun, D., Zhang, T., Su, S., Hao, G., Chen, T., Li, Q. Z., et al. (2019). Body mass index drives changes in DNA methylation: A longitudinal study. *Circ. Res.* 125, 824–833. doi:10.1161/CIRCRESAHA.119.315397
- Vaccaro, M. I., Grasso, D., Ropolo, A., Iovanna, J. L., and Cerquetti, M. C. (2003). VMP1 expression correlates with acinar cell cytoplasmic vacuolization in arginine-induced acute pancreatitis. *Pancreatol.* 3, 69–74. doi:10.1159/000069150
- Varela-Rey, M., Woodhoo, A., Martinez-Chantar, M. L., Mato, J. M., and Lu, S. C. (2013). Alcohol, DNA methylation, and cancer. *Alcohol Res.* 35, 25–35.
- Ventham, N. T., Kennedy, N. A., Adams, A. T., Kalla, R., Heath, S., O'Leary, K. R., et al. (2016). Integrative epigenome-wide analysis demonstrates that DNA methylation may mediate genetic risk in inflammatory bowel disease. *Nat. Commun.* 7, 13507. doi:10.1038/ncomms13507
- Wahl, S., Drong, A., Lehne, B., Loh, M., Scott, W. R., Kunze, S., et al. (2017). Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 541, 81–86. doi:10.1038/nature20784
- Wai, C. T., Greenson, J. K., Fontana, R. J., Kalbfleisch, J. D., Marrero, J. A., Conjeevaram, H. S., et al. (2003). A simple noninvasive index can predict both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology* 38, 518–526. doi:10.1053/jhep.2003.50346
- Xu, K., Zhang, X., Wang, Z., Hu, Y., and Sinha, R. (2018). Epigenome-wide association analysis revealed that SOCS3 methylation influences the effect of cumulative stress on obesity. *Biol. Psychol.* 131, 63–71. doi:10.1016/j.biopsycho.2016.11.001
- Xu, X., Chang, X., Xu, Y., Deng, P., Wang, J., Zhang, C., et al. (2020). SAMD14 promoter methylation is strongly associated with gene expression and poor prognosis in gastric cancer. *Int. J. Clin. Oncol.* 25, 1105–1114. doi:10.1007/s10147-020-01647-4
- Zakhari, S. (2013). Alcohol metabolism and epigenetics changes. *Alcohol Res.* 35, 6–16.
- Zannas, A. S., Wiechmann, T., Gassen, N. C., and Binder, E. B. (2016). Gene-stress-epigenetic regulation of FKBP5: Clinical and translational implications. *Neuropsychopharmacology* 41, 261–274. doi:10.1038/npp.2015.235
- Zhang, X., Asllanaj, E., Amiri, M., Portilla-Fernandez, E., Bramer, W. M., Nano, J., et al. (2021). Deciphering the role of epigenetic modifications in fatty liver disease: A systematic review. *Eur. J. Clin. Invest.* 51, e13479. doi:10.1111/eci.13479
- Zhang, X., Hu, Y., Justice, A. C., Li, B., Wang, Z., Zhao, H., et al. (2017). DNA methylation signatures of illicit drug injection and hepatitis C are associated with HIV frailty. *Nat. Commun.* 8, 2243. doi:10.1038/s41467-017-02326-1
- Zhao, Y., Xue, F., Sun, J., Guo, S., Zhang, H., Qiu, B., et al. (2014). Genome-wide methylation profiling of the different stages of Hepatitis B virus-related hepatocellular carcinoma development in plasma cell-free DNA reveals potential biomarkers for early detection and high-risk monitoring of hepatocellular carcinoma. *Clin. Epigenetics* 6, 30. doi:10.1186/1868-7083-6-30
- Zhao, Y. X., Sun, Y. Y., Huang, A. L., Li, X. F., Huang, C., Ma, T. T., et al. (2018). MicroRNA-200a induces apoptosis by targeting ZEB2 in alcoholic liver disease. *Cell Cycle* 17, 250–262. doi:10.1080/15384101.2017.1417708



## OPEN ACCESS

## EDITED BY

Xiaofan Lu,  
INSERM U964 Institut de Génétique et  
de Biologie Moléculaire et Cellulaire  
(IGBMC), France

## REVIEWED BY

Jialin Meng,  
First Affiliated Hospital of Anhui Medical  
University, China  
Huan Chen,  
Zhuzhou Central Hospital, China

## \*CORRESPONDENCE

Meiping Guan,  
mpguan@163.com

<sup>†</sup>These authors have contributed equally  
to this work and share first authorship

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 21 August 2022

ACCEPTED 23 September 2022

PUBLISHED 14 October 2022

## CITATION

Wu P, Guo L, Li X, Du Y, Lin X, Ma X, Lin Y,  
Wen C, Yang C, Liu N, Feng Q, Xue Y and  
Guan M (2022), Comprehensive analysis  
of epigenomics and transcriptome data  
to identify potential target genes  
associated with obesity.  
*Front. Genet.* 13:1024300.  
doi: 10.3389/fgene.2022.1024300

## COPYRIGHT

© 2022 Wu, Guo, Li, Du, Lin, Ma, Lin,  
Wen, Yang, Liu, Feng, Xue and Guan.  
This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/)  
(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.

# Comprehensive analysis of epigenomics and transcriptome data to identify potential target genes associated with obesity

Peili Wu<sup>1†</sup>, Lei Guo<sup>1†</sup>, Xuelin Li<sup>1</sup>, Yuejun Du<sup>2</sup>, Xiaochun Lin<sup>1</sup>,  
Xiaoqin Ma<sup>1</sup>, Yingbei Lin<sup>1</sup>, Churan Wen<sup>1</sup>, Chuyi Yang<sup>1</sup>,  
Nannan Liu<sup>1</sup>, Qijian Feng<sup>1</sup>, Yaoming Xue<sup>1</sup> and Meiping Guan<sup>1\*</sup>

<sup>1</sup>Department of Endocrinology and Metabolism, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, China, <sup>2</sup>Department of Urology, Nanfang Hospital, Southern Medical University, Guangzhou, Guangdong, China

DNA methylation is closely related to the occurrence and development of many diseases, but its role in obesity is still unclear. This study aimed to find the potential differentially methylated genes associated with obesity occurrence and development. By combining methylation and transcriptome analysis, we identified the key genes in adipose tissue affecting the occurrence and development of obesity and revealed the possible molecular mechanisms involved in obesity pathogenesis. We first screened 14 methylation-related differential genes and verified their expression in adipose tissue by quantitative polymerase chain reaction (qPCR). Seven genes with the same expression pattern were identified as key genes, namely, *CCRL2*, *GPT*, *LGALS12*, *PC*, *SLC27A2*, *SLC4A4*, and *TTC36*. Then, the immune microenvironment of adipose tissue was quantified by CIBERSORT, and we found that the content of M0 macrophages and T follicular helper cells in adipose tissue was significantly increased and decreased, respectively, in the obese group. Furthermore, the relationship between key genes and the immune microenvironment was analyzed. Additionally, the metabolic pathway activity of each sample was calculated based on the ssGSEA algorithm, and the key gene–metabolic network was constructed. Moreover, we performed a CMAP analysis based on the differential genes in adipose tissue to screen out drugs potentially effective in obesity treatment. In conclusion, we identified seven methylation-related key genes closely related to obesity pathogenesis and explored the potential mechanism of their role in obesity. This study provided novel insights into the molecular mechanisms and management of obesity.

## KEYWORDS

obesity, methylation, immune infiltration, metabolism, comprehensive analysis

## Introduction

Obesity is a chronic metabolic disease. Currently, the incidence rate of obesity and obesity-based metabolic diseases is increasing worldwide (Ng et al., 2014), greatly reducing the quality of life. Therefore, it is increasingly important to explore and develop effective treatments for obesity. Obesity is caused by excessive accumulation and abnormal distribution of body fat due to long-term energy intake exceeding consumption. Adipose tissue is the core of the body's energy balance and metabolic homeostasis. Targeting adipose tissue metabolic activity is one of the most attractive strategies to effectively reduce weight and improve metabolism.

Adipose tissue is mainly divided into two categories, brown adipose tissue (BAT) and white adipose tissue (WAT). Among them, BAT resists cold and increases energy consumption by increasing heat production, while WAT, as the main site for nutrition storage, secretes important endocrine hormones to communicate with the central nervous system and other peripheral tissues, which jointly maintain the changing nutritional environment (Kajimura et al., 2015; Chen et al., 2017). Additionally, WAT can also coordinate with immune cells and play an important role in integrating immune and metabolic signals (Martinez-Santibanez and Lumeng, 2014; Crewe et al., 2017), revealing a potential new target for preventing and treating obesity and its related complications by regulating the immune microenvironment.

DNA methylation is associated with causing genetic changes without changing the DNA sequence, which is one of the most stable epigenetic mechanisms. DNA methylation is widely involved in the occurrence and development of various characteristics and diseases. Methylation of DNA is generally associated with elevated gene expression, whereas unmethylated DNA is associated with repressed gene expression (Parle-McDermott and Ozaki, 2011). The mechanism is believed to be linked with transcriptional silencing involving the interference of RNA polymerase complex and associated transcription factors (Klose and Bird, 2006; Parle-McDermott and Ozaki, 2011). In obesity, studies have shown that DNA methylation can regulate the homeostasis of energy balance in WAT. Previous studies have reported that methylation changes of certain genes, such as ATP binding cassette subfamily G member 1 (*ABCG1*), sterol regulatory element binding transcription factor (*SREBP1*), and carnitine palmitoyl transferase 1A (*CPT1A*), were related to obesity, and these DNA methylation-related markers might act as predictors of obesity and related diseases. However, due to the unsatisfactory duplication or validation results of DNA methylation markers in obesity, clinical application of DNA methylation-related genes is still relatively limited. Hence, further studies on the key genes of DNA methylation are needed in the future to lay a foundation for them as predictors of obesity and related diseases.

Differential genes and methylation profiles between the control and obese groups were analyzed by combining methylation and transcriptome data to seek potential differentially methylated genes associated with obesity. Furthermore, the selected candidate genes were subsequently verified in human adipose tissue samples. The immune microenvironment and metabolic regulation pathway were further explored, and the relationship between them and the key genes was also analyzed. Additionally, drug prediction was performed based on differential genes of adipose tissue, and potential small molecules for obesity treatment were screened. This study provided potential obesity-related therapeutic targets for further research and obesity management in clinical practice.

## Materials and methods

### Data acquisition

The GEO database (<https://www.ncbi.nlm.nih.gov/geo>) stores gene expression data from all over the world (Barrett et al., 2007). In this study, three datasets based on subcutaneous adipose analysis were enrolled. The 450-k methylation matrix within GSE67024 (GPL13534) was downloaded from the GEO database, including 15 obese patients (BMI  $41.36 \pm 4.54 \text{ kg/m}^2$ ) and 14 never-obese control women (BMI  $25.11 \pm 2.49 \text{ kg/m}^2$ ). The expression profile of GSE174475 (GPL18573) was downloaded, including 14 obese patients (BMI  $34.31 \pm 3.17 \text{ kg/m}^2$ ) and 29 non-obese control women (BMI  $26.45 \pm 2.47 \text{ kg/m}^2$ ). The expression profile of GSE156906 (GPL24676) was acquired, including 52 obese patients with an average BMI near  $38.45 \text{ kg/m}^2$  and 14 non-obese control with an average BMI of  $22.90 \text{ kg/m}^2$  (Fuchs et al., 2021). Age and gender between obese and non-obese patients in the datasets show no statistical difference (more details in Supplementary Tables S1–S3).

### Differential gene expression analysis and enrichment analysis

The limma package is a commonly used method to identify differential genes (Ritchie et al., 2015). In this study, the R package “limma” was used to analyze the differentially expressed genes between control and obese samples. Filter criteria of differentially expressed genes were set as  $|\log_2 \text{ fold change}| > 1$  and a  $p < 0.05$ . Differential gene volcano plots were drawn using the R package “ggplot2” (<https://link.springer.com/book/10.1007/978-3-319-24277-4>), and differential gene heatmaps were drawn using the “pheatmap” package. The “clusterProfiler” (Yu et al., 2012) package was used to perform GO and KEGG pathway enrichment analysis, and the possible pathways of differential genes involved in the progression of obesity were comprehensively explored. GO analysis included three different categories, namely, biological process (BP),

molecular function (MF), and cellular component (CC). Enriched pathways with both P and Q values less than 0.05 were considered to be significant.

## DNA methylation analysis

ChAMP is a comprehensive methylation analysis pipeline including functions of low-quality probe filtering, differential methylation position detection, methylated genomic block detection, and so on. In this study, the methylation data in the GSE67024 dataset were analyzed by the R package “ChAMP” (Morris et al., 2014; Tian et al., 2017), and the methylation levels of adipose tissue genes between the control and the obese groups were compared. The differentially methylated genes were screened. Filter criteria of differentially methylated genes were  $|\log_2 \text{ fold change}| > 0.2$  and an adjusted P of  $< 0.05$ . Lower-expression genes with hypermethylation and higher-expression genes with hypomethylation were identified as key genes in disease development.

## Patients and quantitative real-time PCR

Patients who met the inclusion criteria were enrolled: a. aged 18–65; b. subcutaneous adipose tissue and clinical data available; c. informed consent. Subcutaneous adipose tissue was obtained from 14 men and 10 women with diseases such as renal cysts, kidney stones, adrenal adenoma, and primary aldosteronism, who underwent laparoscopic adrenalectomy (more details in [Supplementary Materials](#)). According to the Chinese standard of obesity, in this study, the obese group was set with a BMI of no less than  $28 \text{ kg/m}^2$ . The average BMI in the obese group was significantly different compared to that of the control group ( $30.01 \pm 1.48 \text{ kg/m}^2$  vs.  $22.43 \pm 2.81 \text{ kg/m}^2$ ,  $p < 0.001$ ). No significant difference in age and gender between the two groups was observed (more details in [Supplementary Table S4](#)). It was approved by the Ethics Committee of Nanfang Hospital, and all subjects. Total RNA was extracted from subcutaneous adipose tissue, cDNA was reverse-transcribed, and quantitative reverse transcription-polymerase chain reaction (RT-PCR) was conducted on the QuantStudio5 real-time PCR system. All primers were confirmed by PubMed BLAST, and human  $\beta$ -ACTIN served as the internal control. Primer sequences are given in the [Supplementary Materials](#).

## Immune cell infiltration

CIBERSORT is a deconvolution method that effectively links the gene expression data with the content of immune cells and quantifies the composition of immune cells through gene expression profiles (Chen et al., 2018). For each case, the total of predicted cell subgroups was equal to 1, which indicates the

enrichment of immune cell infiltration. This study quantified the immune microenvironment of adipose tissue gene expression data of GSE174475 by the CIBERSORT algorithm, inferred the relative content of immune cells in adipose tissue, and compared the relative level of immune cell content in different groups. Additionally, Pearson’s correlation analysis was used to explore the correlation between key genes and the content of immune cells, and  $p < 0.05$  was considered statistically significant.

## Connectivity map analysis of drugs

The connectivity map (CMap) is a database for drug prediction based on differential gene expression (Lamb et al., 2006). It contains 6100 examples of 1309 small-molecule drugs and is mainly used to reveal the functional relationship between genes, drugs, and diseases. In the present study, we predicted potential molecular compounds for obesity treatment based on the CMAP database. We first analyzed the differential genes between the control and obese groups in adipose tissue and then input the differential genes into the CMAP database for drug prediction. The correlation between drugs and differential genes was displayed by a score from  $-1$  to  $1$ , with a negative number indicating the gene expression pattern of the corresponding interference, suggesting that this interference might have a potential therapeutic effect on obesity.

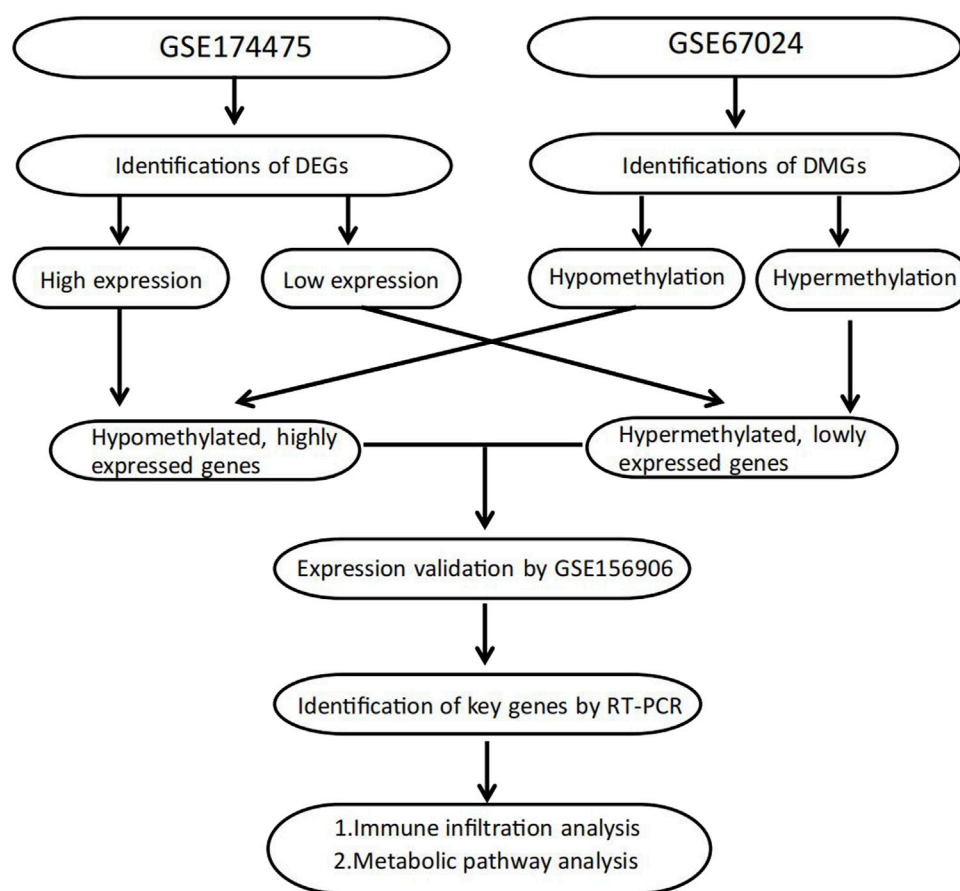
## Statistical analysis

Data analysis was conducted by the R framework (version 4.1.2). All statistical tests were bilateral, and a  $p$ -value  $< 0.05$  was considered statistically significant.

## Results

### Differential expression analysis of expression profiles and methylation, gene ontology, and Kyoto Encyclopedia of Genes and Genomes Enrichment Analysis

The study design of this research is displayed in [Figure 1](#). In this study, GSE174475, including 43 cases of adipose tissue transcriptome data, was downloaded from the NCBI GEO public database. According to the expression profile data of adipose tissue, differential analysis was conducted to screen the differential genes between the control and obesity groups. A total of 742 differential genes were selected, including 614 upregulated genes and 128 downregulated genes ([Figures 2A,B](#)). We further performed pathway enrichment analysis on these differential genes. The Gene Ontology (GO) results showed that the differential genes were mainly enriched in the signaling pathways such as leucocyte-mediated



**FIGURE 1**  
Flowchart of study design.

immunity, adaptive immune response based on somatic, T cell activation, and lymphocyte-mediated immunity (Figure 2C). The Kyoto Encyclopedia of Genes and Genomes (KEGG) results displayed that the differential genes were mainly enriched in the chemical signaling pathway, natural killer cell-mediated cytotoxicity, Th17 cell differentiation, and other signaling pathways (Figure 2D). On the other hand, we downloaded the 450-k data of GSE67024 from the NCBI GEO public database, including the methylation gene profiles in adipose tissue of 29 cases, and analyzed the differential methylation sites with the ChAMP package. A total of 1186 significant differentially methylated probes were screened, including corresponding 81 hypomethylated genes and 647 hypermethylated genes (Figures 3A,B).

## Combined analysis of DNA methylation and transcriptome in adipose tissue

Genes that were lower expressed in transcriptome analysis while being hypermethylated in the methylation dataset or higher

expression genes with hypomethylation were further selected. The Venn plots showed that 14 genes were identified, namely, *AGBL4*, *CCRL2*, *CHST11*, *FASN*, *GPT*, *LAPTM5*, *LGALS12*, *LRIG1*, *PC*, *SLC27A2*, *SLC4A4*, *TRPV2*, and *TTC36* (Figures 4A,B). Additionally, the present study verified the expression pattern of these genes through another adipose tissue dataset, GSE156906. The results showed that 13 of the 14 genes had the same expression pattern, namely, *AGBL4*, *CCRL2*, *CHST11*, *FASN*, *GPT*, *LAPTM5*, *LGALS12*, *LRIG1*, *PC*, *SLC27A2*, *SLC4A4*, *TRPV2*, and *TTC36* (Figures 4C,D). Moreover, we also verified the expression of key genes in adipose tissue between nine obese and 15 non-obese patients

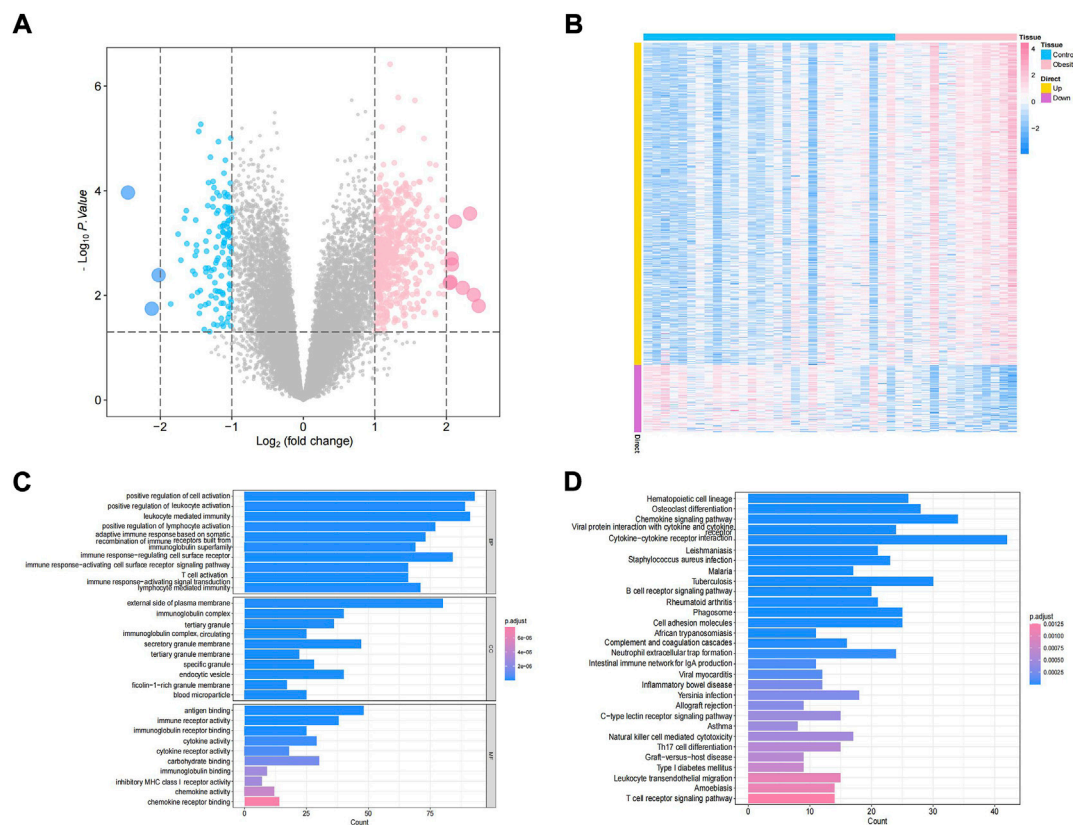


FIGURE 2

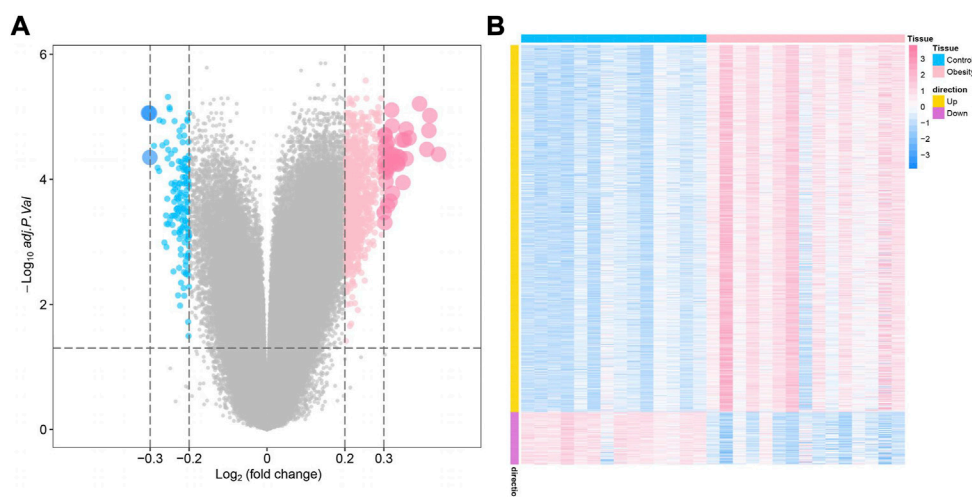
Identification of obesity differential genes. (A) Volcano plot of differential gene expression. Blue indicates differentially expressed downregulated genes, and pink indicates differentially expressed upregulated genes. (B) Differential gene expression heatmap, where blue indicates low-expression genes and pink indicates high-expression genes. (C,D) GO/KEGG enrichment pathway of differential genes.

by RT-qPCR and found that seven genes remained in the same expression pattern in obese patients, namely, *CCRL2*, *GPT*, *LGALS12*, *PC*, *SLC27A2*, *SLC4A4*, and *TTC36* (all  $p$  value < 0.05) (Figure 4E). Therefore, these seven genes were regarded as potential key genes in obesity, and further analysis in exploring their potential molecular mechanisms in adipose tissue was conducted.

## Analysis of the immune microenvironment in adipose tissue

In addition to storing nutrients, adipose tissue is also an important immune organ in humans. It contains many immune cells, such as T cells, B cells, and macrophages. Under different microenvironments or stimulation, adipose tissue can secrete cytokines to recruit immune cells and further secrete various inflammatory factors, ultimately leading to chronic inflammation and insulin resistance. This study further explored the immune microenvironment

of adipose tissue and the correlation between key genes and immune cell content in obesity based on the dataset GSE174475. We first quantified the content of immune cells in each sample by the CIBERSORT algorithm (Figures 5A,B). The results showed that the content of M0 macrophages was significantly increased in the obese group compared to that of the control group, while the content of T follicular helper cells was significantly decreased in the obese group (Figure 5C). The immunomicroenvironmental analysis based on GSE156906 also showed that the number of M0 macrophages was significantly increased in the obese group (Supplementary Figures S1A–C), which was consistent with the abovementioned results. Furthermore, we analyzed the correlation between the expression of key genes and the content of immune cells. The results showed that the seven genes were strongly correlated with the content of immune cells, suggesting that the key genes might affect the occurrence and development of obesity by influencing the immune microenvironment of adipose tissue (Figures 6A–G).



**FIGURE 3**

Identification of obesity differentially methylated genes. **(A)** Volcano plot of differentially methylated genes, where blue indicates differentially methylated downregulated genes and pink indicates differentially methylated upregulated genes. **(B)** Differential methylation heatmap, where blue indicates hypomethylated genes and pink indicates hypermethylated genes.

## Metabolic activity analysis of adipose tissue

Adipose tissue can sense the energy metabolism of the body and then secrete adipokines to regulate energy homeostasis. The dysfunction of adipose tissue is manifested by the aggravation of inflammatory factors and insulin resistance. The dysfunction of adipose tissue and dysregulated glucolipid metabolism accelerate each other. In this study, the metabolic pathway score of each sample was quantified by the ssGSEA algorithm. In the metabolic pathway heatmap, significant difference in scores for lipid metabolism-related signatures and other metabolism signatures between the control and obese samples was observed (Figure 7A). Additionally, we also further quantified the correlation between key genes and metabolic pathways and constructed a key gene–metabolic pathway network (Figures 7B,C).

## Prediction of potential therapeutic drugs for obese patients

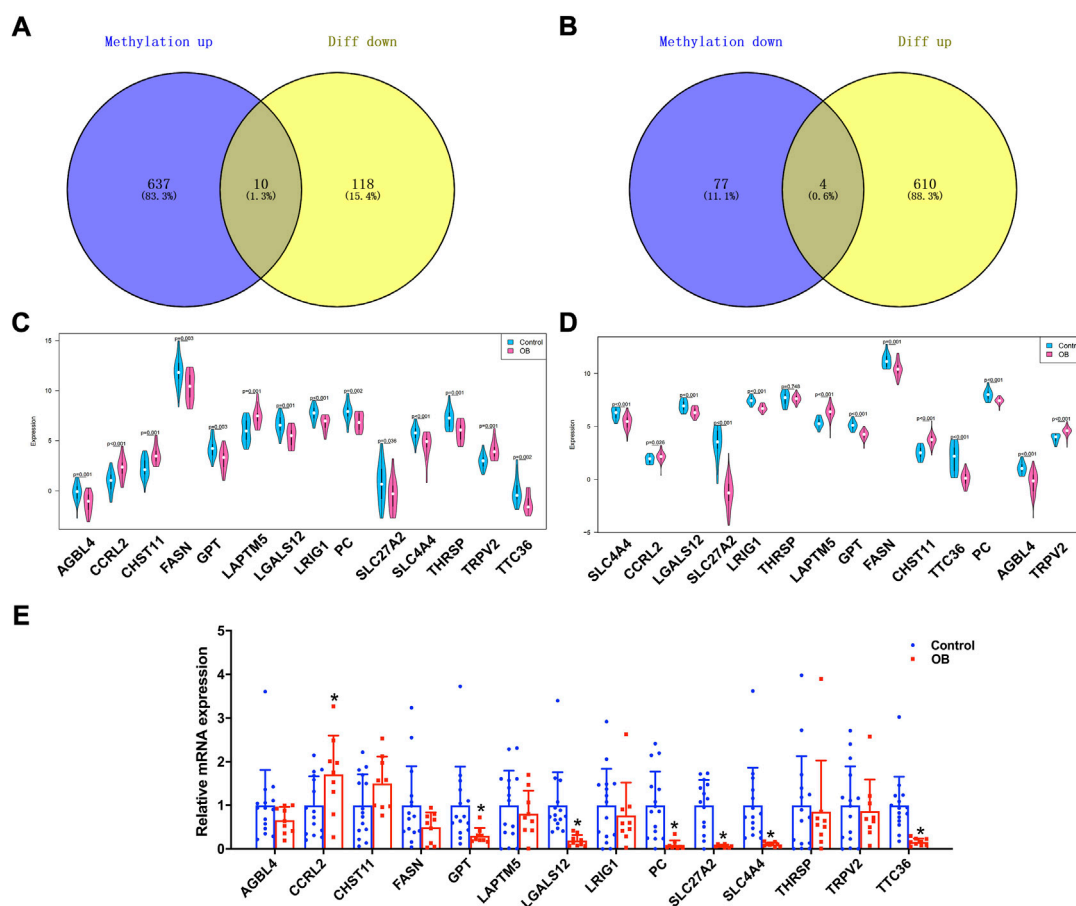
To identify potential drugs for the targeted treatment of obesity, we used the CMap database to evaluate the potential of small-molecule drugs in obesity treatment. We uploaded differentially expressed genes of adipose tissue to the CMap database and identified drugs related to obesity treatment. Among them, vemurafenib, NTNCB, dilazep, and PD-0325901 ((with available structure in PubChem database) were highly negatively correlated with obesity progression, indicating

that these compounds might have therapeutic effects on obesity (Figure 8A). The mechanism of action (MOA) and drug targets of these drugs are identified using the CMap database to explore the potential mechanism of their treatment of obesity (Figure 8A), and the 2D visualization of these four drugs is shown (Figure 8B).

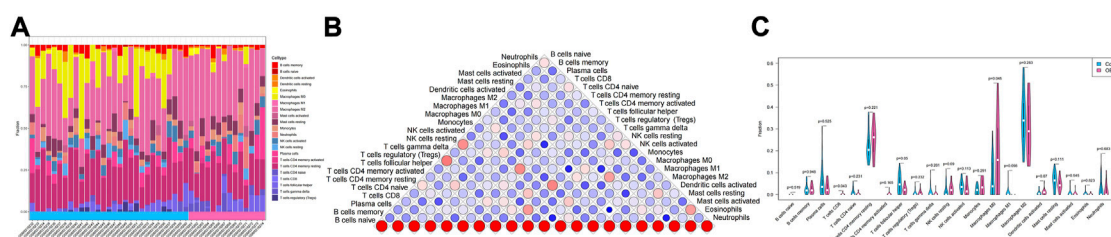
## Discussion

In this study, by analyzing three GEO datasets and verifying in 24 patients, we identified that *CCRL2*, *GPT*, *LGALS12*, *PC*, *SLC27A2*, *SLC4A4*, and *TTC36* might be the key genes that play an important role in obesity pathogenesis. Furthermore, these seven genes showed association with immune microenvironment and metabolic pathways in adipose tissue of obese patients. Additionally, we found that, compared to the control, M0 macrophage content was higher, while T follicular helper cell content was lower in adipose tissues in the obese group. Also, possible therapeutic drugs for obese patients were shown in our study. This study provided potential obesity-related therapeutic targets for further research and obesity management in clinical practice.

An increasing number of studies have shown that methylated genes are closely related to obesity. Simone et al. have identified 278 BMI-associated CpG sites in blood from epigenomic associations in European and Indian Asian populations, of which 120 CpG sites were consistently associated with BMI in both adipose tissue and blood. The epigenome-wide association study (EWAS) in the Asian population has also identified



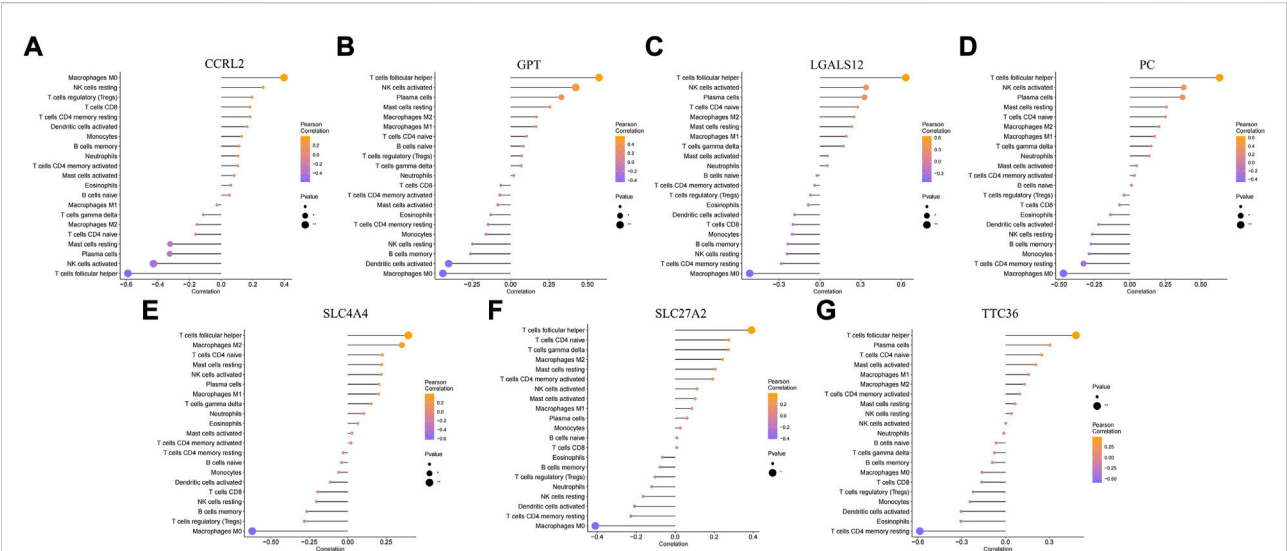
**FIGURE 4**  
Identification of obesity hub genes. **(A,B)** Venn diagram of hypermethylated-low expression and hypomethylated-high expression genes. **(C,D)** Expression patterns of genes in GSE174475 and GSE156906. **(E)** qPCR experiments verified the expression of 14 genes.



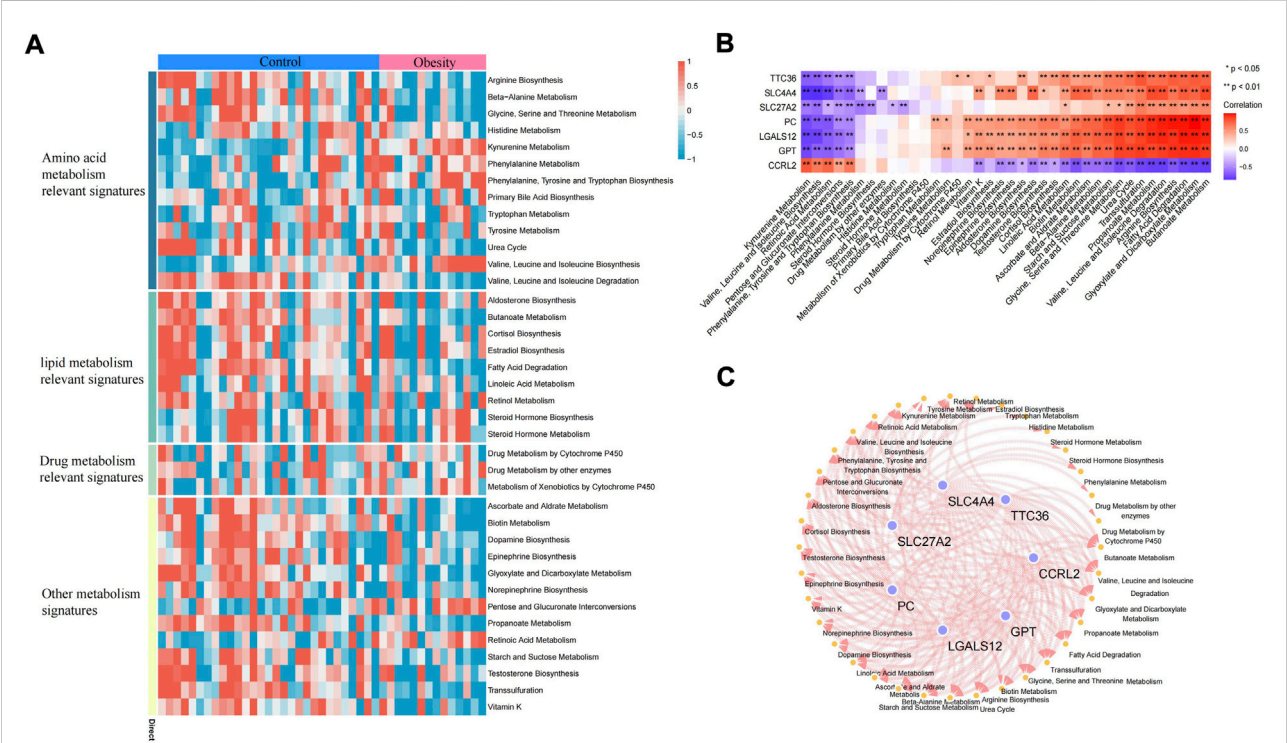
**FIGURE 5**  
Immune infiltration of adipose tissue. **(A)** Relative percentage of 22 immune cell subpopulations in the sample. **(B)** Correlation among 22 immune cells, where blue indicates positive correlation and red indicates positive correlation. **(C)** Difference in immune cell content between normal samples and obese samples, with blue indicating normal samples and pink indicating obese samples.  $p < 0.05$  was considered statistically significant.

116 BMI-associated CpGs, of which 108 were consistent with those of previous studies. In addition to being biomarkers, methylated genes have also been useful for predicting obesity

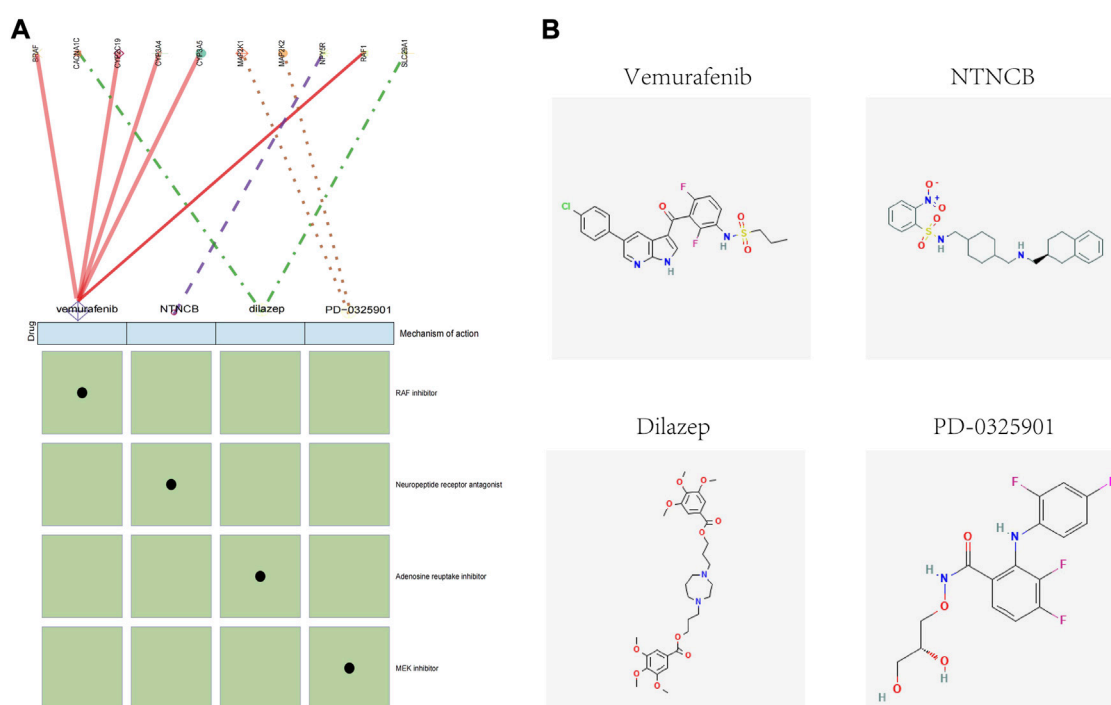
and its related metabolic diseases (Wahl et al., 2017; Gallardo-Escribano et al., 2020). However, although many studies have suggested that excessive accumulation of fat caused most of the



**FIGURE 6** Correlation between key genes and immune cells. (A–G) Relationship between key genes and immune cells, and the genes are *CCRL2*, *GPT*, *LGALS12*, *PC*, *SLC4A4*, *SLC27A2*, and *TTC36*, respectively.



**FIGURE 7** Metabolic pathway of adipose tissue. (A) Heatmap of the metabolic pathway of adipose tissue. (B) Correlation analysis between key genes and metabolic pathways. (C) Relationship between key genes and metabolic pathways (the correlation coefficient is greater than 0.3).

**FIGURE 8**

CMap analysis of the therapeutic drugs for obesity. **(A)** Molecular action and targets of potentially therapeutic drugs. **(B)** 2D visualization of these four drugs.

methylation changes (Wahl et al., 2017; Chen et al., 2021), others have insisted that methylation might be the cause of obesity (Mendelson et al., 2017). Additionally, methylated genes might also be associated with inflammation and lipoprotein-related biological pathways, further aggravating metabolic disorders (Wahl et al., 2017; Chen et al., 2021). These findings still need more experimental studies for further confirmation. In this study, seven candidate obesity-related methylation genes were identified, and our results suggested that these genes might participate in immune infiltration and in glucolipid metabolism.

