frontiers | Research Topics

Epigenetic regulation and non-histone post-translational modification in cancer

Edited by
Yan Gong, Zhongxiao Wang, Yingming Sun, Wei Wei and Xuehui Yang

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-8325-2226-4
DOI 10.3389/978-2-8325-2226-4

About Frontiers

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

Frontiers journal series

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

Dedication to quality

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

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Epigenetic regulation and non-histone post-translational modification in cancer

Topic editors

Yan Gong — Wuhan University, China
Zhongxiao Wang — Biogen Idec, United States
Yingming Sun — Fujian Medical University, China
Wei Wei — Zhongnan Hospital of Wuhan University, China
Xuehui Yang — Maine Medical Center Research Institute

Citation

Gong, Y., Wang, Z., Sun, Y., Wei, W., Yang, X., eds. (2023). *Epigenetic regulation and non-histone post-translational modification in cancer*.
Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2226-4

Table of contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 06 April 2023
DOI 10.3389/fgene.2023.1176174

OPEN ACCESS

EDITED AND REVIEWED BY
Anton A. Buzdin,
European Organisation for Research and Treatment of Cancer, Belgium

*CORRESPONDENCE
Yan Gong,
yan.gong@whu.edu.cn
Yingming Sun,
yingmingsun@fjmu.edu.cn

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

RECEIVED 28 February 2023
ACCEPTED 31 March 2023
PUBLISHED 06 April 2023

CITATION
Luo J, Huang Z, Wei W, Sun Y and Gong Y (2023), Editorial: Epigenetic regulation and non-histone post-translational modification in cancer.
Front. Genet. 14:1176174.
doi: 10.3389/fgene.2023.1176174

Editorial: Epigenetic regulation and non-histone post-translational modification in cancer

Jiang Luo[1], Zhengrong Huang[1], Wei Wei[2], Yingming Sun[3]* and Yan Gong[1]*

[1]Department of Biological Repositories, Tumor Precision Diagnosis and Treatment Technology and Translational Medicine, Hubei Engineering Research Center, Zhongnan Hospital of Wuhan University, Wuhan, China, [2]Brain Research Center, Department of Neurosurgery, Zhongnan Hospital of Wuhan University, Wuhan, China, [3]Department of Radiation and Medical Oncology, Affiliated Sanming First Hospital of Fujian Medical University, Sanming, China

KEYWORDS

epigenetic regulation, non-histone post-translational modification, cancer, DNA methylation, m6A modifications, non-coding RNAs

Editorial on the Research Topic
Epigenetic regulation and non-histone post-translational modification in cancer

Epigenetic changes are essentially involved in both normal organism function and disease progression (Deans and Maggert, 2015). Epigenetic regulation includes DNA methylation or demethylation, chromatin remodeling, histone modifications and non-coding RNAs, which are broadly reported to dysfunction in cancer. Of noted, it is increasingly clear that epigenetic regulation parallels with gene expression modulation. Currently, significant progress has been made in the development of drugs targeting key enzymes involved in epigenetic regulation and post-translational modification without histone. Several drugs have been approved for therapeutic application, and many more are in clinical and preclinical testing (Ferreira and Esteller, 2018; Lu et al., 2020). The aim of this Research Topic is to provide an overview of the current understanding and fundamental findings in the field of epigenetic regulation and non-histone post-translational modifications in cancer. We collected 14 articles including the effects of m6A modifications, non-coding RNAs and SELENBP1 in cancer progression.

DNA methylation is strongly associated with cancer, and hypermethylation of some genes in the promoter region interferes with the reading of DNA information thereby altering epigenetics, thus it has the potential to be a promising target for cancer therapy (Smith and Meissner, 2013). Liexi Xu et al. used the methylation and clinical data of lung adenocarcinoma (LUAD) patients from TCGA. They found 11 differential methylation genes and established a methylation scoring model to assess prognosis, suggesting that these genes could be used as biomarkers of methylation in LUAD. Dong-Mei Hu et al. focused on the relationship between Forkhead box P (FOXP) family DNA methylation and immune-related factors in non-small cell lung cancer (NSCLC) patients. FOXP family is widely involved in regulating immune molecules and influencing immune infiltration in NSCLC, and FOXP family DNA methylation is associated with NSCLC prognosis. In addition to lung

cancer, immunosuppression and immune cells dysfunction are critical to the development of colonic adenocarcinoma (COAD) (Gatenbee et al., 2022). Salem Baldi et al. analyzed the relationship between ARID1B expression, DNA methylation and prognosis in COAD patients based on TCGA, and found that differences in immune cell infiltration were associated with ARID1B expression, and that ARID1B was hypermethylated in COAD tissues. Since the ARID1B methylation levels were negatively associated with mRNA levels, low ARID1B expression was an important indicator of poor prognosis, and ARID1B hypermethylation could be an early diagnostic biomarker in COAD. Furthermore, Xingyu Liu et al. obtained 4 alcohol-related cancer samples from TCGA and GEO databases, and identified a total of 193 differentially methylated probes. By enrichment analysis of differential genes, they concluded that the alcohol might facilitate transcriptional dysfunction *via* inducing the methylation status of transcriptional regulators, leading to tumor development. They also identified the hypermethylated CpG island (chr19:58220189-58220517), which regulated the transcriptional activity of zinc-finger protein 154 as a potential therapeutic biomarker.

In addition to DNA modifications, RNA modifications are prevalent in mammalian cells, and provide a new dimension to regulate gene expression. Among them, N6-methyladenosine (m6A) modifications are the most common in eukaryotic mRNA. Shaojie Li et al. obtained sample data of head and neck squamous cell carcinoma (HNSCC) from TCGA and GEO databases. Via analyzing the correlation between m6A regulator expression and immune scores, they found that the HNSCC patients could be divided into 2 groups based on m6A reader genes (IGF2BP2 and YTHDF1). Low YTHDF1 and IGF2BP2 expressing patients have more immune cells enriched in TME and better prognosis. Mengying Zhou et al. reviewed the role of different m6A-related enzymes in breast cancer, and concluded that m6A-related genes could be used as not only markers for diagnosis and prognosis prediction, but also effective targets for breast cancer treatment.

Moreover, non-coding RNA transcriptional modifications are also crucial (Yao et al., 2022). With the continuous researches on long non-coding RNA (lncRNA), it was found that lncRNA play important roles in various biological regulatory processes. Jing Huang et al. downloaded mRNA expression data from the TCGA database of HNSCC patients, identified 1,117 lncRNAs associated with necrosis, of which 55 ones were associated with patient survival. They selected 24 genes that positively regulated necroptosis to establish a new risk scoring model for assessing HNSCC patient prognosis. Jing Hu et al. obtained information about autophagy-related genes in NSCLC patients from TCGA and HADb, and identified 7 autophagy-related lncRNAs whose composition of risk models could accurately predict the prognosis of NSCLC patients. ABALON deficiency in A549 and NCI-H292 significantly inhibited the proliferation and metastasis of NSCLC cells and promoted autophagy. Wei Yu et al. identified 298 lncRNAs associated with cuproptosis using TCGA data of gastric adenocarcinoma patients, including 13 lncRNAs associated with survival, and further identified 9 lncRNAs by LASSO regression method. Based on these 9 lncRNAs, they established a risk assessment model to evaluate the prognosis and sensitivity of patients to therapeutic drugs.