Adipose tissue is also an important immune organ, containing many immune cells that function to maintain immune homeostasis. Infiltration and activation of immune cells during obesity is a powerful mechanism for adipose tissue remodeling (Engin, 2017), highlighting their potential as immunotherapeutic targets for the prevention and treatment of metabolic diseases. Studies in humans and mice have shown that the content of macrophages in adipose tissue is proportional to BMI (Weisberg et al., 2003). With obesity, adipose tissue is in a low-intensity inflammatory state, in which the percentage of macrophage infiltration in adipose tissue is significantly increased to 41% (Weisberg et al., 2003) compared with the normal state, which is consistent with the results of our study in terms that macrophages were significantly increased in the tissue of obese patients. More importantly, our analysis found that T

follicular helper cells were significantly decreased in obesity, which has been rarely reported in previous studies. No reports have been found on the correlation between T follicular helper cells and adipose tissue or obesity. It should be noted that the BCL6 transcription repressor (Bcl6) has been shown to play an important role in regulating lipid metabolism in adipose tissue. Mice lacking Bcl6 showed multiple characteristics of lipid metabolism disorders, accompanied by adipose tissue dysplasia (LaPensee et al., 2014; Hu et al., 2016). Bcl6 is essential for the differentiation and development of T follicular helper cells (Yu et al., 2009), suggesting that the decrease in T follicular helper cells might be involved in the functional imbalance of adipose tissue in obesity.

In this study, consistent with previous studies (Xu et al., 2022), we found that *CCRL2* was significantly overexpressed in the subcutaneous tissues of obese subjects. However, this result needs to be interpreted with more caution, as one study suggested that the high expression of *CCRL2* in adipose tissue might be a compensatory result. Moreover, *CCRL2* lacks a signal transduction structure to transmit ligand signals into cells; hence, it might affect macrophage infiltration by competitively binding ligands (Xu et al., 2022). *GPT*, its protein that catalyzes reversible reactions of alanine and 2-keto glutaric acid, plays a key role in intermediate metabolism between glucose and amino acid metabolism. Furthermore,

enzyme activity in the serum is usually used as a biomarker of liver injury, but it seems to be rarely investigated in the adipose tissue. Our result demonstrated the potential of *GPT* for subsequent basic research on glucolipid metabolism. Although some studies have shown that *LGALS12* was involved in blood glucose improvement (Corbi et al., 2017), many studies have revealed that it promoted adipose differentiation and development (Yang et al., 2004; Wu et al., 2018). Additionally, *LGALS12* knockdown accelerates lipid catabolism (Baum, 2011; Yang et al., 2011). These results suggest that *LGALS12* acts as a negative regulator in obesity, which is inconsistent with our results, and this inconsistency might be explained by the possibility of adaptive regulation. As a human adipocyte-specific lncRNA, *ADIPINT* regulates lipid metabolism in adipocytes by regulating PC protein abundance and enzymatic activity (Kerr et al., 2022). Because the effects of *PC* often correlate with its activity, the reduced transcript levels of *PC* found in obese individuals in this study were not fully representative of its functional characteristics. Some studies have shown that *SLC4A4* may be involved in the pathogenesis of obesity and also participate in immune infiltration (Li et al., 2015; Li et al., 2022), and the methylation of the solute carrier protein (*SLC*) gene has been closely related to BMI and waist circumference (Mendelson et al., 2017; Sayols-Baixeras et al., 2017). In this study, not only *SLC4A4* but also *SLC27A2* were both significantly downregulated in obese patients, which to a certain extent supports previous studies and suggests that these two genes might be key genes requiring more attention in the future exploration of the obesity mechanism. Although there are no relevant reports on obesity or adipose tissue for *TTC36*, few studies have found that *TTC36* was negatively regulated by *TTC36* methylation in liver cancer tissues (Jing et al., 2022), and it is involved in the process of immune infiltration. It suggests that the role of *TTC36* in adipose tissue might be worthy of further investigation, especially in immune infiltration. In short, there are currently limited reports about these seven key genes, and more experimental research and clinical trials are needed. However, these genes may serve as potential target genes for subsequent basic research.

Our results showed that vemurafenib, NTNCB, dilazep, and PD-0325901 might have therapeutic effects on obesity. Unfortunately, for now, the direct effect of these drugs on obesity was hardly reported. Interestingly, according to the molecular action of potential therapeutic drugs analysis, *SLC29A* was indicated as a potential target in the possible mechanism of dilazep in treating obesity. Recent research suggested *SLC29A* was considered to regulate inosine levels in brown adipose tissue and consequently enhanced thermogenesis (Niemann et al., 2022); hence, these may be the indirect evidence supporting dilazep as the potential therapeutic drug. However,

caution is needed until more evidence on the anti-obesity of these drugs is found.

In summary, we successfully identified seven key genes, namely, *CCRL2*, *GPT*, *LGALS12*, *PC*, *SLC27A2*, *SLC4A4*, and *TTC36*, which are involved in obesity occurrence and development likely through the immune microenvironment. Meanwhile, M0 macrophage content was higher, while T follicular helper cell content was lower in adipose tissues in the obese group. Our results also indicated that vemurafenib, NTNCB, dilazep, and PD-0325901 might have therapeutic effects on obesity.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://www.ncbi.nlm.nih.gov/geo/>, GSE67024, <https://www.ncbi.nlm.nih.gov/geo/>, GSE174475, <https://www.ncbi.nlm.nih.gov/geo/>, GSE156906.

## Ethics statement

The studies involving human participants were reviewed and approved by the Medical Ethics Committee of Nanfang Hospital of Southern Medical University. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

MG: conceived and designed the study and revised the manuscript. PW: analyzed, interpreted, and wrote the manuscript. LG: interpreted and wrote the manuscript. Other authors collaborated on tissue sample collection and data interpretation. All authors have provided the final approval of the submitted version.

## Funding

This work was supported by the National Natural Science Foundation of China Grants 81870612 and the Science and Technology Planning Project of Guangdong Province 2019A1515011997.

## Acknowledgments

We thank those involved in this work.

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1024300/full#supplementary-material>

## References

- Barrett, T., Troup, D. B., Wilhite, S. E., Ledoux, P., Rudnev, D., Evangelista, C., et al. (2007). NCBI GEO: Mining tens of millions of expression profiles--database and tools update. *Nucleic Acids Res.* 35 (1), D760–D765. doi:10.1093/nar/gkl887
- Baum, L. G. (2011). Burn control, an adipocyte-specific function for galectin-12. *Proc. Natl. Acad. Sci. U. S. A.* 108 (46), 18575–18576. doi:10.1073/pnas.1115738108
- Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1\_12
- Chen, Y., Kassam, I., Lau, S. H., Kooner, J. S., Wilson, R., Peters, A., et al. (2021). Impact of BMI and waist circumference on epigenome-wide DNA methylation and identification of epigenetic biomarkers in blood: An EWAS in multi-ethnic asian individuals. *Clin. Epigenetics* 13 (1), 195. doi:10.1186/s13148-021-01162-x
- Chen, Z., Wang, G. X., Ma, S. L., Jung, D. Y., Ha, H., Altamimi, T., et al. (2017). Nrg4 promotes fuel oxidation and a healthy adipokine profile to ameliorate diet-induced metabolic disorders. *Mol. Metab.* 6 (8), 863–872. doi:10.1016/j.molmet.2017.03.016
- Corbi, S. C. T., Bastos, A. S., Nepomuceno, R., Cirelli, T., Dos Santos, R. A., Takahashi, C. S., et al. (2017). Expression profile of genes potentially associated with adequate glycemic control in patients with type 2 diabetes mellitus. *J. Diabetes Res.* 2017, 2180819. doi:10.1155/2017/2180819
- Crewe, C., An, Y. A., and Scherer, P. E. (2017). The ominous triad of adipose tissue dysfunction: Inflammation, fibrosis, and impaired angiogenesis. *J. Clin. Invest.* 127 (1), 74–82. doi:10.1172/JCI88883
- Engin, A. B. (2017). Adipocyte-macrophage cross-talk in obesity. *Adv. Exp. Med. Biol.* 960, 327–343. doi:10.1007/978-3-319-48382-5\_14
- Fuchs, A., Samovski, D., Smith, G. I., Cifarelli, V., Farabi, S. S., Yoshino, J., et al. (2021). Associations among adipose tissue immunology, inflammation, exosomes and insulin sensitivity in people with obesity and nonalcoholic fatty liver disease. *Gastroenterology* 161 (3), 968–981.e12. doi:10.1053/j.gastro.2021.05.008
- Gallardo-Escribano, C., Buonaiuti, V., Ruiz-Moreno, M. I., Vargas-Candela, A., Vilches-Perez, A., Benitez-Porres, J., et al. (2020). Epigenetic approach in obesity: DNA methylation in a prepubertal population which underwent a lifestyle modification. *Clin. Epigenetics* 12 (1), 144. doi:10.1186/s13148-020-00935-0
- Hu, X., Zhou, Y., Yang, Y., Peng, J., Song, T., Xu, T., et al. (2016). Identification of zinc finger protein Bcl6 as a novel regulator of early adipose commitment. *Open Biol.* 6 (6), 160065. doi:10.1098/rsob.160065
- Jing, W., Peng, R., Li, X., Lv, S., Duan, Y., and Jiang, S. (2022). Study on the prognostic values of TTC36 correlated with immune infiltrates and its methylation in hepatocellular carcinoma. *J. Immunol. Res.* 2022, 7267131. doi:10.1155/2022/7267131
- Kajimura, S., Spiegelman, B. M., and Seale, P. (2015). Brown and beige fat: Physiological roles beyond heat generation. *Cell Metab.* 22 (4), 546–559. doi:10.1016/j.cmet.2015.09.007
- Kerr, A. G., Wang, Z., Wang, N., Kwok, K. H. M., Jalkanen, J., Ludzki, A., et al. (2022). The long noncoding RNA ADIPINT regulates human adipocyte metabolism via pyruvate carboxylase. *Nat. Commun.* 13 (1), 2958. doi:10.1038/s41467-022-30620-0
- Klose, R. J., and Bird, A. P. (2006). Genomic DNA methylation: The mark and its mediators. *Trends biochem. Sci.* 31 (2), 89–97. doi:10.1016/j.tibs.2005.12.008
- Lamb, J., Crawford, E. D., Peck, D., Modell, J. W., Blat, I. C., Wrobel, M. J., et al. (2006). The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313 (5795), 1929–1935. doi:10.1126/science.1132939
- LaPensee, C. R., Lin, G., Dent, A. L., and Schwartz, J. (2014). Deficiency of the transcriptional repressor B cell lymphoma 6 (Bcl6) is accompanied by dysregulated lipid metabolism. *PLoS One* 9 (6), e97090. doi:10.1371/journal.pone.0097090
- Li, J., Zhou, C., Li, J., Su, Z., Sang, H., Jia, E., et al. (2015). Global correlation analysis for microRNA and gene expression profiles in human obesity. *Pathol. Res. Pract.* 211 (5), 361–368. doi:10.1016/j.prp.2014.11.014
- Li, Z., Li, X., Jin, M., Liu, Y., He, Y., Jia, N., et al. (2022). Identification of potential biomarkers and their correlation with immune infiltration cells in schizophrenia using combinative bioinformatics strategy. *Psychiatry Res.* 314, 114658. doi:10.1016/j.psychres.2022.114658
- Martinez-Santibanez, G., and Lumeng, C. N. (2014). Macrophages and the regulation of adipose tissue remodeling. *Annu. Rev. Nutr.* 34, 57–76. doi:10.1146/annurev-nutr-071812-161113
- Mendelson, M. M., Marioni, R. E., Joeannes, R., Liu, C., Hedman, A. K., Aslibekyan, S., et al. (2017). Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: A mendelian randomization approach. *PLoS Med.* 14 (1), e1002215. doi:10.1371/journal.pmed.1002215
- Morris, T. J., Butcher, L. M., Feber, A., Teschendorff, A. E., Chakravarthy, A. R., Wojdacz, T. K., et al. (2014). ChAMP: 450k chip analysis methylation pipeline. *Bioinformatics* 30 (3), 428–430. doi:10.1093/bioinformatics/btt684
- Ng, M., Fleming, T., Robinson, M., Thomson, B., Graetz, N., Margono, C., et al. (2014). Global, regional, and national prevalence of overweight and obesity in children and adults during 1980–2013: A systematic analysis for the global burden of disease study 2013. *Lancet* 384 (9945), 766–781. doi:10.1016/S0140-6736(14)60460-8
- Niemann, B., Haufs-Brusberg, S., Puetz, L., Feickert, M., Jaekstein, M. Y., Hoffmann, A., et al. (2022). Apoptotic Brown adipocytes enhance energy expenditure via extracellular inosine. *Nature* 609 (7926), 361–368. doi:10.1038/s41586-022-05041-0
- Parle-McDermott, A., and Ozaki, M. (2011). The impact of nutrition on differential methylated regions of the genome. *Adv. Nutr.* 2 (6), 463–471. doi:10.3945/an.111.001008
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Sayols-Baixeras, S., Subirana, I., Fernandez-Sanles, A., Senti, M., Lluís-Ganella, C., Marrugat, J., et al. (2017). DNA methylation and obesity traits: An epigenome-wide association study. The REGICOR study. *Epigenetics* 12 (10), 909–916. doi:10.1080/15592294.2017.1363951
- Tian, Y., Morris, T. J., Webster, A. P., Yang, Z., Beck, S., Feber, A., et al. (2017). ChAMP: Updated methylation analysis pipeline for illumina BeadChips. *Bioinformatics* 33 (24), 3982–3984. doi:10.1093/bioinformatics/btx513
- Wahl, S., Drong, A., Lehne, B., Loh, M., Scott, W. R., Kunze, S., et al. (2017). Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 541 (7635), 81–86. doi:10.1038/nature20784
- Weisberg, S. P., McCann, D., Desai, M., Rosenbaum, M., Leibel, R. L., and Ferrante, A. W., Jr (2003). Obesity is associated with macrophage accumulation in adipose tissue. *J. Clin. Invest.* 112 (12), 1796–1808. doi:10.1172/JCI19246

Wu, W., Yin, Y., Xu, K., Peng, Y., and Zhang, J. (2018). Knockdown of LGALS12 inhibits porcine adipocyte adipogenesis via PKA-Erk1/2 signaling pathway. *Acta Biochim. Biophys. Sin.* 50 (10), 960–967. doi:10.1093/abbs/gmy099

Xu, M., Wang, Y. M., Li, W. Q., Huang, C. L., Li, J., Xie, W. H., et al. (2022). Ccr12 deficiency deteriorates obesity and insulin resistance through increasing adipose tissue macrophages infiltration. *Genes Dis.* 9 (2), 429–442. doi:10.1016/j.gendis.2020.08.009

Yang, R. Y., Hsu, D. K., Yu, L., Chen, H. Y., and Liu, F. T. (2004). Galectin-12 is required for adipogenic signaling and adipocyte differentiation. *J. Biol. Chem.* 279 (28), 29761–29766. doi:10.1074/jbc.M401303200

Yang, R. Y., Yu, L., Graham, J. L., Hsu, D. K., Lloyd, K. C., Havel, P. J., et al. (2011). Ablation of a galectin preferentially expressed in adipocytes increases lipolysis, reduces adiposity, and improves insulin sensitivity in mice. *Proc. Natl. Acad. Sci. U. S. A.* 108 (46), 18696–18701. doi:10.1073/pnas.1109065108

Yu, D., Rao, S., Tsai, L. M., Lee, S. K., He, Y., Sutcliffe, E. L., et al. (2009). The transcriptional repressor Bcl-6 directs T follicular helper cell lineage commitment. *Immunity* 31 (3), 457–468. doi:10.1016/j.immuni.2009.07.002

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16 (5), 284–287. doi:10.1089/omi.2011.0118



## OPEN ACCESS

## EDITED BY

Xiaofan Lu,  
INSERM U964 Institut de Génétique et  
de Biologie Moléculaire et Cellulaire  
(IGBMC), France

## REVIEWED BY

Dan Zhang,  
Institute of Genetics and Developmental  
Biology (CAS), China  
Sun Zhaoyang,  
Fudan University, China  
Zhijia Xia,  
Ludwig Maximilian University of Munich,  
Germany

## \*CORRESPONDENCE

XinZhong Xu,  
xuxinzhong@ahmu.edu.cn

<sup>†</sup>These authors have contributed equally  
to this work and share first authorship

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 22 August 2022

ACCEPTED 29 September 2022

PUBLISHED 21 October 2022

## CITATION

Wang J, Wu Z, Zheng M, Yu S, Zhang X  
and Xu X (2022), CD146 is closely  
associated with the prognosis and  
molecular features of osteosarcoma:  
Guidance for personalized  
clinical treatment.  
*Front. Genet.* 13:1025306.  
doi: 10.3389/fgene.2022.1025306

## COPYRIGHT

© 2022 Wang, Wu, Zheng, Yu, Zhang  
and Xu. 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.

# CD146 is closely associated with the prognosis and molecular features of osteosarcoma: Guidance for personalized clinical treatment

Jingkun Wang<sup>†</sup>, Zhonghan Wu<sup>†</sup>, Meige Zheng, Shuisheng Yu,  
Xin Zhang and XinZhong Xu<sup>\*</sup>

Department of Orthopaedics, The Second Affiliated Hospital of Anhui Medical University, Hefei, China

**Background:** Osteosarcoma (OSA), a focus for orthopedic surgeons, always results in severe death due to metastasis. CD146 is severely expressed in several tumors, indicating its potential as a biomarker for OSA.

**Method:** Two OSA cohorts were enrolled in this study. A Therapeutically Applicable Research to Generate Effective Treatments-Osteosarcoma (TARGET-OS) cohort was used as a training cohort, and GSE21257 was used as the external validation cohort. The R package “limma” was used to discriminate the differentially expressed genes among CD146-high and CD146-low patients and was further annotated by the enriched signaling pathways. The R package MOVICS was used to evaluate immune infiltration and the response to chemotherapy and immunotherapy. All statistical analyses were performed by R version 4.0.2, and  $p < 0.05$  was considered statistically significant.

**Result:** CD146 plays an important role in promoting the progression, invasion, and metastasis of several tumors. In the current study, we first revealed an integrative unfavorable prognosis in patients with tumors ( $p < 0.01$ , HR: 1.10, 95% CI: 1.07–1.14). CD146 is tightly correlated with m5C RNA methylation modification genes in OSA. Furthermore, we revealed that CD146 acts as an oncogene in OSA patients and is linked to poor prognosis in both the TARGET-OS cohort ( $p = 0.019$ , HR: 2.61, 95% CI: 1.171–5.834) and the GSE21257 cohort ( $p = 0.005$ , HR: 3.61, 95% CI: 1.474–8.855), with a total of 137 patients, regardless of whether they were adjusted for clinical pathological features. Highly-expressed CD146 impacts the signaling pathways of cytokine–cytokine receptor interactions and is associated with the high infiltration of immunocytes. Moreover, patients with high CD146 expression were more likely to be sensitive to anti-PD-1 immunotherapy, while patients with low expression of CD146 were more likely to be sensitive to cisplatin and doxorubicin chemotherapy.

**Conclusion:** Overall, CD146 is an independent prognostic factor for OSA patients and can help doctors select clinical treatment strategies.

## KEYWORDS

osteosarcoma, prognosis, microenvironment, immunotherapy, personalized treatment

## Background

Osteosarcoma (OSA), which derives from primitive bone-forming mesenchymal cells, is the most common primary bone malignancy (Valery et al., 2015). The incidence rate of OSA changes by age, sex, and race. There are bimodal age peaks of OSA; one is at the age of 0–24 years old, and the other is over 65 years old (Mirabello et al., 2009a; Rojas et al., 2021). It was considered that the incidence of OSA in males was higher than that in females. Moreover, the incidence in females is higher than that in males in the group aged less than 15 years old (Nie and Peng, 2018). In addition, OSA is more likely to be observed in black people (6.8 per million persons per year) and Hispanics (6.5 per million persons per year) than in white people (4.6 per million persons per year) (Mirabello et al., 2009b; Ottaviani and Jaffe, 2009). Many factors can significantly affect the survival outcome of OSA, such as metastatic status, percentage of necrosis, tumor size, Enneking stage, local recurrence, and treatment therapeutic strategy (Pakos et al., 2009). Multidisciplinary approaches are applied for the treatment of patients with OSA—including surgery, radiotherapy, polychemotherapy, and immunomodulation—as 80%–90% of local OSA will develop metastasis (Ritter and Bielack, 2010). Although multidisciplinary treatment strategies have reduced mortality in OSA patients, the prognosis remains poor, and the 5-year and 10-year survival rates are only 56.31% and 22.33%, respectively (Yasin et al., 2020). Therefore, it is crucial to find robust biomarkers and prognostic factors to predict the progression, invasion, and metastasis of OSA.

Melanoma cell adhesion molecule (MCAM/CD146), which was reported in 1987 by Johnson and colleagues, is an integral membrane glycoprotein and belongs to the immunoglobulin superfamily. It consists of a signal peptide, an extracellular fragment with five immunoglobulin-like domains, a transmembrane region, and a short cytoplasmic tail (Lehmann et al., 1989; Johnson et al., 1993). At present, the pathological processes and physiological function of CD146 have been revealed, for instance, in tissue regeneration, inflammation, infections, signal transduction, cell migration, the cell cycle, mesenchymal stem cell differentiation, angiogenesis, and the immune response (Wang and Yan, 2013). As reported, CD146 is a cell surface receptor and can combine with many ligands involved in proliferation-related signaling pathways (Li et al., 2003). In addition, it can induce tumor angiogenesis (So et al., 2010). These biological characteristics promote tumor progression, invasion, and metastasis, especially in melanoma (Melnikova et al., 2009).

Currently, the focus on OSA has expanded from the tumor cell itself to the tumor microenvironment; the infiltrated immunocytes play an important role in tumor proliferation and migration or are resistant to chemo drugs (Corre et al., 2020). Tumor-associated macrophages account for

approximately 50% of tumor volume of OSA (Huang et al., 2021), and M2 type can accelerate the process of tumor metastasis, which can be treated with all-trans retinoic acid to inhibit the polarization of M2 macrophage (Zhou et al., 2017). Yoshida et al. reported that anti-PD-1 therapy can inhibit the infiltration of Tregs with the murine LM8 cell osteosarcoma model, resulting in reduced tumor volume and prolonged survival (Yoshida et al., 2020). CD146 is highly expressed in advanced primary or metastatic melanoma cells but is rarely detected in normal cells, indicating its potential as a biomarker for predicting tumor initialization and prognosis. Beyond melanoma, increased expression of CD146 has been reported in more than ten types of tumors to date (Zabouo et al., 2009; Feng et al., 2012; Zeng et al., 2012; Tian et al., 2013; Wang and Yan, 2013; Li et al., 2014; Liang et al., 2017; Sechler et al., 2017). These studies have demonstrated that CD146 is a novel metastasis biomarker and prognostic factor and that it can also be used as a therapeutic target (Xie et al., 1997; Zabouo et al., 2009; Wu et al., 2012).

Similar results have been reported for OSA and CD146 (Schiano et al., 2012; Westrøm et al., 2016). In the current study, we confirmed the high expression of CD146 in most cancers, and the expression levels were significantly associated with the corresponding prognosis. Univariate and multivariate analyses were performed to reveal that CD146 acts as an independent prognostic factor for OSA patients. The critical pathways were identified by GO enrichment and HALLMARK analysis. Furthermore, we found that CD146 is a biomarker for anti-PD1 therapy. Our goal was to prove the prognostic value of CD146 in OSA and provide a novel tool for predicting immunotherapy efficacy.

## Materials and methods

### Data collection

Two OSA cohorts were enrolled in this study. The Therapeutically Applicable Research to Generate Effective Treatments (TARGET) osteosarcoma (TARGET-OS) cohort was derived from the UCSC Xena platform (<http://xena.ucsc.edu/>). It contains 84 samples of gene expression profiles and corresponding clinical information. The GENCODE27 annotation file was applied to annotate the gene symbols of mRNA. Another cohort, GSE21257, which contains 53 OSA patients, was downloaded from the Gene Expression Omnibus platform (GEO, <http://www.ncbi.nlm.nih.gov/geo/>). The TARGET-OS cohort was used as a training cohort, and GSE21257 was used as the external validation cohort. We conducted  $\log_2(\text{TPM}+1)$  processing on the expression data of CD146 to scale it.

The median values of CD146 expression were set as threshold values to divide patients into low- and high-expression groups. In addition, we compared the expression of CD146 in normal and tumor tissues based on data from The Cancer Genome Atlas Program (TCGA) project (<https://docs.gdc.cancer.gov/>) and the GTEx project (<https://www.gtexportal.org/home/>). The evaluation of the CD146 prognostic value to overall survival (OS) time in pancreatic cancer was conducted on an online website (<http://sangerbox.com/>).

## Functional enrichment analysis

We used the R package “limma” to discriminate the differentially expressed genes (DEGs) among CD146-high and CD146-low patients (Ritchie et al., 2015). A fold-change > 0.4 and an adjusted *p* value < 0.01 were set as the cut-off values to filter the DEGs. The packages “org.Hs.eg.db” and “msigdb” were used to perform Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and HALLMARK enrichment analyses (Ashburner et al., 2000; Liberzon et al., 2015); clusterProfiler was applied to further elucidate potential gene functional interpretation and pathway enrichment (Yu et al., 2012).

## Immunocyte infiltration assessment

To evaluate the immunocyte infiltration status and the difference between the low- and high-expression groups, we collected 28 immunocyte signatures from prior research (Yoshihara et al., 2013). Subsequently, the runGSVA module in MOVICS was used to perform single sample gene set variation analysis (GSVA) to estimate the normalized enrichment score (NES) for the 28 signatures in each OSA patient (Lu et al., 2020).

## Efficacy prediction of chemotherapy and immunotherapy

We applied the comDrugen module in MOVICS to the clinical data of the TARGET-OS cohort to explore the chemotherapy efficacy difference between the low and high CD146-expression groups. Four chemotherapeutic drugs are shown in the violin plots. We also predicted the chemotherapeutic response to EBFs based on the largest publicly available pharmacogenomics database, the Genomics of Drug Sensitivity in Cancer (GDSC), with the use of the GSCA online website (Liu et al., 2018). To compare the likelihood of the CD146 groups with the immunotherapy

subgroups, we performed subclass mapping analysis in the samples from patients who had received anti-PD-1 or anti-CTLA4 checkpoint therapy (Wang et al., 2016; Thanindrarn et al., 2019).

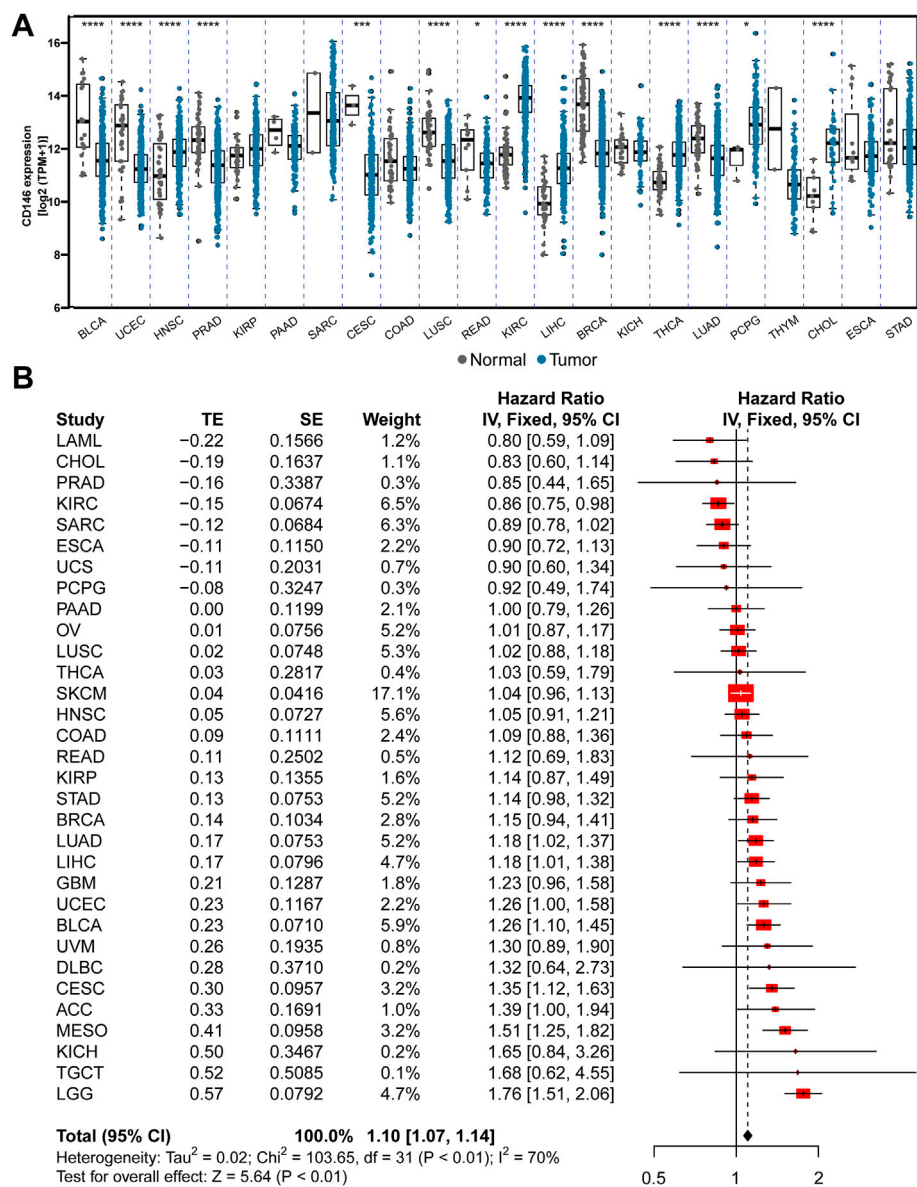
## Statistical analyses

All statistical analyses were performed by R version 4.0.2. Student’s *t*-test was applied to compare the two groups if the data were normally distributed for continuous data; otherwise, the Wilcoxon rank-sum test was used. For categorical data, the chi-square test and Fisher’s exact test were conducted. Kaplan–Meier curves were generated to compare OS based on the log-rank test. A nomogram was constructed by the R package “regplot” with the results from Cox regression analysis. Calibration curves were plotted to assess the calibration ability of the nomogram, and decision curve analysis (DCA) was performed to show the clinical usefulness of the nomogram. The receiver operating characteristic (ROC) area under the curve (AUC) was calculated to assess the stability of prediction. To identify the independent risk factors, univariate and multivariate analyses were performed. We also established the Cox regression model to calculate the hazard ratio (HR) values and the 95% confidence interval (95% CI). The correlation was determined by the Pearson correlation test, with *p* < 0.05 being considered statistically significant.

## Results

### CD146 is highly expressed in pan-cancer tissues but expressed at low levels in normal tissues

Overexpression of CD146 has been reported in various cancers. However, the comparison of CD146 mRNA levels in different tumors is still unclear. Although all tumors have a certain CD146 mRNA level, the expression levels in sarcoma (SARC) and kidney renal clear cell sarcoma (KIRC) are higher than those in other tumors. The results also demonstrated that CD146 is highly expressed in tumor tissues relative to normal tissues in many tumors, including head and neck squamous cell carcinoma (HNSC), prostate adenocarcinoma (PRAD), kidney renal papillary cell carcinoma (KIRP), KIRC, liver hepatocellular carcinoma (LIHC), thyroid carcinoma (THCA), and pheochromocytoma and paraganglia (PCPG) (Figure 1A). Furthermore, we conducted a meta-analysis to evaluate the connection between CD146 expression and OS time across cancers. The results showed that CD146 acted as an oncogene and was linked to poor prognosis in most tumor types (*p* < 0.01,



**FIGURE 1** CD146 is increased in tumor tissue and predicts poor prognosis among pan-cancer. (A) The differential expression of CD146 in 22 types of tumor and adjacent normal tissues. (B) Meta-analysis revealed the integrative hazard risk of CD146 to pan-cancer.

HR: 1.10, 95% CI: 1.07-1.14, Figure 1B). To validate the conclusion, we performed a log-rank test via Kaplan–Meier curves, which was consistent with the meta-analysis (Supplementary Figure S1).

It is widely reported that posttranscriptional modification plays a key role in tumorigenesis, including the methylation modification methods of m1A, m5C, and m6A. Zhang et al. (2022) reported that METTL3 upregulates COPS5 expression by promoting COPS5 methylation in an m6A-related manner and results in the promotion of OSA progression. Li et al. (2022) also

demonstrated the METTL14-mediated epitranscriptome modification of MN1 mRNA and the promotion of OSA tumorigenicity. Therefore, we assessed the correlation between RNA methylation modification genes and CD146 levels. We observed that the CD146 expression levels in most types of tumor positively correlated with RNA methylation modification based on the pan-cancer data from the TCGA project, especially for lymphoid neoplasms, diffuse large B-cell lymphoma, cholangiocarcinoma, ovarian serous cystadenocarcinoma, and thymoma (Figure 2A). For RNA

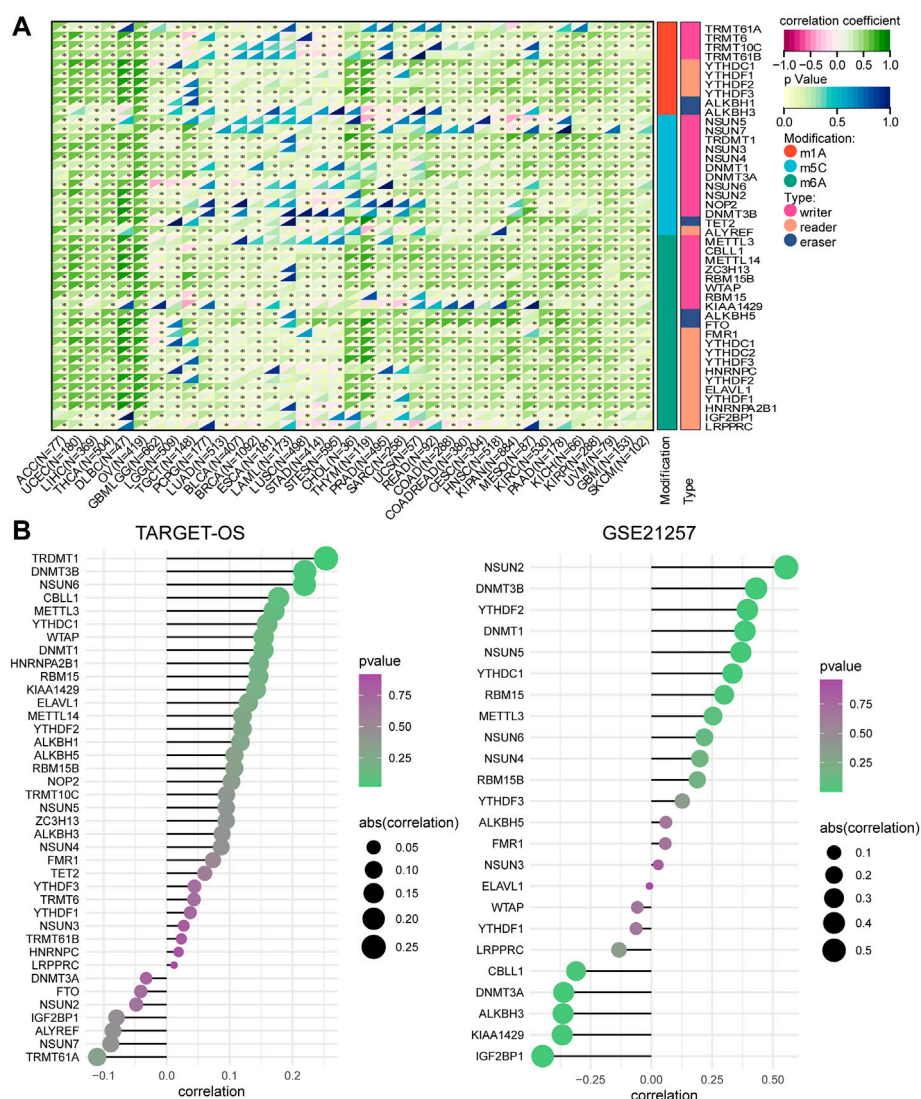


FIGURE 2

CD146 is regulated by RNA methylation. **(A)** Correlation of CD146 expression and RNA methylation regulators in m1A, m5C, and m6A among pan-cancer. **(B)** Correlation of CD146 expression and RNA methylation regulators among pan-cancer OSA datasets.

methylation genes in OSA, we evaluated their correlation with CD146 in the TARGET-OS and GSE21257 cohorts and found similar results, especially for the m5C genes DNMT3B, NSUN2, NSUN5, NSUN6, and DNMT1 (Figure 2B).

## CD146 can act as a prognostic predictor and an independent prognostic factor for OSA

To explore the association between CD146 and OSA, we divided the TARGET-OS cohort into two subgroups: high CD146 expression (HEXP) and low CD146 expression

(LEXP). The clinical feature distribution in HEXP and LEXP showed no difference, except for the first tumor event, and HEXP patients experienced more relapse ( $p = 0.014$ , Table 1). The Kaplan–Meier curve for OS time is shown in Figure 3A. We found a statistically significant difference in OS time among patients in the HEXP and LEXP subgroups ( $p = 1.019$ , HR = 2.61, 95% CI: 1.71–5.834). As the 1-year, 3-year, and 5-year AUCs were 0.573, 0.668, and 0.730, respectively, we confirmed that CD146 is responsible for the poor prognosis in OSA (Figure 3B). Univariate analysis showed that the metastatic-diagnosis ( $p < 0.001$ , HR: 4.764, 95% CI: 2.221–10.221) and CD146 subgroups ( $p = 0.019$ , HR: 2.614, 95% CI: 1.171–5.834) were related to the OS time of OSA (Figure 3C). Multivariable analysis revealed that

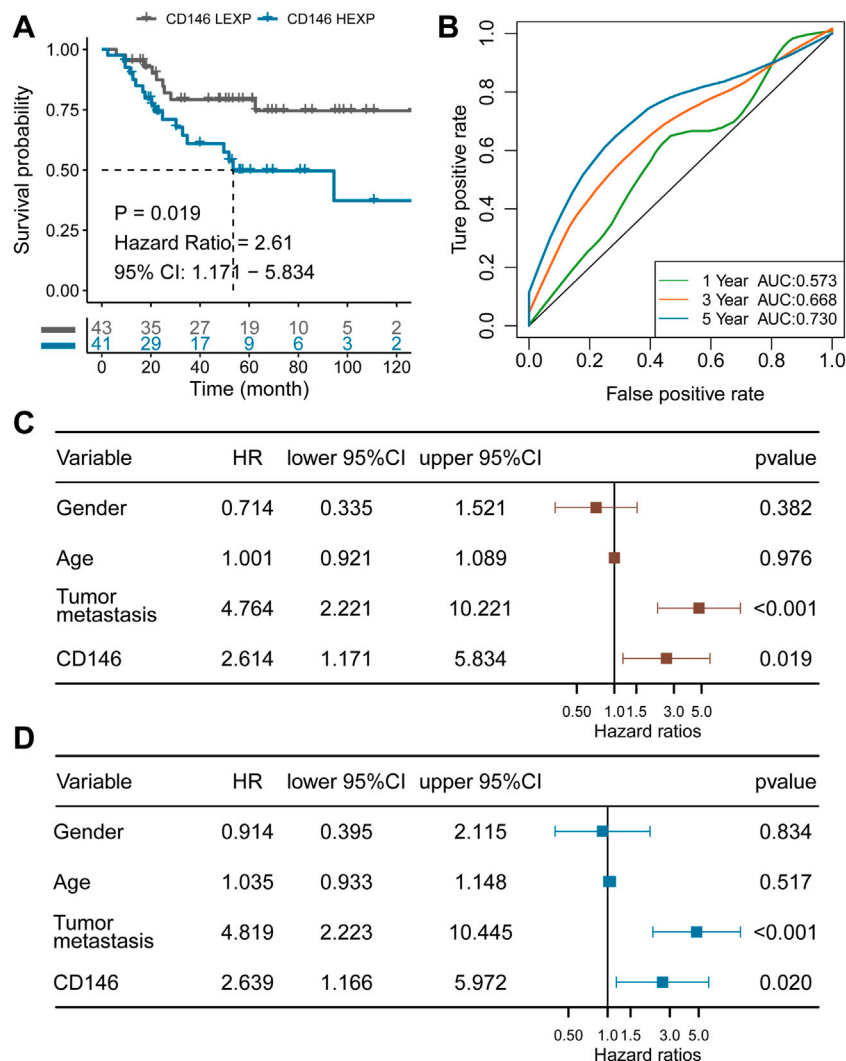
TABLE 1 Clinical features of the CD146 low and high groups in the TARGET-OS cohort.

| Clinicopathological features |                           | CD146 Low    | CD146 high   | p Value |
|------------------------------|---------------------------|--------------|--------------|---------|
| Gender (%)                   | Female                    | 15 (34.9%)   | 22 (53.7%)   | 0.130   |
|                              | Male                      | 28 (65.1%)   | 19 (46.3%)   |         |
| Age, years                   | —                         | 15.55 ± 5.49 | 14.40 ± 4.00 | 0.273   |
| Metastatic-Diag (%)          | No                        | 33 (76.7%)   | 30 (73.2%)   | 0.900   |
|                              | Yes                       | 10 (23.3%)   | 11 (26.8%)   |         |
| Race (%)                     | Asian                     | 2 (4.7%)     | 4 (9.8%)     | 0.477   |
|                              | Black or African American | 4 (9.3%)     | 3 (7.3%)     |         |
|                              | Unknown                   | 8 (18.6%)    | 12 (29.3%)   |         |
|                              | White                     | 29 (67.4%)   | 22 (53.7%)   |         |
| First Event (%)              | Censored                  | 6 (14.0%)    | 5 (12.2%)    | 0.014*  |
|                              | Death                     | 1 (2.3%)     | 1 (2.4%)     |         |
|                              | No event                  | 22 (51.2%)   | 9 (22.0%)    |         |
|                              | Relapse                   | 13 (30.2%)   | 25 (61.0%)   |         |
|                              | SMN                       | 0 (0.0%)     | 1 (2.4%)     |         |
|                              | Unknown                   | 1 (2.3%)     | 0 (0.0%)     |         |
| Metastasis site (%)          |                           | 32 (74.4%)   | 30 (73.2%)   | 0.796   |
|                              | Bone and lung             | 2 (4.7%)     | 3 (7.3%)     |         |
|                              | Bone only                 | 0 (0.0%)     | 1 (2.4%)     |         |
|                              | Lung only                 | 9 (20.9%)    | 7 (17.1%)    |         |
| Primary tumor site (%)       | Arm/hand                  | 3 (7.0%)     | 3 (7.3%)     | 0.657   |
|                              | Leg/foot                  | 30 (69.8%)   | 32 (78.0%)   |         |
|                              | Leg/Foot                  | 8 (18.6%)    | 6 (14.6%)    |         |
|                              | Pelvis                    | 2 (4.7%)     | 0 (0.0%)     |         |
| Specific tumor site (%)      | Arm NOS                   | 0 (0.0%)     | 1 (2.4%)     | 0.097   |
|                              | Femur                     | 14 (32.6%)   | 24 (58.5%)   |         |
|                              | Fibula                    | 5 (11.6%)    | 3 (7.3%)     |         |
|                              | Foot NOS                  | 0 (0.0%)     | 1 (2.4%)     |         |
|                              | Humerus                   | 2 (4.7%)     | 2 (4.9%)     |         |
|                              | Ilium                     | 1 (2.3%)     | 0 (0.0%)     |         |
|                              | Leg NOS                   | 4 (9.3%)     | 2 (4.9%)     |         |
|                              | Pelvis                    | 0 (0.0%)     | 1 (2.4%)     |         |
|                              | Pelvis - ilium            | 0 (0.0%)     | 1 (2.4%)     |         |
|                              | Pelvis/Sacrum             | 1 (2.3%)     | 0 (0.0%)     |         |
|                              | Radius                    | 1 (2.3%)     | 0 (0.0%)     |         |
|                              | Tibia                     | 15 (34.9%)   | 6 (14.6%)    |         |

the same variables were related to poor prognosis, which proved that CD146 is an independent prognostic factor for OSA ( $p = 0.020$ , HR: 2.639, 95% CI: 1.166–5.972, Figure 3D). At present, the impact of age and sex on the prognosis of OSA is controversial (Pan et al., 2019; Ding et al., 2020). It is worth noting that we validated the impact of age and sex on prognosis to be small.