Immunotherapy combined with radiotherapy is one of the best combinations for oncotherapy. Increasing researches suggested that lncRNA might be associated with the responses to immunotherapy and radiotherapy. Chuanhao Zhang et al. acquired RNA-seq data and clinical characteristics of 594 LUAD patients from TCGA, identified 2,093 N7-methylguanosine related lncRNAs, and finally constructed a risk-prognosis model *via* screening 6 prognosis-related lncRNAs, which not only accurately predicted the patient survival, but also reflected the immune characteristics of LUAD patients and provided better guidance for individualized patient treatment. Jianqing Zheng et al. screened 26 immune-related lncRNAs (ir-lncRNAs) differentially expressed in radiation-resistant esophageal squamous cell carcinoma by the GEO database and paired the differentially expressed lncRNA with each other in the GSE45670 dataset to construct 325 ir-lncRNA pairs, they established a prognostic risk model based on 3 pairs of ir-lncRNA, and suggested that macrophage infiltration and differential expression of ir-lncRNA are potential mechanisms of resistance to radiotherapy. Moreover, Linghui Jia et al. collected 58 patients with oral squamous cell carcinoma, and found that the expression of CircPUM1 (a circular RNA) was significantly increased in oral squamous cell carcinoma. CircPUM1 downregulation induced mir-580, which inhibited STAT3 expression, induced apoptosis and enhanced radiosensitivity.

Histones, as essential components of nucleosomes, play an important role in the structure of chromosomes. Histone modifications are considered important epigenetic mechanisms for gene expression regulation, and small molecule inhibitors have been developed to detect the effects of these modifications on cellular proteins (Buuh et al., 2018). SETD2 is the major methyltransferase catalyzing histone H3K36. Zihang Zeng et al. collected data on LUAD patients from GEO and TCGA, and through multi-omics analysis identified that SETD2 was associated with radiosensitivity. SETD2 downregulation attenuated proliferation and migration, and enhanced the apoptosis and radiosensitivity of LUAD cells. They also found that reducing m6A-related genes (RBM2 or YTHDF15) could enhance the protective effect of SETD2 on patient prognosis. Yue Zhang et al. reviewed the function and regulatory mechanisms of SELENBP1 (a selenium-binding protein) during cancer progression, and also discussed potential cancer treatment strategies targeting SELENBP1 epigenetic modifications.

Cancer is not only a genetic disease, but also an epigenetic disease. Epigenetic mechanisms are engaged in the regulation of many aspects of cancer biology (Garcia-Martinez et al., 2021). With the development of mass spectrometry-based proteomics technologies, some non-histone modifications (e.g., lysine acetylation, lactylation) also play key roles in cell growth, metabolism, and signal transduction (Narita et al., 2019; Yang et al., 2023). This Research Topic focused on the regulation of epigenetic and non-histone post-translational modifications, as well as their impacts on cancer development and progression. Patterns of epigenetic regulation and non-histone modifications may be potential predictors of cancer patient prognosis and survival, providing novel insights into the oncotherapy. We hope more researches devote to this field in the future and look forward to their early translation into clinical treatments.

Author contributions

JL and YS drafted the editorial. ZH, WW, and YG critically reviewed and revised the manuscript. All authors approved the submitted version.

Acknowledgments

We thank all of the authors for their valuable contribution and the referees for their rigorous review to this Research Topic. We also thank Nancy Qin, Shirley Deng and the Frontiers team for their kind help.

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

References

Buuh, Z. Y., Lyu, Z., and Wang, R. E. (2018). Interrogating the roles of post-translational modifications of non-histone proteins: Miniperspective. *J. Med. Chem.* 61, 3239–3252. doi:10.1021/acs.jmedchem.6b01817

Deans, C., and Maggert, K. A. (2015). What do you mean, "epigenetic". *Genetics* 199, 887–896. doi:10.1534/genetics.114.173492

Ferreira, H. J., and Esteller, M. (2018). Non-coding RNAs, epigenetics, and cancer: Tying it all together. *Cancer Metastasis Rev.* 37, 55–73. doi:10.1007/s10555-017-9715-8

Garcia-Martinez, L., Zhang, Y., Nakata, Y., Chan, H. L., and Morey, L. (2021). Epigenetic mechanisms in breast cancer therapy and resistance. *Nat. Commun.* 12, 1786. doi:10.1038/s41467-021-22024-3

Gatenbee, C. D., Baker, A.-M., Schenck, R. O., Strobl, M., West, J., Neves, M. P., et al. (2022). Immunosuppressive niche engineering at the onset of human colorectal cancer. *Nat. Commun.* 13, 1798. doi:10.1038/s41467-022-29027-8

Lu, Y., Chan, Y.-T., Tan, H.-Y., Li, S., Wang, N., and Feng, Y. (2020). Epigenetic regulation in human cancer: The potential role of epi-drug in cancer therapy. *Mol. Cancer* 19, 79. doi:10.1186/s12943-020-01197-3

Narita, T., Weinert, B. T., and Choudhary, C. (2019). Functions and mechanisms of non-histone protein acetylation. *Nat. Rev. Mol. Cell Biol.* 20, 156–174. doi:10.1038/s41580-018-0081-3

Smith, Z. D., and Meissner, A. (2013). DNA methylation: Roles in mammalian development. *Nat. Rev. Genet.* 14, 204–220. doi:10.1038/nrg3354

Yang, Z., Yan, C., Ma, J., Peng, P., Ren, X., Cai, S., et al. (2023). Lactylome analysis suggests lactylation-dependent mechanisms of metabolic adaptation in hepatocellular carcinoma. *Nat. Metab.* 5, 61–79. doi:10.1038/s42255-022-00710-w

Yao, Z., Yang, Y., Sun, M., He, Y., Liao, L., Chen, K., et al. (2022). New insights into the interplay between long non-coding RNAs and RNA-binding proteins in cancer. *Cancer Commun.* 42, 117–140. doi:10.1002/cac2.12254

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 12 May 2022
doi: 10.3389/fgene.2022.903634

Identification of Two m6A Readers *YTHDF1* and *IGF2BP2* as Immune Biomarkers in Head and Neck Squamous Cell Carcinoma

Shaojie Li[†], *Qiuji Wu*[†], *Jia Liu and Yahua Zhong*[*]

Hubei Key Laboratory of Tumor Biological Behaviors, Department of Radiation and Medical Oncology, Hubei Cancer Clinical Study Center, Zhongnan Hospital of Wuhan University, Wuhan, China

Background: N6-methyladenosine (m6A) is the most abundant internal modification pattern in mammals that a plays critical role in tumorigenesis and immune regulations. However, the effect of m6A modification on head and neck squamous cell carcinoma (HNSCC) has not been clearly studied.

Methods: We screened m6A regulators that were significantly correlated with tumor immune status indicated by ImmuneScore using The Cancer Genome Atlas (TCGA) dataset and obtained distinct patient clusters based on the expression of these m6A regulators with the R package "CensusClusterPlus." We then performed gene set enrichment analysis (GSEA), CIBERSORT, and single-sample gene set enrichment analysis (ssGSEA) to assess the differences in gene function enrichment and tumor immune microenvironment (TIME) among these clusters. We further conducted differently expressed gene (DEG) analysis and weighted gene co-expression network analysis (WGCNA) and constructed a protein–protein interaction (PPI) network to determine hub genes among these clusters. Finally, we used the GSE65858 dataset as an external validation cohort to confirm the immune profiles related to the expression of m6A regulators.

Results: Two m6A readers, *YTHDF1* and *IGF2BP2*, were found to be significantly associated with distinct immune status in HNSCC. Accordingly, patients were divided into two clusters with Cluster 1 showing high expression of *YTHDF1* and *IGF2BP2* and Cluster 2 showing low expression levels of both genes. Clinicopathologically, patients from Cluster 1 had more advanced T stage and pathological grades than those from Cluster 2. GSEA showed that Cluster 1 was closely related to the RNA modification process and Cluster 2 was significantly correlated with immune regulations. Cluster 2 had a more active

OPEN ACCESS

Edited by:
Zhongxiao Wang,
Biogen Idec, United States

Reviewed by:
Yuanji Xu,
Fujian Provincial Cancer Hospital, China
Dan Ou,
Fudan University, China

****Correspondence:***
Yahua Zhong
doctorzyh73@163.com

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

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: *24 March 2022*
Accepted: *13 April 2022*
Published: *12 May 2022*

Citation:
Li S, Wu Q, Liu J and Zhong Y (2022) Identification of Two m6A Readers YTHDF1 and IGF2BP2 as Immune Biomarkers in Head and Neck Squamous Cell Carcinoma. Front. Genet. 13:903634. doi: 10.3389/fgene.2022.903634

Abbreviations: CPS, combined positive score; DEG, differently expressed gene; EBV, Epstein–Barr virus; FAP, fibroblast activation protein alpha; GEO, the Gene Expression Omnibus; GSEA, gene set enrichment analysis; HNSCC, head and neck squamous cell carcinoma; HPV, human papillomavirus; ICIs, immune checkpoint inhibitors; m6A, N6-Methyladenosine; MCODE, the Molecular Complex Detection; mRNA, messenger RNA; PCA, principal component analysis; PPI, protein-protein interaction; ssGSEA, single sample GSEA; TCGA, The Cancer Genome Atlas; TGF-β, transforming growth factor-β; TIME, tumor immune microenvironment; TME, tumor microenvironment; WGCNA, weighted gene co-expression network analysis.

TIME characterized by a more relative abundance of $CD8^+$ T cells and $CD4^+$ T cells and higher levels of MHC I and MHC II molecules. We constructed a PPI network composed of 16 hub genes between the two clusters, which participated in the T-cell receptor signaling pathway. These results were externally validated in the GSE65858 dataset.

Conclusions: The m6A readers, *YTHDF1* and *IGF2BP2,* were potential immune biomarkers in HNSCC and could be potential treatment targets for cancer immunotherapy.

Keywords: YTHDF1, IGF2BP2, HNSCC, m6A modification, immune microenvironment, immunotherapy

INTRODUCTION

M6A is a critical and abundant internal epigenetic modification on both messenger RNA (mRNA) and non-coding RNAs in mammals. M6A modification is mainly found in the 3′ untranslated regions of the RNA. The reversible and dynamic regulation of m6A is mainly mediated by three different kinds of regulators, namely, writers, erasers, and readers (Roundtree et al., 2017; Nombela et al., 2021). Writers and erasers are methyltransferases (such as METTL3, METTL16, and WTAP) and demethylases (such as ALKBH5 and FTO) that methylate and demethylate RNA adenosine at specific N6 positions, respectively. Readers are a group of RNA binding proteins that recognize m6A sites and initiate downstream events such as RNA splicing, maturation, degradation, and translation (Li Y et al., 2019; Zaccara et al., 2019). These proteins include members of the EIF3, IGF2BP family and YTH family. By modulating RNA export, RNA stability, protein expression, and other biological activities, m6A modification plays an essential role in cancer development. In line with these findings, emerging drugs targeting m6A modification, such as a selective inhibitor of FTO, METTL3, and YTHDF2 have shown promising anti-cancer effects (Huang et al., 2015; Visvanathan et al., 2018; Dixit et al., 2021).

The detailed mechanism by which m6A modification impacts cancer pathogenesis remains unclear. A recent study indicated that m6A also has a dual role in tumorigenesis. Liu et al. reported that YTHDF1 promoted ovarian cancer progression *via* augmenting EIF3C translation (Liu et al., 2020). Li et al. reported that IGF2BP2 prevented SOX2 degradation, leading to colorectal cancer pathogenesis and progression (Li T et al., 2019). Meanwhile, Zhong et al. reported that m6A helped suppress hepatocellular carcinoma through YTHDF2-directed degradation of EGFR (Zhong et al., 2019). Importantly, m6A modification also has a nonnegligible impact on anti-tumor immunity. Overexpression of YTHDF1 enhanced the stability of RNA lysosomal proteases, which led to the degradation of tumor antigens in dendritic cells, disabled $CD8^+$ T cells to bring about immunosurveillance and abolished the effect of immune checkpoint inhibitors (ICIs) (Han et al., 2019). Suppression of METTL3 and METTL14 increased infiltration of $CD8^+$ T cells and secretion of IFN-γ, CXCL9, and CXCL10 in TIME and promoted the response to ICIs in melanoma and pMMR-MSI^{low} colorectal cancer (Wang et al., 2020). However, research aiming to explore the effect of m6A modification on immune profiles in HNSCC is inadequate.

HNSCC is the sixth most common malignant tumor worldwide. Its occurrence is closely linked to carcinogen exposure and viral infection, especially human papillomavirus (HPV) and Epstein–Barr virus (EBV) (Siegel et al., 2021). HNSCC is a group of heterogeneous cancers, and the majority of patients are presented with locally advanced or metastatic stage, leading to poor prognosis (Argiris et al., 2008; Chow, 2020). ICIs-based immunotherapy has prominently improved the efficacy and survival of advanced HNSCC. However, only a small subset of patients could benefit from immunotherapy. Although a combined positive score (CPS) of PD-L1 expression is mostly used to guide immunotherapy, currently no satisfactory predictive biomarker is available for HNSCC.

Here, we attempted to explore the role of m6A regulators in the immune modulation of HNSCC and tried to identify potential m6A-associated biomarkers of immunotherapy in HNSCC. This study might provide a new way to improve the effect of immunotherapy in HNSCC.

METHODS AND MATERIALS

Data Sources and Processing

We downloaded the transcriptomic data (HTSeq-FPKM and HTSeq-Counts) and clinical information of an HNSCC cohort from the TCGA-database (https://portal.gdc.cancer.gov/). Data of HTSeq-Counts was used to analyze DEGs, and HTSeq-FPKM was used to conduct ESTIMATE, clustering, CIBERSORT, ssGSEA, and WGCNA. Mann-Whitney *U* test was performed to compare age and gene expression between two groups. Chi-square test was utilized to compare gender, T stage, N stage, and pathologic stage. Spearman's coefficient was used to conduct correlation analysis. And a p-value < 0.05 (two-sided) was considered statistically significant. Meanwhile, we conducted an external validation dataset by downloading the expression profiling data of the GSE65858 array from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/).

Estimation of Stromal and Immune Cells

We employed the ESTIMATE tool embedded in the R package "estimate" that used gene expression signatures to infer the fraction of stromal and immune cells in the tumor samples and to estimate the elements of tumor microenvironment (TME), including StromalScore, ImmuneScore, ESTIMATEScore, and TumorPurity (Yoshihara et al., 2013).

TABLE 1 | Summary of common m6A regulators.

Type	Regulator	Function
Writer	CBLL1	Stabilizes several subunits of the methyltransferase complex
	METTL3	Catalyzes m6A modification
	METTL14	Provides help of target recognition for METTL3
	METTL16	Catalyzes m6A modification
	RBM15/RBM15B	Binds the m6A methylation complex and recruit it to specific sites in RNA
	WTAP	Helps localization of METTL3-METTL14 into nuclear speckles
	ZC3H13	Anchors WTAP, Virilizer, and Hakai in the nucleus to facilitate m6A methylation and regulate mESC self-renewal
	ZCCHC4	Catalyzes m6A modification of 28S ribosomal RNA
Eraser	ALKBH5	Removes m6A modification
	FTO	Removes m6A modification
Reader	EIF3	Promotes mRNA translation
	HNRNPA2B1	Mediates effects of m6A modification on primary microRNA processing and alternative splicing
	HNRNPC	Affects mRNA abundance and alternative splicing
	IGF2BPs	Enhances mRNA stability
	YTHDC1	Enhances RNA splicing and export
	YTHDC2	Promotes the translation of target RNA and reduce its abundance
	YTHDF1	Promotes mRNA translation
	YTHDF2	Promotes mRNA degradation
	YTHDF3	Interacts with YTHDF1 and YTHDF2 to enhance their effect