Furthermore, we also constructed a prediction nomogram with the clinical features (Figure 4A). We revealed that the CD146 expression group acted as an assistant that enhanced the traditional pathological method. The total AUC value increased to 0.765 (95% CI: 0.660–0.871, Figure 4B), indicating the preferable prognostic value of the nomogram—also confirmed by the

calibration curve. There was no significant difference between the nomogram-predicted survival probability and the actual survival probability ( $p = 0.243$  for 3-year,  $p = 0.416$  for 5-year, Figure 4C). We also observed that the nomogram performed with better prediction effectiveness than single metastasis or CD146 expression (Figure 4D). We calculated the prognostic C-index value of clinical features and observed that the nomogram contained the highest value of 0.781, as compared with age (C-index: 0.490), gender (C-index: 0.561), and metastasis status (C-index: 0.696) (Figure 4E). Moreover, we randomly selected a portion of patients ( $n = 60$ ) to be re-assessed for the prognostic stability of nomogram ten times and observed that all the 10 C-



**FIGURE 3** CD146 is an independent prognostic factor for OSA identified in the TARGET-OS cohort. **(A)** K-M plot showing the diverse clinical overall survival outcomes of the CD146 low and high groups; **(B)** ROC curve showing the prognostic value of the CD146 expression level; **(C)** results from univariate Cox regression analysis of the CD146 group and clinical features; **(D)** results from multivariate Cox regression analysis of the CD146 group and clinical features.

indexes were higher than 0.75 (Figure 4F), indicating that the nomogram is accurate and stable.

### Mechanisms by which CD146 impacts OSA

Enrichment analysis was carried out to investigate the potential mechanism by which CD146 impacts OSA. We first obtained 580 DEGs between the HEXP and LEXP subgroups along with the preset cut-off values mentioned above. The 580 DEGs were further enriched and annotated in different biological pathways. According to the biological process (BP) GO terms, pathways

involving vasculature development and positive anion transport were the most relevant to CD146. In molecular function (MF) analysis, CD146 was obviously relevant to signaling receptor activator activity, receptor-ligand activity, cytokine receptor binding, and cytokine activity (Figure 5A). In KEGG analysis, numerous DEGs were enriched in the neuroactive ligand-receptor interaction and cytokine-cytokine receptor interaction pathways (Figure 5B). For HALLMARK analysis, complement and TNFA signaling via NFkB had the most impact on CD146 (Figure 5C). The above enriched signaling pathways indicated that CD146 played an important role in the immune system in OSA. Therefore, we were concerned about the link between immunocytes and CD146 expression levels. As shown in Figure 5D, 11 types of

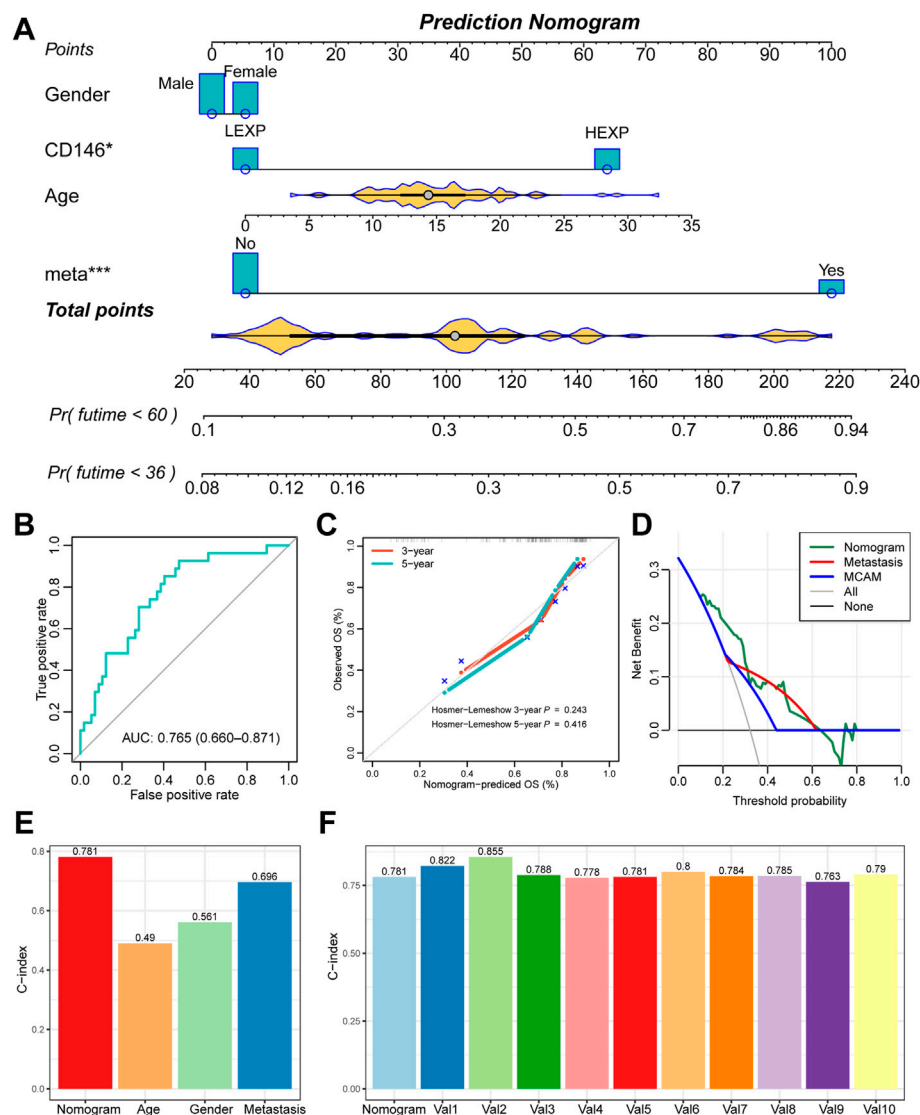


FIGURE 4

The nomogram further enhanced the prognostic value of CD146 in OSA. (A) Prognostic nomogram of 3-year and 5-year overall survival of OSA patients; (B) ROC curves for the nomogram of overall death events; (C) calibration plot to compare the nomogram prediction and the actual death events; (D) DCA curve for the nomogram, metastasis, and CD146 expression group; (E) C-index of the nomogram, age, gender, and metastasis status; (F) 10-times randomization C-index test to assess the stability of the prognostic nomogram.

immunocytes, including activated dendritic cells, eosinophils, gamma cells, immature dendritic cells, mast cells, monocytes, neutrophils, regulatory T cells, T follicular helper cells, type-17 T helper cells, and myeloid-derived suppressor cells were more highly infiltrative in the HEXP group.

## CD146 level used to choose the suitable precise clinical treatment for OSA patients

Immunotherapy and chemotherapy have developed essential treatment strategies for excellent therapeutic effects on malignant

tumors (Furue, 2003; Rodig et al., 2018). However, the current evidence shows that most checkpoint inhibitors have little effect on OSA and that chemotherapeutic drug resistance increases each year for multiple causes (Chatterjee and Bivona, 2019; Chen et al., 2021). Therefore, it is crucial to validate predictive biomarkers to optimize the choice of treatment strategy. In this study, we set the half maximal inhibitory concentration ( $IC_{50}$ ) as the observation index and compared the therapeutic effect of four commonly-used chemotherapy items in the LEXP and HEXP subgroups. We found a significant difference in the  $IC_{50}$  value of LEXP and HEXP to cisplatin (Figure 6A,  $p = 0.002$ ) and that LEXP patients can respond to doxorubicin treatment

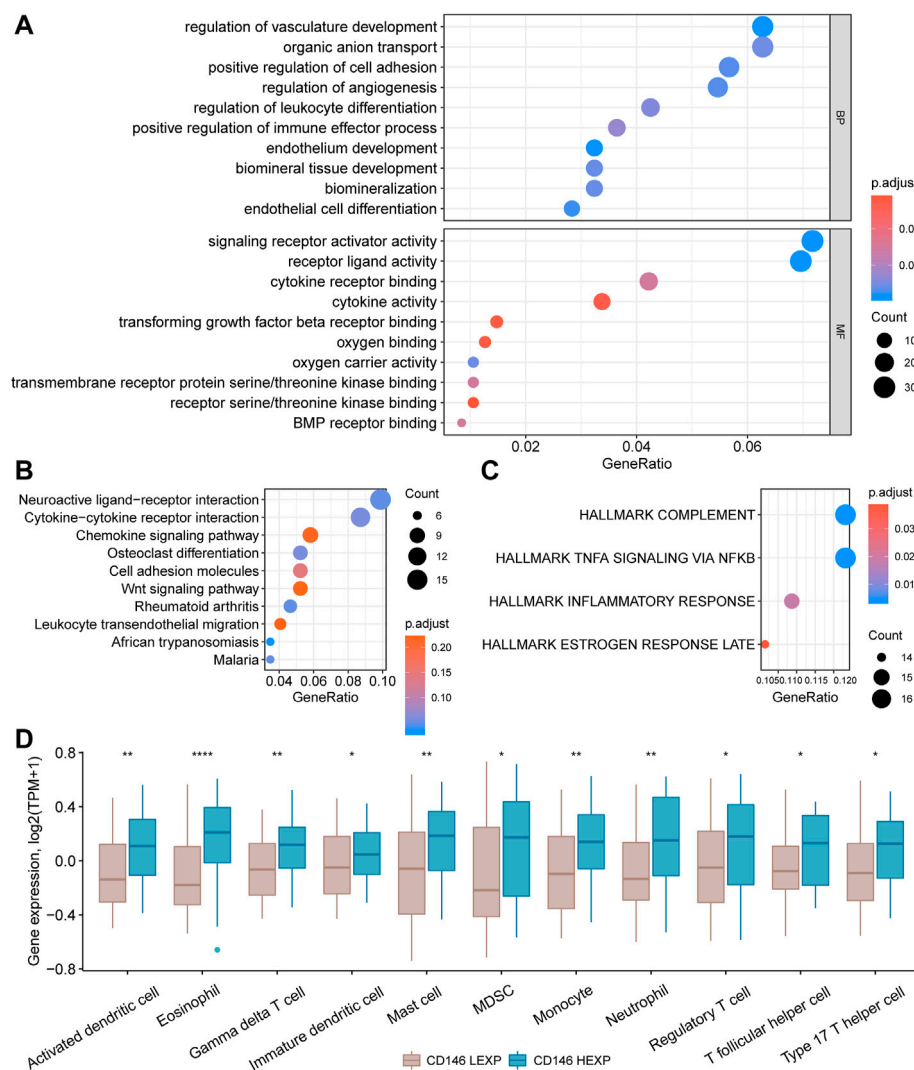


FIGURE 5

Enrichment of signaling pathways and differential infiltration of immunocytes in the CD146 high and low groups. Enrichment of 580 DEGs between the HEXP and LEXP subgroups by GO terms (A), KEGG terms (B), and HALLMARK terms (C). (D) Differential infiltration of 12 immunocytes between the CD146 high and low groups.

( $p = 0.054$ ) but not to etoposide ( $p = 0.21$ ) or bleomycin ( $p = 0.44$ ) (Figure 6A). Furthermore, we also tried to select potential chemotherapy drugs from the GSCA online website and revealed that patients with high levels of CD146 were sensitive to chemotherapy with PLX4720, dabrafenib, SE590885, staurosporine, BRD-K99006945, vemurafenib, and several other proposed drugs (Table 2). To predict the immunotherapy response for each patient with OSA, we subsequently conducted SubMap analysis with an OSA cohort which included both patients receiving and not receiving anti-PD-1 or anti-CTLA4 therapy. There was no significant difference in CTAL-4 treatment response between the LEXP and HEXP subgroups; however, HEXP patients presented a better treatment

response to anti-PD1 therapy than LEXP (Figure 6B, Bonferroni corrected  $p < 0.05$ ).

## Validation of the prognostic value in the extra cohort

To validate the value of prognostic prediction, the GSE21257 cohort containing 53 OSA patients was enrolled. We set the median value of CD146 expression as a cut-off value and divided the 53 OSA patients into HEXP ( $n = 26$ ) and LEXP ( $n = 27$ ) subgroups. The clinicopathological features among the two subgroups showed no difference, except a little

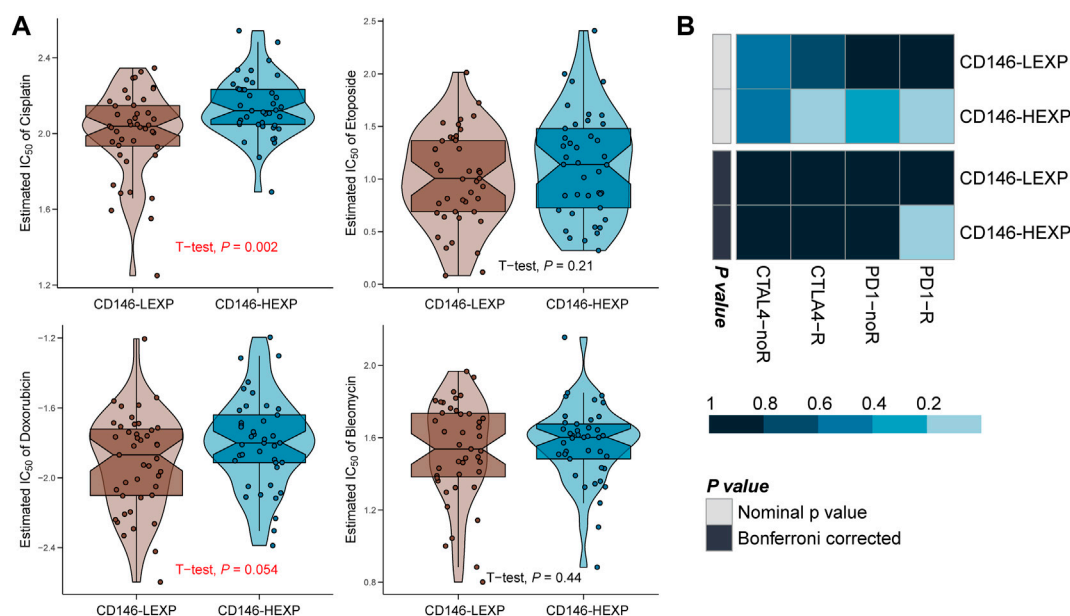


FIGURE 6

The CD146 level reflected the response to immunotherapy and chemotherapy. (A) The response of OSA patients to four common chemotherapy drugs, cisplatin, etoposide, doxorubicin, and bleomycin. (B) The response of OSA patients to anti-PD-1 and anti-CTLA4 immunotherapy.

among the metastasis status of tumor patients with high CD146 expression who met more tumor metastasis whether at or after diagnosis ( $p = 0.042$ , Table 3). Consistent with the results of the training cohort, the LEXP group showed higher OS, and the HEXP group had a 3.61-fold HR compared to the LEXP group (Figure 7A, 95% CI: 1.474–8.855,  $p = 0.005$ ). The 1-year, 2-year, and 5-year AUCs were 0.763, 0.801, and 0.734, respectively, indicating reliability (Figure 7B). After removing the interference of confounding factors, we confirmed that CD146 was an independent prognostic factor (Figures 7C,D). For the prognostic nomogram generated in the TARGET-OS cohort, we calculated the point for each patient in the GSE21257 cohort; the ROC curve showed a high AUC value at 0.868 (95% CI: 0.770–0.966, Figure 7E). There was no significant difference between the nomogram-predicted survival probability and the actual survival probability ( $p = 0.316$ , Figure 7F).

## Discussion

Osteosarcoma (OSA) is one of the focuses of orthopedic surgeons. Although it is the most frequent bone cancer, the incidence rate among all tumor types is rare, with an annual incidence of 3–4 patients per million (Smeland et al., 2019). The metaphyses of long bones are vulnerable to OSA, including the distal femur, proximal tibia, and proximal humerus. Once OSA develops, patients typically present with swelling, pain, localized

enlargement, and pathologic fracture. As a result of the low incidence rate and remarkable heterogeneity, the pathogenesis mechanism of OSA and prognostic factors remain unclear (Yang and Zhang, 2013). However, most patients are diagnosed at an advanced stage and die due to metastasis. Therefore, signatures with high accuracy and sensitivity for detecting metastasis and predicting survival are needed.

CD146 has been reported to be highly expressed in a variety of cancers in prior studies. CD146 plays an important role in promoting the progression, invasion, and metastasis of melanoma, gallbladder adenocarcinoma, and breast cancer. It has been confirmed as a predictor of poor survival in gastric cancer, lung adenocarcinoma, malignant pleural mesothelioma, and non-small cell lung cancer (Kristiansen et al., 2003; Sato et al., 2010; Liu et al., 2012; Jiang et al., 2016). Our study analyzed the expression of CD146 in 22 types of tumor and adjacent normal tissues and the correlation between CD146 mRNA levels and 32 types of tumor. We discovered that the mRNA level of CD146 was correlated with the corresponding prognosis. A high CD146 mRNA concentration indicated a poor survival time, especially in brain lower-grade glioma (LGG), testicular germ cell tumors (TGCT), kidney chromophobe (KICH), ACC, cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), and uveal melanoma (UVM)—consistent with previous studies. In addition, the expression levels of CD146 on OSA cell lines have been observed to be higher

TABLE 2 Prediction of potential drugs for CD146-high patients.

## GDSC

| Symbol | Drug              | Correlation | FDR value |
|--------|-------------------|-------------|-----------|
| CD146  | PLX4720           | -0.33871    | 9.43E-25  |
| CD146  | Dabrafenib        | -0.29988    | 1.24E-17  |
| CD146  | SB590885          | -0.25098    | 1.81E-11  |
| CD146  | (5Z)-7-Oxozeanol  | -0.24425    | 4.24E-12  |
| CD146  | MG-132            | -0.24331    | 0.000315  |
| CD146  | selumetinib       | -0.22606    | 3.49E-11  |
| CD146  | AZ628             | -0.21711    | 0.000396  |
| CD146  | Bortezomib        | -0.21643    | 0.000581  |
| CD146  | Trametinib        | -0.20994    | 1.84E-09  |
| CD146  | PD-0325901        | -0.20486    | 2.98E-08  |
| CD146  | RDEA119           | -0.20124    | 4.27E-09  |
| CD146  | TGX221            | -0.17799    | 0.0027    |
| CD146  | CHIR-99021        | -0.17036    | 1.43E-06  |
| CD146  | CGP-082996        | -0.16761    | 0.019578  |
| CD146  | CI-1040           | -0.16183    | 2.99E-05  |
| CD146  | Dasatinib         | -0.15004    | 0.015527  |
| CD146  | Docetaxel         | -0.14687    | 6.23E-05  |
| CD146  | 17-AAG            | -0.1446     | 6.98E-05  |
| CD146  | Z-LLNle-CHO       | -0.14257    | 0.028161  |
| CD146  | Bleomycin (50 uM) | -0.12835    | 0.000295  |
| CD146  | FH535             | -0.12834    | 0.001947  |
| CD146  | Cisplatin         | -0.12058    | 0.005306  |
| CD146  | piperlongumine    | -0.11333    | 0.003932  |
| CD146  | Elesclomol        | -0.10764    | 0.008985  |

## CTRP

|       |               |          |          |
|-------|---------------|----------|----------|
| CD146 | staurosporine | -0.22964 | 0.017592 |
| CD146 | BRD-K99006945 | -0.21083 | 0.002421 |
| CD146 | vemurafenib   | -0.20254 | 3.17E-05 |
| CD146 | PLX-4720      | -0.15145 | 0.005725 |
| CD146 | CIL41         | -0.13905 | 0.033183 |
| CD146 | simvastatin   | -0.13224 | 0.009889 |
| CD146 | lovastatin    | -0.12506 | 0.004992 |
| CD146 | TGX-221       | -0.12449 | 0.024848 |
| CD146 | AZD6482       | -0.12091 | 0.013324 |
| CD146 | CAL-101       | -0.11769 | 0.018051 |
| CD146 | XL765         | -0.11361 | 0.024965 |
| CD146 | niclosamide   | -0.10575 | 0.011619 |
| CD146 | TG-100-115    | -0.10306 | 0.028828 |
| CD146 | fluvastatin   | -0.10033 | 0.036151 |

than those on normal osteoblast cells (OST) via confocal images (Schiano et al., 2012), implying its predictive value in OSA. We found that high CD146 expression was related to poor clinical outcomes in the training cohort: this result was validated in the

GSE21257 cohort. Thus, we speculated that CD146 can act as a diagnostic biomarker to distinguish OSA from benign lesions. It is important to explore the mechanism of CD146 function in OSA. Lei et al. (2021) reported that CD146 promoted OSA growth by regulating angiogenesis and nourishment, endothelial cell proliferation, permeability, migration, vascular number, and diameter, which were similar to other tumors. However, the underlying mechanisms require further research. We enriched 580 DEGs in different pathways; as a result, vascular development, organic anion transport, cell adhesion, and ligand–receptor interactions are likely involved in the development of OSA. Immunocyte infiltration has been reported in many cancers and is the signature of hot tumors, also named “immune-inflamed tumors”, which are more sensitive to immune checkpoint inhibitors (ICIs) than cold tumors (Meng et al., 2019). The most prominent characteristic of hot tumors is high T-cell infiltration (Liu and Sun, 2021). According to our study, OSA patients with high CD146 expression show abundant and diverse T-cell infiltration. Thus, we speculate that high CD146 expression is responsible for the sensitivity to immunotherapy. The hypothesis was confirmed through SubMap analysis. OSA patients with high CD146 expression were more vulnerable to anti-PD1 therapy. T-cell infiltration has also been reported as a prognostic factor. T follicular helper cells were proven to be capable of predicting pathological complete response after chemotherapy in breast cancer (Gu-Trantien et al., 2013). Nevertheless, we found that low CD146 was related to low T-cell infiltration but was also related to a high response to cisplatin therapy. More research is required in this field. Therefore, the high expression of CD146 groups represent subgroups with high immunocyte infiltration, low response to cisplatin but high response to anti-PD1 therapy, and poor prognosis; the low expression of CD146 groups represents subgroups with low immunocyte infiltration, high response to cisplatin but low response to anti-PD1 therapy, and good prognosis.

Previous studies have found many related factors affecting the prognosis of OSA patients. Age and sex were controversial in different studies. Some studies showed that the 5-year OS at 0–14 years old was lower than that in other age groups, while a meta-analysis showed no statistical correlation between age and the survival of OSAs for sex. Some studies reported that the prognosis of male patients was worse than that of female patients, whereas some studies showed that sex had no correlation with prognosis. (Mirabello et al., 2009a; Joo et al., 2015; Ding et al., 2020). By performing robust statistical analysis, our study showed no correlation between age and the survival of OSA or sex. Multivariate analysis revealed that the CD146 mRNA level was an independent prognostic factor for OSA.

In summary, the role of CD146 in OSA was systematically analyzed. CD146 can act as a novel biomarker in predicting the prognosis of OSA and as an independent prognostic factor. We

TABLE 3 Clinical features of CD146 low and high groups in GSE21257 cohort.

| Clinicopathological features |                                     | CD146 Low     | CD146 high    | p Value            |
|------------------------------|-------------------------------------|---------------|---------------|--------------------|
| Gender (%)                   | Female                              | 11 (40.7%)    | 8 (30.8%)     | 0.638              |
|                              | Male                                | 16 (59.3%)    | 18 (69.2%)    |                    |
| Age, years                   | —                                   | 18.43 ± 10.63 | 19.01 ± 13.85 | 0.866              |
| HUVOS grade (%)              | Grade 1                             | 6 (22.2%)     | 7 (26.9%)     | 0.914              |
|                              | Grade 2                             | 9 (33.3%)     | 7 (26.9%)     |                    |
|                              | Grade 3                             | 7 (25.9%)     | 6 (23.1%)     |                    |
|                              | Grade 4                             | 3 (11.1%)     | 2 (7.7%)      |                    |
|                              | Unknown                             | 2 (7.4%)      | 4 (15.4%)     |                    |
| Metastasis (%)               | After diagnosis                     | 7 (25.9%)     | 13 (50.0%)    | 0.042 <sup>a</sup> |
|                              | At diagnosis                        | 6 (22.2%)     | 8 (30.8%)     |                    |
|                              | No                                  | 14 (51.9%)    | 5 (19.2%)     |                    |
| Location (%)                 | Diaphysis of left femur             | 0 (0.0%)      | 1 (3.8%)      | 0.816              |
|                              | Distal femur                        | 0 (0.0%)      | 1 (3.8%)      |                    |
|                              | Femur                               | 3 (11.1%)     | 2 (7.7%)      |                    |
|                              | Humerus                             | 2 (7.4%)      | 2 (7.7%)      |                    |
|                              | Left distal femur                   | 6 (22.2%)     | 5 (19.2%)     |                    |
|                              | Left femur                          | 0 (0.0%)      | 1 (3.8%)      |                    |
|                              | Left proximal femur                 | 1 (3.7%)      | 0 (0.0%)      |                    |
|                              | Left proximal fibula                | 0 (0.0%)      | 1 (3.8%)      |                    |
|                              | Left proximal humerus               | 1 (3.7%)      | 1 (3.8%)      |                    |
|                              | Left proximal tibia                 | 5 (18.5%)     | 2 (7.7%)      |                    |
|                              | Right distal femur                  | 4 (14.8%)     | 2 (7.7%)      |                    |
|                              | Right distal tibia                  | 1 (3.7%)      | 0 (0.0%)      |                    |
|                              | Right humerus                       | 0 (0.0%)      | 1 (3.8%)      |                    |
|                              | Right proximal femur                | 0 (0.0%)      | 1 (3.8%)      |                    |
|                              | Right proximal fibula               | 1 (3.7%)      | 0 (0.0%)      |                    |
|                              | Right proximal humerus              | 1 (3.7%)      | 0 (0.0%)      |                    |
|                              | Right proximal tibia                | 1 (3.7%)      | 4 (15.4%)     |                    |
|                              | Tibia                               | 1 (3.7%)      | 1 (3.8%)      |                    |
|                              | Unknown                             | 0 (0.0%)      | 1 (3.8%)      |                    |
| Histological (%)             | Anaplastic                          | 1 (3.7%)      | 1 (3.8%)      | 0.362              |
|                              | Chondroblastic                      | 2 (7.4%)      | 4 (15.4%)     |                    |
|                              | Fibroblastic                        | 4 (14.8%)     | 1 (3.8%)      |                    |
|                              | Giant cell rich                     | 0 (0.0%)      | 1 (3.8%)      |                    |
|                              | Osteoblastic                        | 15 (55.6%)    | 17 (65.4%)    |                    |
|                              | Pleomorphic                         | 1 (3.7%)      | 0 (0.0%)      |                    |
|                              | Possibly chondromyxoid fibroma like | 0 (0.0%)      | 1 (3.8%)      |                    |
|                              | Sclerosing                          | 1 (3.7%)      | 1 (3.8%)      |                    |
|                              | Telangiectatic                      | 3 (11.1%)     | 0 (0.0%)      |                    |

<sup>a</sup>,  $p < 0.05$ .

revealed the mechanisms by which CD146 promotes the growth, development, and metastasis of OSA. Contrary to cisplatin therapy, patients with high CD146 expression were more likely to be sensitive to anti-PD-1 therapy. The evidence above and prior research prove the potential value in individual survival prediction, in checkpoint therapy guidance, and as a novel therapeutic target (Stalin et al., 2017).

However, there are several limitations to our study. The samples we collected in the training and validation cohorts were insufficient. CD146 is a reliable prognosis-related factor but its role in the early diagnosis of OSA remains unclear. CD146 mRNA overexpression has been found in many cancers; therefore, the specificity for detection is poor. More research is needed in this field.

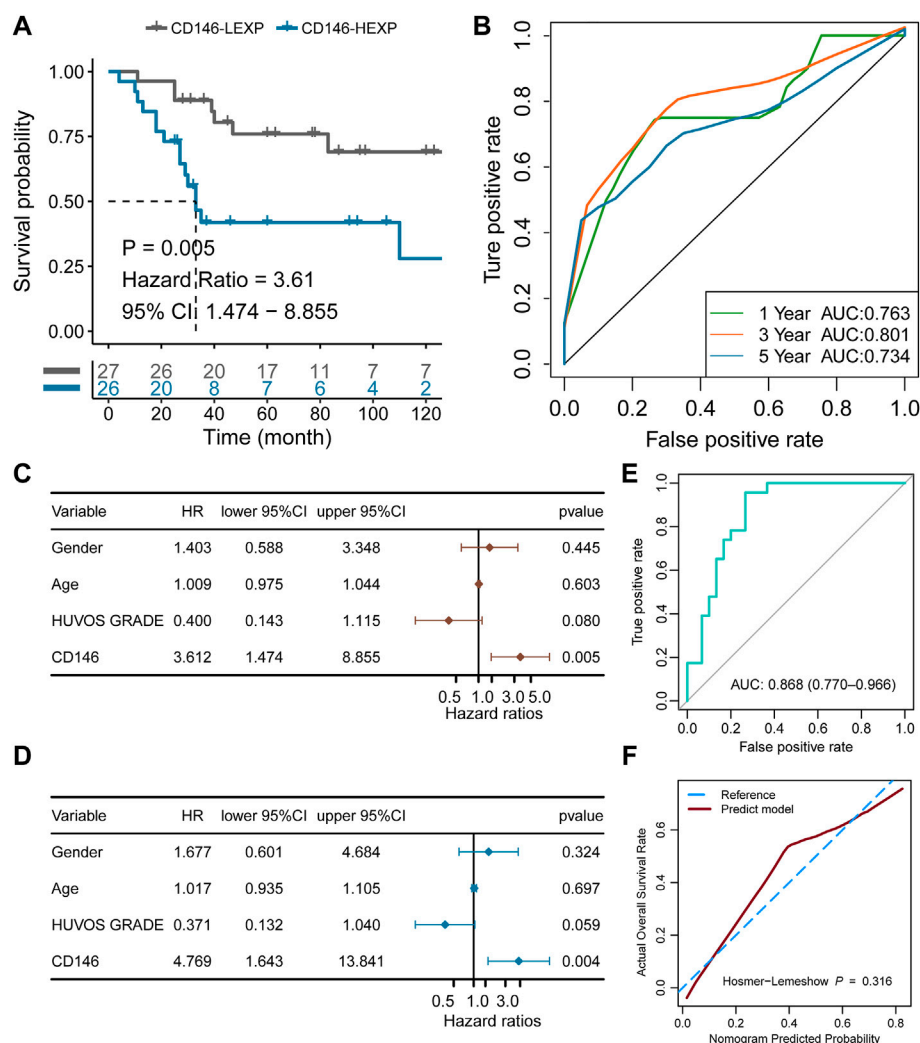


FIGURE 7

Validation of the independent prognostic value of CD146 in the external GSE21257 cohort. **(A)** K-M plot showing the diverse clinical overall survival outcomes of the CD146 low and high groups. **(B)** ROC curve showing the prognostic value of the CD146 expression level. **(C)** Results from univariate Cox regression analysis of the CD146 group and clinical features. **(D)** Results from multivariate Cox regression analysis of the CD146 group and clinical features. **(E)** Validation of the nomogram by ROC curve in GSE21257 cohort. **(F)** Validation of the nomogram by calibration curve in GSE21257 cohort.

## Conclusion

We reveal that CD146 acts as an oncogene in the prognosis of several types of tumor and illustrate its function well in OSA. CD146 is an independent prognostic factor for OSA patients after adjusting for age, sex, and metastatic status at diagnosis. In addition, the expression of CD146 also indicates clinical treatment strategies. Patients with low expression of CD146 are more suitable for chemotherapy with cisplatin and doxorubicin, while patients with high expression of CD146 respond better to anti-PD-1 immunotherapy.

## Data availability statement

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

## Author contributions

JW, ZW, and XX outlined the study design. JW, XZ, and MZ performed implementation and data collection. JW, MZ, and SY contributed to the data analysis. JW and XX contributed to both

the draft and final versions of the manuscript. All authors read and approved the final manuscript.

## Funding

The current study is supported by the Key Research and Development Project in Anhui Province (202104j07020026).

## 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.

## References

- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The gene Ontology consortium. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556
- Chatterjee, N., and Bivona, T. G. (2019). Polytherapy and targeted cancer drug resistance. *Trends Cancer* 5 (3), 170–182. doi:10.1016/j.trecan.2019.02.003
- Chen, C., Xie, L., Ren, T., Huang, Y., Xu, J., and Guo, W. (2021). Immunotherapy for osteosarcoma: Fundamental mechanism, rationale, and recent breakthroughs. *Cancer Lett.* 500, 1–10. doi:10.1016/j.canlet.2020.12.024
- Corre, I., Verrecchia, F., Crenn, V., Redini, F., and Trichet, V. (2020). The osteosarcoma microenvironment: A complex but targetable ecosystem. *Cells* 9 (4), E976. doi:10.3390/cells9040976
- Ding, W. Z., Liu, K., Li, Z., and Chen, S. R. (2020). A meta-analysis of prognostic factors of osteosarcoma. *Eur. Rev. Med. Pharmacol. Sci.* 24 (8), 4103–4112. doi:10.26355/eurrev\_202004\_20989
- Feng, G., Fang, F., Liu, C., Zhang, F., Huang, H., and Pu, C. (2012). CD146 gene expression in clear cell renal cell carcinoma: a potential marker for prediction of early recurrence after nephrectomy. *Int. Urol. Nephrol.* 44 (6), 1663–1669. doi:10.1007/s12555-012-0255-4
- Furue, H. (2003). Chemotherapy cancer treatment during the past sixty years. *Gan Kagaku Ryoho.* 30 (10), 1404–1411.
- Gu-Trantien, C., Loi, S., Garaud, S., Equeter, C., Libin, M., de Wind, A., et al. (2013). CD4<sup>+</sup> follicular helper T cell infiltration predicts breast cancer survival. *J. Clin. Invest.* 123 (7), 2873–2892. doi:10.1172/jci67428
- Huang, Q., Liang, X., Ren, T., Huang, Y., Zhang, H., Yu, Y., et al. (2021). The role of tumor-associated macrophages in osteosarcoma progression - therapeutic implications. *Cell. Oncol.* 44 (3), 525–539. doi:10.1007/s13402-021-00598-w
- Jiang, G., Zhang, L., Zhu, Q., Bai, D., Zhang, C., and Wang, X. (2016). CD146 promotes metastasis and predicts poor prognosis of hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* 35, 38. doi:10.1186/s13046-016-0313-3
- Johnson, J. P., Rothbacher, U., Sers, C., and Rothbacher, U. (1993). The progression associated antigen MUC18: a unique member of the immunoglobulin supergene family. *Melanoma Res.* 3 (5), 337–340. doi:10.1097/00008390-199310000-00006
- Joo, M. W., Shin, S. H., Kang, Y. K., Kawai, A., Kim, H. S., Asavamongkolkul, A., et al. (2015). Osteosarcoma in asian populations over the age of 40 Years: A multicenter study. *Ann. Surg. Oncol.* 22 (11), 3557–3564. doi:10.1245/s10434-015-4414-6
- Kristiansen, G., Yu, Y., Schlüns, K., Sers, C., Dietel, M., and Petersen, I. (2003). Expression of the cell adhesion molecule CD146/MCAM in non-small cell lung cancer. *Anal. Cell. Pathol.* 25 (2), 77–81. doi:10.1155/2003/574829
- Lehmann, J. M., Riethmüller, G., and Johnson, J. P. (1989). MUC18, a marker of tumor progression in human melanoma, shows sequence similarity to the neural cell adhesion molecules of the immunoglobulin superfamily. *Proc. Natl. Acad. Sci. U. S. A.* 86 (24), 9891–9895. doi:10.1073/pnas.86.24.9891
- Lei, X., Wang, K., Wang, W., Jin, H., Gu, W., Chen, Z., et al. (2021). Recognize the role of CD146/MCAM in the osteosarcoma progression: an *in vitro* study. *Cancer Cell Int.* 21 (1), 300. doi:10.1186/s12935-021-02006-7
- Li, G., Kalabis, J., Xu, X., Meier, F., Oka, M., Bogenrieder, T., et al. (2003). Reciprocal regulation of MelCAM and AKT in human melanoma. *Oncogene* 22 (44), 6891–6899. doi:10.1038/sj.onc.1206819
- Li, H. B., Huang, G., Tu, J., Lv, D. M., Jin, Q. L., Chen, J. K., et al. (2022). METTL14-mediated epitranscriptome modification of MN1 mRNA promote tumorigenicity and all-trans-retinoic acid resistance in osteosarcoma. *EBioMedicine* 82, 104142. doi:10.1016/j.ebiom.2022.104142
- Li, Y., Yu, J. M., Zhan, X. M., Liu, L. L., Jin, N., and Zhang, Y. X. (2014). Correlation of CD146 expression and clinicopathological characteristics in esophageal squamous cell carcinoma. *Oncol. Lett.* 8 (2), 859–863. doi:10.3892/ol.2014.2227
- Liang, Y. K., Zeng, D., Xiao, Y. S., Wu, Y., Ouyang, Y. X., Chen, M., et al. (2017). MCAM/CD146 promotes tamoxifen resistance in breast cancer cells through induction of epithelial-mesenchymal transition, decreased ERα expression and AKT activation. *Cancer Lett.* 386, 65–76. doi:10.1016/j.canlet.2016.11.004
- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004
- Liu, C. J., Hu, F. F., Xia, M. X., Han, L., Zhang, Q., and Guo, A. Y. (2018). GSCALite: a web server for gene set cancer analysis. *Bioinformatics* 34 (21), 3771–3772. doi:10.1093/bioinformatics/bty411
- Liu, W. F., Ji, S. R., Sun, J. J., Zhang, Y., Liu, Z. Y., Liang, A. B., et al. (2012). CD146 expression correlates with epithelial-mesenchymal transition markers and a poor prognosis in gastric cancer. *Int. J. Mol. Sci.* 13 (5), 6399–6406. doi:10.3390/ijms13056399
- Liu, Y. T., and Sun, Z. J. (2021). Turning cold tumors into hot tumors by improving T-cell infiltration. *Theranostics* 11 (11), 5365–5386. doi:10.7150/thno.58390
- Lu, X., Meng, J., Zhou, Y., Jiang, L., and Yan, F. (2020). MOVICS: an R package for multi-omics integration and visualization in cancer subtyping. *Bioinformatics* 36, 5539–5541. doi:10.1093/bioinformatics/btaa1018
- Melnikova, V. O., Balasubramanian, K., Villares, G. J., Dobroff, A. S., Zigler, M., Wang, H., et al. (2009). Crosstalk between protease-activated receptor 1 and platelet-activating factor receptor regulates melanoma cell adhesion molecule (MCAM/MUC18) expression and melanoma metastasis. *J. Biol. Chem.* 284 (42), 28845–28855. doi:10.1074/jbc.M109.042150
- Meng, J., Liu, Y., Guan, S., Fan, S., Zhou, J., Zhang, M., et al. (2019). The establishment of immune infiltration based novel recurrence predicting nomogram in prostate cancer. *Cancer Med.* 8 (11), 5202–5213. doi:10.1002/cam4.2433
- Mirabello, L., Troisi, R. J., and Savage, S. A. (2009a). International osteosarcoma incidence patterns in children and adolescents, middle ages and elderly persons. *Int. J. Cancer* 125 (1), 229–234. doi:10.1002/ijc.24320

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1025306/full#supplementary-material>

- Mirabello, L., Troisi, R. J., and Savage, S. A. (2009b). Osteosarcoma incidence and survival rates from 1973 to 2004: data from the surveillance, epidemiology, and end results Program. *Cancer* 115 (7), 1531–1543. doi:10.1002/cncr.24121
- Nie, Z., and Peng, H. (2018). Osteosarcoma in patients below 25 years of age: An observational study of incidence, metastasis, treatment and outcomes. *Oncol. Lett.* 16 (5), 6502–6514. doi:10.3892/ol.2018.9453
- Ottaviani, G., and Jaffe, N. (2009). The epidemiology of osteosarcoma. *Cancer Treat. Res.* 152, 3–13. doi:10.1007/978-1-4419-0284-9\_1
- Pakos, E. E., Nearchou, A. D., Grimer, R. J., Koumoullis, H. D., Abudu, A., Bramer, J. A., et al. (2009). Prognostic factors and outcomes for osteosarcoma: an international collaboration. *Eur. J. Cancer* 45 (13), 2367–2375. doi:10.1016/j.ejca.2009.03.005
- Pan, Y., Chen, D., Hu, T., Lv, G., and Dai, Z. (2019). Characteristics and prognostic factors of patients with osteosarcoma older than 60 Years from the SEER database. *Cancer Control* 26 (1), 1073274819888893. doi:10.1177/1073274819888893
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Ritter, J., and Bielack, S. S. (2010). Osteosarcoma. *Ann. Oncol.* 21, vii320–325. doi:10.1093/annonc/mdq276
- Rodig, S. J., Gusenleitner, D., Jackson, D. G., Gjini, E., Giobbie-Hurder, A., Jin, C., et al. (2018). MHC proteins confer differential sensitivity to CTLA-4 and PD-1 blockade in untreated metastatic melanoma. *Sci. Transl. Med.* 10 (450), eaar3342. doi:10.1126/scitranslmed.aar3342
- Rojas, G. A., Hubbard, A. K., Diessner, B. J., Ribeiro, K. B., and Spector, L. G. (2021). International trends in incidence of osteosarcoma (1988–2012). *Int. J. Cancer* 149 (5), 1044–1053. doi:10.1002/ijc.33673
- Sato, A., Torii, I., Okamura, Y., Yamamoto, T., Nishigami, T., Kataoka, T. R., et al. (2010). Immunocytochemistry of CD146 is useful to discriminate between malignant pleural mesothelioma and reactive mesothelium. *Mod. Pathol.* 23 (11), 1458–1466. doi:10.1038/modpathol.2010.134
- Schiano, C., Grimaldi, V., Casamassimi, A., Infante, T., Esposito, A., Giovane, A., et al. (2012). Different expression of CD146 in human normal and osteosarcoma cell lines. *Med. Oncol.* 29 (4), 2998–3002. doi:10.1007/s12032-012-0158-3
- Sechler, M., Parrish, J. K., Birks, D. K., and Jedlicka, P. (2017). The histone demethylase KDM3A, and its downstream target MCAM, promote Ewing Sarcoma cell migration and metastasis. *Oncogene* 36 (29), 4150–4160. doi:10.1038/onc.2017.44
- Smeland, S., Bielack, S. S., Whelan, J., Bernstein, M., Hogendoorn, P., Krailo, M. D., et al. (2019). Survival and prognosis with osteosarcoma: outcomes in more than 2000 patients in the EURAMOS-1 (European and American osteosarcoma study) cohort. *Eur. J. Cancer* 109, 36–50. doi:10.1016/j.ejca.2018.11.027
- So, J. H., Hong, S. K., Kim, H. T., Jung, S. H., Lee, M. S., Choi, J. H., et al. (2010). Gicerin/Cd146 is involved in zebrafish cardiovascular development and tumor angiogenesis. *Genes Cells* 15 (11), 1099–1110. doi:10.1111/j.1365-2443.2010.01448.x
- Stalin, J., Nollet, M., Dignat-George, F., Bardin, N., and Blot-Chaubaud, M. (2017). Therapeutic and diagnostic antibodies to CD146: Thirty years of research on its potential for detection and treatment of tumors. *Antibodies (Basel)* 6 (4), E17. doi:10.3390/antib6040017
- Thanindratarn, P., Dean, D. C., Nelson, S. D., Hornicek, F. J., and Duan, Z. (2019). Advances in immune checkpoint inhibitors for bone sarcoma therapy. *J. Bone Oncol.* 15, 100221. doi:10.1016/j.jbo.2019.100221
- Tian, B., Zhang, Y., and Li, N. (2013). CD146 protein as a marker to predict postoperative liver metastasis in colorectal cancer. *Cancer Biother. Radiopharm.* 28 (6), 466–470. doi:10.1089/cbr.2012.1426
- Valery, P. C., Laversanne, M., and Bray, F. (2015). Bone cancer incidence by morphological subtype: a global assessment. *Cancer Causes Control* 26 (8), 1127–1139. doi:10.1007/s10552-015-0607-3
- Wang, S. D., Li, H. Y., Li, B. H., Xie, T., Zhu, T., Sun, L. L., et al. (2016). The role of CTLA-4 and PD-1 in anti-tumor immune response and their potential efficacy against osteosarcoma. *Int. Immunopharmacol.* 38, 81–89. doi:10.1016/j.intimp.2016.05.016
- Wang, Z., and Yan, X. (2013). CD146, a multi-functional molecule beyond adhesion. *Cancer Lett.* 330 (2), 150–162. doi:10.1016/j.canlet.2012.11.049
- Westrom, S., Bønsdorff, T. B., Abbas, N., Bruland Ø, S., Jonasdottir, T. J., Mælandsmo, G. M., et al. (2016). Evaluation of CD146 as target for radioimmunotherapy against osteosarcoma. *PLoS One* 11 (10), e0165382. doi:10.1371/journal.pone.0165382
- Wu, Z., Wu, Z., Li, J., Yang, X., Wang, Y., Yu, Y., et al. (2012). MCAM is a novel metastasis marker and regulates spreading, apoptosis and invasion of ovarian cancer cells. *Tumour Biol.* 33 (5), 1619–1628. doi:10.1007/s13277-012-0417-0
- Xie, S., Luca, M., Huang, S., Gutman, M., Reich, R., Johnson, J. P., et al. (1997). Expression of MCAM/MUC18 by human melanoma cells leads to increased tumor growth and metastasis. *Cancer Res.* 57 (11), 2295–2303.
- Yang, J., and Zhang, W. (2013). New molecular insights into osteosarcoma targeted therapy. *Curr. Opin. Oncol.* 25 (4), 398–406. doi:10.1097/CCO.0b013e3283622c1b
- Yasin, N. F., Abdul Rashid, M. L., and Ajit Singh, V. (2020). Survival analysis of osteosarcoma patients: A 15-year experience. *J. Orthop. Surg.* 28 (1), 2309499019896662. doi:10.1177/2309499019896662
- Yoshida, K., Okamoto, M., Sasaki, J., Kuroda, C., Ishida, H., Ueda, K., et al. (2020). Anti-PD-1 antibody decreases tumour-infiltrating regulatory T cells. *BMC Cancer* 20 (1), 25. doi:10.1186/s12885-019-6499-y
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612
- Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16 (5), 284–287. doi:10.1089/omi.2011.0118
- Zabouo, G., Imbert, A. M., Jacquemier, J., Finetti, P., Moreau, T., Esterni, B., et al. (2009). CD146 expression is associated with a poor prognosis in human breast tumors and with enhanced motility in breast cancer cell lines. *Breast Cancer Res.* 11 (1), R1. doi:10.1186/bcr2215
- Zeng, Q., Li, W., Lu, D., Wu, Z., Duan, H., Luo, Y., et al. (2012). CD146, an epithelial-mesenchymal transition inducer, is associated with triple-negative breast cancer. *Proc. Natl. Acad. Sci. U. S. A.* 109 (4), 1127–1132. doi:10.1073/pnas.1111053108
- Zhang, C., Wan, J., Liu, Q., Long, F., Wen, Z., and Liu, Y. (2022). METTL3 upregulates COPS5 expression in osteosarcoma in an m(6)A-related manner to promote osteosarcoma progression. *Exp. Cell Res.* 420 (2), 113353. doi:10.1016/j.yexcr.2022.113353
- Zhou, Q., Xian, M., Xiang, S., Xiang, D., Shao, X., Wang, J., et al. (2017). All-trans retinoic acid prevents osteosarcoma metastasis by inhibiting M2 polarization of tumor-associated macrophages. *Cancer Immunol. Res.* 5 (7), 547–559. doi:10.1158/2326-6066.CIR-16-0259



## OPEN ACCESS

## EDITED BY

Haitao Wang,  
Center for Cancer Research, National  
Cancer Institute (NIH), United States

## REVIEWED BY

Dawei Zhou,  
The Ohio State University, United States  
Bo Zhu,  
University of Texas MD Anderson  
Cancer Center, United States

## \*CORRESPONDENCE

Sanjie Jiang,  
jiangsanjie@bgi.com  
Jingjing Li,  
jingjingli.md.phd@gmail.com

<sup>†</sup>These authors have contributed equally  
to this work

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Cell and  
Developmental Biology

RECEIVED 28 September 2022

ACCEPTED 24 October 2022

PUBLISHED 03 November 2022

## CITATION

Li P, Wang Y, Sun Y, Jiang S and Li J  
(2022), *N*<sup>6</sup>-methyladenosine RNA  
methylation: From regulatory  
mechanisms to potential  
clinical applications.  
*Front. Cell Dev. Biol.* 10:1055808.  
doi: 10.3389/fcell.2022.1055808

## COPYRIGHT

© 2022 Li, Wang, Sun, Jiang 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.