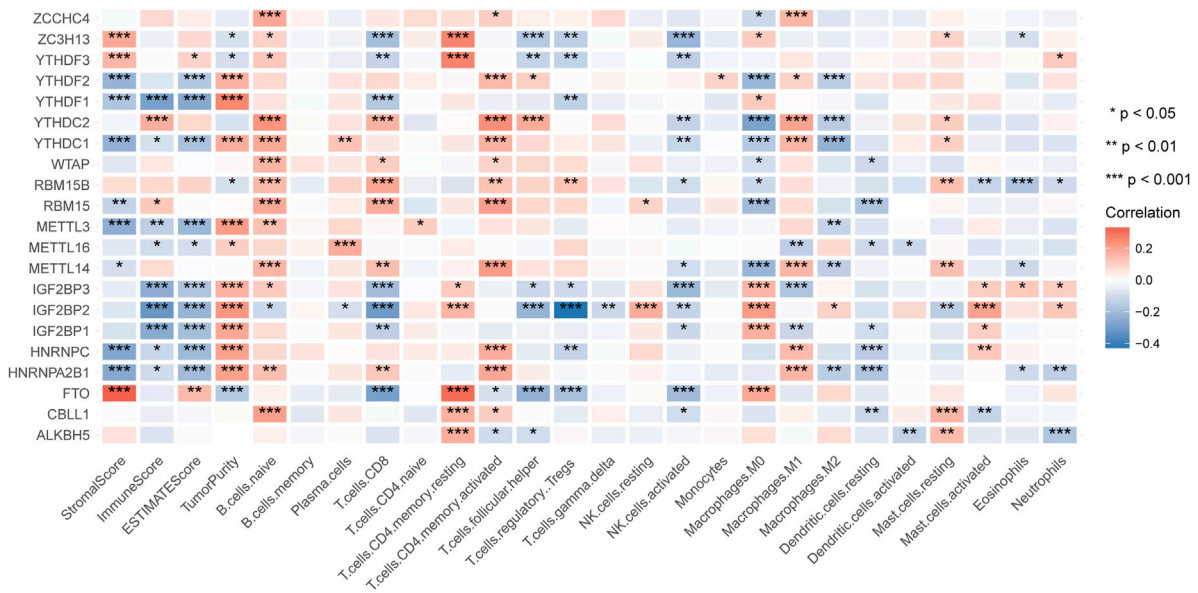


FIGURE 1 | The correlation of m6A regulators with the results of ESTIMATE and CIBERSORT.

Consensus Clustering

To explore the influence of m6A modification on immune profiles of HNSCC, we calculated the correlation between the expression of m6A modification regulators and ESTIMATE results with Spearman's coefficient. And then we performed consensus clustering of tumor samples based on the expression of *YTHDF1* and *IGF2BP2*. We accomplished consensus clustering and result visualization with the R package "ConsensusClusterPlus" (Wilkerson and Hayes, 2010). And we examined the efficacy of the above consensus clustering by principal component analysis (PCA) with the R package "factoextra." The Kaplan–Meier method and log-rank test were utilized to compare overall survival between the two clusters.

Gene Set Enrichment Analysis

We employed the software GSEA (https://www.gsea-msigdb.org/gsea/) to determine different pathways enriched in the two clusters based on the default defined set of genes (Mootha et al., 2003; Subramanian et al., 2005). We selected "c5.go.cc.v7.4.symbols.gmt" from MSigDB Collection as pre-defined ontology gene set, and considered a pathway as significantly enriched pathway with the absolute normalized enrichment score > 1 (|NES| >1) and p value < 0.05.

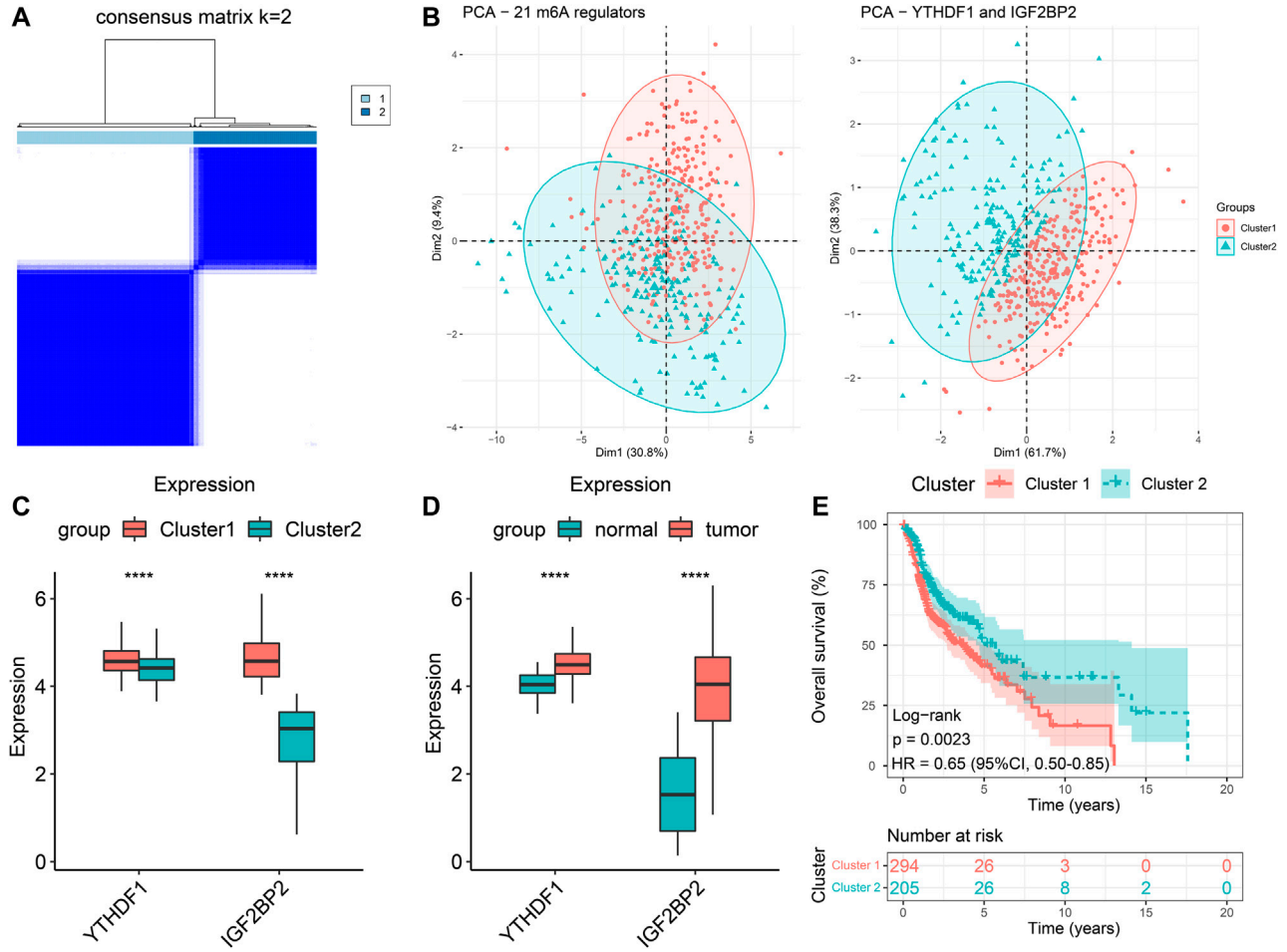


FIGURE 2 | Clustering of patients with HNSCC in TCGA cohort based on expression of YTHDF1 and IGF2BP2. **(A)** Consensus clustering matrix for k = 2. **(B)** The results of PCA of clustering based on 21 m6A regulators as well as *YTHDF1* and *IGF2BP2*. **(C)** Comparison the expression levels of *YTHDF1* and *IGF2BP2* between Cluster 1 and Cluster 2. **(D)** Comparison of the expression levels of *YTHDF1* and *IGF2BP2* between tumor samples and normal samples. **(E)** Kaplan–Meier curves of the overall survival in two clusters. ****$p < 0.0001$.

Immune Microenvironment Analysis

CIBERSORT, a method excelling in decreasing noise and unknown mixtures and identifying similar cell types, was conducted to recognize the cell composition of solid tumors by using gene expression profiles (Newman et al., 2015). ssGSEA, an extension of GSEA, was used to calculate separate enrichment scores for each pairing of a sample and gene set (Hänzelmann et al., 2013). We applied both strategies to explore the difference and relation of TIME among the clusters of patients with HNSCC.