# *N*<sup>6</sup>-methyladenosine RNA methylation: From regulatory mechanisms to potential clinical applications

Peipei Li<sup>1,2†</sup>, Yuntao Wang<sup>1†</sup>, Yiwen Sun<sup>2</sup>, Sanjie Jiang<sup>2\*</sup> and  
Jingjing Li<sup>1\*</sup>

<sup>1</sup>Department of Oncology, Weifang Medical University, Weifang, China, <sup>2</sup>BGI Genomics, BGI-Shenzhen, Shenzhen, China

Epitranscriptomics has emerged as another level of epigenetic regulation similar to DNA and histone modifications. *N*<sup>6</sup>-methyladenosine (m<sup>6</sup>A) is one of the most prevalent and abundant posttranscriptional modifications, widely distributed in many biological species. The level of *N*<sup>6</sup>-methyladenosine RNA methylation is dynamically and reversibly regulated by distinct effectors including methyltransferases, demethylases, histone modification and metabolites. In addition, *N*<sup>6</sup>-methyladenosine RNA methylation is involved in multiple RNA metabolism pathways, such as splicing, localization, translation efficiency, stability and degradation, ultimately affecting various pathological processes, especially the oncogenic and tumor-suppressing activities. Recent studies also reveal that *N*<sup>6</sup>-methyladenosine modification exerts the function in immune cells and tumor immunity. In this review, we mainly focus on the regulatory mechanisms of *N*<sup>6</sup>-methyladenosine RNA methylation, the techniques for detecting *N*<sup>6</sup>-methyladenosine methylation, the role of *N*<sup>6</sup>-methyladenosine modification in cancer and other diseases, and the potential clinical applications.

## KEYWORDS

RNA methylation, *N*<sup>6</sup>-methyladenosine, regulatory mechanisms, cancer, tumor therapy

## 1 Introduction

With the development of epigenetics, epitranscriptomics has emerged as another level of epigenetic regulation and has recently become a research hotspot. The epitranscriptome refers to the relevant functional changes of the transcriptome without any alteration of the RNA sequence. Conceptually, the epitranscriptome covers all the chemical modifications of RNA dynamically regulated by the removal and addition of various chemical groups in cells (Saletore et al., 2012). To date, over 170 RNA chemical modifications have been identified, including *N*<sup>6</sup>-methyladenosine (m<sup>6</sup>A), *N*<sup>1</sup>-methyladenosine (m<sup>1</sup>A), 5-hydroxymethylcytosine (hm<sup>5</sup>C), 5-methylcytidine (m<sup>5</sup>C), ribose 2'-O-methylation (Nm), 1-methylguanine (m<sup>1</sup>G), 6-methylguanine (m<sup>6</sup>G),

7-methylguanine ( $m^7G$ ),  $N^4$ -acetylcytidine ( $ac^4C$ ) and pseudouridine ( $w$ ) (Wiener and Schwartz, 2021), but most of their functions are largely unknown. Among them, 72 variants of methyl group modifications are conjugated at distinct positions in RNA bases. Since the first discovery of  $m^6A$  RNA methylation in 1974, it has been identified as one of the most prevalent and abundant posttranscriptional modifications, widely distributed in many biological species, such as mammals (Desrosiers et al., 1974; Liu et al., 2022), plants (Yue et al., 2019; Yu Q. et al., 2021), zebrafish (Zhao et al., 2017), insects (Yang et al., 2021), yeast (Yadav and Rajasekharan, 2018), bacteria (Deng et al., 2015) and viruses (Bayoumi et al., 2020), accounting for approximately 50% of total methylated ribonucleotides in total RNA content (Wei et al., 1975). It is estimated that more than 7,000 mRNAs with  $m^6A$  modification are distributed in mammalian cells, with a frequency of 0.1–0.6% of adenosines (Ke et al., 2015). In addition,  $m^6A$  deposition also exists in other types of RNA, including rRNA, tRNA, small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), long noncoding RNA (lncRNA), microRNA (miRNA), and circular RNA (circRNA) (Pendleton et al., 2017; van Tran et al., 2019; Dai et al., 2020).

In 2012, several decades after the first discovery of  $m^6A$  RNA methylation, utilizing  $m^6A$ -specific antibodies, two groups independently conducted fragmented RNA immunoprecipitation and subsequent high-throughput RNA deep sequencing (termed “MeRIP-seq” or “ $m^6A$ -seq”) to map  $m^6A$  throughout the transcriptome in humans and mice (Dominissini et al., 2012; Meyer et al., 2012). The results first revealed that  $m^6A$  modification was widely distributed in mRNA, additionally,  $m^6A$  modification mainly occurred in the common motif RRACH (R = G or A, H = A, C or U), but only 1–5% of these sites were methylated in cellular RNA. Notably, most  $m^6A$  peaks were evolutionarily conserved between the human and mouse transcriptomes. More surprisingly,  $m^6A$  modification on the RRACH motif was preferentially enriched in 3'-untranslated regions (3'-UTRs) and near stop codons of coding sequences (CDS) (Meyer et al., 2012), indicating that the RRACH motif is not sufficient for the determination of  $m^6A$  modification. The  $m^6A$  levels vary in distinct cell contexts and are involved in multiple RNA metabolism pathways, such as splicing, localization, translation efficiency, stability and degradation (Kasowitz et al., 2018; Liu et al., 2020; He and He, 2021), ultimately affecting various physiological and pathological processes.

$m^6A$  RNA methylation is dynamic and reversible and is also tightly regulated by three types of proteins, methyltransferases (“writers”), demethylases (“erasers”) and  $m^6A$  binding proteins (“readers”). The  $m^6A$  methylase complex was first purified in the 1990s (Bokar et al., 1994). METTL3 is identified as a predominant component and contains a catalytically active subunit. Another methyltransferase, METTL14, is essential for structural stability to facilitate the catalysis of  $m^6A$  methylation. The larger methyltransferase holocomplex is composed of

WTAP, HAKAI, RBM15, RBM15B, VIRMA and ZC3H13 and is approximately 1,000 kDa in size (Oerum et al., 2021). FTO was the first  $m^6A$  demethylase identified in 2011, followed by ALKBH5 (Niu et al., 2013). These two enzymes can remove  $m^6A$  methylation from RNA, posing a novel research field of regulation for epitranscriptomics. The functions of  $m^6A$  RNA methylation are mediated by different  $m^6A$  “readers” that selectively recognize  $m^6A$  in a direct or indirect manner and conduct distinct functions (Shi et al., 2019). The  $m^6A$  binding proteins include the YTH family, the heterogeneous nuclear ribonuclease (HNRNP) family and FMRP (Figure 1).

In this review, we will address the regulatory mechanisms of  $m^6A$  RNA methylation, the techniques for detecting  $m^6A$  methylation, the role of  $m^6A$  modification in cancer as well as the potential clinical applications.

## 2 Regulation of $N^6$ -methyladenosine RNA modification

$m^6A$  RNA methylation is functionally important and tightly modulated by several molecular mechanisms in eukaryotes. As described above, catalytic enzymes, methyltransferases and demethylases dynamically and reversibly direct the addition and removal of  $m^6A$  RNA methylation, and are termed “ $m^6A$  writers” and “ $m^6A$  erasers”, respectively. Some binding proteins (“ $m^6A$  readers”) recognize and function by decoding  $m^6A$  methylation as well as recruiting downstream functional protein complexes to mediate biological activities. Histone modification also guides  $m^6A$  deposition in stop codons of CDSs and 3'-UTRs. The dynamics of  $m^6A$  regulation can be achieved by some transcription factors that recruit the  $m^6A$  methyltransferase complex to specific RNA loci in distinct cellular contexts. In addition, nutrition and metabolites can reverse and modulate  $m^6A$  methylation patterns (Figure 2).

### 2.1 RNA $N^6$ -methyladenosine machinery

$m^6A$  methylation is catalysed by a multicomponent  $m^6A$  methyltransferase complex that is composed of two predominant proteins, methyltransferase-like 3 (METTL3) and methyltransferase-like 14 (METTL14), and their cofactors WTAP, HAKAI, RBM15, RBM15B, VIRMA and ZC3H13 (Knuckles et al., 2018; Yue et al., 2018; Bawankar et al., 2021). Although both METTL3 and METTL14 have methyltransferase domains, only METTL3 contains a catalytically active subunit, which requires S-adenosylmethionine (SAM) as a substrate to mediate catalytic activity. The SAM binding pocket is distributed on one side of the central  $\beta$ -sheet and is enclosed by the catalytic site loop (Wang et al., 2016). METTL14 is associated with the stabilization of the conformation between METTL3 and the RNA

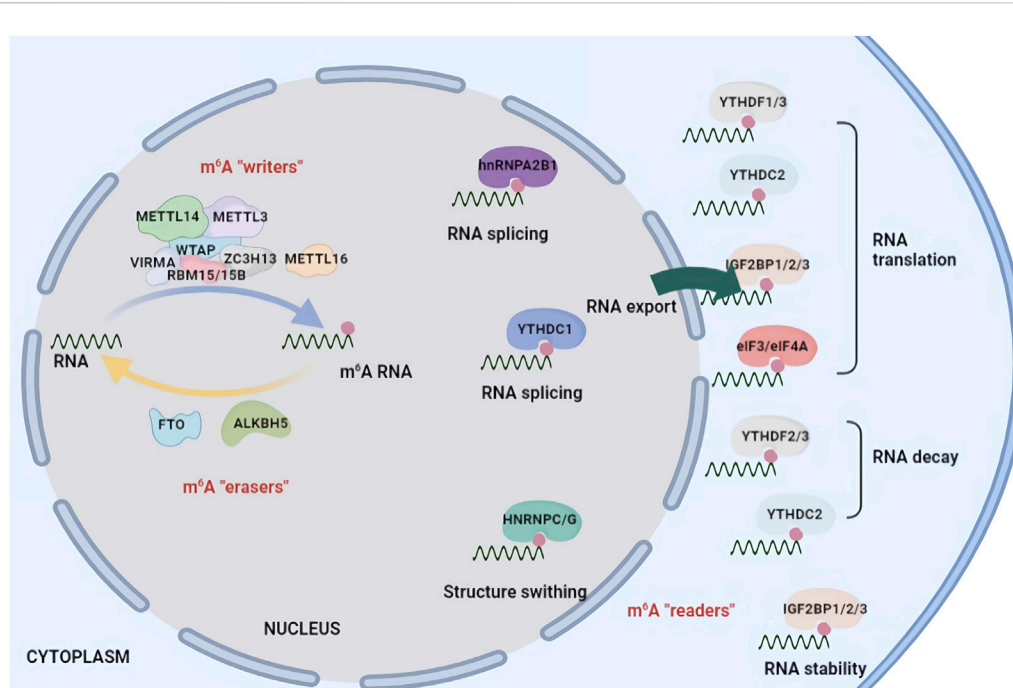


FIGURE 1

Chemical basis and molecular functions of m<sup>6</sup>A machinery. m<sup>6</sup>A methylation is catalysed by a multicomponent m<sup>6</sup>A methyltransferase complex ("m<sup>6</sup>A writers") that is composed of two predominant proteins, METTL3 and METTL14, as well as their cofactors WTAP, HAKAI, RBM15, RBM15B, VIRMA and ZC3H13. m<sup>6</sup>A methylation can be removed by demethylases ("m<sup>6</sup>A erasers") including FTO and ALKBH5. m<sup>6</sup>A modification affects RNA fate by recruiting m<sup>6</sup>A-binding proteins ("m<sup>6</sup>A readers") such as YTHDF1/2/3, YTHDC1/2, IGF2BP1/2/3 and HNRNPC/A2B1. m<sup>6</sup>A methylation is involved in RNA splicing, localization, translation efficiency, stability and degradation.

substrate (Zhou H. et al., 2021). The METTL3-METTL14 complex is formed in the cytoplasm and is located in the nucleus, and it induces m<sup>6</sup>A methylation (Scholler et al., 2018). Both *in vitro* methylation assays and CLIP combined with photoactivatable ribonucleoside-enhanced crosslinking (PAR-CLIP) suggested that the METTL3-METTL14 complex efficiently catalyses m<sup>6</sup>A methylation on the GGACU or GGAC motifs of RNAs, consistent with the RRACH motif of a previous study (Liu et al., 2014). The depletion of *Mettl3* and/or *Mettl14* could greatly reduce the peak numbers and the enrichment of m<sup>6</sup>A in the global transcriptome (Vu et al., 2017; Weng et al., 2018). Except for these two modulators, other m<sup>6</sup>A writers lack methyltransferase activity. WTAP guides METTL3 and METTL14 into nuclear speckles to efficiently methylate target RNAs (Ping et al., 2014). RBM15 and RBM15B have been confirmed to interact with WTAP by coimmunoprecipitation and bind to specific RNA regions that are adjacent to the DRACH sequence, suggesting that RBM15 and RBM15B can recruit the METTL3-WTAP complex and direct these methyltransferases to DRACH consensus sequence sites for m<sup>6</sup>A modification (Patil et al., 2016; Shi et al., 2019). RBM15 and RBM15B targets induce X-chromosome inactivation and gene silencing by binding to lncRNA XIST, and VIRMA prefers to mediate alternative polyadenylation and

mRNA methylation near the 3'-UTRs and stop codon regions (Yue et al., 2018; Zhu et al., 2021). ZC3H13 complexes with WTAP or other cofactors to regulate nuclear m<sup>6</sup>A RNA methylation (Wen et al., 2018). Recently, METTL16, as an independent RNA methyltransferase, was shown to catalyse the m<sup>6</sup>A methylation of U6 spliceosomal RNA (snRNA) and U6-like hairpins of *Mat2a* mRNA (Shima et al., 2017). ZCCHC4, a new m<sup>6</sup>A methyltransferase, was found to specifically recognize the AAC motif associated with rRNA methylation (Ma et al., 2019).

After deposition, m<sup>6</sup>A methylation is reversible and can be removed by demethylases ("m<sup>6</sup>A erasers"). FTO belongs to the nonheme Fe(II)- and  $\alpha$ -KG-dependent dioxygenase AlkB family. FTO was the first enzyme reported to modulate m<sup>6</sup>A demethylation in 2011. In addition, m<sup>6</sup>Am RNA, m<sup>1</sup>A RNA, m<sup>3</sup>T single-stranded DNA and m<sup>3</sup>U single-stranded RNA modifications can be demethylated by FTO (Wei et al., 2018). Mauer et al. found that the catalytic activity of FTO towards m<sup>6</sup>Am was approximately 10 times greater than that towards m<sup>6</sup>A (Mauer et al., 2017). It has been reported that snRNA and snoRNA are targets of FTO (Mauer et al., 2019). Another m<sup>6</sup>A demethylase, ALKBH5, is a member of the ALKB family and seems to be specific for m<sup>6</sup>A RNA methylation (Yu F. et al., 2021).

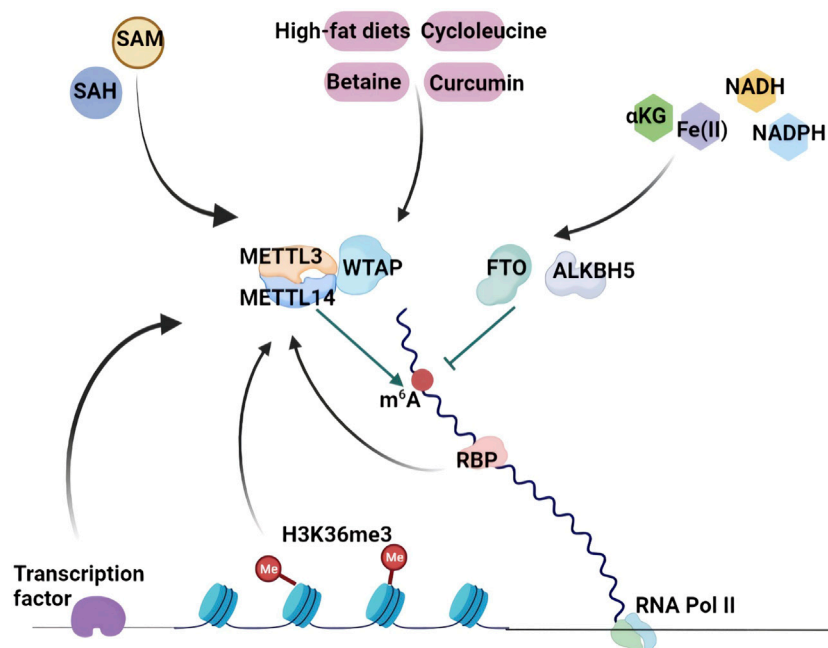


FIGURE 2

The biogenesis and regulatory mechanisms of RNA  $m^6A$  Methylation. Methyltransferases and demethylases dynamically and reversibly direct the addition and removal of  $m^6A$  RNA methylation. H3K36me3 guides  $m^6A$  deposition in stop codons of CDSs and 3'-UTRs. The dynamics of  $m^6A$  regulation can be achieved by some transcription factors that recruit the  $m^6A$  methyltransferase complex to specific RNA loci in distinct cellular contexts. In addition,  $m^6A$  methylation patterns are modulated by nutrition and metabolites including SAM, SAH, cyclolucine, betaine, curcumin, high-fat diets,  $\alpha$ KG, NADH and NADPH.

$m^6A$  modification affects RNA fate by recruiting  $m^6A$ -binding proteins ( $m^6A$  'readers') such as YTH domain-containing proteins, insulin-like growth factor 2 mRNA-binding proteins IGF2BP1-3 and the heterogeneous nuclear ribonuclease (HNRNP) family (Zhao et al., 2020). In mammals, YTH domain-containing proteins contain five members: YTHDC1, YTHDC2, YTHDF1, YTHDF2, and YTHDF3. YTHDC1 play roles in alternative splicing events, nuclear export of RNAs into the cytoplasm and mRNA decay (Xiao et al., 2016). YTHDC2 promotes target mRNA translation (Mao et al., 2019). YTHDF1 interacts with the translation initiation factors eIF3 and eIF4A3 to enhance the translation efficiency of  $m^6A$ -modified mRNAs (Liu et al., 2020; Cai et al., 2022). YTHDF2 and YTHDF3 are involved in the degradation of target mRNAs associated with p-bodies, and the depletion of *Ythdf2* and *Ythdf3* causes a considerable increase in  $m^6A$  mRNA abundance in cells. hnRNPA2B1, hnRNPC and hnRNPG are related to mRNA splicing (Bi et al., 2019). In contrast, IGF2BP1/2/3, FMRP, and PRRC2A are essential for the stabilization of  $m^6A$ -modified transcripts in a  $m^6A$ -dependent manner (Huang et al., 2018; Edens et al., 2019; Wu et al., 2019).

## 2.2 Histone modification guides $N^6$ -methyladenosine deposition

Histone modification has an effect on  $m^6A$  deposition. H3 lysine 36 trimethylation (H3K36me3) is a classical transcription activator that shows a similar  $m^6A$  distribution. H3K36me3 chromatin immunoprecipitation (ChIP)-seq analysis indicated that approximately 70% of H3K36me3 sites overlapped with  $m^6A$  peaks, suggesting a close connection between H3K36me3 and  $m^6A$  modification (Huang et al., 2019) (Figure 2). SETD2 and KDM4A are H3K36me3 methyltransferase and H3K36me3 demethylase, respectively (Mar et al., 2017). More surprisingly, changes in H3K36me3 levels by utilizing dCas9-SETD2 or dCas9-KDM4A can significantly alter  $m^6A$  abundance in human and mouse transcriptomes, revealing that H3K36me3 is a modulator of  $m^6A$  deposition. Of note, most H3K36me3-dependent  $m^6A$  sites are targeted by METTL3, METTL14 and WTAP, demonstrating the association between H3K36me3 and  $m^6A$  modification. In addition, METTL14 can directly bind to H3K36me3, which leads to the recruitment of other  $m^6A$  methyltransferases to activate RNA Pol II and controls  $m^6A$  methylation on mRNA (Huang et al., 2019; Zhou X. L. et al., 2021). Therefore, the

relationship between H3K36me3 and m<sup>6</sup>A provides a new way to enrich multiple aspects of gene expression regulation.

## 2.3 Transcription factors affect N<sup>6</sup>-methyladenosine deposition

Transcription factors also recruit m<sup>6</sup>A methyltransferases to regulate RNA modification in specific cell contexts. Zinc-finger protein 217 (ZFP217) is a transcriptional activator of some key pluripotency genes that are essential for maintaining self-renewal in mESCs. METTL3 can be bound and sequestered by ZFP217, preventing the formation of the m<sup>6</sup>A methyltransferase complex and m<sup>6</sup>A methylation on ZFP217 target transcripts (Aguilo et al., 2015). In contrast, the transcription factors SMAD2 and SMAD3 preferentially recruit the METTL3-METTL14-WTAP methyltransferase complex to their target transcripts and increase the m<sup>6</sup>A modification of target transcripts (Bertero et al., 2018). Different functions of m<sup>6</sup>A deposition on the target transcripts determine the distinct roles of ZFP217 and SMAD2/3 in ESCs. Another study reported that METTL3 could interact with the CAATT-box binding protein CEBPZ on target transcripts and mediate m<sup>6</sup>A modification to promote the translation of target mRNAs that maintain the leukaemic state in acute myeloid leukaemia (AML) cells (Barbieri et al., 2017). Frequently, these transcription factors manipulate m<sup>6</sup>A deposition on a subset of target transcripts in specific cellular contexts to implement dynamic regulation of gene expression (Figure 2).

## 2.4 Nutritional metabolism and metabolites regulate N<sup>6</sup>-methyladenosine deposition

Evidence has shown that nutritional challenge and metabolites play crucial roles in the manipulation of m<sup>6</sup>A deposition. Cycloleucine is a competitive inhibitor of methionine adenosyltransferase that decreases m<sup>6</sup>A RNA methylation levels by reducing SAM concentrations (Kang et al., 2018). Betaine, as a methyl donor for SAM synthesis, prompts m<sup>6</sup>A methylation by suppressing FTO expression in the adipose tissues of high-fat diet-fed mice, which increases the expression of the mitochondrial protein PGC-1 $\alpha$  to improve metabolic disorder (Zhou et al., 2015). It has been reported that curcumin enhances m<sup>6</sup>A modification by decreasing ALKBH5 and increasing METTL3 and METTL14 expression in the livers of piglets (Ding et al., 2016). Undoubtedly, high-fat diets affect m<sup>6</sup>A modification in various models and tissues (Tung et al., 2010; Wu et al., 2020) (Figure 2).

SAM, a common methyl donor, is involved in most cellular methylation processes. The change in cellular SAM concentration affects DNA and histone methylation as well as

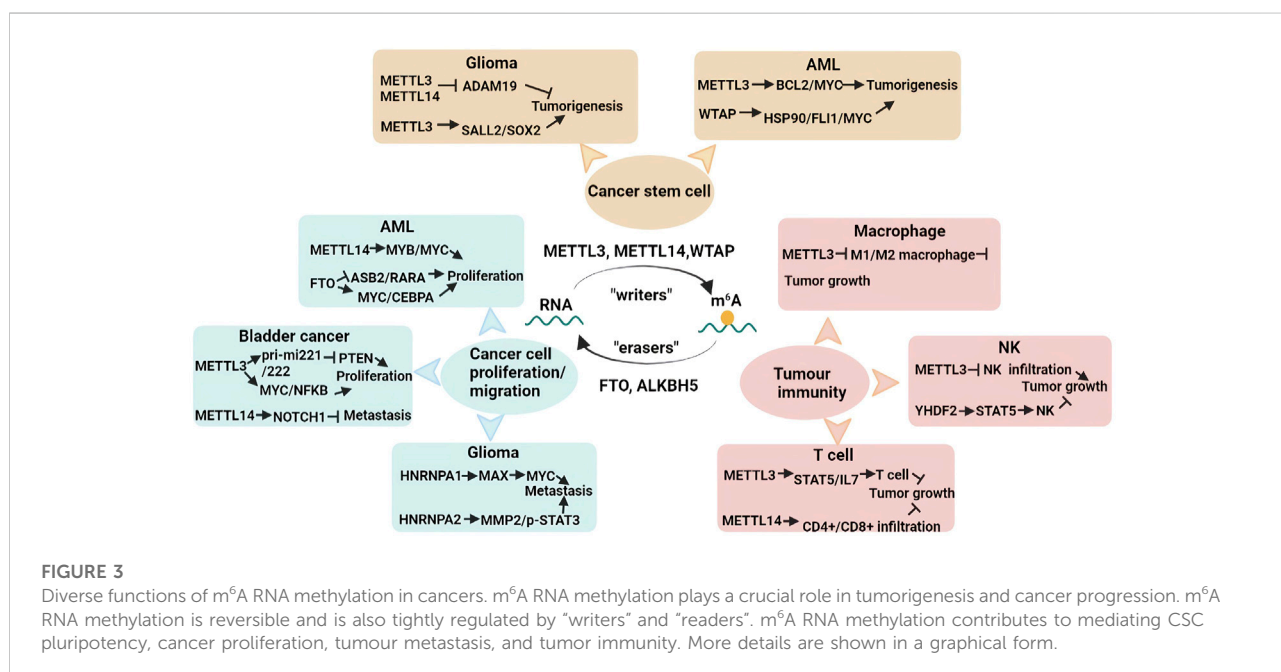
RNA methylation (Duncan et al., 2013). METTL3 requires SAM as a substrate to mediate catalytic activity and m<sup>6</sup>A writing. Interestingly, the SAM binding affinity of METTL3 is regulated by substrate RNA availability. S-adenosyl homocysteine (SAH) is the metabolite of SAM during the methylation reaction that can strongly inhibit METTL3 methyltransferase activity (Li F. et al., 2016; Selberg et al., 2019). Demethylases, FTO and ALKBH5 are 2-oxoglutarate ( $\alpha$ KG)- and Fe(II) dependent. The FTO and ALKBH5 mutants of the  $\alpha$ KG-Fe(II) oxygenase domain lost the catalytic activities of m<sup>6</sup>A demethylation (Feng et al., 2014; Zhang X. et al., 2019). Recently, NADH and NADPH were identified as the direct binding partners of FTO by using a fluorescence quenching assay. Both NADH and NADPH could enhance FTO demethylase activity, indicating that reducing NADPH and NADH may attenuate demethylation reactions. Conversely, the induction of NADPH by glucose injection or a high-fat diet suppressed m<sup>6</sup>A modification (Figure 2). In contrast, the depletion of G6P dehydrogenase (G6PD) or NAD kinase (NADK) enhanced cellular m<sup>6</sup>A abundance, which was reversed by NADPH supplementation (Wang L. et al., 2020).

## 3 Approaches for detecting N<sup>6</sup>-methyladenosine RNA methylation

Several techniques have been developed for detecting m<sup>6</sup>A RNA methylation (Table 1). Although immuno-northern blot and m<sup>6</sup>A dot blot facilitate easier and faster observation of global m<sup>6</sup>A levels, the disadvantages are obvious with lower sensitivity and are semiquantitative accuracy (Nagarajan et al., 2019). High-performance liquid chromatography–mass spectrometry (HPLC–MS/MS) is used for quantifying m<sup>6</sup>A levels with high sensitivity; however, this approach cannot provide details about RNA sequence and localization information (Thuring et al., 2017). Site-specific cleavage and radioactive labelling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET) is suitable for stoichiometric quantification, but it is very tedious and is only used to validate the known m<sup>6</sup>A changes at a given site (Liu and Pan, 2016). Above all, these techniques are not appropriate for widespread identification and localization of modified sites; subsequently, high-throughput sequencing methods have emerged and have been rapidly developed. m<sup>6</sup>A-specific antibody-based high-throughput sequencing strategies are widely used for the identification of m<sup>6</sup>A, which include m<sup>6</sup>A-Seq, MeRIPSeq, PA-m<sup>6</sup>A-Seq, m<sup>6</sup>A-CLIP and m<sup>6</sup>A individual-nucleotide resolution cross-linking and immunoprecipitation (miCLIP) (Dominissini et al., 2012; Meyer et al., 2012; Chen et al., 2015; Ke et al., 2015; Linder et al., 2015). MeRIP-seq was first identified in 2012, allowing for m<sup>6</sup>A analysis with 100- to 200-nucleotide resolution. In terms of MeRIP-seq, mRNA was fragmented into 100-nucleotide lengths

TABLE 1 Approaches for the detection of m<sup>6</sup>A RNA methylation.

| Approaches                 | Principle                         | Advantages                   | Limitations   | References                  |
|----------------------------|-----------------------------------|------------------------------|---|-----------------------------|
| m <sup>6</sup> A dot blot  | Antibody based immunoblot         | Easy, fast                   | Low sensitive, semi-quantitative                    | Nagarajan et al. (2019)     |
| HPLC-MS/MS                 | Mass spectrum                     | High sensitive, quantitative | Lack of RNA sequence and localization               | Thuring et al. (2017)       |
| SCARLET                    | Thin-layer chromatography         | Quantitative                 | Complicated, low-throughput                         | Liu and Pan, (2016)         |
| m <sup>6</sup> A-Seq       | Antibody based sequencing         | High-throughput              | 100–200 nucleotide resolution                       | Dominissini et al. (2012)   |
| MeRIPSeq                   | Antibody based sequencing         | High-throughput              | 100–200 nucleotide resolution, antibody specificity | Meyer et al. (2012)         |
| miCLIP                     | Antibody based sequencing         | High-throughput, single site | Low cross-linking efficiency, antibody specificity  | Linder et al. (2015)        |
| MAZTER-Seq                 | Endoribonuclease based sequencing | High-throughput, single site | Preference of the enzyme                            | Garcia-Campos et al. (2019) |
| m <sup>6</sup> A-REF-Seq   | Endoribonuclease based sequencing | High-throughput, single site | Preference of the enzyme                            | Zhang Z. et al. (2019)      |
| m <sup>6</sup> A-SEAL      | Chemical labeling sequencing      | High-throughput, single site | Chemical labeling efficiency                        | Wang Y. et al. (2020)       |
| m <sup>6</sup> A-label-seq | Chemical labeling sequencing      | High-throughput, single site | Chemical labeling efficiency                        | Shu et al. (2020)           |
| m <sup>6</sup> A-SAC-seq   | Chemical labeling sequencing      | High-throughput, single site | Chemical labeling efficiency                        | Hu et al. (2022)            |



and immunoprecipitated by a m<sup>6</sup>A-specific antibody with the combination of high-throughput deep sequencing. miCLIP enables the detection of m<sup>6</sup>A residues at precise positions with single-nucleotide resolution. However, the number of identified m<sup>6</sup>A peaks is limited due to the low cross-linking efficiency of this method. m<sup>6</sup>A-specific antibody-based sequencing approaches have some obvious drawbacks. m<sup>6</sup>A antibodies are not strikingly specific for m<sup>6</sup>A, may also bind to m<sup>6</sup>Am sites and other non-m<sup>6</sup>A-specific sequences. Distinct commercial m<sup>6</sup>A antibodies show differences in affinity for m<sup>6</sup>A (Hausmann et al., 2016). Endoribonuclease-based techniques include antibody-free m<sup>6</sup>A sequencing methods such as

m<sup>6</sup>A-REF-Seq and MAZTER-Seq and rely on the endoribonuclease activity of MazF (Zhang Z. et al., 2019; Garcia-Campos et al., 2019). Therefore, the motif preference of endoribonuclease determines the limitation, and these methods detect a portion of the m<sup>6</sup>A sites. Recently, chemical labelling strategies have been reported, including m<sup>6</sup>A-SEAL, m<sup>6</sup>A-label-seq and m<sup>6</sup>A-SAC-seq (Wang Y. et al., 2020; Shu et al., 2020; Hu et al., 2022). Nevertheless, improvements in the chemical labelling efficiency are needed. A specific method for m<sup>6</sup>A detection in the global transcriptome should be validated by other techniques to obtain a more accurate m<sup>6</sup>A landscape.

## 4 Role of $N^6$ -methyladenosine RNA methylation in human diseases

$m^6A$  RNA methylation is involved in multiple RNA metabolism pathways, such as splicing, localization, translation efficiency, stability and degradation, ultimately affecting various physiological and pathological processes. For instance,  $m^6A$  methylation has been demonstrated to regulate the haematopoietic system, the central nervous system, immunity stemness, mammalian spermatogenesis and brain development. The dysregulation of  $m^6A$  methylation is associated with various diseases, especially cancer (Figure 3).

### 4.1 $N^6$ -methyladenosine and cancer

#### 4.1.1 $N^6$ -methyladenosine in cancer stem cells

Cancer stem cells (CSCs) are a type of cells possessing a stem cell-like capacity to self-renew, differentiate and survive and give rise to many types of cancers (Nassar and Blanpain, 2016). CSCs lead to the tolerance of standard therapeutics and CSCs, tumour recurrence, and distant metastasis (Walcher et al., 2020). Cui et al. found that  $m^6A$  methylation played a role in the tumorigenesis of glioma stem cells (GSCs). *Mettl3/14* knockdown prominently promoted GSC self-renewal and tumorigenesis by decreasing  $m^6A$  levels, whereas *METTL3* overexpression exerted negative effects (Cui et al., 2017; Zhang et al., 2017). MeRIP-seq revealed that silencing *Mettl3/14* altered  $m^6A$  enrichment and that  $m^6A$  RNA methylation of *ADAM19* regulated GSC self-renewal. Similarly, Visvanathan et al. confirmed that *METTL3* was essential for glioma cell differentiation and GSC maintenance. Furthermore, *METTL3* is also involved in radiosensitivity and DNA repair through the *SOX2* axis in GSCs (Visvanathan et al., 2018).

Wang et al. provided the first evidence of the interplay between  $m^6A$  methylation and osteosarcoma stem cells (OSCs). Compared with non-OSCs, *METTL14* and *FTO* were significantly reduced in OSCs. Meanwhile, MeRIP-seq and RNA-seq analyses of OSCs and non-OSCs revealed that differentially expressed genes containing differentially methylated  $m^6A$  peaks were associated with the Wnt pathway and the pluripotency of stem cells (Wang et al., 2019).

Ly P Vu and others revealed that *METTL3* expression was elevated in leukaemia cells compared with normal haematopoietic cells. miCLIP analysis showed that *METTL3* enhanced the  $m^6A$  methylation of target genes such as *Bcl2*, *Myc* and *Pten* in the human acute myeloid leukaemia (AML) MOLM-13 cell line, which promoted the mRNA translation of these genes, thereby retaining pluripotency properties and inhibiting cell differentiation. Notably, silencing of *Mettl3* in human myeloid leukemia cell lines promoted cell differentiation and cell apoptosis (Vu et al.,

2017). *WTAP* is associated with haematopoietic stem cell (HSC) homeostasis and haematopoietic regeneration (Wang H. et al., 2018; Li et al., 2018). Recent research conducted by Bansal et al. suggested a role for  $m^6A$  modification in myeloid leukaemia. *WTAP* expression was increased in AML cells derived from patients with AML, while knockdown of *Wtap* led to the repression of cell proliferation, the activation of cell differentiation and apoptosis in a leukaemia cell line (Bansal et al., 2014) (Figure 3).

#### 4.1.2 $N^6$ -methyladenosine in cancer cell proliferation and migration

Numerous reports have elucidated that  $m^6A$  modification is involved in cancer cell proliferation and tumour metastasis in different types of cancers. Recent findings revealed that *METTL14* exerted an oncogenic function by increasing the expression levels of targets such as *MYB* and *MYC* in AML, while *SPI* downregulated the expression level of *MEETL14* (Weng et al., 2018). *FTO* demethylase is elevated and plays an oncogenic role in AML. It has been demonstrated that a high level of *FTO* prompts cell proliferation and viability, whereas it reduces cell apoptosis and global  $m^6A$  methylation by repressing the expression of *ASB2* and *RARA* (Li Z. et al., 2017). Su et al. observed that R-2-hydroxyglutarate (R-2HG) attenuated *FTO* activity and augmented global  $m^6A$  modification in R-2HG-sensitive AML cells, which decreased the stability of *Cebpa/Myc* mRNA and the activities of relevant cell signalling pathways (Su et al., 2018).

In bladder cancer, *METTL3* interacts with *DGCR8* to facilitate pri-miR221/222 maturation, which leads to a decrease in *PTEN* and ultimately promotes cell proliferation (Han et al., 2019). Similarly, another study showed that *METTL3* was significantly elevated in bladder cancer, and *METTL3* knockdown dramatically suppressed cancer cell proliferation, cell invasion, and tumour formation through the *AFF4/MYC/NF- $\kappa$ B* axis cell signalling pathway (Cheng et al., 2019). Xie et al. revealed that *METTL3* also binds to *YTHDF2*, which induces the degradation of target tumour suppressor mRNAs, including *Klf4* and *Setd7*, regulating the progression of bladder cancer (Xie et al., 2020). Nevertheless, *METTL14* has been confirmed to be decreased in bladder cancer, and depletion of *Mettl14* accelerated cell proliferation, tumour metastasis and self-renewal by decreasing the stability of  $m^6A$ -modified *Notch1* transcripts (Gu et al., 2019).

Evidence indicates that the *hnRNPA1* expression level is elevated by *EGFRvIII*, leading to increased glycolytic gene expression in gliomas. Meanwhile, *hnRNPA1* promotes *Max* mRNA splicing and then induces the generation of Delta *Max*, which promotes *Myc*-dependent cell transformation (Babic et al., 2013). Another study showed that silencing *hnRNPA2* represses cancerous cell viability, cell invasion, tumour metastasis and chemoresistance by decreasing the expression of *MMP-2* and phospho-*STAT3*. Notably, *hnRNPA2* has been regarded as an

oncogenic driver in gliomas (Deng et al., 2016) (Figure 3). Numerous studies have described how m<sup>6</sup>A methylation contributes to cell proliferation, cell invasion, and tumour metastasis in other cancers, including breast cancer, ovarian cancer, cervical cancer, prostate cancer, lung cancer, hepatocellular carcinoma, gastric carcinoma, pancreatic cancer and colorectal cancer (Hu et al., 2019; Yang et al., 2019; Ma et al., 2020; Yang et al., 2020; Guo et al., 2021; Zhang and Xu, Forthcoming 2022).

#### 4.1.3 N<sup>6</sup>-methyladenosine in tumour immunity

Recent studies have revealed that m<sup>6</sup>A RNA methylation induces the activation and infiltration of various immune cells into the tumour microenvironment (TME), influencing the efficacy of cancer immunotherapy. Macrophages are closely associated with tumour initiation and progression. Yin et al. elucidated that METTL3 in macrophages regulates tumour development. Silencing of *Mettl3* in macrophages facilitated tumour growth and lung metastasis. The TME was reshaped by inducing regulatory T (Treg) cells into tumour sites and promoting the infiltration of M1-and M2-like tumour-associated macrophages (Yin et al., 2021). *Mettl14* knockdown in macrophages suppressed the antitumour activity of CD8<sup>+</sup> T cells and improved tumour growth (Dong et al., 2021).

Natural killer (NK) cells play an important role in cancer immune surveillance and can directly recognize and kill cancer cells. YTHDF2 was critical for modulating NK-cell maturation, NK-cell homeostasis, IL-15-mediated survival, and antitumor activity due to the regulation of downstream target genes such as *Stat5*, *Eomes* and *Tarbp* (Ma et al., 2021). Song et al. found that METTL3 expression was decreased in tumour-infiltrating NK cells of cancer patients. In mice, they observed that depletion of *Mettl3* enhanced NK-cell responsiveness to IL-15 and promoted tumour progression and metastasis by targeting SHP-2 (Song et al., 2021).

Silencing of *Mettl3* in CD4<sup>+</sup> T cells destroyed T-cell differentiation and homeostasis by repressing the activation of IL-7-mediated STAT5/suppressor of cytokine signalling (Li H. B. et al., 2017). Yao et al. revealed that conditional depletion of *Mettl3* in CD4<sup>+</sup> T cells inhibited T follicular helper differentiation and maturation, thereby preventing the antibody response of B cells by promoting the degradation of m<sup>6</sup>A-modified *Tcf7* mRNA (Yao et al., 2021). In breast cancer, the expression levels of METTL14 have a positive correlation with the infiltration of CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells, dendritic cells, macrophages and neutrophils, but they negatively correlated with Treg cells in breast cancer (Gong et al., 2020) (Figure 3).

## 4.2 N<sup>6</sup>-methyladenosine and other human diseases

Emerging evidence has demonstrated that m<sup>6</sup>A RNA methylation is closely related to other human diseases, including

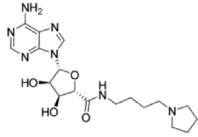
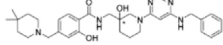
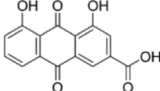
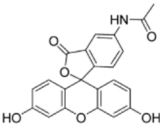
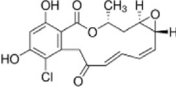
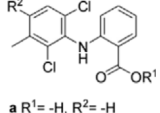
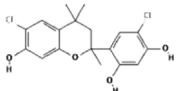
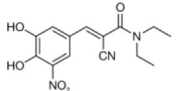
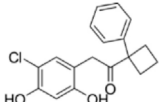
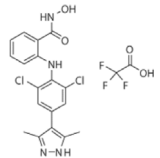
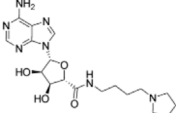
cardiovascular disease, metabolic syndrome, psychiatric disorders and autoinflammatory disorders. Dorn et al. (Dorn et al., 2019) showed that the global m<sup>6</sup>A level of cardiomyocytes was significantly elevated in response to hypertrophic stimulation and that METTL3 played a vital role in cardiomyocyte hypertrophy. Notably, upregulated m<sup>6</sup>A modification resulted in compensated cardiac hypertrophy, whereas downregulated m<sup>6</sup>A levels led to eccentric cardiomyocyte remodelling and dysfunction. Overexpression of *Mettl3* increased the expression levels of mitogen-activated protein (MAP)3K6, MAP4K5, and MAPK14 in cardiomyocytes, which was positively correlated with cardiomyocyte size, revealing that METTL3 is sufficient to drive cardiomyocyte hypertrophy. However, no histopathologic changes were observed in *Mettl3*-overexpressing mice.

Zhou et al. identified that YTHDC2 was significantly repressed in nonalcoholic fatty liver disease (NAFLD) patients, and *Ythdc2*-depleted hepatocytes led to the accumulation of excessive triglycerides (TGs) by reducing the expression levels of lipogenic genes, including fatty acid synthase, sterol regulatory element-binding protein 1c, and acetyl-CoA carboxylase 1 (Zhou B. et al., 2021). Furthermore, m<sup>6</sup>A sequencing was performed in human type 2 diabetes islets, and sequencing analysis showed that multiple hypomethylated transcripts were associated with insulin secretion, the insulin/IGF1 signalling pathway and cell cycle progression. *Mettl14* knockout in mouse  $\beta$ -cells caused a reduction in global m<sup>6</sup>A levels, giving rise to a similar islet phenotype in human T2D (De Jesus et al., 2019).

In Alzheimer's disease mouse models, the global m<sup>6</sup>A level was increased in the hippocampus and the cortex compared to C57BL/6 mice, and the interaction of FTO and APOE contributed to the increase in Alzheimer's disease risk. The overall m<sup>6</sup>A level was elevated in the cortex and the hippocampus of APP/PS1 (Alzheimer's disease) mice compared to C57BL/6 control mice, and FTO was found to interact with APOE, which was associated with Alzheimer's disease risk in a prospective cohort study (Keller et al., 2011; Han et al., 2020). Huang et al. found that ALKBH5/FAAH enhanced the expression of circSTAG1, which attenuated astrocyte dysfunction and depressive-like behaviours *in vitro* and *in vivo* (Huang et al., 2020).

m<sup>6</sup>A methylation also has a contributory effect on autoinflammatory disorders. Luo et al. identified that the expression levels of METTL14, ALKBH5 and YTHDF2 were downregulated in peripheral blood mononuclear cells of systemic lupus erythematosus (SLE) patients (Luo et al., 2020). In addition, multivariate logistic regression analysis showed that repression of ALKBH5 and YTHDF2 was considered a risk factor for SLE. However, direct mechanistic data should be provided for the function of m<sup>6</sup>A modification in SLE progression. Another study found that *Mettl14* knockdown inhibited the activation of Treg cells, which impaired the balance between Th17 and Treg cells, leading to the development of spontaneous colitis (Lu et al., 2020).

TABLE 2 Overview of the small-molecule drugs and m<sup>6</sup>A-related factor inhibitors described in the text.

| Drug/inhibitor         | Molecular structure  | m <sup>6</sup> A proteins involved | Function         | References                                |
|------------------------|--|------------------------------------|------------------|---|
| SAM mimic              |   | METTL3                             | METTL3 inhibitor | <a href="#">Bedi et al. (2020)</a>        |
| UZH1a                  |   | METTL3                             | METTL3 inhibitor | <a href="#">Moroz-Omori et al. (2021)</a> |
| Rhein                  |   | FTO                                | FTO inhibitor    | <a href="#">Li Q. et al. (2016)</a>       |
| Fluorescein derivative |   | FTO                                | FTO inhibitor    | <a href="#">Wang et al. (2015)</a>        |
| Radicicol              |   | FTO                                | FTO inhibitor    | <a href="#">Wang R. et al. (2018)</a>     |
| Meclofenamic acid (MA) | <br>a R <sup>1</sup> = -H, R <sup>2</sup> = -H | FTO                                | FTO inhibitor    | <a href="#">Huang et al. (2015)</a>       |
| CHTB                   |   | FTO                                | FTO inhibitor    | <a href="#">Qiao et al. (2016)</a>        |
| Entacapone             |   | FTO                                | FTO inhibitor    | <a href="#">Peng et al. (2019)</a>        |
| N-CDPCB                |   | FTO                                | FTO inhibitor    | <a href="#">Qiao et al. (2017)</a>        |
| Dac51                  |   | FTO                                | FTO inhibitor    | <a href="#">Liu et al. (2021)</a>         |
| MV1035                 |   | ALKBH5                             | ALKBH5 inhibitor | <a href="#">Malacrida et al. (2020)</a>   |

## 5 Potential clinical applications of $N^6$ -methyladenosine RNA modification

Due to the major role of  $m^6A$  RNA modification in tumour and other disease progression,  $m^6A$ -associated proteins can be developed as potential therapeutic targets for tumours and other diseases (Table 2). METTL3 attenuates the sensitivity of colon cancer cells to chemotherapy of L-OHP and CPT-11 by upregulating the expression level of CBX8 in a  $m^6A$ -dependent manner (Zhang Y. et al., 2019). Meanwhile, *Ythdf1*-depleted colon cancer cells are more sensitive to 5-FU and L-OHP (Nishizawa et al., 2018). Another study identified that gemcitabine drove the apoptosis of pancreatic cancer cells with low METTL3 expression (Taketo et al., 2018). The combination of  $m^6A$  methylation and chemotherapeutic drugs contributes to resolving drug resistance in tumours. Cas13-directed methyltransferase has been used for cancer treatment, targeting  $m^6A$  of specific RNA loci (Wilson et al., 2020; Lo et al., 2022).

Specific inhibitors based on  $m^6A$ -related enzymes have been studied. Bedi et al. (Bedi et al., 2020) performed a virtual screening assay to identify potential METTL3 inhibitors from 4,000 adenosine derivatives. One compound, a SAM mimic, was found to be the first inhibitor of METTL3. However, the therapeutic value of this compound is still somewhat limited due to cell penetration issues and nonspecific targets. They also identified another METTL3 inhibitor, UZH1a, that decreased the  $m^6A/A$  ratios in three different cell lines, MOLM-13, HEK293T and U2OS, but the specificity of UZH1a needs further improvement (Moroz-Omori et al., 2021).

Currently, more papers are focused on FTO inhibitors. Rhein, a competitive inhibitor of FTO, binds to the catalytic domain and blocks the catalytic activity of FTO (Li Q. et al., 2016). Fluorescein derivatives, radicicol, IOX3 and MA (a nonsteroidal anti-inflammatory drug) were subsequently identified to repress FTO expression (Huang et al., 2015; Wang et al., 2015; Wang R. et al., 2018). Similar to MA, MA2, an isomer of MA, interacts with FTO to increase the  $m^6A$  level but possesses better cell penetration ability. In addition, CHTB, entacapone and N-CDPCB also repress the demethylase activity of FTO (Qiao et al., 2016; Qiao et al., 2017; Peng et al., 2019). Dac51 inhibits the activity of FTO, promotes T cell response and enhances the anti-PD-1 therapy (Liu et al., 2021). Small-molecule FTO inhibitors have been developed as potential drugs not only for tumour therapy but also for the treatment of neurological diseases, cardiovascular disease, metabolic syndrome and autoinflammatory disorders. To date, little is known about specific ALKBH5 inhibitors. However, the imidazobenzoxazin-5-thione MV1035 was found to be a potential candidate ALKBH5 inhibitor (Malacrida et al., 2020). ALK-04 as an inhibitor of ALKBH5 reduces the infiltration of myeloid-derived suppressor cells and Treg cells and suppresses tumour

growth by enhancing the therapeutic effect of anti-PD-1 treatment (Li et al., 2020). (Table 2).

## 6 Conclusions and future prospects

Despite the initial discovery of  $m^6A$  in 1974,  $m^6A$  modification has received much less attention than histone and DNA epigenetic modifications. Since the establishment of MeRIP-Seq for mapping  $m^6A$  deposition at a transcriptome-wide level in 2012, interest in studying  $m^6A$  methylation has grown rapidly. The level of  $m^6A$  is dynamically and reversibly regulated by various effectors termed “writers”, “erasers” and “readers”. The METTL3-14-WTAP methyltransferase complex is the core constituent of most  $m^6A$  writers, whereas FTO and ALKBH5, as  $m^6A$  demethylases, catalyse the removal of  $m^6A$ . Currently, emerging evidence has shown that transcription factor and histone modification signatures together shape  $m^6A$  deposition, suggesting that the contribution of transcription factor histone modification contributes to the modulation of  $m^6A$ . The interplay between histone H3K36me3 and  $m^6A$  modification provides a novel layer of gene expression regulation. Further crosstalk between  $m^6A$  RNA modification and other epigenetic modifications should be carefully elucidated in the future. Nutritional metabolism and metabolites also influence the regulation of  $m^6A$ . It is possible that tissue-specific  $m^6A$  levels partially correlate with the metabolic activities of specific organs. Overall, the precise regulatory mechanisms of  $m^6A$  are still in their infancy and need to be further investigated. Although high-throughput sequencing-based techniques have greatly promoted the research field of  $m^6A$  methylation, none of the available approaches have simultaneously reached the achievements with single-base resolution, quantification of  $m^6A$  disposition, and low input of RNA. A future method for the direct sequencing of  $m^6A$  RNA will be better for illustrating its functions and dynamics *in vivo*.

As mentioned above, abnormal  $m^6A$  methylation is closely related to various diseases, including cancer, cardiovascular disease, metabolic syndrome, psychiatric disorders and autoinflammatory disorders. Accordingly,  $m^6A$  methylation is a double-edged sword for tumour: a lack of  $m^6A$  modification on specific genes accelerates tumour development, while over- $m^6A$  modification of other genes also induces tumour progression. In fact, many major challenges remain in elucidating the relationships between  $m^6A$  and cancer. First, whether the multiple roles proposed in cancer actually rely on  $m^6A$  modification should be considered. Second, we should not ignore the notion that  $m^6A$ -related effectors may mediate tumour development and the progression in a  $m^6A$ -independent manner. Third, it is worth noting that we elucidate the regulatory association of noncoding RNAs and  $m^6A$  methylation in tumour. Fourth, the effects of  $m^6A$  regulators on tumor cells and immune cells are complicated and need to be

carefully concerned. Fifth, over 170 RNA modifications have been identified, it is worth evaluating whether other RNA modifications in the same RNA transcripts affects the role of m<sup>6</sup>A methylation in human diseases. Finally, further studies should be carried out to assess the clinical value of m<sup>6</sup>A in diseases.

Many studies have demonstrated that m<sup>6</sup>A regulatory factors are suitable as therapeutic targets, and some inhibitors of m<sup>6</sup>A-related factors, especially FTO, have been discovered. Nevertheless, no clinical trials using m<sup>6</sup>A inhibitors for the treatment of cancer and other diseases have been reported yet. We also provide several reasonable strategies for driving m<sup>6</sup>A-based therapy: 1) Nanoparticles can specifically deliver m<sup>6</sup>A modification molecules to target immune cells for tumour immunotherapeutic treatment. 2) A programmable m<sup>6</sup>A gene-editing system by dCas13 or dRCas9 provides a potential tool for the treatment of diseases. 3) Chimeric antigen receptor (CAR) immune cells with lentivirus-mediated gene delivery of m<sup>6</sup>A effectors are beneficial to cancer immunotherapy. 4) It is feasible to treat cancer and other diseases by using a m<sup>6</sup>A inhibitor combined with other therapies. Further studies are urgently required for the understanding of RNA m<sup>6</sup>A modifications and clinical applications.

## Author contributions

JL and JS made substantial contributions to conception and design. PL involved in drafting and revising the manuscript, YW

revised the manuscript and draw the figures, YS drafted the tables.

## Funding

This work was supported by the National Natural Science Foundation of China (NSFC 82104289), Weifang Science and Technology Bureau Plan Project (2021YX081), Science and Technology Innovation Plan from Weifang Medical College (041004), Yuandu Scholar Grant of Weifang City to JL and Postdoctoral Research Grant of Shenzhen to LPP (2237PT).

## 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.

## References

- Aguilo, F., Zhang, F., Sancho, A., Fidalgo, M., Di Cecilia, S., Vashisht, A., et al. (2015). Coordination of m(6)A mRNA methylation and gene transcription by ZFP217 regulates pluripotency and reprogramming. *Cell Stem Cell* 17 (6), 689–704. doi:10.1016/j.stem.2015.09.005
- Babic, I., Anderson, E. S., Tanaka, K., Guo, D., Masui, K., Li, B., et al. (2013). EGFR mutation-induced alternative splicing of Max contributes to growth of glycolytic tumors in brain cancer. *Cell Metab.* 17 (6), 1000–1008. doi:10.1016/j.cmet.2013.04.013
- Bansal, H., Yihua, Q., Iyer, S. P., Ganapathy, S., Proia, D. A., Penalba, L. O., et al. (2014). WTAP is a novel oncogenic protein in acute myeloid leukemia. *Leukemia* 28 (5), 1171–1174. doi:10.1038/leu.2014.16
- Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millan-Zambrano, G., Robson, S. C., et al. (2017). Promoter-bound METTL3 maintains myeloid leukaemia by m(6)A-dependent translation control. *Nature* 552 (7683), 126–131. doi:10.1038/nature24678
- Bawankar, P., Lence, T., Paolantoni, C., Haussmann, I. U., Kazlauskienė, M., Jacob, D., et al. (2021). Hakai is required for stabilization of core components of the m(6)A mRNA methylation machinery. *Nat. Commun.* 12 (1), 3778. doi:10.1038/s41467-021-23892-5
- Bayoumi, M., Rohaim, M. A., and Munir, M. (2020). Structural and virus regulatory insights into avian N6-methyladenosine (m6A) machinery. *Front. Cell Dev. Biol.* 8, 543. doi:10.3389/fcell.2020.00543
- Bedi, R. K., Huang, D., Eberle, S. A., Wiedmer, L., Sledz, P., and Caflisch, A. (2020). Small-molecule inhibitors of METTL3, the major human epitranscriptomic writer. *ChemMedChem* 15 (9), 744–748. doi:10.1002/cmdc.202000011
- Bertero, A., Brown, S., Madrigal, P., Osnato, A., Ortmann, D., Yiangou, L., et al. (2018). The SMAD2/3 interactome reveals that TGFβ controls m6A mRNA methylation in pluripotency. *Nature* 555 (7695), 256–259. doi:10.1038/nature25784
- Bi, Z., Liu, Y., Zhao, Y., Yao, Y., Wu, R., Liu, Q., et al. (2019). A dynamic reversible RNA N(6)-methyladenosine modification: Current status and perspectives. *J. Cell. Physiol.* 234 (6), 7948–7956. doi:10.1002/jcp.28014
- Bokar, J. A., Rath-Shambaugh, M. E., Ludwiczak, R., Narayan, P., and Rottman, F. (1994). Characterization and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. *J. Biol. Chem.* 269 (26), 17697–17704. doi:10.1016/s0021-9258(17)32497-3
- Cai, J., Chen, Z., Zhang, Y., Wang, J., Zhang, Z., Wu, J., et al. (2022). CircRHBD1 augments metabolic rewiring and restricts immunotherapy efficacy via m(6)A modification in hepatocellular carcinoma. *Mol. Ther. Oncolytics* 24, 755–771. doi:10.1016/j.omto.2022.02.021
- Chen, K., Luo, G. Z., and He, C. (2015). High-resolution mapping of N<sup>6</sup>-Methyladenosine in transcriptome and genome using a photo-crosslinking-assisted strategy. *Methods Enzymol.* 560, 161–185. doi:10.1016/bs.mie.2015.03.012
- Cheng, M., Sheng, L., Gao, Q., Xiong, Q., Zhang, H., Wu, M., et al. (2019). The m6A methyltransferase METTL3 promotes bladder cancer progression via AFF4/NF-κB/MYC signaling network. *Oncogene* 38 (19), 3667–3680. doi:10.1038/s41388-019-0683-z
- Cui, Q., Shi, H., Ye, P., Li, L., Qu, Q., Sun, G., et al. (2017). m(6)A RNA methylation regulates the self-renewal and tumorigenesis of glioblastoma stem cells. *Cell Rep.* 18 (11), 2622–2634. doi:10.1016/j.celrep.2017.02.059

- Dai, F., Wu, Y., Lu, Y., An, C., Zheng, X., Dai, L., et al. (2020). Crosstalk between RNA m(6)A modification and non-coding RNA contributes to cancer growth and progression. *Mol. Ther. Nucleic Acids* 22, 62–71. doi:10.1016/j.omtn.2020.08.004
- De Jesus, D. F., Zhang, Z., Kahraman, S., Brown, N. K., Chen, M., Hu, J., et al. (2019). m(6)A mRNA methylation regulates human beta-cell Biology in physiological states and in type 2 diabetes. *Nat. Metab.* 1 (8), 765–774. doi:10.1038/s42255-019-0089-9
- Deng, J., Chen, S., Wang, F., Zhao, H., Xie, Z., Xu, Z., et al. (2016). Effects of hnRNP A2/B1 knockdown on inhibition of glioblastoma cell invasion, growth and survival. *Mol. Neurobiol.* 53 (2), 1132–1144. doi:10.1007/s12035-014-9080-3
- Deng, X., Chen, K., Luo, G. Z., Weng, X., Ji, Q., Zhou, T., et al. (2015). Widespread occurrence of N6-methyladenosine in bacterial mRNA. *Nucleic Acids Res.* 43 (13), 6557–6567. doi:10.1093/nar/gkv596
- Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U. S. A.* 71 (10), 3971–3975. doi:10.1073/pnas.71.10.3971
- Ding, L., Li, J., Song, B., Xiao, X., Zhang, B., Qi, M., et al. (2016). Curcumin rescues high fat diet-induced obesity and insulin sensitivity in mice through regulating SREBP pathway. *Toxicol. Appl. Pharmacol.* 304, 99–109. doi:10.1016/j.taap.2016.05.011
- Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* 485 (7397), 201–206. doi:10.1038/nature11112
- Dong, L., Chen, C., Zhang, Y., Guo, P., Wang, Z., Li, J., et al. (2021). The loss of RNA N(6)-adenosine methyltransferase Mettl14 in tumor-associated macrophages promotes CD8(+) T cell dysfunction and tumor growth. *Cancer Cell* 39 (7), 945–957. e910. doi:10.1016/j.ccell.2021.04.016
- Dorn, L. E., Lasman, L., Chen, J., Xu, X., Hund, T. J., Medvedovic, M., et al. (2019). The N(6)-methyladenosine mRNA methylase METTL3 controls cardiac homeostasis and hypertrophy. *Circulation* 139 (4), 533–545. doi:10.1161/CIRCULATIONAHA.118.036146
- Duncan, T. M., Reed, M. C., and Nijhout, H. F. (2013). The relationship between intracellular and plasma levels of folate and metabolites in the methionine cycle: A model. *Mol. Nutr. Food Res.* 57 (4), 628–636. doi:10.1002/mnfr.201200125
- Edens, B. M., Vissers, C., Su, J., Arumugam, S., Xu, Z., Shi, H., et al. (2019). FMRP modulates neural differentiation through m(6)a-dependent mRNA nuclear export. *Cell Rep.* 28 (4), 845–854. doi:10.1016/j.celrep.2019.06.072
- Feng, C., Liu, Y., Wang, G., Deng, Z., Zhang, Q., Wu, W., et al. (2014). Crystal structures of the human RNA demethylase Alkbh5 reveal basis for substrate recognition. *J. Biol. Chem.* 289 (17), 11571–11583. doi:10.1074/jbc.M113.546168
- Garcia-Campos, M. A., Edelheit, S., Toth, U., Safra, M., Shachar, R., Viukov, S., et al. (2019). Deciphering the “m(6)A code” via antibody-independent quantitative profiling. *Cell* 178 (3), 731–747. e716. doi:10.1016/j.cell.2019.06.013
- Gong, P. J., Shao, Y. C., Yang, Y., Song, W. J., He, X., Zeng, Y. F., et al. (2020). Analysis of N6-methyladenosine methyltransferase reveals METTL14 and ZC3H13 as tumor suppressor genes in breast cancer. *Front. Oncol.* 10, 578963. doi:10.3389/fonc.2020.578963
- Gu, C., Wang, Z., Zhou, N., Li, G., Kou, Y., Luo, Y., et al. (2019). Mettl14 inhibits bladder TIC self-renewal and bladder tumorigenesis through N(6)-methyladenosine of Notch1. *Mol. Cancer* 18 (1), 168. doi:10.1186/s12943-019-1084-1
- Guo, J., Zheng, J., Zhang, H., and Tong, J. (2021). RNA m6A methylation regulators in ovarian cancer. *Cancer Cell Int.* 21 (1), 609. doi:10.1186/s12935-021-02318-8
- Han, J., Wang, J. Z., Yang, X., Yu, H., Zhou, R., Lu, H. C., et al. (2019). METTL3 promote tumor proliferation of bladder cancer by accelerating pri-miR221/222 maturation in m6A-dependent manner. *Mol. Cancer* 18 (1), 110. doi:10.1186/s12943-019-1036-9
- Han, M., Liu, Z., Xu, Y., Liu, X., Wang, D., Li, F., et al. (2020). Abnormality of m6A mRNA methylation is involved in alzheimer's disease. *Front. Neurosci.* 14, 98. doi:10.3389/fnins.2020.00098
- Hausmann, I. U., Bodi, Z., Sanchez-Moran, E., Mongan, N. P., Archer, N., Fray, R. G., et al. (2016). m(6)A potentiates Sxl alternative pre-mRNA splicing for robust *Drosophila* sex determination. *Nature* 540 (7632), 301–304. doi:10.1038/nature20577
- He, P. C., and He, C. (2021). m(6) A RNA methylation: from mechanisms to therapeutic potential. *EMBO J.* 40 (3), e105977. doi:10.15252/embj.2020105977
- Hu, B. B., Wang, X. Y., Gu, X. Y., Zou, C., Gao, Z. J., Zhang, H., et al. (2019). N(6)-methyladenosine (m(6)A) RNA modification in gastrointestinal tract cancers: Roles, mechanisms, and applications. *Mol. Cancer* 18 (1), 178. doi:10.1186/s12943-019-1099-7
- Hu, L., Liu, S., Peng, Y., Ge, R., Su, R., Senevirathne, C., et al. (2022). m(6)A RNA modifications are measured at single-base resolution across the mammalian transcriptome. *Nat. Biotechnol.* 40, 1210–1219. doi:10.1038/s41587-022-01243-z
- Huang, H., Weng, H., Sun, W., Qin, X., Shi, H., Wu, H., et al. (2018). Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat. Cell Biol.* 20 (3), 285–295. doi:10.1038/s41556-018-0045-z
- Huang, H., Weng, H., Zhou, K., Wu, T., Zhao, B. S., Sun, M., et al. (2019). Histone H3 trimethylation at lysine 36 guides m(6)A RNA modification co-transcriptionally. *Nature* 567 (7748), 414–419. doi:10.1038/s41586-019-1016-7
- Huang, R., Zhang, Y., Bai, Y., Han, B., Ju, M., Chen, B., et al. (2020). N(6)-Methyladenosine modification of fatty acid amide hydrolase messenger RNA in circular RNA STAG1-regulated astrocyte dysfunction and depressive-like behaviors. *Biol. Psychiatry* 88 (5), 392–404. doi:10.1016/j.biopsych.2020.02.018
- Huang, Y., Yan, J., Li, Q., Li, J., Gong, S., Zhou, H., et al. (2015). Meclofenamic acid selectively inhibits FTO demethylation of m6A over ALKBH5. *Nucleic Acids Res.* 43 (1), 373–384. doi:10.1093/nar/gku1276
- Kang, H., Zhang, Z., Yu, L., Li, Y., Liang, M., and Zhou, L. (2018). FTO reduces mitochondria and promotes hepatic fat accumulation through RNA demethylation. *J. Cell. Biochem.* 119 (7), 5676–5685. doi:10.1002/jcb.26746
- Kasowitz, S. D., Ma, J., Anderson, S. J., Leu, N. A., Xu, Y., Gregory, B. D., et al. (2018). Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLoS Genet.* 14 (5), e1007412. doi:10.1371/journal.pgen.1007412
- Ke, S., Alemu, E. A., Mertens, C., Gantman, E. C., Fak, J. J., Mele, A., et al. (2015). A majority of m6A residues are in the last exons, allowing the potential for 3' UTR regulation. *Genes Dev.* 29 (19), 2037–2053. doi:10.1101/gad.269415.115
- Keller, L., Xu, W., Wang, H. X., Winblad, B., Fratiglioni, L., and Graff, C. (2011). The obesity related gene, FTO, interacts with APOE, and is associated with alzheimer's disease risk: A prospective cohort study. *J. Alzheimers Dis.* 23 (3), 461–469. doi:10.3233/JAD-2010-101068
- Knuckles, P., Lence, T., Hausmann, I. U., Jacob, D., Kreim, N., Carl, S. H., et al. (2018). Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m(6)A machinery component Wtap/Fl(2)d. *Genes Dev.* 32 (5-6), 415–429. doi:10.1101/gad.309146.117
- Li, F., Kennedy, S., Hajian, T., Gibson, E., Seitova, A., Xu, C., et al. (2016a). A radioactivity-based assay for screening human m6A-RNA methyltransferase, METTL3-METTL14 complex, and demethylase ALKBH5. *J. Biomol. Screen.* 21 (3), 290–297. doi:10.1177/1087057115623264
- Li, H. B., Tong, J., Zhu, S., Batista, P. J., Duffy, E. E., Zhao, J., et al. (2017a). m(6)A mRNA methylation controls T cell homeostasis by targeting the IL-7/STAT5/SOCS pathways. *Nature* 548 (7667), 338–342. doi:10.1038/nature23450
- Li, N., Kang, Y., Wang, L., Huff, S., Tang, R., Hui, H., et al. (2020). ALKBH5 regulates anti-PD-1 therapy response by modulating lactate and suppressive immune cell accumulation in tumor microenvironment. *Proc. Natl. Acad. Sci. U. S. A.* 117 (33), 20159–20170. doi:10.1073/pnas.1918986117
- Li, Q., Huang, Y., Liu, X., Gan, J., Chen, H., and Yang, C. G. (2016b). Rhein inhibits AlkB repair enzymes and sensitizes cells to methylated DNA damage. *J. Biol. Chem.* 291 (21), 11083–11093. doi:10.1074/jbc.M115.711895
- Li, Z., Qian, P., Shao, W., Shi, H., He, X. C., Gogol, M., et al. (2018). Suppression of m(6)A reader Ythdf2 promotes hematopoietic stem cell expansion. *Cell Res.* 28 (9), 904–917. doi:10.1038/s41422-018-0072-0
- Li, Z., Weng, H., Su, R., Weng, X., Zuo, Z., Li, C., et al. (2017b). FTO plays an oncogenic role in acute myeloid leukemia as a N(6)-methyladenosine RNA demethylase. *Cancer Cell* 31 (1), 127–141. doi:10.1016/j.ccell.2016.11.017
- Linder, B., Grozhik, A. V., Olarerin-George, A. O., Meydan, C., Mason, C. E., and Jaffrey, S. R. (2015). Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* 12 (8), 767–772. doi:10.1038/nmeth.3453
- Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* 10 (2), 93–95. doi:10.1038/nchembio.1432
- Liu, L., Li, H., Hu, D., Wang, Y., Shao, W., Zhong, J., et al. (2022). Insights into N6-methyladenosine and programmed cell death in cancer. *Mol. Cancer* 21 (1), 32. doi:10.1186/s12943-022-01508-w
- Liu, N., and Pan, T. (2016). Probing N<sup>6</sup>-methyladenosine (m<sup>6</sup>A) RNA modification in total RNA with SCARLET. *Methods Mol. Biol.* 1358, 285–292. doi:10.1007/978-1-4939-3067-8\_17
- Liu, T., Wei, Q., Jin, J., Luo, Q., Liu, Y., Yang, Y., et al. (2020). The m6A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation. *Nucleic Acids Res.* 48 (7), 3816–3831. doi:10.1093/nar/gkaa048

- Liu, Y., Liang, G., Xu, H., Dong, W., Dong, Z., Qiu, Z., et al. (2021). Tumors exploit FTO-mediated regulation of glycolytic metabolism to evade immune surveillance. *Cell Metab.* 33 (6), 1221–1233. e1211. doi:10.1016/j.cmet.2021.04.001
- Lo, N., Xu, X., Soares, F., and He, H. H. (2022). The basis and promise of programmable RNA editing and modification. *Front. Genet.* 13, 834413. doi:10.3389/fgene.2022.834413
- Lu, T. X., Zheng, Z., Zhang, L., Sun, H. L., Bissonnette, M., Huang, H., et al. (2020). A new model of spontaneous colitis in mice induced by deletion of an RNA m(6A) methyltransferase component METTL14 in T cells. *Cell. Mol. Gastroenterol. Hepatol.* 10 (4), 747–761. doi:10.1016/j.jcmgh.2020.07.001
- Luo, Q., Fu, B., Zhang, L., Guo, Y., Huang, Z., and Li, J. (2020). Decreased peripheral blood ALKBH5 correlates with markers of autoimmune response in systemic lupus erythematosus. *Dis. Markers* 2020, 8193895. doi:10.1155/2020/8193895
- Ma, H., Wang, X., Cai, J., Dai, Q., Natchiar, S. K., Lv, R., et al. (2019). N(6-) Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. *Nat. Chem. Biol.* 15 (1), 88–94. doi:10.1038/s41589-018-0184-3
- Ma, S., Yan, J., Barr, T., Zhang, J., Chen, Z., Wang, L. S., et al. (2021). The RNA m6A reader YTHDF2 controls NK cell antitumor and antiviral immunity. *J. Exp. Med.* 218 (8), e20210279. doi:10.1084/jem.20210279
- Ma, X. X., Cao, Z. G., and Zhao, S. L. (2020). m6A methyltransferase METTL3 promotes the progression of prostate cancer via m6A-modified LEF1. *Eur. Rev. Med. Pharmacol. Sci.* 24 (7), 3565–3571. doi:10.26355/eurrev\_202004\_20817
- Malacrida, A., Rivara, M., Di Domizio, A., Cislighi, G., Miloso, M., Zuliani, V., et al. (2020). 3D proteome-wide scale screening and activity evaluation of a new ALKBH5 inhibitor in U87 glioblastoma cell line. *Bioorg. Med. Chem.* 28 (4), 115300. doi:10.1016/j.bmc.2019.115300
- Mao, Y., Dong, L., Liu, X. M., Guo, J., Ma, H., Shen, B., et al. (2019). m(6A) in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat. Commun.* 10 (1), 5332. doi:10.1038/s41467-019-13317-9
- Mar, B. G., Chu, S. H., Kahn, J. D., Krivtsov, A. V., Koche, R., Castellano, C. A., et al. (2017). SETD2 alterations impair DNA damage recognition and lead to resistance to chemotherapy in leukemia. *Blood* 130 (24), 2631–2641. doi:10.1182/blood-2017-03-775569
- Mauer, J., Luo, X., Blanjoie, A., Jiao, X., Grozhik, A. V., Patil, D. P., et al. (2017). Reversible methylation of m(6)Am in the 5' cap controls mRNA stability. *Nature* 541 (7637), 371–375. doi:10.1038/nature21022
- Mauer, J., Sindelar, M., Despic, V., Guez, T., Hawley, B. R., Vasseur, J. J., et al. (2019). FTO controls reversible m(6)Am RNA methylation during snRNA biogenesis. *Nat. Chem. Biol.* 15 (4), 340–347. doi:10.1038/s41589-019-0231-8
- Meyer, K. D., Saitore, Y., Zumbo, P., Elemento, O., Mason, C. E., and Jaffrey, S. R. (2012). Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* 149 (7), 1635–1646. doi:10.1016/j.cell.2012.05.003
- Moroz-Omori, E. V., Huang, D., Kumar Bedi, R., Cheriyanunel, S. J., Bochenkova, E., Dolbois, A., et al. (2021). METTL3 inhibitors for epitranscriptomic modulation of cellular processes. *ChemMedChem* 16 (19), 3035–3043. doi:10.1002/cmdc.202100291
- Nagarajan, A., Janostiak, R., and Wajapeyee, N. (2019). Dot blot analysis for measuring global N(6)-methyladenosine modification of RNA. *Methods Mol. Biol.* 1870, 263–271. doi:10.1007/978-1-4939-8808-2\_20
- Nassar, D., and Blanpain, C. (2016). Cancer stem cells: Basic concepts and therapeutic implications. *Annu. Rev. Pathol.* 11, 47–76. doi:10.1146/annurev-pathol-012615-044438
- Nishizawa, Y., Konno, M., Asai, A., Koseki, J., Kawamoto, K., Miyoshi, N., et al. (2018). Oncogene c-Myc promotes epitranscriptome m(6A) reader YTHDF1 expression in colorectal cancer. *Oncotarget* 9 (7), 7476–7486. doi:10.18632/oncotarget.23554
- Niu, Y., Zhao, X., Wu, Y. S., Li, M. M., Wang, X. J., and Yang, Y. G. (2013). N6-methyl-adenosine (m6A) in RNA: An old modification with a novel epigenetic function. *Genomics Proteomics Bioinforma.* 11 (1), 8–17. doi:10.1016/j.gpb.2012.12.002
- Oerum, S., Meynier, V., Catala, M., and Tisne, C. (2021). A comprehensive review of m6A/m6Am RNA methyltransferase structures. *Nucleic Acids Res.* 49 (13), 7239–7255. doi:10.1093/nar/gkab378
- Patil, D. P., Chen, C. K., Pickering, B. F., Chow, A., Jackson, C., Guttman, M., et al. (2016). m(6A) RNA methylation promotes XIST-mediated transcriptional repression. *Nature* 537 (7620), 369–373. doi:10.1038/nature19342
- Pendleton, K. E., Chen, B., Liu, K., Hunter, O. V., Xie, Y., Tu, B. P., et al. (2017). The U6 snRNA m(6A) methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* 169 (5), 824–835. e814. doi:10.1016/j.cell.2017.05.003
- Peng, S., Xiao, W., Ju, D., Sun, B., Hou, N., Liu, Q., et al. (2019). Identification of entacapone as a chemical inhibitor of FTO mediating metabolic regulation through FOXO1. *Sci. Transl. Med.* 11 (488), eaau7116. doi:10.1126/scitranslmed.aau7116
- Ping, X. L., Sun, B. F., Wang, L., Xiao, W., Yang, X., Wang, W. J., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. *Cell Res.* 24 (2), 177–189. doi:10.1038/cr.2014.3
- Qiao, Y., Yang, Q., Song, C., and Chang, J. (2017). Computational insights into the origin of decrease/increase in potency of N-CDPCB analogues toward FTO. *J. Biomol. Struct. Dyn.* 35 (8), 1758–1765. doi:10.1080/07391102.2016.1193445
- Qiao, Y., Zhou, B., Zhang, M., Liu, W., Han, Z., Song, C., et al. (2016). A novel inhibitor of the obesity-related protein FTO. *Biochemistry* 55 (10), 1516–1522. doi:10.1021/acs.biochem.6b00023
- Saitore, Y., Meyer, K., Korlach, J., Vilfan, I. D., Jaffrey, S., and Mason, C. E. (2012). The birth of the epitranscriptome: Deciphering the function of RNA modifications. *Genome Biol.* 13 (10), 175. doi:10.1186/gb-2012-13-10-175
- Scholler, E., Weichmann, F., Treiber, T., Ringle, S., Treiber, N., Flatley, A., et al. (2018). Interactions, localization, and phosphorylation of the m(6A) generating METTL3-METTL14-WTAP complex. *RNA* 24 (4), 499–512. doi:10.1261/rna.064063.117
- Selberg, S., Blokhina, D., Aatonen, M., Koivisto, P., Siltanen, A., Mervala, E., et al. (2019). Discovery of small molecules that activate RNA methylation through cooperative binding to the METTL3-14-WTAP complex active site. *Cell Rep.* 26 (13), 3762–3771. e3765. doi:10.1016/j.celrep.2019.02.100
- Shi, H., Wei, J., and He, C. (2019). Where, when, and how: Context-dependent functions of RNA methylation writers, readers, and erasers. *Mol. Cell* 74 (4), 640–650. doi:10.1016/j.molcel.2019.04.025
- Shima, H., Matsumoto, M., Ishigami, Y., Ebina, M., Muto, A., Sato, Y., et al. (2017). S-adenosylmethionine synthesis is regulated by selective N(6)-adenosine methylation and mRNA degradation involving METTL16 and YTHDC1. *Cell Rep.* 21 (12), 3354–3363. doi:10.1016/j.celrep.2017.11.092
- Shu, X., Cao, J., Cheng, M., Xiang, S., Gao, M., Li, T., et al. (2020). A metabolic labeling method detects m(6A) transcriptome-wide at single base resolution. *Nat. Chem. Biol.* 16 (8), 887–895. doi:10.1038/s41589-020-0526-9
- Song, H., Song, J., Cheng, M., Zheng, M., Wang, T., Tian, S., et al. (2021). METTL3-mediated m(6A) RNA methylation promotes the anti-tumour immunity of natural killer cells. *Nat. Commun.* 12 (1), 5522. doi:10.1038/s41467-021-25803-0
- Su, R., Dong, L., Li, C., Nachtergaele, S., Wunderlich, M., Qing, Y., et al. (2018). R-2HG exhibits anti-tumor activity by targeting FTO/m(6A)/MYC/CEBPA signaling. *Cell* 172 (1–2), 90–105. doi:10.1016/j.cell.2017.11.031
- Taketo, K., Konno, M., Asai, A., Koseki, J., Toratani, M., Satoh, T., et al. (2018). The epitranscriptome m6A writer METTL3 promotes chemo- and radioresistance in pancreatic cancer cells. *Int. J. Oncol.* 52 (2), 621–629. doi:10.3892/ijo.2017.4219
- Thuring, K., Schmid, K., Keller, P., and Helm, M. (2017). LC-MS analysis of methylated RNA. *Methods Mol. Biol.* 1562, 3–18. doi:10.1007/978-1-4939-6807-7\_1
- Tung, Y. C., Ayuso, E., Shan, X., Bosch, F., O'Rahilly, S., Coll, A. P., et al. (2010). Hypothalamic-specific manipulation of Fto, the ortholog of the human obesity gene FTO, affects food intake in rats. *PLoS One* 5 (1), e8771. doi:10.1371/journal.pone.0008771
- van Tran, N., Ernst, F. G. M., Hawley, B. R., Zorbas, C., Ulryck, N., Hackert, P., et al. (2019). The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. *Nucleic Acids Res.* 47 (15), 7719–7733. doi:10.1093/nar/gkz619
- Visvanathan, A., Patil, V., Arora, A., Hegde, A. S., Arivazhagan, A., Santosh, V., et al. (2018). Essential role of METTL3-mediated m(6A) modification in glioma stem-like cells maintenance and radioresistance. *Oncogene* 37 (4), 522–533. doi:10.1038/onc.2017.351
- Vu, L. P., Pickering, B. F., Cheng, Y., Zaccara, S., Nguyen, D., Minuesa, G., et al. (2017). The N(6)-methyladenosine (m(6A))-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. *Nat. Med.* 23 (11), 1369–1376. doi:10.1038/nm.4416
- Walcher, L., Kistenmacher, A. K., Suo, H., Kitte, R., Dłuczek, S., Strauss, A., et al. (2020). Cancer stem cells-origins and biomarkers: Perspectives for targeted personalized therapies. *Front. Immunol.* 11, 1280. doi:10.3389/fimmu.2020.01280
- Wang, H., Zuo, H., Liu, J., Wen, F., Gao, Y., Zhu, X., et al. (2018a). Loss of YTHDF2-mediated m(6A)-dependent mRNA clearance facilitates hematopoietic stem cell regeneration. *Cell Res.* 28 (10), 1035–1038. doi:10.1038/s41422-018-0082-y
- Wang, L., Song, C., Wang, N., Li, S., Liu, Q., Sun, Z., et al. (2020a). NADP modulates RNA m(6A) methylation and adipogenesis via enhancing FTO activity. *Nat. Chem. Biol.* 16 (12), 1394–1402. doi:10.1038/s41589-020-0601-2
- Wang, P., Dextader, K. A., and Nam, Y. (2016). Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol. Cell* 63 (2), 306–317. doi:10.1016/j.molcel.2016.05.041