DEGs and Weighted Gene Co-Expression Network Analysis

In order to explore the hub genes that contributed to biological divergences among different patient clusters, we performed an analysis of DEGs and WGCNA successively. First, the R package "limma" enabled us to compare transcriptome data (HTSeq-Counts) to locate DEGs. The screening thresholds were set as |log2FoldChange | > 0.6 and *p*-value of < 0.05, and the results were visualized by volcano plot and heat map. Then, we conducted WGCNA on the DEGs with the R package "WGCNA," which was a system biology method to gather closely related genes in special modules and calculate the relationship between the modules and external sample traits (Zhang and Horvath, 2005; Langfelder and Horvath, 2008). We explored the connection of DEGs with clustering and four aspects of ESTIMATE. Finally, we selected the module of genes tightly related to both clustering and ImmuneScore for subsequent analysis.

Functional Enrichment and Protein–Protein Interaction Network Analysis

In order to investigate the above module of genes ulteriorly, we uploaded the gene list to Metascape (http://metascape.org/) for pathway and process enrichment analysis and PPI enrichment analysis (Zhou et al., 2019). Functional enrichment analysis was carried out in various ontology sources, including GO Biological Processes, Reactome Gene Sets, KEGG Pathway, Canonical Pathways, and WikiPathways. PPI enrichment analysis was also

TABLE 2 | Summary of clinical characteristics of patients in the two clusters.

	Cluster 1 (*n* = 245)	Cluster 2 (*n* = 148)	*p* value
Age, median (range)	60 (19–90)	61 (24–87)	0.28
Gender, *n* (%)			0.47
Female	63 (25.7)	43 (29.1)	—
Male	182 (74.3)	105 (70.9)	—
T stage, *n* (%)			0.0077
T1	18 (7.4)	23 (15.5)	—
T2	59 (24.1)	42 (28.4)	—
T3	66 (26.9)	23 (15.6)	—
T4	102 (41.6)	60 (40.5)	—
N stage, *n* (%)			0.50
N0	101 (41.2)	68 (45.9)	—
N1	39 (15.9)	24 (16.2)	—
N2	99 (40.4)	55 (37.2)	—
N3	6 (2.4)	1 (0.7)	—
Pathological stage, *n* (%)			0.14
Stage I	12 (4.9)	15 (10.1)	—
Stage II	30 (12.2)	19 (12.8)	—
Stage III	51 (20.8)	22 (14.9)	—
Stage IV	152 (62.0)	92 (62.2)	—
Grade, *n* (%)			0.029
G1	22 (9.0)	28 (18.9)	—
G2	162 (66.1)	84 (56.8)	—
G3	60 (24.5)	36 (24.3)	—
G4	1 (0.4)	0 (0)	—

performed, and if the number of proteins in the network fell between 3 and 500, the Molecular Complex Detection (MCODE) algorithm would be carried out to separate proteins to build interaction networks more precisely (Bader and Hogue, 2003).

RESULTS

Identification of m6A Regulators Associated With HNSCC Immune Profiles

After excluding repeated samples and those without adequate survival information, we got 499 patients of HNSCC with unique samples for the following analysis. To explore whether the expression of m6A regulators impacted HNSCC immune profiles, we extracted the expression of 21 m6A modification regulators (**Table 1**) and applied the ESTIMATE tool and CIBERSORT algorithm to calculate the ESTIMATE scores and immune cells infiltration of 499 HNSCC patients. By analyzing the correlation between m6A regulators' expression and the ImmuneScore, we found that *YTHDC2* and *RBM15* were positively correlated with ImmuneScore, while *YTHDF1*, *YTHDC1*, *METTL3*, *METTL16*, *IGF2BP1-3*, *HNRNPC*, and *HNRNPA2B1* were negatively correlated with ImmuneScore (**Figure 1**). Next, we sorted the absolute values of the ImmuneScores correlated with the 21 m6A regulators (**Supplementary Material S1**). We selected the first two regulators with the highest ImmuneScores, *YTHDF1* and *IGF2BP2*, to construct an immune-associated signature.

Consensus Clustering of Patients With HNSCC Based on YTHDF1 and IGF2BP2

We extracted the expression data of *YTHDF1* and *IGF2BP2* of the 499 HNSCC patients and performed consensus clustering, and obtained two clusters of patients (**Figure 2A**). There were 294 patients in Cluster 1 and 205 patients in Cluster 2. After excluding 106 patients without tumor stage and tumor grade information, the clinical characteristics of the remaining 393 patients were summarized in **Table 2**. PCA plot indicated the above clustering had good efficiency of distinction (**Figure 2B**). Cluster 1 had higher expression of *YTHDF1* and *IGF2BP2* (**Figure 2C**), and a more advanced T stage and pathological grade than Cluster 2 (**Table 2**). Expression levels of *YTHDF1* and *IGF2BP2* were also compared between tumor tissue and normal tissue, and we found that both of them were higher expressed in tumor tissue (**Figure 2D**). Kaplan–Meier curve (**Figure 2E**) showed that patients in Cluster 2 had better overall survival than their counterparts in Cluster 1 (HR = 0.65, 95%CI [0.50–0.85], p = 0.0023).

Immune Profiles of YTHDF1- and IGF2BP2-Based Clusters

GSEA was performed to compare pathway enrichment between the two clusters. We found that biological pathways related to m6A modification including negative regulation of DNA repair, regulation of mRNA catabolic progress, and nuclear export, were enriched in Cluster 1. On the other side, immune-related biological pathways, such as humoral immune response, positive regulation of NK cell-mediated cytotoxicity, and regulation of inflammatory response to an antigenic stimulus, were more enriched in Cluster 2 (**Figure 3A**). These results indicated that Cluster 2 was closely associated with immune modulation of head and neck cancers. In order to comprehend the difference in immune infiltration profiles between Cluster 1 and Cluster 2, we performed CIBERSORT,