- Wang, R., Han, Z., Liu, B., Zhou, B., Wang, N., Jiang, Q., et al. (2018b). Identification of natural compound radicicol as a potent FTO inhibitor. *Mol. Pharm.* 15 (9), 4092–4098. doi:10.1021/acs.molpharmaceut.8b00522
- Wang, T., Hong, T., Huang, Y., Su, H., Wu, F., Chen, Y., et al. (2015). Fluorescein derivatives as bifunctional molecules for the simultaneous inhibiting and labeling of FTO protein. *J. Am. Chem. Soc.* 137 (43), 13736–13739. doi:10.1021/jacs.5b06690
- Wang, Y., Xiao, Y., Dong, S., Yu, Q., and Jia, G. (2020b). Antibody-free enzyme-assisted chemical approach for detection of N(6)-methyladenosine. *Nat. Chem. Biol.* 16 (8), 896–903. doi:10.1038/s41589-020-0525-x
- Wang, Y., Zeng, L., Liang, C., Zan, R., Ji, W., Zhang, Z., et al. (2019). Integrated analysis of transcriptome-wide m(6)A methylome of osteosarcoma stem cells enriched by chemotherapy. *Epigenomics* 11 (15), 1693–1715. doi:10.2217/epi-2019-0262
- Wei, C. M., Gershowitz, A., and Moss, B. (1975). Methylated nucleotides block 5' terminus of HeLa cell messenger RNA. *Cell* 4 (4), 379–386. doi:10.1016/0092-8674(75)90158-0
- Wei, J., Liu, F., Lu, Z., Fei, Q., Ai, Y., He, P. C., et al. (2018). Differential m(6)A, m(6)Am, and m(1)A demethylation mediated by FTO in the cell nucleus and cytoplasm. *Mol. Cell* 71 (6), 973–985. e975. doi:10.1016/j.molcel.2018.08.011
- Wen, J., Lv, R., Ma, H., Shen, H., He, C., Wang, J., et al. (2018). Zc3h13 regulates nuclear RNA m(6)A methylation and mouse embryonic stem cell self-renewal. *Mol. Cell* 69 (6), 1028–1038. e1026. doi:10.1016/j.molcel.2018.02.015
- Weng, H., Huang, H., Wu, H., Qin, X., Zhao, B. S., Dong, L., et al. (2018). METTL14 inhibits hematopoietic stem/progenitor differentiation and promotes leukemogenesis via mRNA m(6)A modification. *Cell Stem Cell* 22 (2), 191–205. e199. doi:10.1016/j.stem.2017.11.016
- Wiener, D., and Schwartz, S. (2021). The epitranscriptome beyond m(6)A. *Nat. Rev. Genet.* 22 (2), 119–131. doi:10.1038/s41576-020-00295-8
- Wilson, C., Chen, P. J., Miao, Z., and Liu, D. R. (2020). Programmable m(6)A modification of cellular RNAs with a Cas13-directed methyltransferase. *Nat. Biotechnol.* 38 (12), 1431–1440. doi:10.1038/s41587-020-0572-6
- Wu, J., Li, Y., Yu, J., Gan, Z., Wei, W., Wang, C., et al. (2020). Resveratrol attenuates high-fat diet induced hepatic lipid homeostasis disorder and decreases m(6)A RNA methylation. *Front. Pharmacol.* 11, 568006. doi:10.3389/fphar.2020.568006
- Wu, R., Li, A., Sun, B., Sun, J. G., Zhang, J., Zhang, T., et al. (2019). A novel m(6)A reader Prrc2a controls oligodendroglial specification and myelination. *Cell Res.* 29 (1), 23–41. doi:10.1038/s41422-018-0113-8
- Xiao, W., Adhikari, S., Dahal, U., Chen, Y. S., Hao, Y. J., Sun, B. F., et al. (2016). Nuclear m(6)A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* 61 (4), 507–519. doi:10.1016/j.molcel.2016.01.012
- Xie, H., Li, J., Ying, Y., Yan, H., Jin, K., Ma, X., et al. (2020). METTL3/YTHDF2 m(6)A axis promotes tumorigenesis by degrading SETD7 and KLF4 mRNAs in bladder cancer. *J. Cell. Mol. Med.* 24 (7), 4092–4104. doi:10.1111/jcmm.15063
- Yadav, P. K., and Rajasekharan, R. (2018). The m(6)A methyltransferase Ime4 and mitochondrial functions in yeast. *Curr. Genet.* 64 (2), 353–357. doi:10.1007/s00294-017-0758-8
- Yang, F., Jin, H., Que, B., Chao, Y., Zhang, H., Ying, X., et al. (2019). Dynamic m(6)A mRNA methylation reveals the role of METTL3-m(6)A-CDP1 signaling axis in chemical carcinogenesis. *Oncogene* 38 (24), 4755–4772. doi:10.1038/s41388-019-0755-0
- Yang, X., Wei, X., Yang, J., Du, T., Yin, C., Fu, B., et al. (2021). An unexpected role for Dicer as a reader of the unacetylated DNA binding domain of p53 in transcriptional regulation. *Sci. Adv.* 7 (19), eabi6684. doi:10.1126/sciadv.abi6684
- Yang, Y., Wei, Q., Tang, Y., Yuanyuan, W., Luo, Q., Zhao, H., et al. (2020). Loss of hnRNP2B1 inhibits malignant capability and promotes apoptosis via down-regulating Lin28B expression in ovarian cancer. *Cancer Lett.* 475, 43–52. doi:10.1016/j.canlet.2020.01.029
- Yao, Y., Yang, Y., Guo, W., Xu, L., You, M., Zhang, Y. C., et al. (2021). METTL3-dependent m(6)A modification programs T follicular helper cell differentiation. *Nat. Commun.* 12 (1), 1333. doi:10.1038/s41467-021-21594-6
- Yin, H., Zhang, X., Yang, P., Zhang, X., Peng, Y., Li, D., et al. (2021). RNA m6A methylation orchestrates cancer growth and metastasis via macrophage reprogramming. *Nat. Commun.* 12 (1), 1394. doi:10.1038/s41467-021-21514-8
- Yu, F., Wei, J., Cui, X., Yu, C., Ni, W., Bungert, J., et al. (2021a). Post-translational modification of RNA m6A demethylase ALKBH5 regulates ROS-induced DNA damage response. *Nucleic Acids Res.* 49 (10), 5779–5797. doi:10.1093/nar/gkab415
- Yu, Q., Liu, S., Yu, L., Xiao, Y., Zhang, S., Wang, X., et al. (2021b). RNA demethylation increases the yield and biomass of rice and potato plants in field trials. *Nat. Biotechnol.* 39 (12), 1581–1588. doi:10.1038/s41587-021-00982-9
- Yue, H., Nie, X., Yan, Z., and Weining, S. (2019). N6-methyladenosine regulatory machinery in plants: Composition, function and evolution. *Plant Biotechnol. J.* 17 (7), 1194–1208. doi:10.1111/pbi.13149
- Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., et al. (2018). VIRMA mediates preferential m(6)A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. *Cell Discov.* 4, 10. doi:10.1038/s41421-018-0019-0
- Zhang, Q., and Xu, K. (Forthcoming 2022). The role of regulators of RNA m6A methylation in lung cancer. *Genes & Dis.* doi:10.1016/j.gendis.2021.12.017
- Zhang, S., Zhao, B. S., Zhou, A., Lin, K., Zheng, S., Lu, Z., et al. (2017). m(6)A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. *Cancer Cell* 31 (4), 591–606. e596. doi:10.1016/j.ccell.2017.02.013
- Zhang, X., Wei, L. H., Wang, Y., Xiao, Y., Liu, J., Zhang, W., et al. (2019a). Structural insights into FTO's catalytic mechanism for the demethylation of multiple RNA substrates. *Proc. Natl. Acad. Sci. U. S. A.* 116 (8), 2919–2924. doi:10.1073/pnas.1820574116
- Zhang, Y., Kang, M., Zhang, B., Meng, F., Song, J., Kaneko, H., et al. (2019b). m(6)A modification-mediated CBX8 induction regulates stemness and chemosensitivity of colon cancer via upregulation of LGR5. *Mol. Cancer* 18 (1), 185. doi:10.1186/s12943-019-1116-x
- Zhang, Z., Chen, L. Q., Zhao, Y. L., Yang, C. G., Roundtree, I. A., Zhang, Z., et al. (2019c). Single-base mapping of m(6)A by an antibody-independent method. *Sci. Adv.* 5 (7), eaax0250. doi:10.1126/sciadv.aax0250
- Zhao, B. S., Wang, X., Beadell, A. V., Lu, Z., Shi, H., Kuuspalu, A., et al. (2017). m(6)A-dependent maternal mRNA clearance facilitates zebrafish maternal-to-zygotic transition. *Nature* 542 (7642), 475–478. doi:10.1038/nature21355
- Zhao, Y., Shi, Y., Shen, H., and Xie, W. (2020). m(6)A-binding proteins: the emerging crucial performers in epigenetics. *J. Hematol. Oncol.* 13 (1), 35. doi:10.1186/s13045-020-00872-8
- Zhou, B., Liu, C., Xu, L., Yuan, Y., Zhao, J., Zhao, W., et al. (2021a). N(6)-methyladenosine reader protein YT521-B homology domain-containing 2 suppresses liver steatosis by regulation of mRNA stability of lipogenic genes. *Hepatology* 73 (1), 91–103. doi:10.1002/hep.31220
- Zhou, H., Yin, K., Zhang, Y., Tian, J., and Wang, S. (2021b). The RNA m6A writer METTL14 in cancers: Roles, structures, and applications. *Biochim. Biophys. Acta. Rev. Cancer* 1876 (2), 188609. doi:10.1016/j.bbcan.2021.188609
- Zhou, X., Chen, J., Chen, J., Wu, W., Wang, X., and Wang, Y. (2015). The beneficial effects of betaine on dysfunctional adipose tissue and N6-methyladenosine mRNA methylation requires the AMP-activated protein kinase  $\alpha$ 1 subunit. *J. Nutr. Biochem.* 26 (12), 1678–1684. doi:10.1016/j.jnutbio.2015.08.014
- Zhou, X. L., Huang, F. J., Li, Y., Huang, H., and Wu, Q. C. (2021c). SEDT2/METTL14-mediated m6A methylation awakening contributes to hypoxia-induced pulmonary arterial hypertension in mice. *Aging (Albany NY)* 13 (5), 7538–7548. doi:10.18632/aging.202616
- Zhu, W., Wang, J. Z., Wei, J. F., and Lu, C. (2021). Role of m6A methyltransferase component VIRMA in multiple human cancers (Review). *Cancer Cell Int.* 21 (1), 172. doi:10.1186/s12935-021-01868-1



## OPEN ACCESS

## EDITED BY

Xiaofan Lu,  
INSERM U964 Institut de Génétique et  
de Biologie Moléculaire et Cellulaire  
(IGBMC), France

## REVIEWED BY

Wenyi Jin,  
City University of Hong Kong, Hong  
Kong SAR, China  
Travis Steele Johnson,  
Indiana University Bloomington,  
United States

## \*CORRESPONDENCE

Xin Liang,  
liangxin0314@163.com

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 22 September 2022

ACCEPTED 04 November 2022

PUBLISHED 18 November 2022

## CITATION

Liang X, Yin Y and Li N (2022), GOLM1 is  
related to the inflammatory/immune  
nature of uveal melanoma and acts as a  
promising indicator for prognosis and  
immunotherapy response.  
*Front. Genet.* 13:1051168.  
doi: 10.3389/fgene.2022.1051168

## COPYRIGHT

© 2022 Liang, Yin 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.

# GOLM1 is related to the inflammatory/immune nature of uveal melanoma and acts as a promising indicator for prognosis and immunotherapy response

Xin Liang<sup>1\*</sup>, Yu Yin<sup>2</sup> and Ning Li<sup>3</sup>

<sup>1</sup>Department of Ophthalmology, Shanghai General Hospital, Shanghai, China, <sup>2</sup>Department of Pathology, Anhui Medical University, Hefei, China, <sup>3</sup>Department of Ophthalmology, Anhui Medical University, Hefei, China

**Purpose:** Inflammatory/immune-related features are associated with the immunotherapy and prognosis of uveal melanoma (UVM). In this study, we systematically analyzed the correlation between GOLM1 and the inflammatory/immune nature of UVM and explored its potential value in predicting prognosis and guiding immunotherapy for UVM patients.

**Methods:** A total of 143 UVM patients were enrolled in the current study. The differentially expressed genes between the GOLM1-low expression (LEXP) and GOLM1-high expression (HEXP) subgroups were calculated by the “limma” package and further annotated to reveal the key pathways by the “ClusterProfiler” package. Immunocyte infiltration was evaluated by single-sample gene set enrichment analysis, while the potential response to immunotherapy was realized by subclass mapping analysis. Moreover, tumor tissue sections from 23 UVM patients were collected and stained for GOLM1 (1:300; cat# DF8100, Affinity Biosciences), PD-L1 (1:250; cat# ab213524, Abcam), PD-1 (1:100; cat# ab52587, Abcam), CTLA-4 (1:300; cat# DF6793, Affinity Biosciences), and IFN- $\gamma$  (1:300; cat# DF6045, Affinity Biosciences).

**Results:** We found that higher expression of GOLM1 correlated with an unfavorable prognosis in UVM patients. Multivariate Cox regression analysis suggested that GOLM1 served as a prognostic factor independent of clinicopathological parameters. Notably, we found that the expression of PD-1, PD-L1, IFN- $\gamma$ , and CTLA4 was higher in the GOLM1-high subgroup than in the GOLM1-low expression subgroup at the mRNA level and was subsequently validated at the protein level by immunohistochemistry. Gene pattern and SubMap analyses confirmed the indicator role of GOLM1 in predicting immunotherapy response in UVM.

**Conclusion:** Taken together, GOLM1 is a novel prognostic marker, and it can be employed to predict the overall survival outcomes and treatment responses of anti-PD-1/PD-L1 and anti-CTLA4 therapies for UVM patients.

## KEYWORDS

GOLM1, inflammation, tumor microenvironment, immunotherapy, prognosis

## Background

It has been reported that immune infiltrates are detectable in the majority of primary and metastatic uveal melanoma (UVM) (Krishna et al., 2017; Qin et al., 2020). The progression and metastasis of UVM are associated with inflammatory infiltration. Some UVMs with genetic aberrations, such as loss of one copy of chromosome 3 and BAP1 protein, produce inflammatory mediators. CD8<sup>+</sup>, CD4<sup>+</sup>, Foxp3<sup>+</sup> T cells and macrophages are recruited and activated, resulting in the generation of more inflammatory mediators and a cancer-related tumor-promoting inflammatory microenvironment (Bronkhorst and Jager, 2012). The activation of nuclear factor- $\kappa$ B (NF- $\kappa$ B) was influenced, and additional proinflammatory multifunctional chemokines and cytokines were released, which affected infiltrating T lymphocytes (TILs) and tumor-associated macrophages (TAMs) in UVM. These driving T lymphocyte and macrophage infiltration are tightly linked to tumor recurrence and survival (Gezgin et al., 2017).

GOLM1 is a Golgi membrane protein mainly expressed in epithelial cells and upregulated in response to viral infection, with a coding gene of 3,042 bp in full length (Kladney et al., 2000). As reported, GOLM1 can be expressed in most human tissues, mainly in epithelial cells. GOLM1 is synthesized in the endoplasmic reticulum and is primarily transported to the cis-Golgi and is also a secreted protein that is detected in blood and urine. Until now, the biological function of GOLM1 in cells and tissues has not been entirely clear. Some studies have suggested that GOLM1 may play an important role in supporting normal cell function. In addition, the critical implication of GOLM1 in the prevention, prediction, and treatment of cancer has been established (Sun et al., 2011). Additionally, increased expression of GOLM1 is associated with some malignant biological characteristics and is an independent prognostic factor for poor overall survival (Sun et al., 2011; Bao et al., 2013; Chen et al., 2013). Notably, Kim et al. (2012) found that GOLM1 has a potential impact on inhibiting IL-12 production by dendritic cells, thus promoting tumor progression by preventing T cell responses in hepatocellular carcinoma. Consistently, Zhang et al. (2017a) revealed that host innate immune processes could be repressed by GOLM1, which will enhance the replication of the hepatitis C virus (HCV), thus driving the pathogenesis of hepatocellular carcinoma. Furthermore, increasing evidence suggests that GOLM1 regulates the STAT3 pathway, which is upstream of PD-L1 (Wölflé et al., 2011; Song et al., 2018; Zhang et al., 2019; Yan et al., 2020). Thus, it makes sense to regulate PD-L1 expression through targeting GOLM1. These findings highlight the potential role of GOLM1 in regulating immune responses in cancers.

However, the connection between GOLM1 and UVM, particularly its role in regulating inflammatory/immune responses in this tumor, has not been investigated. Herein, a higher expression level of GOLM1 was found in UVM patients

than in controls. GOLM1 was considered an independent prognostic factor after the adjustment of clinicopathological features. In addition, the highly co-expressed genes of GOLM1 were primarily enriched in inflammatory/immune-related pathways, and patients with higher expression of GOLM1 might be more sensitive to anti-CTLA4 and anti-PD-1/PD-L1 immunotherapies. Our results provide a novel marker for clinical prediction, prognosis, and immunotherapy guidance for UVM patients.

## Materials and methods

### Patient summary

In this study, a total of 193 UVM patients were enrolled for analysis, from four cohorts, TCGA-UVM ( $n = 80$ ), GSE21138 ( $n = 63$ ), GSE84976 ( $n = 27$ ) and a real-world cohort ( $n = 23$ ), the basic information of these four cohorts listed in Table 1. For TCGA-UVM, the GDC platform was used to download the gene expression profile and clinical information via the “TCGAbiolinks” package; the GENCODE27 annotation file realized the further annotation of mRNA gene symbols. For GSE21138 and GSE84976, the gene expression profile and clinical features of whom came from the Gene Expression Omnibus platform (<http://www.ncbi.nlm.nih.gov/geo/>). Furthermore, the expression of GOLM1 in pan-cancer, the comparison of GOLM1 in normal and tumor tissues, and the GOLM1 level to overall survival (OS) for each cancer type were explored via the online website (<http://sangerbox.com/>).

### Signaling pathway annotation

For GOLM1-low (LEXP) and GOLM1-high (HEXP) subgroups, the “limma” package was used to identify differentially expressed genes (DEGs) between them, accompanied by further filtering with the requirement of fold-change  $>1$  and  $p$ -value  $<0.01$ . Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology (GO), and HALLMARK background gene set signatures were downloaded from MSigDB (Subramanian et al., 2005). We further implemented DEG enrichment annotation via the “ClusterProfiler” package (Yu et al., 2012).

### Tumor immunocyte infiltration evaluation

To reveal the immunocyte infiltration status among UVM patients, we collected 28 immunocyte signatures from a previous study (Yoshihara et al., 2013). With the support of the “GSVA” R package, a single sample gene set enrichment analysis (ssGSEA) (Barbie et al., 2009; Hanzelmann et al., 2013) was conducted.

TABLE 1 Basic informal of enrolled cohorts.

|                        | TCGA-UVM (N = 80) | GSE21138 (N = 63) | GSE84976 (N = 27) | Real-world (N = 23) | Method         | p-value |
|------------------------|-------------------|-------------------|-------------------|---------------------|----------------|---------|
| Age, months            |                   |                   |                   |                     | Kruskal–Wallis | 0.929   |
| Mean (SD)              | 61.7 (13.9)       | 61.0 (12.3)       | 62.2 (15.2)       | 61.3 (13.0)         |                |         |
| Median [Min, Max]      | 61.5 [22.0, 86.0] | 62.1 [28.6, 85.0] | 68.0 [28.0, 84.0] | 60.0 [39.0, 85.0]   |                |         |
| Gender                 |                   |                   |                   |                     | Chi-square     | 0.238   |
| Female                 | 35 (43.8%)        | 24 (38.1%)        | —                 | 15 (57.7%)          |                |         |
| Male                   | 45 (56.3%)        | 39 (61.9%)        | —                 | 11 (42.3%)          |                |         |
| Basal diameter, mm     |                   |                   |                   |                     | Kruskal–Wallis | 0.044   |
| Mean (SD)              | 16.9 (3.47)       | 15.4 (3.78)       | —                 | 15.9 (8.88)         |                |         |
| Median [Min, Max]      | 17.0 [7.79, 25.0] | 15.0 [9.00, 23.0] | —                 | 15.0 [4.00, 35.0]   |                |         |
| Missing                | 1 (1.3%)          | 10 (15.9%)        | —                 | 3 (11.5%)           |                |         |
| Thickness, mm          |                   |                   |                   |                     | Kruskal–Wallis | 0.036   |
| Mean (SD)              | 10.4 (2.81)       | 11.7 (2.02)       | —                 | 13.3 (7.96)         |                |         |
| Median [Min, Max]      | 10.5 [4.00, 16.0] | 11.7 [6.00, 17.0] | —                 | 12.0 [3.00, 32.0]   |                |         |
| Missing                | 0 (0%)            | 0 (0%)            | —                 | 3 (11.5%)           |                |         |
| Extrascleral extension |                   |                   |                   |                     | Chi-square     | 1.000   |
| Unknown                | 5 (6.3%)          | 10 (15.9%)        | —                 | —                   |                |         |
| No                     | 68 (85.0%)        | 48 (76.2%)        | —                 | —                   |                |         |
| Yes                    | 7 (8.8%)          | 5 (7.9%)          | —                 | —                   |                |         |

Then, the normalized enrichment score (NES) of 28 immunocyte signatures was calculated for each UVM sample.

## Predicting the response to immunotherapy

For an examination of individuals' possibility of responding to immunotherapy, the expression value of 795 response-specific genes was generated from a melanoma cohort in which anti-PD-1 or anti-CTLA-4 checkpoint therapy was applied (Chen et al., 2016). The comparison between GOLM1 groups and immunotherapy groups was realized by subclass mapping analysis, thereby identifying potential responders in GOLM1 groups (Hoshida et al., 2007).

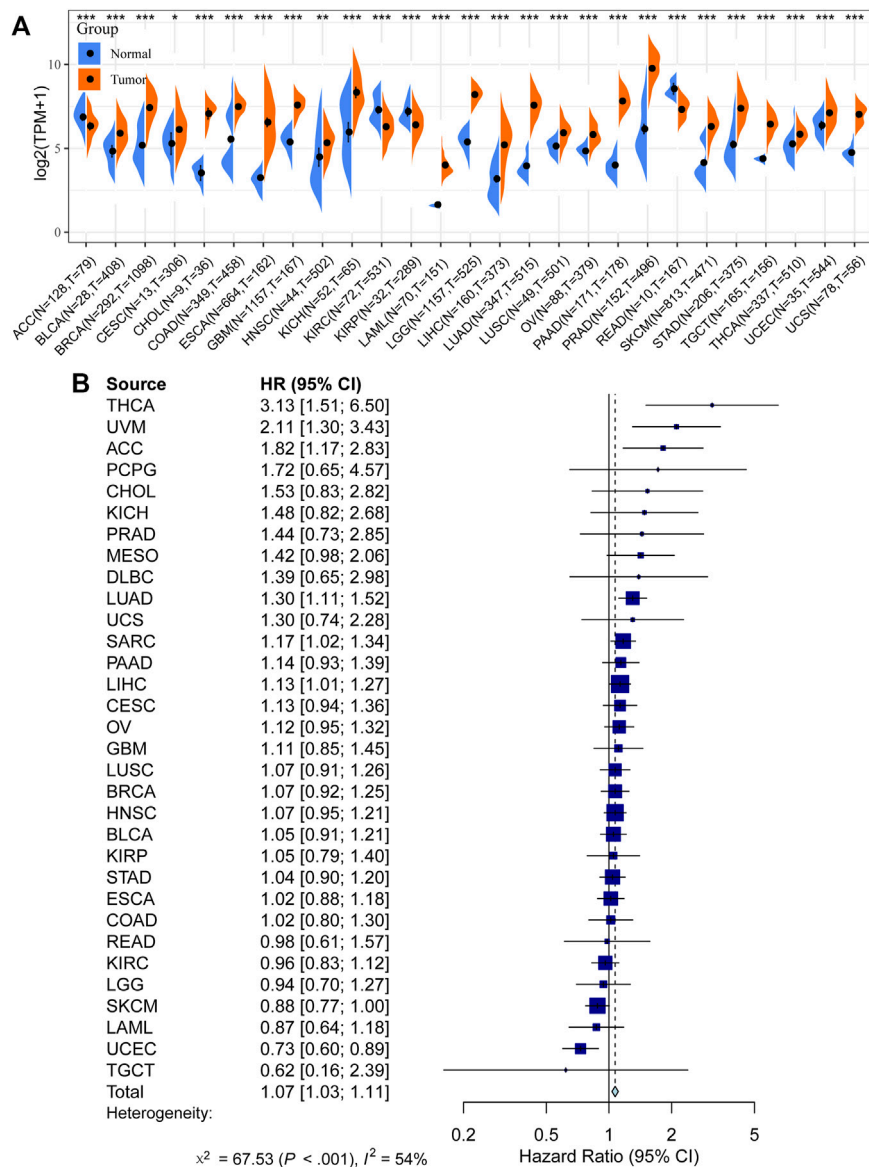
## Immunohistochemistry

We collected basic information and paraffin tissue sections of 23 patients with ocular melanoma from the First and the Second Affiliated Hospital of Anhui Medical University (Supplementary Table S1). The Institutional Review Board of Anhui Medical University approved this study. Paraffin tissues were cut into five- $\mu$ m-thick sections for histological analysis. Xylene and different concentrations of alcohol were used for deparaffinization and dehydration of the sections. Then, the sections were placed in boiling antigen retrieval

solution for 15 min to complete antigen retrieval of those tissues. After the sections were equilibrated to room temperature in antigen retrieval solution, 3% hydrogen peroxide solution was used for catalase blocking, and sections were placed into 0.1% potassium permanganate solution at 37°C for 40 min, followed by placement into 1% oxalic acid solution for decolorization. After washing 3 times with the washing solution, sections were incubated with primary antibodies [PD-L1 (1:250; cat#ab213524, Abcam), PD-1 (1:100; cat#ab52587, Abcam), CTLA-4 (1:300; cat#DF6793, Affinity Biosciences), IFN- $\gamma$  (1:300; cat#DF6045, Affinity Biosciences) and GOLM1 (1:300; cat#DF8100, Affinity Biosciences)] at 37°C for 1 h. Subsequently, after washing three times with the washing solution, the sections were incubated with biotinylated secondary antibody (1:200) for 30 min, followed by adding horseradish peroxidase-labeled streptavidin to the sections for 30 min. Finally, diaminobenzidine staining was used to detect the immunoreactivity of those tissues. The detailed procedures of H-score counting were demonstrated in our previous publication (Yin et al., 2019).

## Statistical analyses

We performed all the statistical analyses *via* R version 4.0.2. For the continuous data, a *t* test was used for comparison if the data is normally distributed; otherwise, the Wilcoxon rank-sum



**FIGURE 1** Prognostic value of GOLM1 across cancers. (A) Differential expression of GOLM1 in normal and tumor tissues; (B) Prognostic value of GOLM1 for overall survival across cancers.

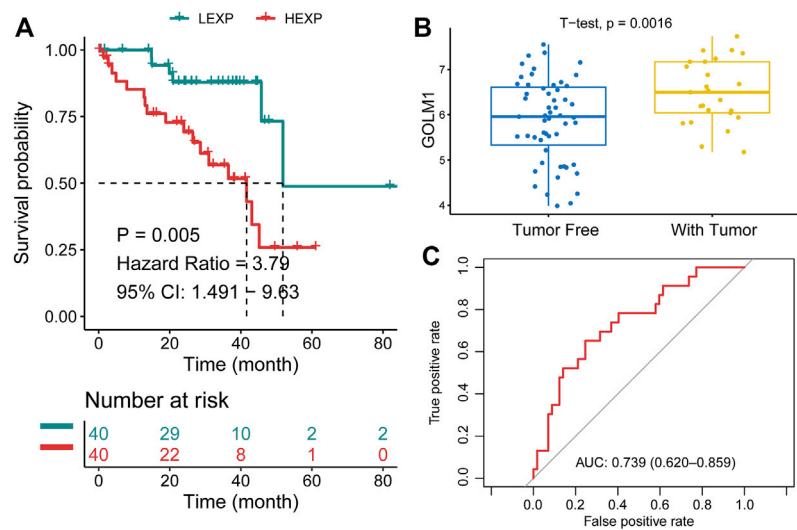
test was conducted. Categorical data were compared by Fisher’s exact test and the Chi-square test. Cox proportional hazard regression for hazard ratio (HR) and a log-rank test Kaplan–Meier curve, with the 95% confidential interval (95% CI), was performed to illustrate the difference in OS. The prediction efficiency of GOLM1 expression for OS was examined based on the time-dependent incident dynamic receiver operating characteristic (ROC) area that was below the curve (AUC) values (having a 60-month survival endpoint) (Blanche et al., 2013). Multivariate Cox regression analyses explored the prognostic value of GOLM1 after clinical feature

adjustment. A two-sided  $p$ -value  $<0.05$  demonstrated statistical significance.

## Results

### GOLM1 acts as an oncogene across cancers

We first compared the expression of GOLM1 in tumor and normal tissues. Tumor tissues were obtained from the TCGA

**FIGURE 2**

Prognostic value of GOLM1 in uveal melanoma. **(A)** Kaplan–Meier curve showing the polarized overall survival in GOLM1-low (LEXP) and GOLM1-high (HEXP) subgroups; **(B)** Differential expression of GOLM1 in patients lived with or without tumors; **(C)** Receiver operating characteristic curve showing the prognostic value of GOLM1.

**TABLE 2** Summarization of the clinical features in TCGA-uveal melanoma cohort.

| Clinical features              | Level      | GOLM1-low       | GOLM1-high      | <i>p</i> -value |
|--------------------------------|------------|-----------------|-----------------|-----------------|
| Number                         |            | 40              | 40              |                 |
| Overall survival time (months) |            | 30.43 ± 18.26   | 22.92 ± 17.33   | 0.063           |
| Age (years)                    |            | 59.00 ± 14.89   | 64.30 ± 12.56   | 0.089           |
| Gender (%)                     | Female     | 18 (45.0)       | 17 (42.5)       | 1.000           |
|                                | Male       | 22 (55.0)       | 23 (57.5)       |                 |
| Stage (%)                      | Stage II   | 25 (62.5)       | 14 (35.0)       | 0.042*          |
|                                | Stage III  | 14 (35.0)       | 22 (55.0)       |                 |
|                                | Stage IV   | 1 (2.5)         | 3 (7.5)         |                 |
|                                | Unknown    | 0 (0.0)         | 1 (2.5)         |                 |
| Tumor status (%)               | Tumor-free | 31 (77.5)       | 24 (60.0)       | 0.148           |
|                                | With tumor | 9 (22.5)        | 16 (40.0)       |                 |
| Basal diameter, cm             |            | 16.29 ± 3.23    | 17.58 ± 3.63    | 0.098           |
| Thickness, cm                  |            | 10.13 ± 2.50    | 10.71 ± 3.09    | 0.357           |
| Tumor volume, cm <sup>3</sup>  |            | 756.77 ± 380.99 | 949.51 ± 576.01 | 0.085           |
| Extrascleral extension (%)     | No         | 37 (92.5)       | 31 (77.5)       | 0.124           |
|                                | Yes        | 1 (2.5)         | 6 (15.0)        |                 |
|                                | Unknown    | 2 (5.0)         | 3 (7.5)         |                 |

TCGA, the cancer genome atlas; \*,  $p < 0.05$ .

project, while normal tissues combined tumor-adjacent tissues from the TCGA project and normal tissues from the GTEx project. Higher expression of GOLM1 in tumor tissues than in normal tissues was observed in most tumor types ( $p < 0.05$ ,

85.19%, 23/27, Figure 1A). Furthermore, GOLM1 acted as an oncogene linked with poor prognosis in the combined meta-analysis of 32 types of cancer ( $p < 0.001$ , HR: 2.11, 95% CI: 1.30–3.43, Figure 1B).

TABLE 3 Multiple cox regression analysis results in two uveal melanoma cohorts.

| Clinical features                       | HR     | 95% CI        | p-value |
|---|--------|---------------|---------|
| TCGA-uveal melanoma                     |        |               |         |
| Age, years                              | 1.050  | 0.988–1.104   | 0.059   |
| Gender, Male vs. Female                 | 1.404  | 0.518–3.804   | 0.505   |
| Stage, III vs. II                       | 0.741  | 0.286–1.915   | 0.535   |
| Stage, IV vs. II                        | 24.025 | 1.965–293.782 | 0.013*  |
| Tumor status, With tumor vs. Tumor-free | 13.098 | 3.705–46.298  | <0.001* |
| GOLM1, HEXP vs. LEXP                    | 3.222  | 1.152–9.010   | 0.026*  |
| GSE21138                                |        |               |         |
| Age, years                              | 1.040  | 1.003–1.08    | 0.036*  |
| Gender, Male vs. Female                 | 1.321  | 0.602–2.902   | 0.488   |
| Extrascleral extension, Yes vs. No      | 2.438  | 0.883–6.727   | 0.085   |
| GOLM1, HEXP vs. LEXP                    | 2.681  | 1.220–5.888   | 0.014*  |
| GSE84976                                |        |               |         |
| Age, years                              | 1.030  | 0.988–1.073   | 0.167   |
| GOLM1, HEXP vs. LEXP                    | 3.819  | 1.262–11.562  | 0.018*  |

TCGA, the cancer genome atlas; HR, hazard ratio; CI, confidential interval; \*,  $p < 0.05$ .

## GOLM1 predicts the prognosis of uveal melanoma patients

In the TCGA-UVM cohort, we assigned patients to GOLM1-low (LEXP) and GOLM1-high (HEXP) subgroups, of which the latter subgroup presented poor prognosis ( $p = 0.005$ , HR: 3.790, 95% CI: 1.491–9.630, Figure 2A). The average OS time for HEXP patients was only  $22.92 \pm 17.33$  months, while that of the other subgroup was  $30.43 \pm 18.26$  months ( $p = 0.063$ , Table 2). Specifically, patients who lived with the tumor displayed a high level of GOLM1 compared to those without tumors ( $p = 0.0016$ , Figure 2B). Afterwards, we evaluated the prognostic value of GOLM1 over time, obtaining the AUC value of 0.739, which 95% CI ranged from 0.620 to 0.859 (Figure 2C), which represented a principal predictive value. After adjusting for age, sex, tumor stage, and tumor status, we revealed GOLM1 as an independent prognostic factor for UVM patients ( $p = 0.026$ , HR: 3.222, 95% CI: 1.152–9.010, Table 3).

## GOLM1 correlates with the tumor inflammatory/immune environment

To better understand how GOLM1 impacts tumor progression in UVM, efforts were made to search for potential interacting signaling pathways. First, we identified 531 DEGs between the LEXP and HEXP subgroups (Figure 3A), and the signaling pathways related to these DEGs were enriched. In the biological process analysis of GO terms, GOLM1 was mostly associated with the antigen processing and presentation pathway and the response to IFN- $\gamma$  pathway; in the molecular function analysis, it exhibited associations with the MHC

protein complex and the endoplasmic structure; in the cellular component analysis, GOLM1 influenced the binding of peptide antigen, extracellular matrix, and MHC protein complex (Figure 3B). And the pathway enrichment was also conducted based on the 50 classical HALLMARK tumor pathways, we observed the activation of interferon gamma response, allograft rejection, complement and epithelial mesenchymal transition (Figure 3C).

These findings demonstrated the possible impact of GOLM1 on the immune infiltration of UVM. We further compared the infiltrated immunocytes in the LEXP and HEXP groups of GOLM1, of which the latter displayed a higher infiltration of type 1 T helper cells, CD8 T cells, natural killer T cells, T follicular helper cells, and CD4 T cells (all  $p < 0.05$ , Figure 4A).

## GOLM1 is positively associated with a favorable response to immunotherapy

Immune checkpoints are targets for immunotherapy in numerous cancers. Regarding the expression of the four main immune checkpoint genes, including PD-1, PD-L1, PD-L2, and CTLA4, the HEXP subgroup presented higher expression levels of these markers than LEXP (all  $p < 0.05$ , Figure 4B). In addition, the GOLM1 expression was positively correlated with the expression of PD-1 ( $p < 0.05$ ,  $R = 0.46$ ), PD-L1 ( $p < 0.05$ ,  $R = 0.45$ ), PD-L2 ( $p < 0.05$ ,  $R = 0.38$ ), and CTLA4 ( $p < 0.05$ ,  $R = 0.38$ ) (Figure 4C). We also predicted the potential responders with the SubMap analysis of a melanoma cohort receiving anti-PD-1 and anti-CTLA4 therapies, demonstrating that patients in the HEXP group are more likely to be sensitive to both anti-CTLA4 and anti-PD-1/PD-L1 therapies (Figure 4D).

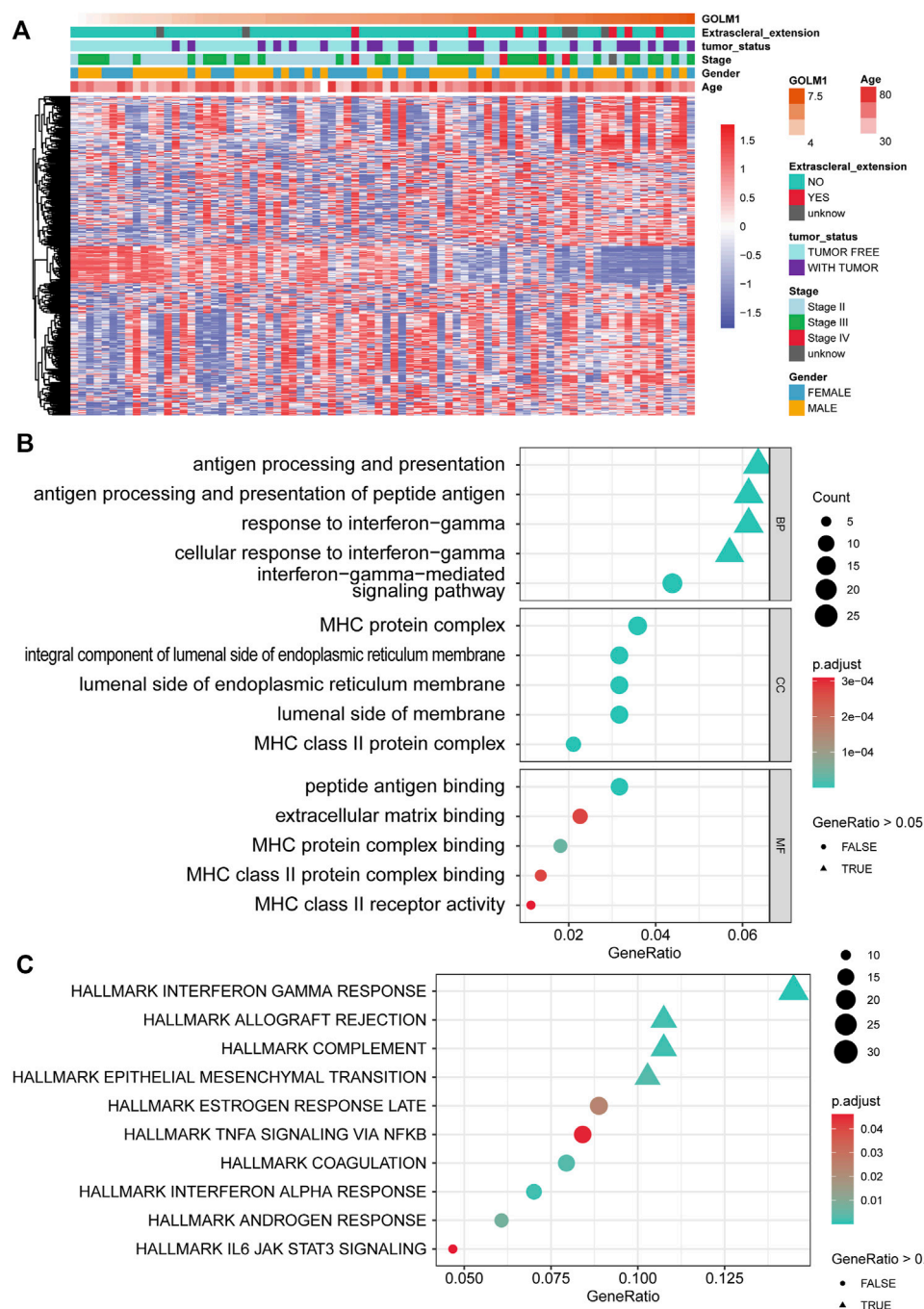


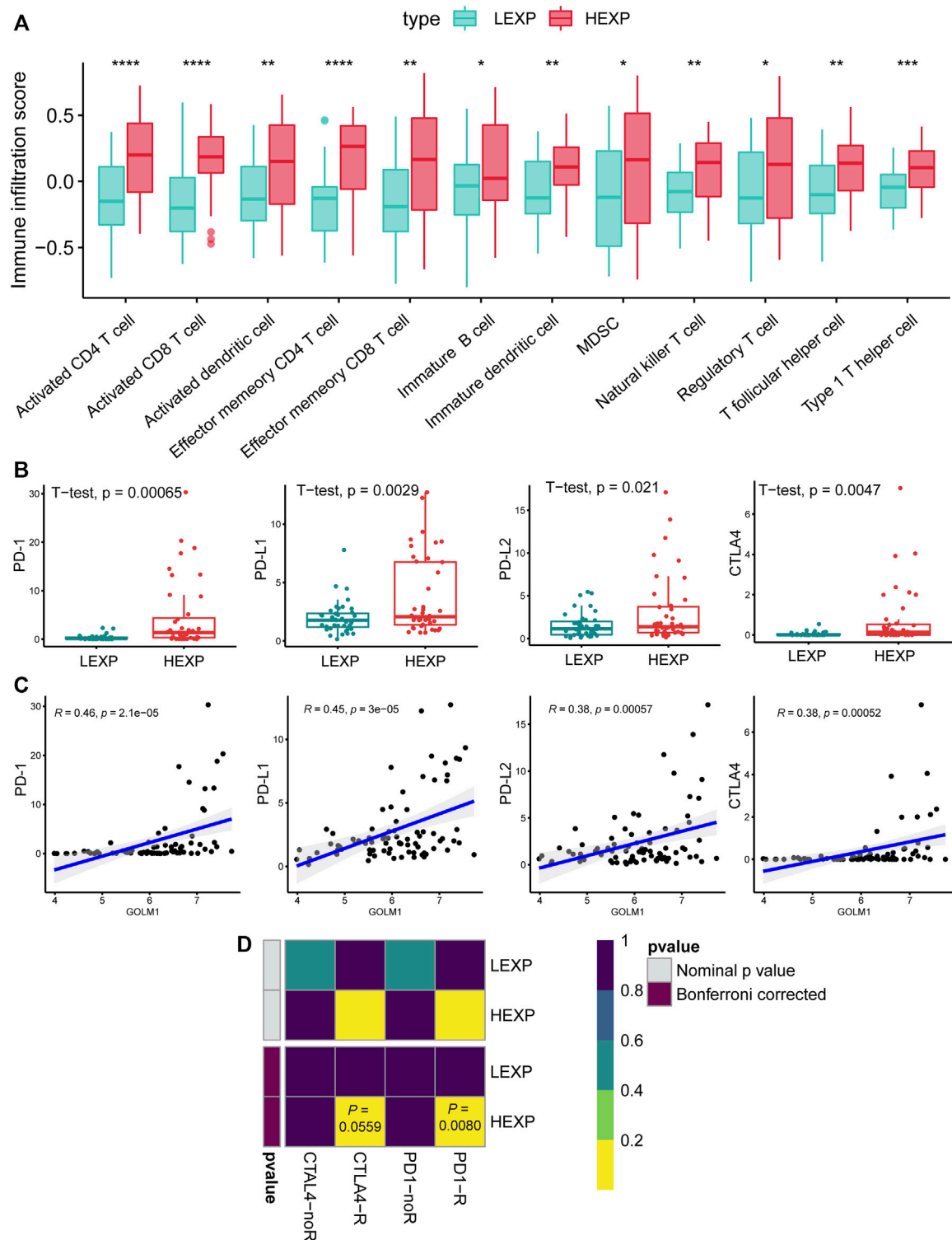
FIGURE 3

Pathway enrichment of GOLM1-associated genes. **(A)** Heatmap showing the expression of 531 differentially expressed genes (DEGs) between the GOLM1-low (LEXP) and GOLM1-high (HEXP) subgroups; **(B)** Pathway enrichment of the gene ontology (GO) terms for these DEGs; **(C)** Pathway enrichment of the HALLMARK terms for these DEGs.