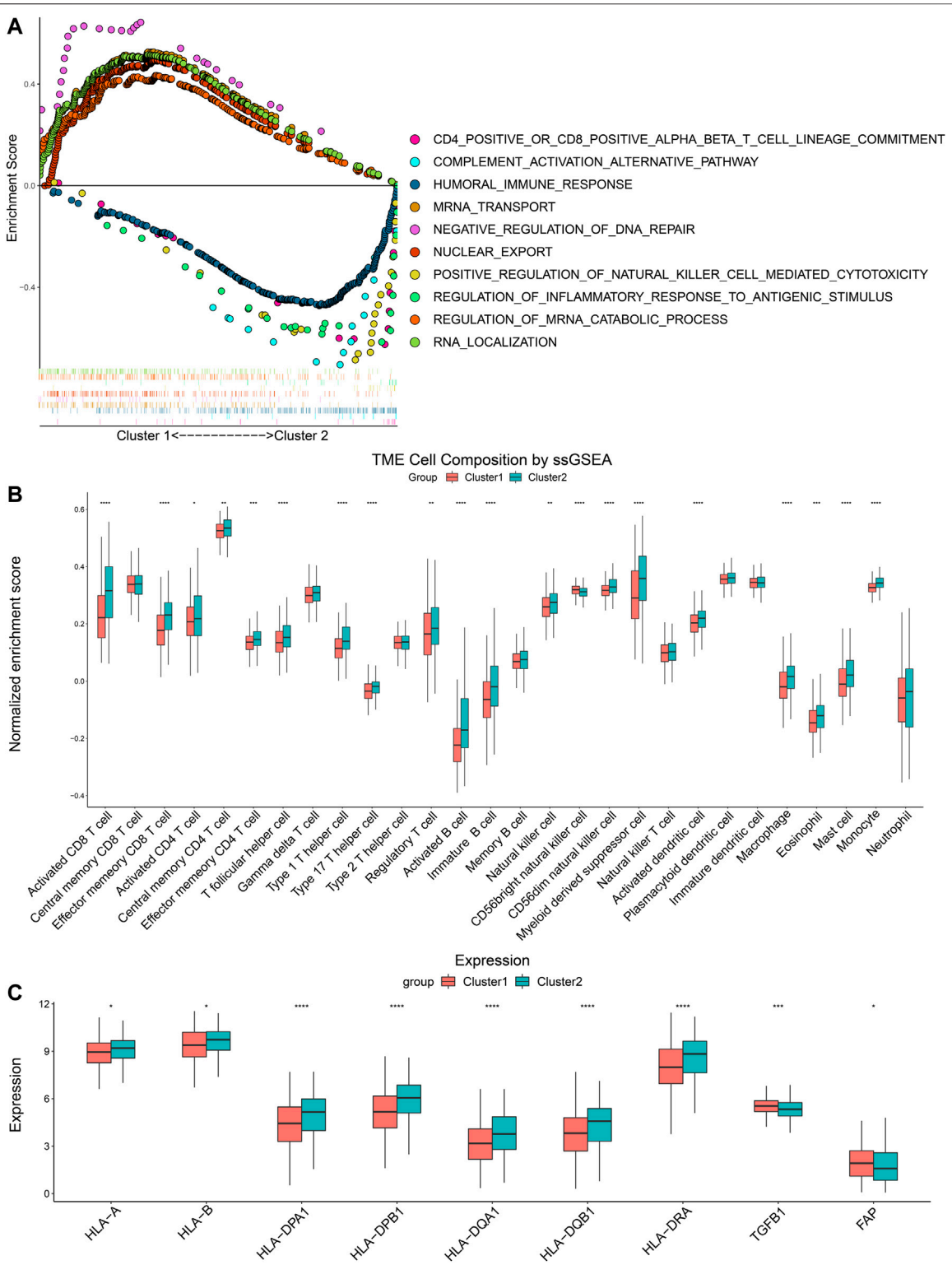


FIGURE 3 | Differences of GSEA and immune cells infiltration between two clusters. **(A)** The tendency of enrichment of biological pathways between two clusters. **(B)** ssGSEA indicated different immune cells infiltration between two clusters. **(C)** Comparison of immune-related molecules between Cluster 1 and Cluster 2. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

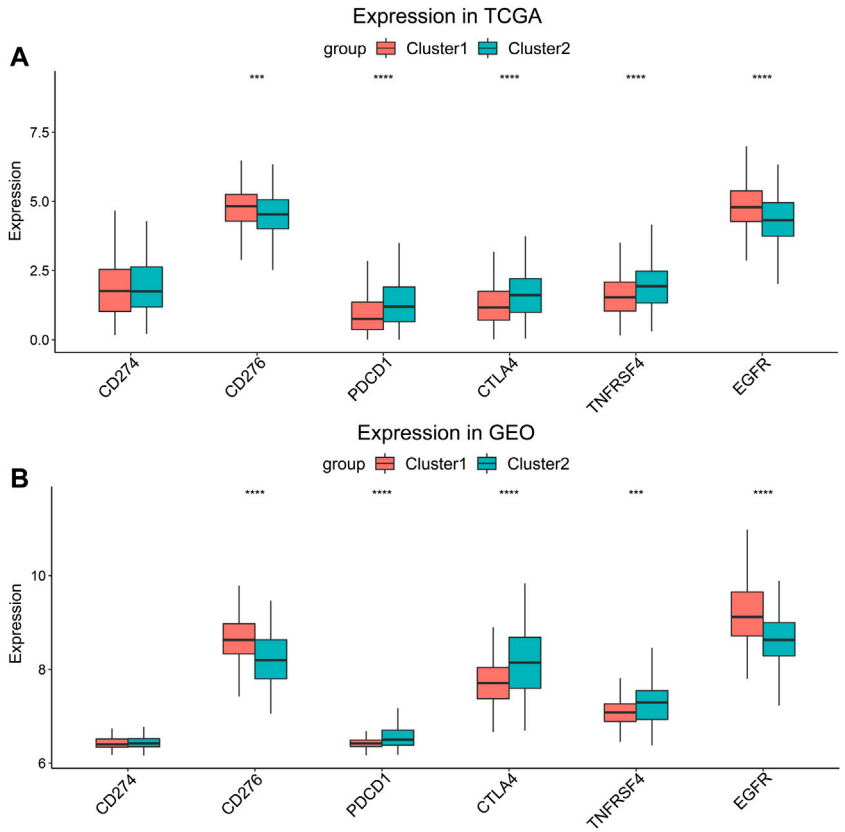


FIGURE 4 | Expression level comparison of HNSCC-related genes. **(A)** TCGA cohort and **(B)** GEO cohort. ***$p < 0.001$, ****$p < 0.0001$.

ssGSEA, and compared the expression of immune-related genes. The results of CIBERSORT indicated that $CD4^+$ T memory resting cells, resting NK cells, M0 macrophages, and activated mast cells had higher percentages in Cluster 1. Cluster 2 highly expressed plasma cells, $CD8^+$ T cells, regulatory T cells, and resting mast cells (**Supplementary Figure S1**). Furthermore, the results of ssGSEA were in parallel with the results of CIBERSORT and demonstrated the majority of immune cell types, including activated $CD8^+$ T cells, activated $CD4^+$ T cells, activated B cells, and natural killer cells, were enriched in the TIME of Cluster 2. Therefore, Cluster 2 manifested more active anti-tumor immune cell gathering (**Figure 3B**). Next, we compared the expression of critical immune-related molecules. Both MHC I and II molecules played a central role in the adaptive immune response. MHC I molecule were encoded by *HLA-A* and *HLA-B* genes, and MHC II molecules were encoded by *HLA-DP*, *HLA-DQ,* and *HLA-DR* genes. Cluster 2 had higher levels of MHC I and II molecules compared with Cluster 1. *TGFB1* encoded transforming growth factor-β (TGF-β) and *FAP* coded fibroblast activation protein alpha (FAP), both of which took part in disabling anti-tumor immune cells and impeding infiltration of immune cells. Cluster 2 had lower levels of TGF-β and FAP than Cluster 1 (**Figure 3C**). These results showed that Cluster 2 had a more immune-stimulatory TIME than Cluster 1.

ASSOCIATION OF YTHDF1- AND IGF2BP2-BASED CLUSTERING WITH HNSCC-RELATED GENES

Several genes were known to influence biological behavior and response to immunotherapy of HNSCC (**Supplementary Material S2**). We compared their expression levels between Cluster 1 and Cluster 2 in the TCGA cohort (**Figure 4A**) and the GEO cohort (**Figure 4B**). *PDCD1*, *CTLA4,* and *TNFRSF4* (encoding OX40) were higher expressed in Cluster 2, while the expression of *CD276* and *EGFR* were higher in Cluster 1. But there was no significant difference in the expression of *CD274* (encoding PD-L1) between the two clusters. The higher expression of *CD276* and *EGFR* suggested that the HNSCC of Cluster 1 might connect with worse biological behavior.

Identifying Hub Genes by DEG Analysis and WGCNA

After removing duplicated samples and extracting mRNA expression from original transcriptome data (HTSeq-Counts), we got a matrix of 499 tumor patients with a unique sample and 18,192 gene expression data. We performed DEG analysis

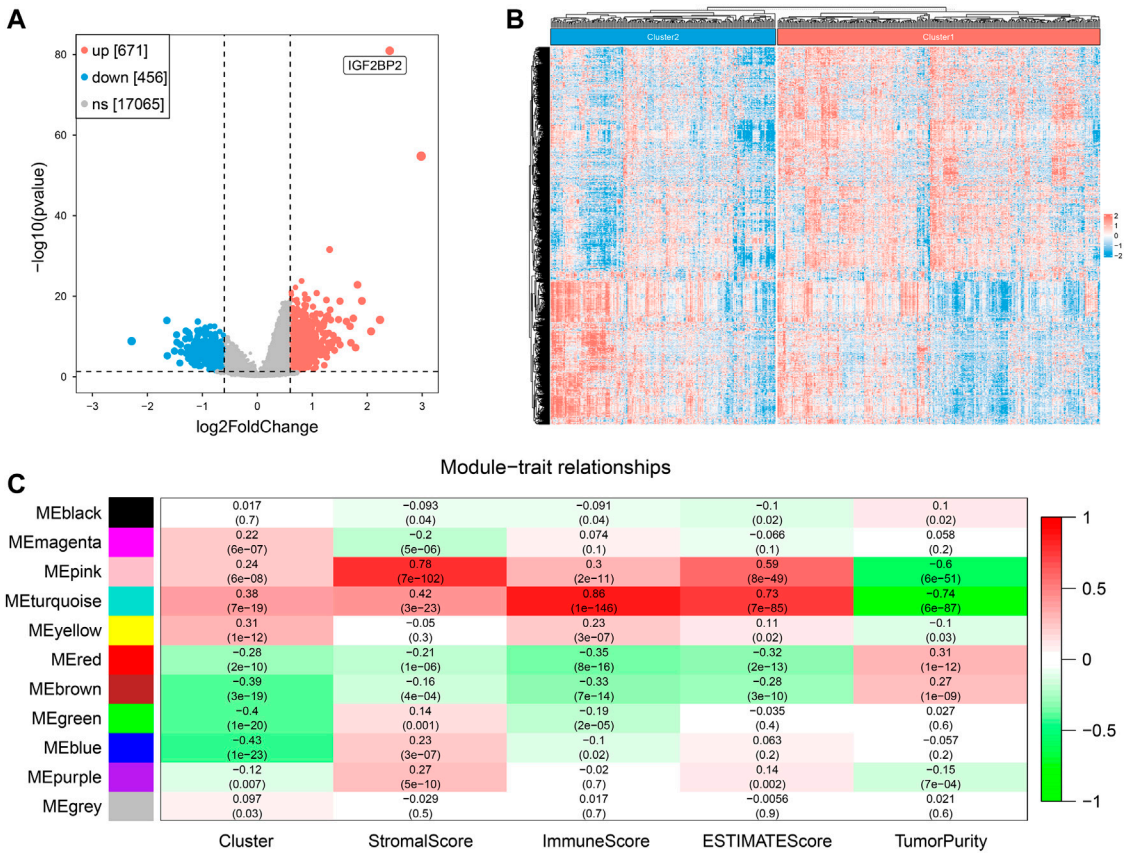


FIGURE 5 | Screening of hub genes related to clustering and ImmuneScore in two clusters. **(A)** The volcano plot for results of differentially expressed genes. **(B)** The heatmap of differentially expressed genes. **(C)** Correlation of gene modules with results of clustering and ESTIMATE.

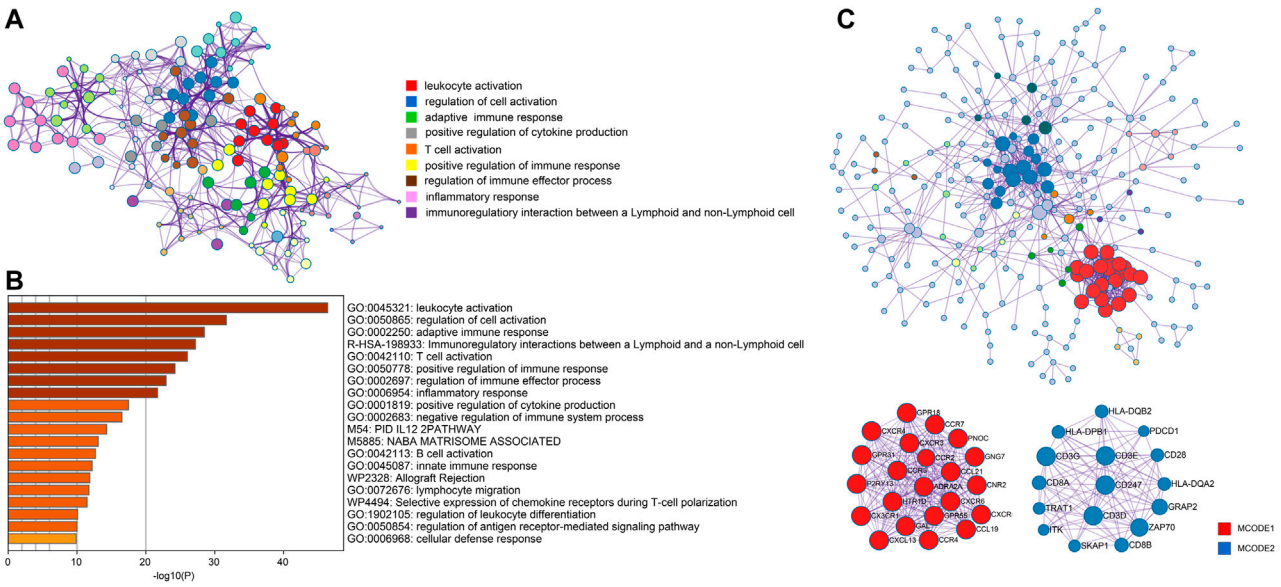


FIGURE 6 | Analysis of functional enrichment and construction of PPI network. **(A)** Selected enriched terms for a network, colored by cluster group ID. **(B)** Functional enrichment analysis in various ontology sources. **(C)** Protein-protein interaction network analysis for whole selected genes and two highlighted MCODE components.