## External validation of GOLM1 prognostic value

To further confirm the prognostic value of GOLM1, we validated these findings in the GSE21138 cohort. Sixty-three

patients were separated into the LEXP ( $n = 32$ ) and HEXP ( $n = 31$ ) subgroups based on the median value of GOLM1 expression. Patients in the HEXP subgroup displayed a 1.98-fold HR compared with the LEXP subgroup, with a 95% CI of 1.005–3.905 ( $p = 0.048$ , Figure 5A). The higher expression of



**FIGURE 4**  
GOLM1 is linked with tumor-infiltrated immunocytes and immunotherapy. **(A)** Differential infiltration of immunocytes in the GOLM1-low (LEXP) group compared to the GOLM1-high (HEXP) group of The Cancer Genome Atlas (TCGA)-uveal melanoma cohort; **(B)** Differential expression of PD-1, PD-L1, PD-L2, and CTLA4 in the LEXP and HEXP subgroups of the TCGA-uveal melanoma cohort; **(C)** Correlation between GOLM1 and four immune checkpoints; **(D)** Prediction of the response to anti-CTLA4 and anti-PD-1/PD-L1 therapy.

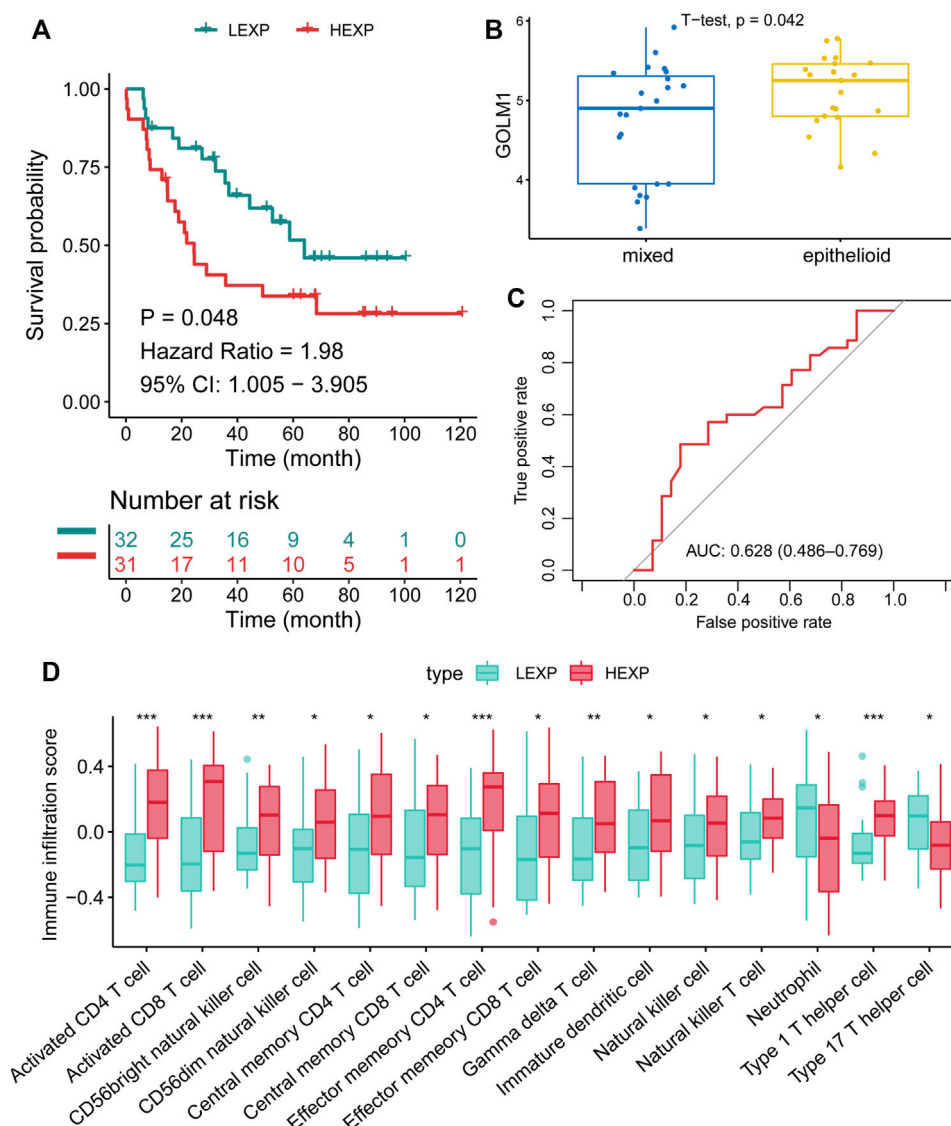


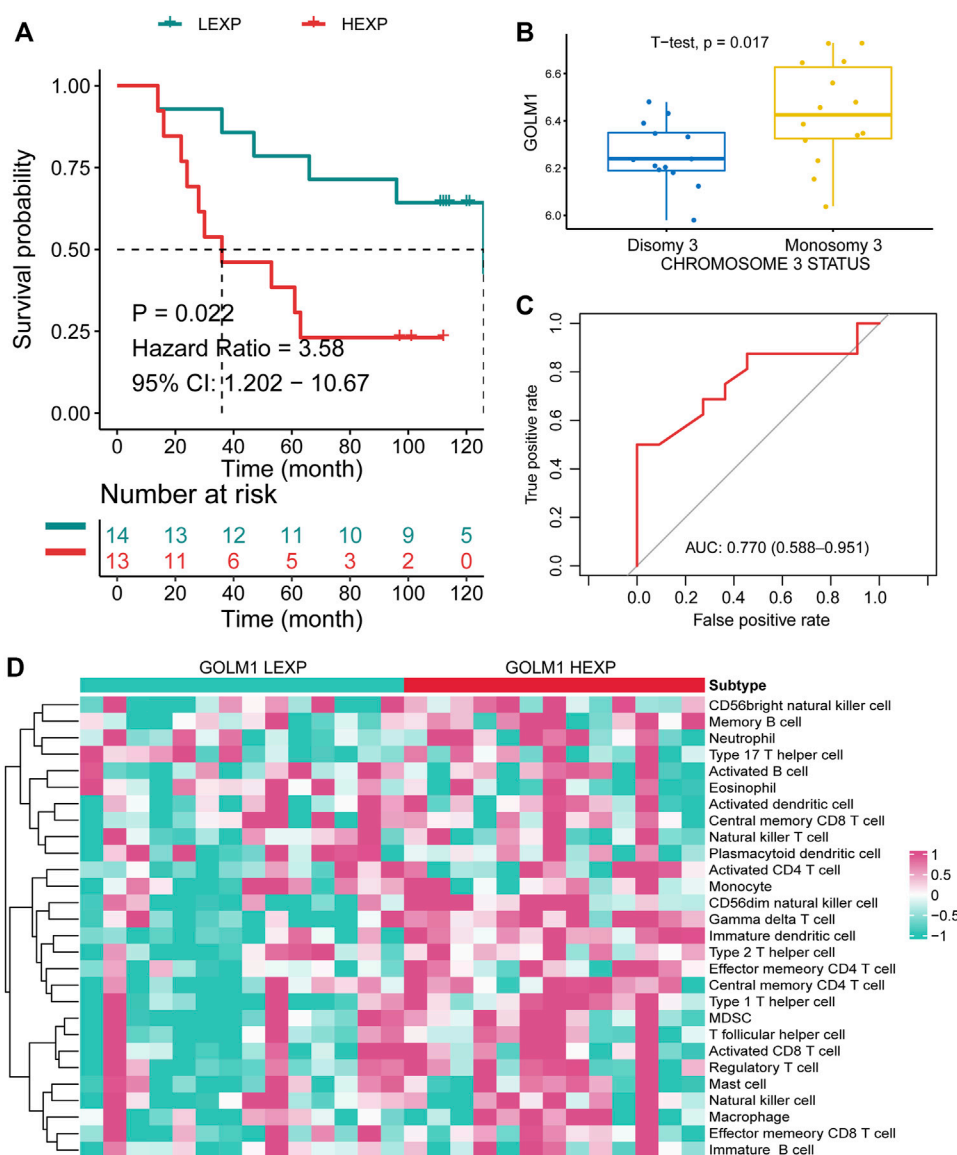
FIGURE 5

Confirmation of the prognostic value of GOLM1 in the GSE21138 cohort. (A) Kaplan–Meier curve showing the polarized overall survival in GOLM1-low (LEXP) and GOLM1-high (HEXP) subgroups; (B) Differential expression of GOLM1 in epithelioid tumors and mixed tumors; (C) Receiver operating characteristic curve showing the prognostic value of GOLM1; (D) Differential infiltration of immunocytes in the LEXP and HEXP groups.

GOLM1 in the epithelioid tumor than in the mixed tumor ( $p = 0.042$ , Figure 5B) was consistent with the worst prognosis of epithelioid UVM. Based on the adjustment of age, sex, and extrascleral extension, we revealed GOLM1 to be an independent prognostic factor ( $p = 0.014$ , HR: 2.681, 95% CI: 1.220–5.888, Table 3). The AUC of GOLM1 in this cohort is 0.628 (95% CI: 0.486–0.769, Figure 5C). In the comparison of tumor-infiltrated immunocytes, the HEXP group of the GSE21138 cohort presented higher infiltrations, especially in the T cells (Figure 5D).

As for GSE84976 cohort, 27 patients were separated into the LEXP ( $n = 14$ ) and HEXP ( $n = 13$ ) subgroups based on the

median value of GOLM1 expression. Patients in the HEXP subgroup displayed a 3.58-fold HR compared with the LEXP subgroup, with a 95% CI of 1.202–10.67 ( $p = 0.022$ , Figure 6A). Monosomy chromosome 3 indicates the poor prognosis of UVM as widely reported, and we observed that patients with monosomy chromosome 3 contained the higher level of GOLM1 than patients with disomy chromosome 3 ( $p = 0.017$ , Figure 6B), which consistent with the above findings. The AUC of GOLM1 in this cohort was 0.770, with 95% CI ranged from 0.588 to 0.951, demonstrating a moderate prognostic value (Figure 6C). We also observed the enrichment of

**FIGURE 6**

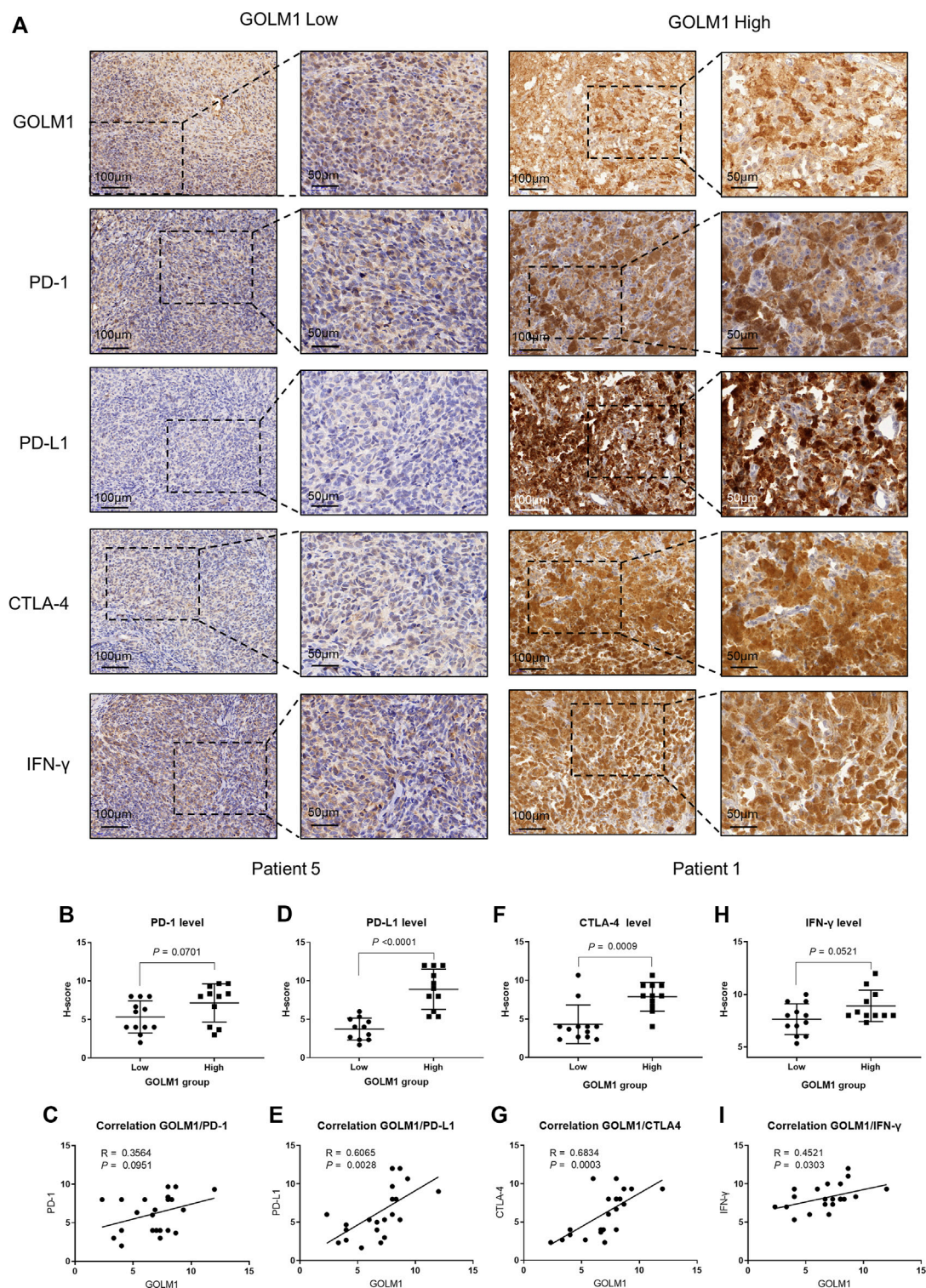
Confirmation of the prognostic value of GOLM1 in the GSE84976 cohort. **(A)** Kaplan–Meier curve showing the polarized overall survival in GOLM1-low (LEXP) and GOLM1-high (HEXP) subgroups; **(B)** Differential expression of GOLM1 in patients with disomy and monosomy chromosome 3; **(C)** Receiver operating characteristic curve showing the prognostic value of GOLM1; **(D)** Differential infiltration of immunocytes in the LEXP and HEXP groups.

immunocytes infiltration in GOLM1 HEXP patients in GSE84976 cohort, especially gamma delta T cell, memory B cell and immature dendritic cell (Figure 6D).

## Validation of the links between GOLM1 and immune markers by IHC

As mentioned above, we revealed the links between GOLM1 and immune checkpoints, as well as immunocytes.

We collected 23 tumor cases and evaluated the protein levels of GOLM1, PD-1, PD-L1, CTLA4, and IFN- $\gamma$  by IHC staining. All the original pictures were put in Supplementary Figure S1, and two representative patients were chosen to display in Figure 7A, patient 1 with positive GOLM1 and patient 5 with negative GOLM1. We first separated the 23 patients into LEXP ( $n = 12$ ) and HEXP ( $n = 11$ ) subgroups according to the H-score. We found that patients belonging to the HEXP subgroup also displayed high levels of PD-1 ( $p = 0.0701$ , Figure 7B), PD-L1 ( $p < 0.001$ , Figure 7D), CTLA-4 ( $p = 0.009$ , Figure 7F), and IFN- $\gamma$  ( $p =$

**FIGURE 7**

Correlation between GOLM1 expression and PD-1, PD-L1, CTLA4, and IFN- $\gamma$ . **(A)** Immunohistochemistry assay revealed the protein expression of GOLM1, PD-1, PD-L1, CTLA4, and IFN- $\gamma$ . **(B,C)** The expression difference of PD-1 among GOLM1-low (LEXP) and GOLM1-high (HEXP) subgroups **(B)** and the correlation between PD-1 and GOLM1 expression **(C)**. **(D,E)** The expression difference of PD-L1 among the LEXP and HEXP groups **(D)** and the correlation between PD-L1 and GOLM1 expression **(E)**. **(F,G)** The expression difference of CTLA4 among the LEXP and HEXP groups **(F)** and the correlation between CTLA4 and GOLM1 expression **(G)**. **(H,I)** The expression difference of IFN- $\gamma$  among the LEXP and HEXP groups **(H)** and the correlation between IFN- $\gamma$  and GOLM1 expression **(I)**.

0.0521, Figure 7H) compared to those with LEXP. Moreover, the increased H-score of GOLM1 was obviously positively correlated with the elevated values of PD-1 ( $R = 0.3564$ ,  $p = 0.0951$ , Figure 7C), PD-L1 ( $R = 0.6065$ ,  $p = 0.0028$ , Figure 7E), CTLA-4 ( $R = 0.6834$ ,  $p = 0.0003$ , Figure 7G) and IFN- $\gamma$  ( $R = 0.4521$ ,  $p = 0.0303$ , Figure 7I). These results from the IHC staining not only supported the new findings revealed in the current study but also suggested that these patients with a high level of GOLM1 commonly had an elevated abundance of immune checkpoint markers, highlighting its potential role in serving as a therapeutic target or immunotherapy indicator.

## Discussion

To identify novel prognostic biomarkers that could be used to reflect the tumor-infiltrated inflammatory/immune microenvironment, we performed the current study. As a result, we found that GOLM1 was able to predict the overall survival outcomes of UVM patients and served as a risk factor independent of clinicopathological features. We analyzed the co-expressed genes of GOLM1, and pathway analysis indicated that these genes mainly belonged to immune-related pathways. We also analyzed the tumor-infiltrated immunocyte proportion differences between the LEXP and HEXP subgroups, with significant variations identified. Tumor microenvironment activity is correlated with the efficacy of immune checkpoint blockade therapy. We performed SubMap analysis and revealed that the UVM patients in the HEXP group were predicted to be more sensitive to PD-1/PD-L1 and CTLA4 therapies than those in the LEXP subgroup. Our findings suggested that GOLM1 was a novel prognostic marker for UVM patients and could be employed to predict immunotherapy treatment efficacy.

The prognostic value of GOLM1 was first tested at the pancancer level, the result of which presented the highest predictive value in thyroid carcinoma, followed by UVM. Upregulation of GOLM1 was observed in multiple cancers and is involved in the mediation of malignant behaviors. Such upregulation in prostate cancer promoted the proliferation and progression of cancer cells by activating PI3K/AKT/mTOR signaling (Yan et al., 2018), thereby being recognized as a biomarker in detecting prostate cancer and evaluating its aggressiveness (Varambally et al., 2008; Li et al., 2012). Similar findings were obtained in hepatocellular carcinoma. Chen et al. (2015) found that abnormal expression of GOLM1 increased the proliferation and progression of hepatocellular carcinoma both *in vitro* and *in vivo*. Furthermore, the prognostic value of GOLM1 in predicting overall survival and recurrence-free survival among hepatocellular carcinoma patients was also proven (Mao et al., 2010; Bao et al., 2013). Moreover, the

oncogenic role and prognostic value of GOLM1 in non-small-cell lung cancer were identified (Zhang et al., 2010; Zhang et al., 2017b). Given the importance of GOLM1 in cancers, its prognostic value, functional role, and potential mechanisms in UVM have aroused wide research interest. In addition to its prognostic value, we performed extensive analyses to reveal the underlying mechanisms regarding the worse survival outcomes of the HEXP subgroup than the LEXP subgroup. Notably, GOLM1 expression was found to be positively associated with the proportions of tumor-infiltrated immunocytes. Patients with higher GOLM1 expression were also predicted to be more sensitive to PD-1/PD-L1 and CTLA4 treatment than those with lower GOLM1 expression. For further validation, we performed an IHC assay and obtained consistent findings that the expression of GOLM1 was positively associated with the expression of PD-1, PD-L1, CTLA4, and IFN- $\gamma$ . This result highlighted its potential role in guiding clinical immune checkpoint blockade therapy.

## Conclusion

In summary, we systematically analyzed the prognostic role of GOLM1 at the pan-cancer level and revealed its high predictive value for the overall survival of UVM. The results also identified it as a risk factor independent of clinicopathological features. The expression of this gene is positively associated with PD-1, PD-L1, CTLA4, and IFN- $\gamma$  expression, and patients with higher expression of GOLM1 are predicted to be more sensitive to anti-CTLA4 as well as anti-PD-1/PD-L1 therapies. These findings offer individual clinical outcome predictions and immune checkpoint blockade therapy guidance for UVM patients.

## 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 the Anhui Medical University Affiliated Hospital.

## Author contributions

XL and NL conceived and designed this study. YY collected the pathological sections. XL did the statistical analyses. XL and

YY wrote the main manuscript. All authors read and approved the manuscript.

## Funding

The authors supported by the Shanghai General Hospital and Anhui Medical University.

## Acknowledgments

We thank NL (The First Affiliated Hospital of Anhui Medical University) for his assistance in conducting this study.

## 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.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1051168/full#supplementary-material>

## References

- Bao, Y. X., Cao, Q., Yang, Y., Mao, R., Xiao, L., Zhang, H., et al. (2013). Expression and prognostic significance of golgiglycoprotein73 (GP73) with epithelial-mesenchymal transition (EMT) related molecules in hepatocellular carcinoma (HCC). *Diagn. Pathol.* 8, 197. doi:10.1186/1746-1596-8-197
- Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462 (7269), 108–112. doi:10.1038/nature08460
- Blanche, P., Dartigues, J. F., and Jacqmin-Gadda, H. (2013). Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat. Med.* 32 (30), 5381–5397. doi:10.1002/sim.5958
- Bronkhorst, I. H., and Jager, M. J. (2012). Uveal melanoma: The inflammatory microenvironment. *J. Innate Immun.* 4 (5–6), 454–462. doi:10.1159/000334576
- Chen, M. H., Jan, Y. H., Chang, P. M., Chuang, Y. J., Yeh, Y. C., Lei, H. J., et al. (2013). Expression of GOLM1 correlates with prognosis in human hepatocellular carcinoma. *Ann. Surg. Oncol.* 20, S616–S624. doi:10.1245/s10434-013-3101-8
- Chen, P. L., Roh, W., Reuben, A., Cooper, Z. A., Spencer, C. N., Prieto, P. A., et al. (2016). Analysis of immune signatures in longitudinal tumor samples yields insight into biomarkers of response and mechanisms of resistance to immune checkpoint blockade. *Cancer Discov.* 6 (8), 827–837. doi:10.1158/2159-8290.Cd-15-1545
- Chen, X., Wang, Y., Tao, J., Shi, Y., Gai, X., Huang, F., et al. (2015). mTORC1 up-regulates GP73 to promote proliferation and migration of hepatocellular carcinoma cells and growth of xenograft tumors in mice. *Gastroenterology* 149 (3), 741–752. doi:10.1053/j.gastro.2015.05.005
- Gezgin, G., Dogrusöz, M., van Essen, T. H., Kroes, W. G. M., Luyten, G. P. M., van der Velden, P. A., et al. (2017). Genetic evolution of uveal melanoma guides the development of an inflammatory microenvironment. *Cancer Immunol. Immunother.* 66 (7), 903–912. doi:10.1007/s00262-017-1991-1
- Hanzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14, 7. doi:10.1186/1471-2105-14-7
- Hoshida, Y., Brunet, J. P., Tamayo, P., Golub, T. R., and Mesirov, J. P. (2007). Subclass mapping: Identifying common subtypes in independent disease data sets. *PLoS One* 2 (11), e1195. doi:10.1371/journal.pone.0001195
- Kim, H. J., Lv, D., Zhang, Y., Peng, T., and Ma, X. (2012). Golgi phosphoprotein 2 in physiology and in diseases. *Cell Biosci.* 2 (1), 31. doi:10.1186/2045-3701-2-31
- Kladney, R. D., Bulla, G. A., Guo, L., Mason, A. L., Tollefson, A. E., Simon, D. J., et al. (2000). GP73, a novel Golgi-localized protein upregulated by viral infection. *Gene* 249 (1–2), 53–65. doi:10.1016/s0378-1119(00)00136-0
- Krishna, Y., McCarthy, C., Kalirai, H., and Coupland, S. E. (2017). Inflammatory cell infiltrates in advanced metastatic uveal melanoma. *Hum. Pathol.* 66, 159–166. doi:10.1016/j.humpath.2017.06.005
- Li, W., Wang, X., Li, B., Lu, J., and Chen, G. (2012). Diagnostic significance of overexpression of Golgi membrane protein 1 in prostate cancer. *Urology* 80 (4), e1–e957. doi:10.1016/j.urology.2012.06.017
- Mao, Y., Yang, H., Xu, H., Lu, X., Sang, X., Du, S., et al. (2010). Golgi protein 73 (GOLPH2) is a valuable serum marker for hepatocellular carcinoma. *Gut* 59 (12), 1687–1693. doi:10.1136/gut.2010.214916
- Qin, Y., Bollin, K., de Macedo, M. P., Carapeto, F., Kim, K. B., Roszik, J., et al. (2020). Immune profiling of uveal melanoma identifies a potential signature associated with response to immunotherapy. *J. Immunother. Cancer* 8 (2), e000960. doi:10.1136/jitc-2020-000960
- Song, T. L., Nairismägi, M. L., Laurensia, Y., Lim, J. Q., Tan, J., Li, Z. M., et al. (2018). Oncogenic activation of the STAT3 pathway drives PD-L1 expression in natural killer/T-cell lymphoma. *Blood* 132 (11), 1146–1158. doi:10.1182/blood-2018-01-829424
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102
- Sun, Y., Yang, H., Mao, Y., Xu, H., Zhang, J., Li, G., et al. (2011). Increased Golgi protein 73 expression in hepatocellular carcinoma tissue correlates with tumor aggression but not survival. *J. Gastroenterol. Hepatol.* 26 (7), 1207–1212. doi:10.1111/j.1440-1746.2011.06733.x
- Varambally, S., Laxman, B., Mehra, R., Cao, Q., Dhanasekaran, S. M., Tomlins, S. A., et al. (2008). Golgi protein GOLM1 is a tissue and urine biomarker of prostate cancer. *Neoplasia* 10 (11), 1285–1294. doi:10.1593/neo.08922
- Wölfe, S. J., Strebovsky, J., Bartz, H., Sähr, A., Arnold, C., Kaiser, C., et al. (2011). PD-L1 expression on tolerogenic APCs is controlled by STAT-3. *Eur. J. Immunol.* 41 (2), 413–424. doi:10.1002/eji.201040979
- Yan, G., Ru, Y., Wu, K., Yan, F., Wang, Q., Wang, J., et al. (2018). GOLM1 promotes prostate cancer progression through activating PI3K-AKT-mTOR signaling. *Prostate* 78 (3), 166–177. doi:10.1002/pros.23461
- Yan, J., Zhou, B., Guo, L., Chen, Z., Zhang, B., Liu, S., et al. (2020). GOLM1 upregulates expression of PD-L1 through EGFR/STAT3 pathway in hepatocellular carcinoma. *Am. J. Cancer Res.* 10 (11), 3705–3720.
- Yin, Y., Xu, L., Chang, Y., Zeng, T., Chen, X., Wang, A., et al. (2019). N-Myc promotes therapeutic resistance development of neuroendocrine prostate cancer by

differentially regulating miR-421/ATM pathway. *Mol. Cancer* 18 (1), 11. doi:10.1186/s12943-019-0941-2

Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4, 2612. doi:10.1038/ncomms3612

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Zhang, F., Gu, Y., Li, X., Wang, W., He, J., and Peng, T. (2010). Up-regulated Golgi phosphoprotein 2 (GOLPH2) expression in lung adenocarcinoma tissue. *Clin. Biochem.* 43 (12), 983–991. doi:10.1016/j.clinbiochem.2010.05.010

Zhang, W., Zhang, J., Zhang, Z., Guo, Y., Wu, Y., Wang, R., et al. (2019). Overexpression of indoleamine 2, 3-dioxygenase 1 promotes epithelial-mesenchymal transition by activation of the IL-6/STAT3/PD-L1 pathway in bladder cancer. *Transl. Oncol.* 12 (3), 485–492. doi:10.1016/j.tranon.2018.11.012

Zhang, X., Zhu, C., Wang, T., Jiang, H., Ren, Y., Zhang, Q., et al. (2017a). GP73 represses host innate immune response to promote virus replication by facilitating MAVS and TRAF6 degradation. *PLoS Pathog.* 13 (4), e1006321. doi:10.1371/journal.ppat.1006321

Zhang, Y., Hu, W., Wang, L., Han, B., Lin, R., and Wei, N. (2017b). Association of GOLPH2 expression with survival in non-small-cell lung cancer: Clinical implications and biological validation. *Biomark. Med.* 11 (11), 967–977. doi:10.2217/bmm-2017-0199



## OPEN ACCESS

## EDITED BY

Xiaofan Lu,  
INSERM U964 Institut de Génétique et  
de Biologie Moléculaire et Cellulaire  
(IGBMC), France

## REVIEWED BY

Hansheng Wu,  
First Affiliated Hospital of Shantou  
University Medical College, China  
Shuai Shao,  
The Ohio State University, United States  
Yundong Zhou,  
Shanghai Medical Innovation Fusion  
Biomedical Research Center, China

## \*CORRESPONDENCE

Ruilin Sun,  
sunruilin213@126.com

## SPECIALTY SECTION

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

RECEIVED 17 August 2022

ACCEPTED 31 October 2022

PUBLISHED 24 November 2022

## CITATION

Gong J, Yang J, He Y, Chen X, Yang G  
and Sun R (2022), Construction of m7G  
subtype classification on heterogeneity  
of sepsis.

*Front. Genet.* 13:1021770.

doi: 10.3389/fgene.2022.1021770

## COPYRIGHT

© 2022 Gong, Yang, He, Chen, Yang and  
Sun. This is an open-access article  
distributed under the terms of the  
[Creative Commons Attribution License](#)  
(CC BY). The use, distribution or  
reproduction in other forums is  
permitted, provided the original  
author(s) 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.

# Construction of m7G subtype classification on heterogeneity of sepsis

Jinru Gong<sup>1</sup>, Jiasheng Yang<sup>1</sup>, Yaowei He<sup>1</sup>, Xiaoxuan Chen<sup>1</sup>,  
Guangyu Yang<sup>1,2</sup> and Ruilin Sun<sup>1\*</sup>

<sup>1</sup>Department of Pulmonary and Critical Care Medicine, Guangdong Second Provincial General Hospital, Guangzhou, China, <sup>2</sup>The Second School of Clinical Medicine, Southern Medical University, Guangzhou, China

Sepsis is a highly heterogeneous disease and a major factor in increasing mortality from infection. N7-Methylguanosine (m7G) is a widely RNA modification in eukaryotes, which involved in regulation of different biological processes. Researchers have found that m7G methylation contributes to a variety of human diseases, but its research in sepsis is still limited. Here, we aim to establish the molecular classification of m7G gene-related sepsis, reveal its heterogeneity and explore the underlying mechanism. We first identified eight m7G related prognostic genes, and identified two different molecular subtypes of sepsis through Consensus Clustering. Among them, the prognosis of C2 subtype is worse than that of C1 subtype. The signal pathways enriched by the two subtypes were analyzed by ssGSEA, and the results showed that the amino acid metabolism activity of C2 subtype was more active than that of C1 subtype. In addition, the difference of immune microenvironment among different subtypes was explored through CIBERSORT algorithm, and the results showed that the contents of macrophages M0 and NK cells activated were significantly increased in C2 subtype, while the content of NK cells resting decreased significantly in C2 subtype. We further explored the relationship between immune regulatory genes and inflammation related genes between C2 subtype and C1 subtype, and found that C2 subtype showed higher expression of immune regulatory genes and inflammation related genes. Finally, we screened the key genes in sepsis by WGCNA analysis, namely NUDT4 and PARN, and verified their expression patterns in sepsis in the datasets GSE131761 and GSE65682. The RT-PCR test further confirmed the increased expression of NUDT4 in sepsis patients. In conclusion, sepsis clustering based on eight m7G-related genes can well distinguish the heterogeneity of sepsis patients and help guide the personalized treatment of sepsis patients.

## KEYWORDS

sepsis, N7-methylguanosine, heterogeneity, prognosis, integrated analysis

## Introduction

Sepsis is a highly heterogeneous disease characterized by life-threatening organ dysfunction caused by dysregulation of host response to infection, which is the main factor to increase infection mortality (Singer et al., 2016). Globally, sepsis remains a serious public health issue. Sepsis affects about 20 million cases per year, and the mortality rate is about 26% (Fleischmann et al., 2016). The treatment of sepsis mainly includes the combination of early antibiotics, fluid resuscitation and symptomatic treatment of vasopressors, which can significantly improve the symptoms of patients. The risk of infection, however, increases with long-term hospitalization and invasive operation. Conversely, frequent drug treatment can easily lead to multiple drug-resistant bacterial infections, leading to multiple organ failure and other symptoms, which poses severe challenges to the treatment and management of sepsis (Zhang and Ning, 2021). In addition, due to the high heterogeneity of sepsis, there is no proven standard or disease classification that can effectively guide the treatment of patients (Purcarea and Sovaila, 2020). Therefore, it is essential to explore the molecular mechanism of disease progression and the classification of related diseases for the treatment of sepsis.

RNA methylation regulates gene expression at the post transcriptional level and is an epigenetic regulatory mode. At present, more than 150 RNA methylation modifications have been found in eukaryotes. Among them, m7G methylation modification is one of the most common base modifications in post transcriptional regulation, which widely occurs in the 5' cap region of tRNA, rRNA and eukaryotic mRNA (Liu et al., 2017; Roundtree et al., 2017). m7G modification is involved in the regulation of various processes, such as mRNA transcription, splicing and translation. More and more studies have found that the occurrence of many human diseases is related to the methylation modification of m7G. Bing et al. found that there were significant differences in m7G mRNA modification in drug-resistant AML cells, and the low methylated m7G modification level was significantly enriched in ABC transporter related mRNA, suggesting that the down-regulation of m7G methylation can actively regulate ABC transporter related genes in AML cells, leading to drug resistance in AML (Zhang et al., 2022). Chen et al., 2022 confirmed that m7G methyltransferase METTL1 promotes the development of HNSCC by regulating PI3K/Akt signaling pathway, and changes the immune microenvironment and intercellular communication between HNSCC tumors and stroma. Epigenetic regulation plays a central role in the pathogenesis of sepsis (Binnie et al., 2020), but the role and mechanism of m7G in sepsis remain unclear.

In order to solve the above problems, this study intends to explore the role and potential molecular mechanism of m7G related genes in sepsis, classify sepsis patients into different

subtypes according to m7G-related genes. The pathway enrichment and immune microenvironment were explored to explain the potential mechanism of heterogeneity of different subtypes, and the key regulatory genes NUDT4 and PARN were identified, which are significantly correlated with the prognosis of septic shock patients. This study provides the basis for the treatment and management of septic patients.

## Materials and methods

### Data acquired

GEO database (Barrett et al., 2013) (<https://www.ncbi.nlm.nih.gov/>) is a gene expression database maintained by NCBI, which stores gene expression data uploaded by research institutions around the world. In this study, GSE65682 dataset was downloaded from GEO database, including 42 healthy control samples and 760 sepsis samples. We extracted 479 samples with complete survival data for classification in sepsis patients. We downloaded GSE131761 dataset from GEO database, including 15 healthy control samples and 81 septic shock samples, to further explore the expression of m7G related regulatory genes. m7G gene was obtained through previous literature.

### Prognostic genes screening and consensus clustering

In order to further classify the prognosis of sepsis patients, we selected m7G-related genes for consensus clustering (Swift et al., 2004). The selection of prognostic genes was based on Kaplan Meier (KM) survival analysis (Goel et al., 2010) ( $p < 0.05$ ). After screening the prognostic m7G-related genes, the consensus clustering was used for subtype clustering. About 80% of the samples were analyzed in each iteration, and a total of 50 iterations were performed. The optimal cluster number is determined by the cumulative distribution function (CDF) curve of the consistency score, the clear difference between groups in the consistency matrix heatmap, and the characteristics of the consistency cumulative distribution function map. KM curve will be used to evaluate the prognosis of different m7G sepsis subtypes with a cutoff value of  $p < 0.05$ .

### WGCNA network construction

WGCNA (Langfelder and Horvath, 2008) is a method to summarize gene expression data into different coexpression modules, which can be used to explore the relationship between different modules and the correlation between modules and clinical symptoms. We took the genes with

the highest variance of 5,000 within the gene expression data to construct the coexpression network.  $\beta$  is a soft threshold, which is related to the independence and average connectivity of modules. The soft threshold of this study is setting as 12. The topological overlap degree matrix (TOM) represents the overlap of network neighbors. The hierarchical clustering method is used to construct the cluster tree structure of TOMmatrix, and the dynamic tree cutting method of pheatmap package is used to cluster graphs. Based on the subtype information of different m7G-related sepsis, the correlation between different modules and clinical phenotype was evaluated. The relationship between the module and the clinical phenotype was calculated by Pearson correlation test.  $p < 0.05$  was defined as significant correlation. The module with the largest correlation was selected for subsequent analysis.

## Go and KEGG functional enrichment

ssGSEA (Barbie et al., 2009) is an extension of the GSEA method, which allows the definition of an enrichment score that represents the absolute enrichment of gene sets in each sample within a given dataset. In this study, the ssGSEA algorithm of GSVA package was used to evaluate the pathway levels of GO and KEGG in different subtypes. The GO items included three categories: biological process (BP), molecular function (MF) and cellular component (CC). The background gene set used for the analysis was from the GSVA database (Liberzon et al., 2015) (C2 and C5).

## Immune gene correlation analysis

CIBERSORT (Chen et al., 2018) deconvolution algorithm can estimate the composition and relative abundance of immune cells in mixed cells based on gene transcriptome data. In this study, CIBERSORT LM22 was used to estimate the expression matrix of immune cell characteristics, and then the CIBERSORT algorithm was used to quantify the relative proportion of immune cell infiltration in different sepsis subtypes, and to compare the differences of immune cells between these two groups. The correlation analysis between the key genes and the content of immune cells is carried out to evaluate the relationship between the key genes and immune cells. It is considered statistically significant if  $p$  is less than 0.05.

## RT-PCR validation of hub genes

A total of 10 participants were recruited from the Department of Pulmonary and Critical Care Medicine of

the Guangdong Second Provincial General Hospital, including 5 sepsis cases and 5 non-septic patients. The study was carried out in accordance with the Helsinki Declaration and was approved by the Ethics Committee of the Second People's Hospital of Guangdong Province. The whole blood samples of each case were collected into tubes with EDTA. Total RNA was isolated using the TRIzol reagent (TIANGEN, CHINA) in line with the manufacturer's instructions. The real-time PCR was conducted using AOPR-1200 detection kit (Genecopies, China). The sequences of primers for the indicated genes were as follows: GAPDH forward (F), CAAGAGCACAAGAGG AAGAGAG and reverse (R), CTACATGGCAACTGTGAG GAG; NUDT4 F, CCTCCTAAAGTGCTGGGATTAC and R, CAAAGTCCTGGGAGAGAAGAAA; PARN F, CAAAGT GTACCAGGCCATAGAG, and R, CTGAAGGTCCATCAC TGATTCC.

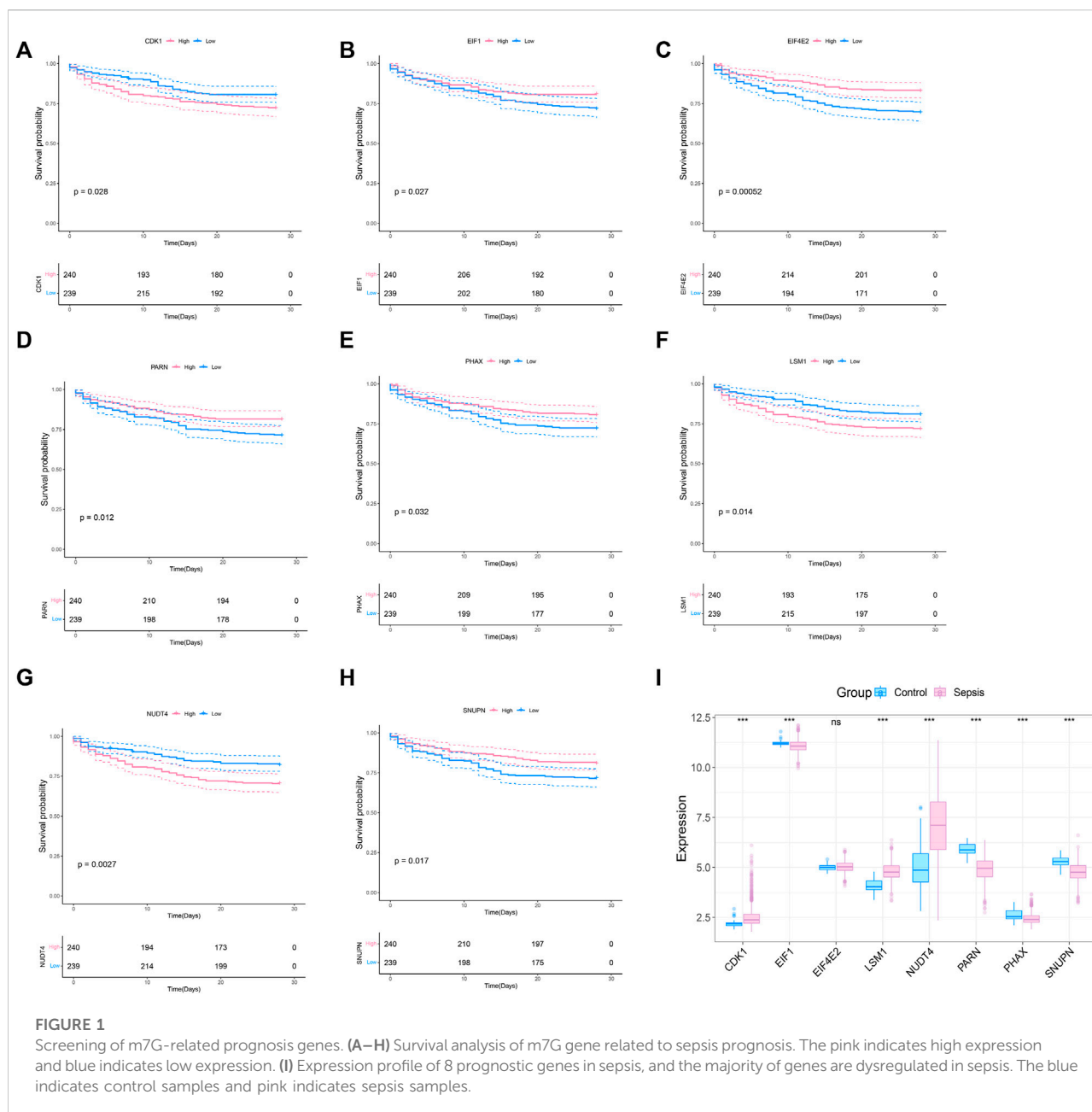
## Statistic method

R program (4.1.2) was used to analyze data and draw diagrams. The differences between subgroups were tested by Wilcoxon test, and all correlations were calculated by Pearson method. Survival curves were generated by Kaplan-Meier method and compared by log rank.  $p < 0.05$  with statistical significance.

## Result

### Prognosis of m7G related genes in sepsis

To explore whether the molecular classification of the m7G related genes in sepsis can explain the heterogeneity of sepsis patients, we acquired the expression data of GSE65682 from the GEO database, with a total of 802 samples, including 42 healthy control samples and 760 sepsis samples. We extracted 479 samples with complete survival status for follow-up analysis. 42 m7G related genes were obtained from the previous literature, and then the list was limited to the genes with available RNA expression data in GSE65682, leaving 33 m7G regulated genes. They are EIF4E2, NUDT4, PARN, LSM1, SNUPN, EIF1, CDK1, PHAX, CYFIP1, CCNB1, NCBP1, EIF4E3, DCPS, NSUN2, EIF4A1, JUND, LARP1, NUDT3, APAF1, EIF3D, XPO1, WDR4, NCBP2, GEMIN5, EIF4E, IFIT5, NUDT16, IPO8, METTL1, DCP2, TGS1, EIF4G1, and EIF4G3. We performed survival analysis on these 33 genes and screened the prognosis-related m7G genes. The results showed that a total of 8 prognosis related genes were screened ( $p < 0.05$ ) (Figures 1A–G). In addition, we analyzed the expression of these 8 m7G genes in normal

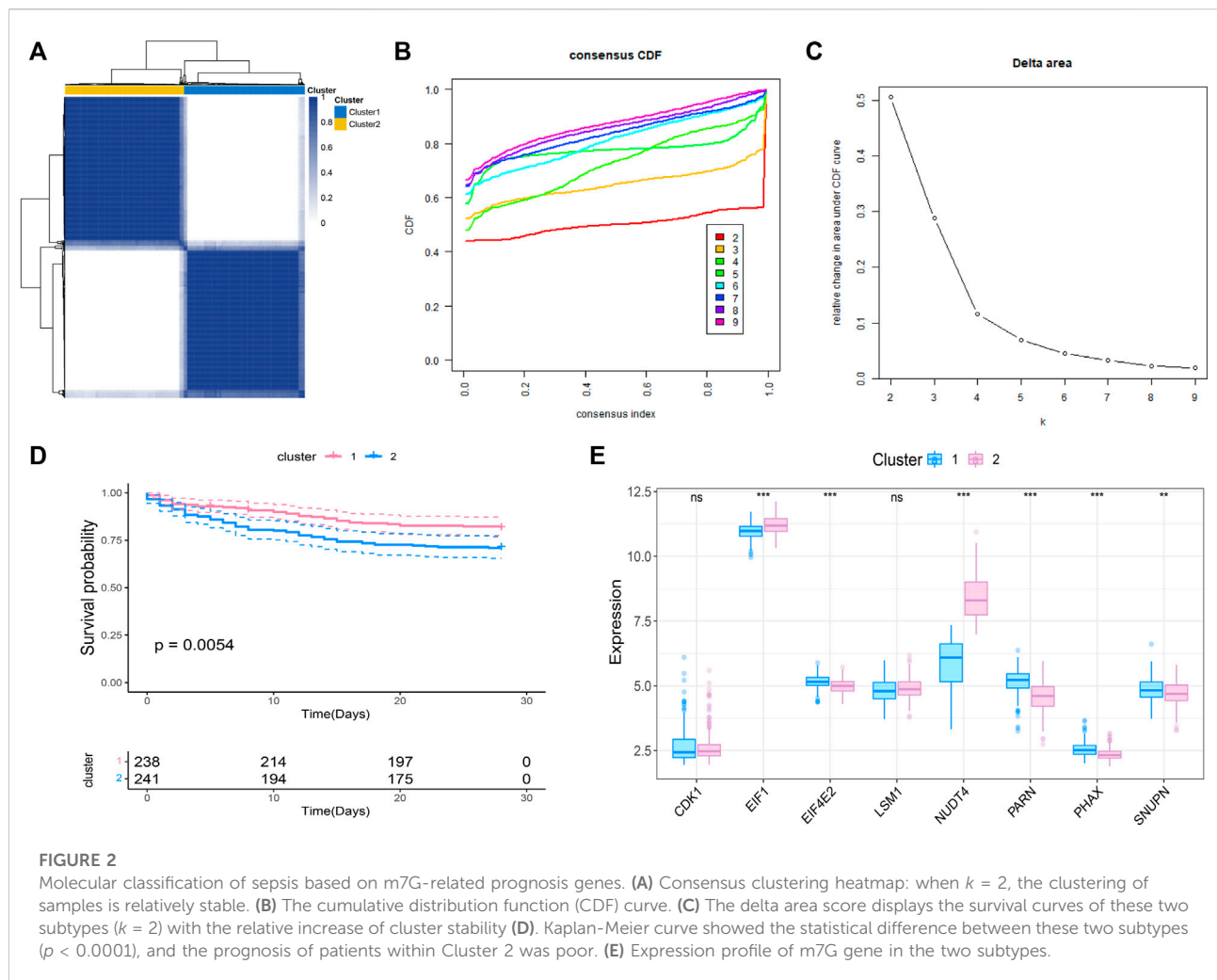


samples and sepsis samples, and the results showed that 7 genes showed different expression in two groups of samples (Figure 1I).

## Molecular classification of prognosis-related m7G genes in sepsis

We further conducted the consensus clustering and performed molecular classification of GSE65682 based on the expression of the m7G related prognostic genes. The

results showed that the boundary between the two subtypes of the sample was clear when  $k = 2$ , so the sepsis was divided into two clusters (Figures 2A–C). In addition, KM survival analysis showed that the survival of these two clusters was significantly different (Figure 2D), suggesting that there were different survival outcomes between the two subgroups. Therefore, it is particularly important to further explore the molecular characteristics of the two subgroups. We also analyzed the expression differences of 8 genes in the two subtypes, and the results are shown in the figure (Figure 2E).



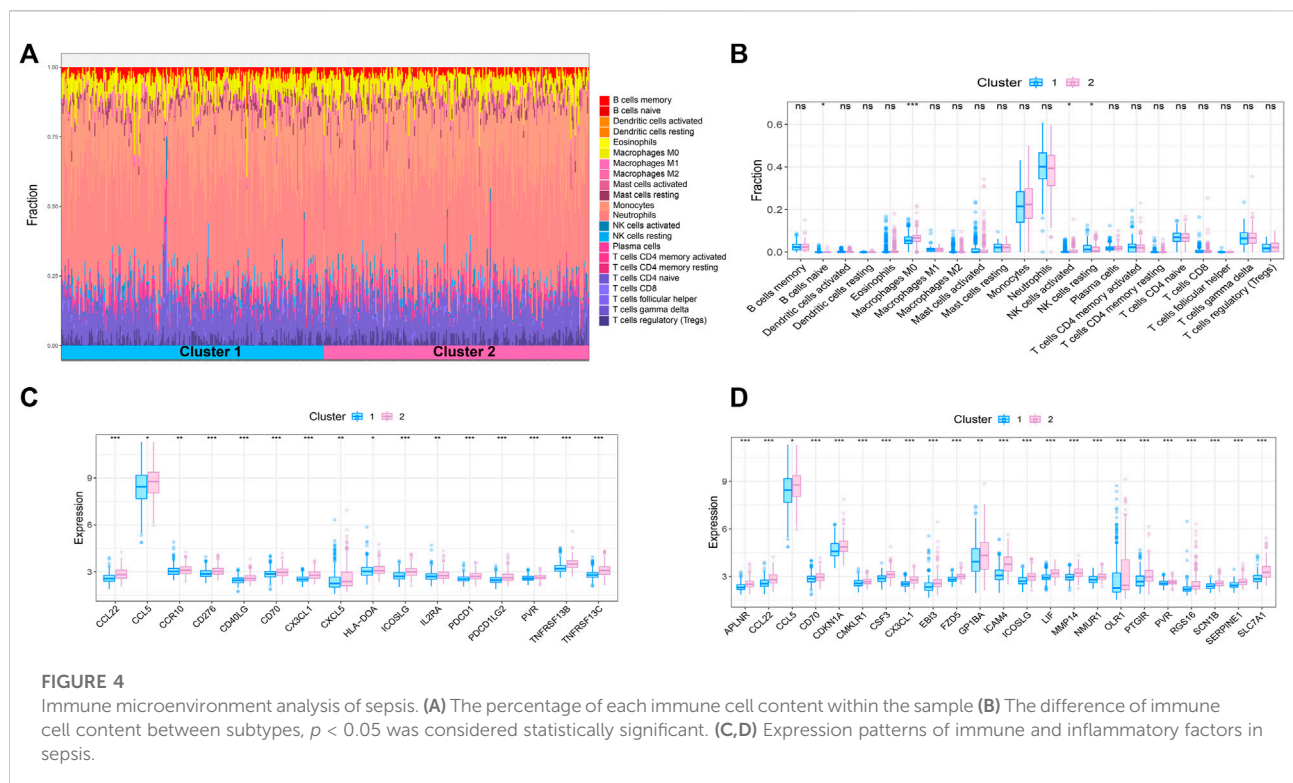
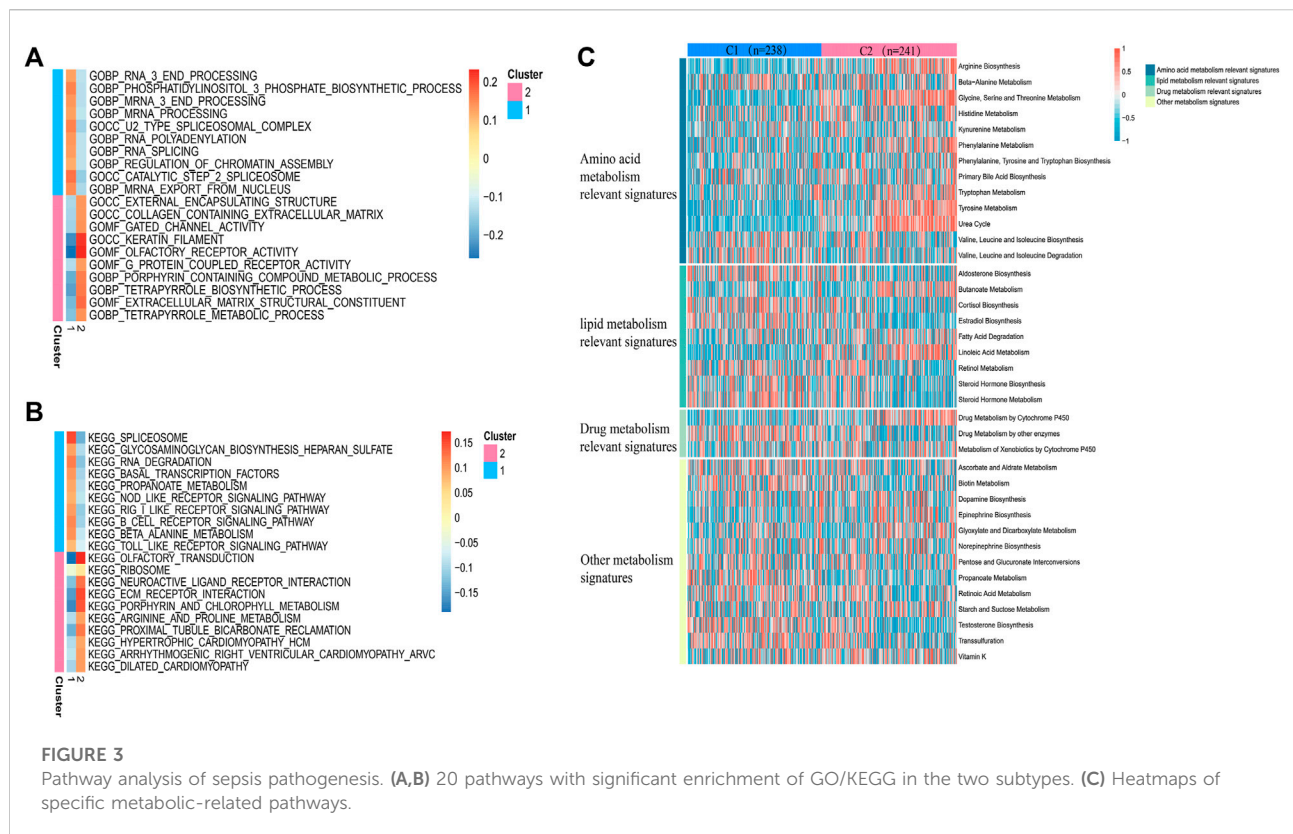
## Functional enrichment among subtypes

In view of the different survival outcomes between C1 subtype and C2 subtype, this study will use ssGSEA to explore the molecular mechanism of different subtypes and clarify the reasons for the different survival outcomes of these patients. We selected the 20 most representative pathways for cluster 1 and cluster 2 to make the visualization, which revealed different pathways enriched in each subtype (Figures 3A,B). The results of GO analysis showed that cluster 2 was significantly related to the pathway of porphyrin containing combined metabolic process, tetrapyrrole biological process extra cellular matrix structural construct, tetrapyrrole metabolic process. KEGG analysis showed that cluster 2 was significantly correlated with ECM receptor interaction, porphyrin and chlorophyll metabolism, and arginine and proline metabolism pathways. The above results suggest that the subtype differences may be closely related to the metabolic activities. We further

quantified four different metabolic activities, and the results suggest that the amino acid metabolic activities in patients with C2 subtype are more active than those of C1 subtype (Figure 3C).

## Heterogeneity of immune status among m7G subtypes

Sepsis is an acute organ dysfunction syndrome caused by physiological and pathophysiological responses to infection. Immune dysfunction is related to abnormal coagulation, endothelial and epithelial barrier destruction, and leads to changes in vascular activity of multiple organ dysfunction. Therefore, this study intends to further explore the difference of immune microenvironment among different subtypes through CIBERSORT algorithm. The results showed that compared with cluster 1, the contents of macrophages M0 and NK cells activated in cluster 2 were



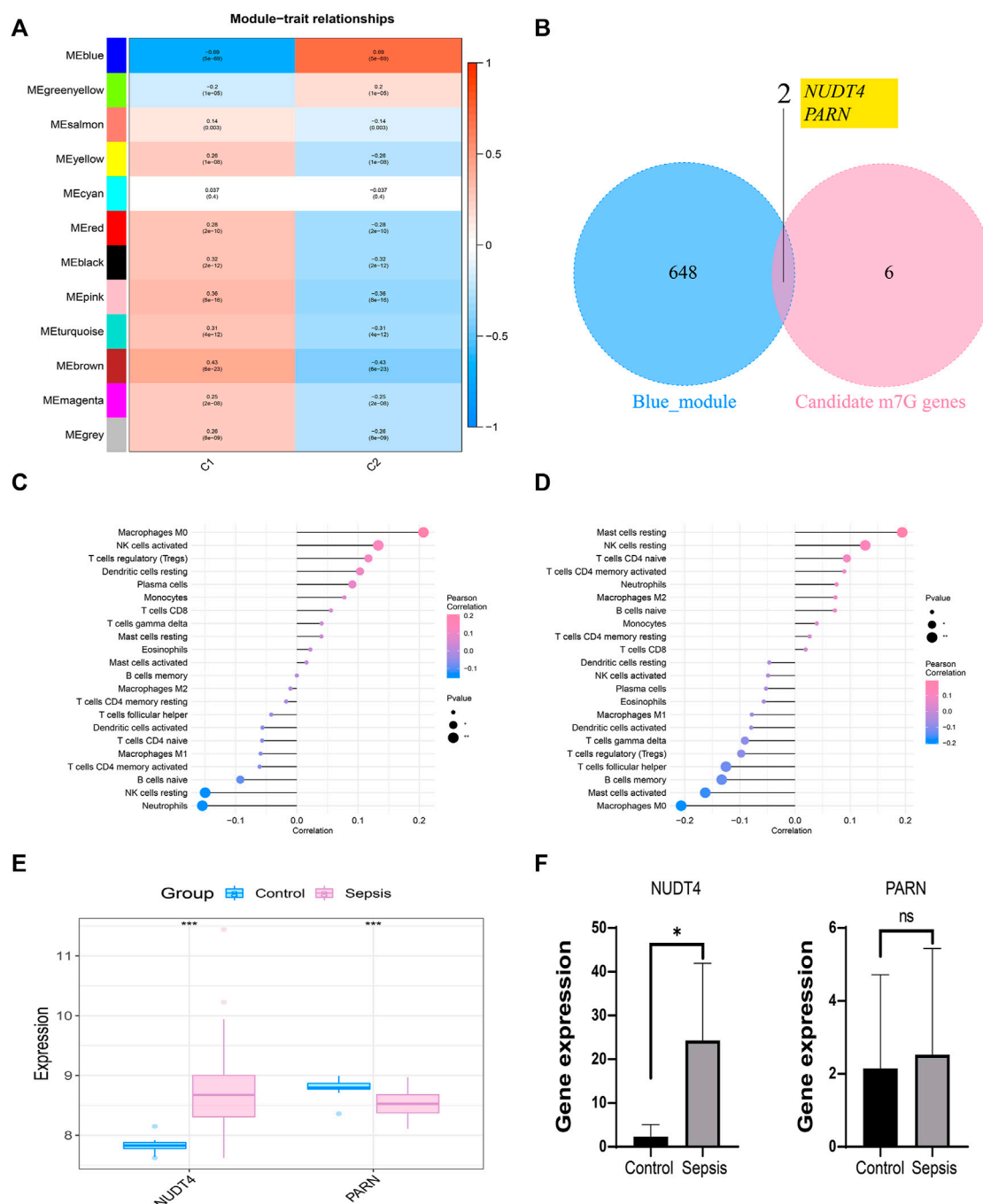
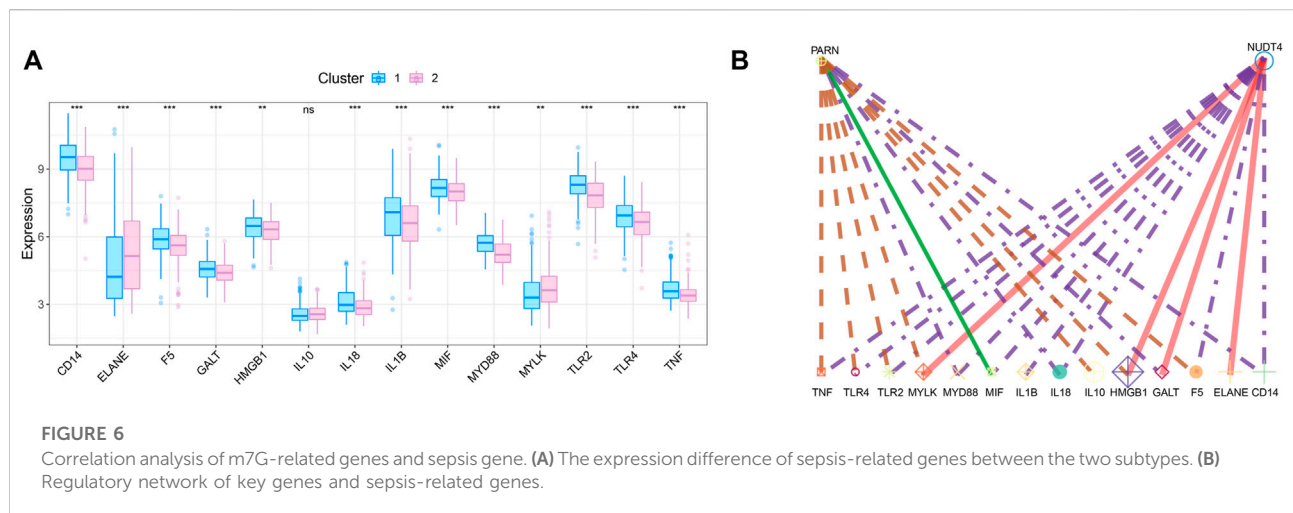


FIGURE 5

Construction of weighted co-expression network and identification of key genes. (A) The heatmap of the correlation between the module characteristic genes and the clinical of sepsis subtypes. The blue module with the highest correlation was selected for subsequent analysis. (B) Identification of key genes in sepsis. (C) Correlation between NUDT4 and immune cells. (D) Correlation between PARN and immune cells. (E) Verification of the expression of key genes in external validation dataset. (F) The RT-PCR validation of key genes expression.

significantly increased, while the content of NK cells resting was significantly decreased in cluster 2 (Figures 4A,B). In addition, some immune related genes and inflammation related genes were obtained from TISIDB database and

GSEA database to verify the expression of immune genes and inflammation genes in the two subtypes, respectively. The results showed that the genes were up-regulated in cluster 2 (Figures 4C,D).



## Construction of WGCNA network based on molecular classification and RT-PCR validation of hub genes

We took the classification status of two different sepsis subpopulations as clinical symptoms, further constructed WGCNA network according to the expression profile data of GSE65682, screened the key genes affecting sepsis progression, and explored the sepsis related gene regulatory network. The soft threshold of WGCNA network was set to 12, and 12 gene modules were co classified. It was found that the blue module had the highest correlation with sepsis subtypes [ $\text{cor} = 0.69$ ,  $p = (5e - 69)$ ] (Figure 5A). We intersected 650 genes of the blue module with 8 prognosis related m7G genes to obtain 2 intersection genes, which are NUDT4 and PARN respectively (Figure 5B). The two genes are strongly correlated with the content of immune cells, among which NUDT4 is positively correlated with Macrophages M0, NK cells activated, T cells regulatory (Tregs), Dendritic cells resting, Plasma cells, and negatively correlated with Neutrophils, NK cells resting, and B cells naive (Figure 5C). PARN is positively correlated with Mast cells resting, NK cells resting, T cells CD4 naive, and negatively correlated with macrophages M0, Mast cells activated, etc. (Figure 5D). We also explored the expression of two key genes through the GSE131761 and GSE65682 datasets, and the results showed that the expression trends of NUDT4 and PARN were consistent in the two datasets (Figure 5E). Additionally, we collected samples from sepsis patients and control cases to explore the expression pattern of hub genes. The results showed that the expression of NUDT4 was significantly increased in sepsis patients (Figure 5F).

## Correlation between key genes and disease regulating genes

We obtained the disease regulatory genes of sepsis according to GENECARD database, and selected the top 20 genes of relevance for difference analysis (Figure 6A). In order to explore the relationship between key genes and sepsis disease regulation, we performed correlation analysis on key genes and 20 disease regulation genes. The regulatory network of m7G regulatory genes and sepsis related genes is shown in Figure 6B.

## Discussion

Sepsis is a disease of multiple organ dysfunction caused by the rapid response of the host to infection (Huang et al., 2019). Worldwide, sepsis affects about 30 million people every year and is one of the main causes of death of critically ill patients. In hospitalized patients, any infection may lead to sepsis, and its incidence rate is about 1%–2% (Huang et al., 2019). Therefore, its treatment cost is also the highest (Rocheteau et al., 2015). In recent years, sepsis related research has been developing, but its incidence rate and mortality are still increasing. Sepsis patients are also facing many adverse effects such as physical and psychological (Iwashyna et al., 2012; Gaieski et al., 2013). In addition, due to the different clinical manifestations of sepsis and the obvious heterogeneity among patients (Huang et al., 2019), there are still many challenges in the diagnosis, treatment and management of sepsis.

Epigenetic regulation plays a key role in the occurrence and development of sepsis. Epigenetic changes mainly refer to the changes in gene expression caused by the body in response to external stimuli, which are the payment of environmental factors

and genetic factors. After the patients with sepsis are stimulated by external infectious factors, the body will have an inflammatory storm. The host pathogen interaction will lead to the epigenetic changes of the host key inflammatory regulatory genes (Binnie et al., 2020). In some animal models, treatment with epigenetic modifiers can alleviate sepsis related organ damage and improve its survival (Cao et al., 2014; Shih et al., 2016). m7G is one of epigenetic regulation. It is known that m7G regulation plays an important role in a variety of human diseases (Malbec et al., 2019; Luo et al., 2022), but its role in septic shock is still unclear. We previously analyzed the potential role of m7G in septic shock through the GEO dataset and found that 8 m7G related genes were closely related to the prognosis of septic shock patients. Based on these 8 m7G related genes, two different subtypes (C1 and C2) were identified by consensus clustering, and it was found that patients with C2 subtype had poor prognosis. In addition, the results of pathway analysis showed that the amino acid metabolism of C2 subtype was more active than that of C1 subtype. The results of immune infiltration analysis showed that the contents of macrophages M0 and NK cells activated in C2 subtype were significantly increased, while the content of NK cells resting decreased significantly in C2 subtype. Additionally, C2 patients had higher expression levels of immune related genes and inflammation related genes. Finally, we screened out two genes that were significantly associated with sepsis prognosis, NUDT4 and PARN.

Systemic inflammatory response syndrome sepsis syndrome is associated with hypermetabolism, increased oxygen consumption and energy consumption, activation of peripheral protein catabolism, especially enhanced metabolic activity in the liver and viscera (Dahn et al., 1987; Barton and Cerra, 1989). Under physiological conditions, the liver mainly synthesizes constituent proteins, such as albumin and transferrin, while in sepsis, the synthesis changes from constituent proteins to acute phase proteins, including procalcitonin, C-reactive protein, complement factor, binding globin  $\alpha$  2-macroglobulin and  $\alpha$  1-acid glycoprotein (Lang et al., 2007; Remick, 2007). In addition, in sepsis, citrulline synthesis is reduced, but the decomposition is increased, and the progressive decrease of citrulline content in the body is one of the reasons for the suppression of macrophage function. Citrulline supplementation can improve the synthesis of endogenous arginine and NO, thus improving the prognosis of sepsis (Xiao et al., 2015). Studies have shown that the level of branched chain amino acids is related to the severity of sepsis, and can predict the risk of death of sepsis patients in ICU (Reisinger et al., 2021). Our study also suggests that the amino acid metabolism level of C2 subtype patients is significantly higher than that of C1 subtype. The poor prognosis of C2 subtype patients may be related to the amino acid metabolism level *in vivo*. On the other hand, changes in lipid metabolism during sepsis are protective reactions against infection, and changes in lipid mass spectrometry are directly

related to inflammation (Bermudes et al., 2018). Lipoproteins have the ability to bind and neutralize toxic bacterial substances that regulate cytokine production during inflammation, thereby weakening host response (Murch et al., 2007). Our results also revealed changes in lipid metabolism pathways between different subtypes.

Sepsis is an acute organ dysfunction syndrome caused by physiological and pathophysiological responses to infection. Immune dysfunction may trigger coagulation abnormalities, endothelial and epithelial barrier disruption, and ultimately vascular activity changes leading to multiple organ dysfunction (Kotas and Matthay, 2018). However, the specific cellular and molecular pathways responsible are not fully understood. Our study showed that the M0 level of C2 subtype macrophages was significantly higher than that of C1 subtypes. Macrophages, as the first line of defense against pathogens, are rapidly activated by alveolar macrophages in the early stage to fight against infection and promote the regression of inflammation in the later stage (Hussell and Bell, 2014). Among them, IFN- $\gamma$  can enhance the release of IL-1, IL-6, IL-8, and tumor necrosis factor (TNF), IL-4, IL-13, and IL-10 from monocytes/macrophages exposed to LPS, and inhibit the production of pro-inflammatory factors. And a large number of pro-inflammatory and anti-inflammatory production are related to the occurrence of sepsis (Cavaillon and Adib-Conquy, 2005). In addition, the balance between lung tissue cells and peripheral blood mononuclear cell-derived macrophages may be an important marker of inflammatory balance during lung infection. In patients with severe coronavirus infection, peripheral monocytes/macrophages increased significantly while alveolar macrophages decreased significantly, suggesting that the abnormal number of peripheral monocytes/macrophages and alveolar macrophages may also be an important cause of immune imbalance (Liao et al., 2020). Therefore, in addition to focusing on the production of cytokines by macrophages to participate in the pathogenesis of sepsis, macrophages in different positions also need to be focused. Interestingly, recent studies have shown that M0 macrophages do not conform to standard M1 or M2 models, and they are more similar to M2 macrophages, which may represent another type of TAM. M0 macrophages are one of the cell subsets closely related to the poor prognosis of breast cancer (Ali et al., 2016), prostate cancer (Jairath et al., 2020) and lung adenocarcinoma (Liu et al., 2017). Our results also suggest that M0 macrophages are significantly increased in C2 subtypes with poor prognosis in sepsis. On the other hand, NK cell resting was also significantly different between C1 and C2 subtypes. NK cells are abundant in tissues such as lung, liver, spleen and blood (Grégoire et al., 2007). IFN- $\gamma$ , GM-CSF and TNF- $\alpha$  are major cytokines produced by activated NK cells (Huntington et al., 2007), and have a protective effect during infection but a deleterious effect during aseptic or infectious systemic inflammatory response syndrome. In addition, studies

have also shown that NK cell function can be affected by IL-10 and TGF- $\beta$ -1 (Scott et al., 2006; Ralainirina et al., 2007).

Our study also identified the key genes that significantly affect the prognosis in sepsis, namely NUDT and PARN. NUDT's catalyze the hydrolysis of various nucleoside pyrophosphates associated with other amino acid moieties (Bessman, 2019). In the process of eliminating hydrolytic substrates, NUDT plays a signaling and regulatory role in metabolism (Mildvan et al., 2005). Studies have shown that NUDT4 can be used as a prognostic target related to m7G methylation in gastric cancer, and the prognosis model based on this can better predict the prognosis of gastric cancer patients (Li et al., 2022). However, there is a lack of relevant research on NUDT4 in infectious diseases and even sepsis. PARN, a major mammalian deadenylase, is the only known enzyme that binds both the 5' cap structure and the 3' poly A, thereby increasing the degradation rate and enhancing its sustained synthesis ability. PARN is important in oocyte maturation, embryogenesis, early development, DNA damage and cell cycle progression. This enzyme is also involved in the regulation of nonsense mediated mRNA decay and cytoplasmic polyadenylation (Balatsos et al., 2012). Through whole exome and sequencing of patients with congenital dyskeratosis (DC), the study found that there were compound heterozygous mutations (c.204 g > T and c.178-245del) in PARN. At the same time, B cells and NK blood cells were also detected to be reduced, the ratio of CD4:CD8 was inverted, and naive CD4 and CD8 cells were reduced (Zeng et al., 2020). However, in sepsis, lymphocytopenia and T-cell depletion are often found to be immunopathological characteristics (van der Poll et al., 2017). Therefore, the mechanism of PARN in sepsis deserves further discussion, especially whether PARN mediates immune cell apoptosis in sepsis deserves further attention.

In conclusion, we successfully performed molecular classification based on m7G-related genes in sepsis patients, and there were significant differences in prognosis among different subgroups. We also explored the key events such as signal pathways and immune infiltration within these subtypes. Our results can better explain the heterogeneity of sepsis patients and provide a basis for early intervention of sepsis patients. In addition, we identified two potential therapeutic targets for sepsis, namely NUDT4 and PARN, both of which are closely related to the prognosis of sepsis patients. Collectively, these findings will contribute to a better understanding of the occurrence and development of sepsis.

## References

- Ali, H. R., Chlon, L., Pharoah, P. D., Markowitz, F., and Caldas, C. (2016). Patterns of immune infiltration in breast cancer and their clinical implications: A gene-expression-based retrospective study. *PLoS Med.* 13 (12), e1002194. doi:10.1371/journal.pmed.1002194
- Balatsos, N. A., Maragozidis, P., Anastasakis, D., and Stathopoulos, C. (2012). Modulation of poly(A)-specific ribonuclease (PARN): Current knowledge and

## 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

JG: design, analysis, operations, manuscript writing, and manuscript revision. JY and YH: figures and manuscript writing. XC: document translation. GY: blood samples collection, qPCR test and paper writing. RS: design and manuscript revision.

## Funding

This work was supported by the Youth Research Project Fund (No. YQ2019-012) of Guangdong Second Provincial General Hospital.

## Acknowledgments

We are very grateful to the people who participated in this work, especially the data from the GEO database.

## 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.

perspectives. *Curr. Med. Chem.* 19 (28), 4838–4849. doi:10.2174/092986712803341539

Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462 (7269), 108–112. doi:10.1038/nature08460

- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: Archive for functional genomics data sets--update. *Nucleic Acids Res.* 41, D991–D995. doi:10.1093/nar/gks1193
- Barton, R., and Cerra, F. B. (1989). The hypermetabolism. Multiple organ failure syndrome. *Chest* 96 (5), 1153–1160. doi:10.1378/chest.96.5.1153
- Bermudes, A. C. G., de Carvalho, W. B., Zamberlan, P., Muramoto, G., Maranhão, R. C., and Delgado, A. F. (2018). Changes in lipid metabolism in pediatric patients with severe sepsis and septic shock. *Nutrition* 47, 104–109. doi:10.1016/j.nut.2017.09.015
- Bessman, M. J. (2019). A cryptic activity in the Nudix hydrolase superfamily. *Protein Sci.* 28 (8), 1494–1500. doi:10.1002/pro.3666
- Binnie, A., Tsang, J. L. Y., Hu, P., Carrasqueiro, G., Castelo-Branco, P., and Dos Santos, C. C. (2020). Epigenetics of sepsis. *Crit. Care Med.* 48 (5), 745–756. doi:10.1097/ccm.0000000000004247
- Cao, Q., Wang, X., Jia, L., Mondal, A. K., Diallo, A., Hawkins, G. A., et al. (2014). Inhibiting DNA Methylation by 5-Aza-2'-deoxycytidine ameliorates atherosclerosis through suppressing macrophage inflammation. *Endocrinology* 155 (12), 4925–4938. doi:10.1210/en.2014-1595
- Cavaillon, J. M., and Adib-Conquy, M. (2005). Monocytes/macrophages and sepsis. *Crit. Care Med.* 33, S506–S509. doi:10.1097/01.ccm.0000185502.21012.37
- Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1\_12
- Chen, J., Li, K., Chen, J., Wang, X., Ling, R., Cheng, M., et al. (2022). Aberrant translation regulated by METTL1/WDR4-mediated tRNA N7-methylguanosine modification drives head and neck squamous cell carcinoma progression. *Cancer Commun.* 42 (3), 223–244. doi:10.1002/cac2.12273
- Dahn, M. S., Lange, P., Lobdell, K., Hans, B., Jacobs, L. A., and Mitchell, R. A. (1987). Splanchnic and total body oxygen consumption differences in septic and injured patients. *Surgery* 101 (1), 69–80.
- Fleischmann, C., Scherag, A., Adhikari, N. K., Hartog, C. S., Tsaganos, T., Schlattmann, P., et al. (2016). Assessment of global incidence and mortality of hospital-treated sepsis. Current estimates and limitations. *Am. J. Respir. Crit. Care Med.* 193 (3), 259–272. doi:10.1164/rccm.201504-0781OC
- Gaieski, D. F., Edwards, J. M., Kallan, M. J., and Carr, B. G. (2013). Benchmarking the incidence and mortality of severe sepsis in the United States. *Crit. Care Med.* 41 (5), 1167–1174. doi:10.1097/CCM.0b013e31827c09f8
- Goel, M. K., Khanna, P., and Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. *Int. J. Ayurveda Res.* 1 (4), 274–278. doi:10.4103/0974-7788.76794
- Grégoire, C., Chasson, L., Luci, C., Tomasello, E., Geissmann, F., Vivier, E., et al. (2007). The trafficking of natural killer cells. *Immunol. Rev.* 220 (1), 169–182. doi:10.1111/j.1600-065X.2007.00563.x
- Huang, M., Cai, S., and Su, J. (2019). The pathogenesis of sepsis and potential therapeutic targets. *Int. J. Mol. Sci.* 20 (21), E5376. doi:10.3390/ijms20215376
- Huntington, N. D., Vossenrich, C. A., and Di Santo, J. P. (2007). Developmental pathways that generate natural-killer-cell diversity in mice and humans. *Nat. Rev. Immunol.* 7 (9), 703–714. doi:10.1038/nri2154
- Hussell, T., and Bell, T. J. (2014). Alveolar macrophages: Plasticity in a tissue-specific context. *Nat. Rev. Immunol.* 14 (2), 81–93. doi:10.1038/nri3600
- Iwashyna, T. J., Cooke, C. R., Wunsch, H., and Kahn, J. M. (2012). Population burden of long-term survivorship after severe sepsis in older Americans. *J. Am. Geriatr. Soc.* 60 (6), 1070–1077. doi:10.1111/j.1532-5415.2012.03989.x
- Jairath, N. K., Farha, M. W., Srinivasan, S., Jairath, R., Green, M. D., Dess, R. T., et al. (2020). Tumor immune microenvironment clusters in localized prostate adenocarcinoma: Prognostic impact of macrophage enriched/plasma cell non-enriched subtypes. *J. Clin. Med.* 9 (6), E1973. doi:10.3390/jcm9061973
- Kotas, M. E., and Matthay, M. A. (2018). Mesenchymal stromal cells and macrophages in sepsis: New insights. *Eur. Respir. J.* 51 (4), 1800510. doi:10.1183/13993003.00510-2018
- Lang, C. H., Frost, R. A., and Vary, T. C. (2007). Regulation of muscle protein synthesis during sepsis and inflammation. *Am. J. Physiol. Endocrinol. Metab.* 293 (2), E453–E459. doi:10.1152/ajpendo.00204.2007
- Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9, 559. doi:10.1186/1471-2105-9-559
- Li, X. Y., Wang, S. L., Chen, D. H., Liu, H., You, J. X., Su, L. X., et al. (2022). Construction and validation of a m7G-related gene-based prognostic model for gastric cancer. *Front. Oncol.* 12, 861412. doi:10.3389/fonc.2022.861412
- Liao, M., Liu, Y., Yuan, J., Wen, Y., Xu, G., Zhao, J., et al. (2020). Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* 26 (6), 842–844. doi:10.1038/s41591-020-0901-9
- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004
- Liu, N., Zhou, K. I., Parisien, M., Dai, Q., Diatchenko, L., and Pan, T. (2017a). N6-methyladenosine alters RNA structure to regulate binding of a low-complexity protein. *Nucleic Acids Res.* 45 (10), 6051–6063. doi:10.1093/nar/gkx141
- Liu, X., Wu, S., Yang, Y., Zhao, M., Zhu, G., and Hou, Z. (2017b). The prognostic landscape of tumor-infiltrating immune cell and immunomodulators in lung cancer. *Biomed. Pharmacother.* 95, 55–61. doi:10.1016/j.biopha.2017.08.003
- Luo, Y., Yao, Y., Wu, P., Zi, X., Sun, N., and He, J. (2022). The potential role of N(7)-methylguanosine (m7G) in cancer. *J. Hematol. Oncol.* 15 (1), 63. doi:10.1186/s13045-022-01285-5
- Malbec, L., Zhang, T., Chen, Y. S., Zhang, Y., Sun, B. F., Shi, B. Y., et al. (2019). Dynamic methylome of internal mRNA N(7)-methylguanosine and its regulatory role in translation. *Cell Res.* 29 (11), 927–941. doi:10.1038/s41422-019-0230-z
- Mildvan, A. S., Xia, Z., Azurmendi, H. F., Saraswat, V., Legler, P. M., Massiah, M. A., et al. (2005). Structures and mechanisms of Nudix hydrolases. *Arch. Biochem. Biophys.* 433 (1), 129–143. doi:10.1016/j.abb.2004.08.017
- Murch, O., Collin, M., Hinds, C. J., and Thiemermann, C. (2007). Lipoproteins in inflammation and sepsis. I. Basic science. *Intensive Care Med.* 33 (1), 13–24. doi:10.1007/s00134-006-0432-y
- Purcarea, A., and Sovaila, S. (2020). Sepsis, a 2020 review for the internist. *Rom. J. Intern. Med.* 58 (3), 129–137. doi:10.2478/rjim-2020-0012
- Ralainirina, N., Poli, A., Michel, T., Poos, L., André, E., Hentges, F., et al. (2007). Control of NK cell functions by CD4+CD25+ regulatory T cells. *J. Leukoc. Biol.* 81 (1), 144–153. doi:10.1189/jlb.0606409
- Reisinger, A. C., Posch, F., Hackl, G., Marsche, G., Sourij, H., Bourgeois, B., et al. (2021). Branched-chain amino acids can predict mortality in ICU sepsis patients. *Nutrients* 13 (9), 3106. doi:10.3390/nu13093106
- Remick, D. G. (2007). Pathophysiology of sepsis. *Am. J. Pathol.* 170 (5), 1435–1444. doi:10.2353/ajpath.2007.060872
- Rocheteau, P., Chatre, L., Briand, D., Mebarki, M., Jouvion, G., Bardon, J., et al. (2015). Sepsis induces long-term metabolic and mitochondrial muscle stem cell dysfunction amenable by mesenchymal stem cell therapy. *Nat. Commun.* 6, 10145. doi:10.1038/ncomms10145
- Roundtree, I. A., Evans, M. E., Pan, T., and He, C. (2017). Dynamic RNA modifications in gene expression regulation. *Cell* 169 (7), 1187–1200. doi:10.1016/j.cell.2017.05.045
- Scott, M. J., Hoth, J. J., Turina, M., Woods, D. R., and Cheadle, W. G. (2006). Interleukin-10 suppresses natural killer cell but not natural killer T cell activation during bacterial infection. *Cytokine* 33 (2), 79–86. doi:10.1016/j.cyto.2005.12.002
- Shih, C. C., Liao, M. H., Hsiao, T. S., Hii, H. P., Shen, C. H., Chen, S. J., et al. (2016). Procainamide inhibits DNA methylation and alleviates multiple organ dysfunction in rats with endotoxic shock. *PLoS One* 11 (9), e0163690. doi:10.1371/journal.pone.0163690
- Singer, M., Deutschman, C. S., Seymour, C. W., Shankar-Hari, M., Annane, D., Bauer, M., et al. (2016). The third international consensus definitions for sepsis and septic shock (Sepsis-3). *Jama* 315 (8), 801–810. doi:10.1001/jama.2016.0287
- Swift, S., Tucker, A., Vinciotti, V., Martin, N., Orengo, C., Liu, X., et al. (2004). Consensus clustering and functional interpretation of gene-expression data. *Genome Biol.* 5 (11), R94. doi:10.1186/gb-2004-5-11-r94
- van der Poll, T., van de Veerdonk, F. L., Scicluna, B. P., and Netea, M. G. (2017). The immunopathology of sepsis and potential therapeutic targets. *Nat. Rev. Immunol.* 17 (7), 407–420. doi:10.1038/nri.2017.36
- Xiao, F., Guo, Z., and Yan, Q. (2015). The metabolic change in citrulline and its use in sepsis. *Zhonghua Wei Zhong Bing Ji Jiu Yi Xue* 27 (6), 534–537. doi:10.3760/cma.j.issn.2095-4352.2015.06.024
- Zeng, T., Lv, G., Chen, X., Yang, L., Zhou, L., Dou, Y., et al. (2020). CD8(+) T-cell senescence and skewed lymphocyte subsets in young Dyskeratosis Congenita patients with PARN and DKC1 mutations. *J. Clin. Lab. Anal.* 34 (9), e23375. doi:10.1002/jcla.23375
- Zhang, B., Li, D., and Wang, R. (2022). Transcriptome profiling of N7-methylguanosine modification of messenger RNA in drug-resistant acute myeloid leukemia. *Front. Oncol.* 12, 926296. doi:10.3389/fonc.2022.926296
- Zhang, Y. Y., and Ning, B. T. (2021). Signaling pathways and intervention therapies in sepsis. *Signal Transduct. Target. Ther.* 6 (1), 407. doi:10.1038/s41392-021-00816-9

# Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

## Discover the latest Research Topics

[See more →](#)

### Frontiers

Avenue du Tribunal-Fédéral 34  
1005 Lausanne, Switzerland  
[frontiersin.org](https://frontiersin.org)

### Contact us

+41 (0)21 510 17 00  
[frontiersin.org/about/contact](https://frontiersin.org/about/contact)