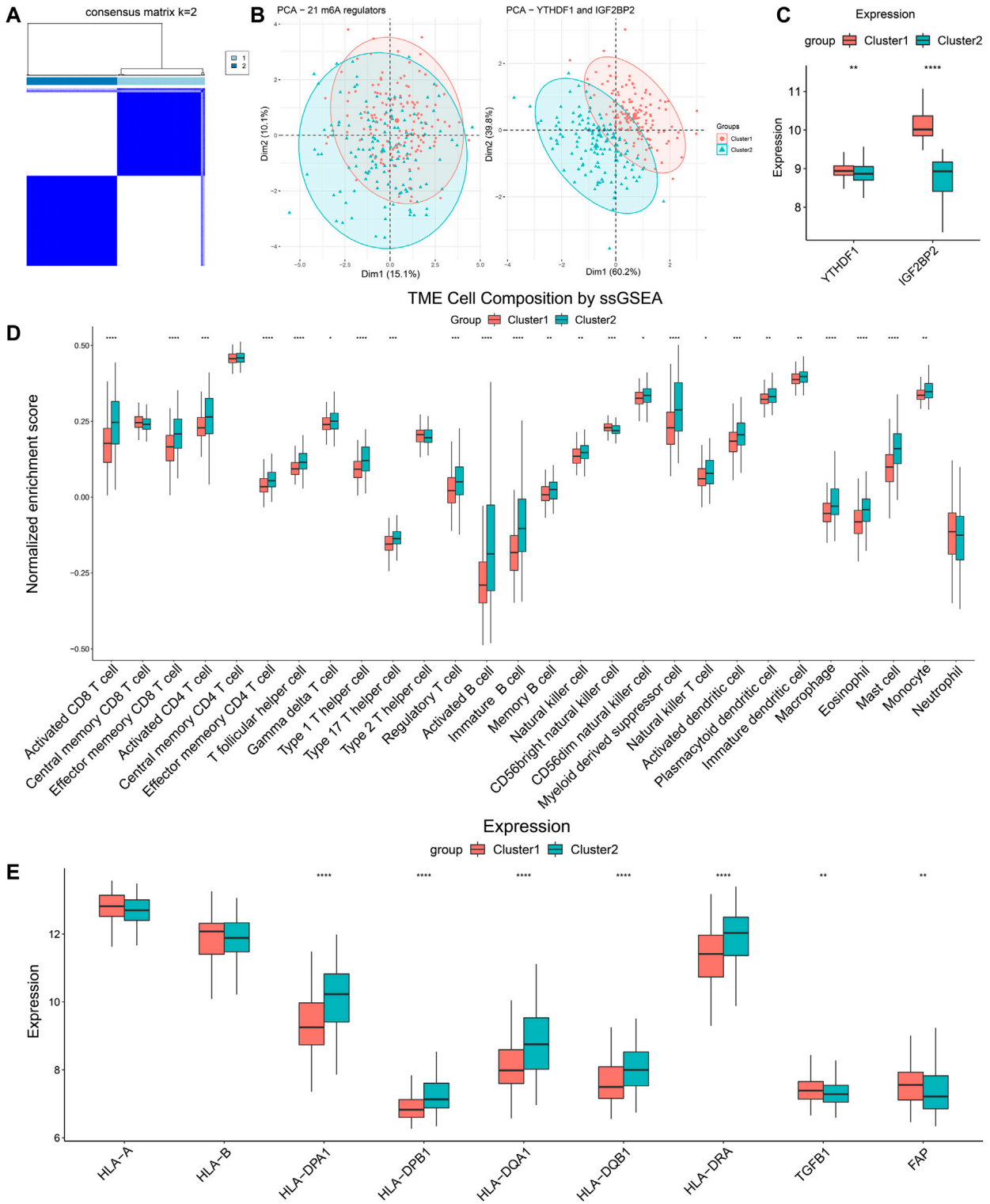


FIGURE 7 | Validation of clustering based on YTHDF1 and IGF2BP2 in GSE65858. **(A)** Consensus clustering matrix for k = 2. **(B)** The results of PCA of clustering based on 21 m6A regulators as well as *YTHDF1* and *IGF2BP2*. **(C)** Comparison of expression level of *YTHDF1* and *IGF2BP2* between Cluster 1 and Cluster 2. **(D)** ssGSEA indicated different immune cells infiltration between two clusters. **(E)** Comparison of immune-related molecules between Cluster 1 and Cluster 2. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

between the two clusters and obtained 1127 DEGs (671 upregulated and 456 downregulated). A volcano plot and heatmap were used to visualize the DEG results (**Figures 5A,B**). We set 3 as the soft power (**Supplementary Figure S2**) and then divided the 1127 DEGs into 11 modules by conducting WGCNA (**Supplementary Figure S3**), and we found that the turquoise module, containing 416 genes, was the most relevant with both clustering ($R = 0.38$, P = 7e−19) and ImmuneScore ($R = 0.86$, P = 1e−146) (**Figure 5C**). These results suggested that DEGs of the turquoise modules played an important role in influencing clustering and immune profiles.

Functional Enrichment Analysis and PPI Network Analysis of Hub Genes

In order to better understand genes in the turquoise model, we uploaded them to Metascape for functional enrichment analysis and constructed the PPI network. The most significant pathways in the functional enrichment analysis were related to immune modulation, including lymphocyte activation, positive regulation of immune response, regulation of immune effector process, and B cell activation (**Figures 6A,B**). In the PPI network analysis, the MCODE algorithm further divided the whole PPI network into two major MCODEs. The MCODE 1 was related to G alpha (i) signaling events and GPCR ligand binding. The MCODE2 contained 16 genes (*PDCD1*, *CD28*, *CD247*, *CD3D*, *CD3E*, *CD3G*, *CD8A*, *CD8B*, *HLA-DPB1*, *HLA-DQA2*, *HLA-DQB2*, *GRAP2*, *TRAT1*, *SKAP1*, *ZAP70,* and *ITK*) was tightly correlated with the generation of second messenger molecules and T-cell receptor signaling pathway (**Figure 6C**).

Verification of Immune Characteristics of Clustering Based on YTHDF1 and IGF2BP2 in the GEO Database

To externally validate the significance of the m6A reader-based clustering, we obtained the expression profiling data of the GSE65858 array from the GEO database. The GSE65858 array involved 270 patients with HNSCC. Similarly, we performed consensus clustering in the 270 patients using the R package "ConsensusClusterPlus," and found that the optimal number of clustering was 2 (**Figure 7A**). PCA plot indicated the above clustering had good efficiency of distinction (**Figure 7B**). There were 137 patients in Cluster 1 and 133 patients in Cluster 2, and the latter had significantly lower expression of *YTHDF1* and *IGF2BP2* (**Figure 7C**). We also performed CIBERSORT and ssGSEA to estimate TME composition. The results of CIBERSORT showed that, compared with Cluster 1, Cluster 2 had significantly more plasma cells, $CD8^+$ T cells, regulatory T cells, gamma delta T cells, and less activated dendritic cells and activated mast cells (**Supplementary Figure S4**). In addition, results obtained from ssGSEA indicated that Cluster 2 highly expressed activated $CD8^+$ T cells, activated $CD4^+$ T cells, T helper cells (Type 1 and 17), and activated B cells, indicating that Cluster 2 had a more favorable TIME than Cluster 1 (**Figure 7D**). A comparison of immune-related molecules was also performed, and we found Cluster 2 had higher expression of MHC II molecule and lower levels of TGF-β and FAP (**Figure 7E**). The above results from the GEO database confirmed that *YTHDF1*- and *IGF2BP2*-based patient clustering showed distinct immune profiles.

DISCUSSION

m6A methylation on mRNA is an abundant internal epigenetic modification that has attracted great attention in recent decades, especially in the tumor research area. Exiting research studies had revealed the complex roles of m6A in cancer by regulating the expression of oncogenes and tumor suppressor genes. This effect was cancer-dependent and also varied among different types of m6A regulators. For instance, the m6A writer, METTL3 was found to promote the translation of c-MYC and BCL2 to accelerate leukemia progression by suppressing differentiation and apoptosis in acute myelocytic leukemia (Vu et al., 2017). On the other hand, the m6A eraser, FTO was found to promote the degradation of BNIP3 and inhibited the proliferation and invasion of breast cancer cells (Niu et al., 2019). And the readers, including EIF3, YTH family, and IGF2BP family, mainly regulated the translation and degradation of targeted RNA to participate in m6A modification. Recent studies demonstrated that m6A modification played an important role in regulating the immune response. Targeting and disabling IGF2BPs through circNDUFB2 could prevent the progression of non-small cell lung cancer and activate anti-tumor immunity (Li et al., 2021). YTHDF1 mediated the increase of lysosomal proteases and tumor antigen degradation in dendritic cells and could weaken anti-tumor response and disable $CD8^+$ T cells (Han et al., 2019). However, research about the effect of m6A modification on HNSCC was inadequate.

We utilized transcriptomic data from the TCGA dataset and GEO dataset to establish an m6A regulator-based immune phenotype of HNSCC. Specifically, we found that two readers of m6A, *IGF2BP2*, and *YTHDF1* could effectively indicate immune-stimulatory and immune-suppressive HNSCC. To this end, we scored every patient in the TCGA cohort with the ESTIMATE tool and CIBERSORT algorithm to calculate the correlation of m6A regulators with ESTIMATE scores and immune cells infiltration. The absolute Spearman's coefficient of ImmuneScore of *IGF2BP1*, *IGF2BP2*, *IGF2BP3* and *YTHDF1* were above 0.2. We selected two with the highest absolute coefficients, *IGF2BP2,* and *YTHDF1* for the following analysis. We found patients with HNSCC could be divided into two clusters with different immune profiles based on the expression of *IGF2BP2* and *YTHDF1*. TME was roughly categorized into three types, namely "infiltrated," "excluded" and "desert" (Hegde and Chen, 2020). The infiltrated type was characterized by sufficient infiltration of $CD8^+$ T cells and a high level of MHC I molecule. Desert type was featured by the absence of $CD8^+$ T cell infiltration, low level of MHC I molecule, and high level of FAP. According to the expression

profiles of immune cells and immune molecules, we tended to consider Cluster 1 as excluded or desert type and Cluster 2 as infiltrated type. In addition, GSEA demonstrated that a number of signal pathways related to immune response were enriched in Cluster 2. Therefore, it was speculated that the TME of Cluster 2 was more favorable for ICIs-based immunotherapy.

Overexpression of *EGFR* could be detected in over 90% of HNSCC, which was an important signal receptor that brings about tumorigenesis, proliferation, and metastasis through downstream pathways, including PI3K/AKT and MAPK (Nicholson et al., 2001; Citri and Yarden, 2006). CD276, also named B7-H3, was one of the immune checkpoint molecules, which was upregulated in HNSCC and helped tumor cells evade immunological surveillance. High expression of *CD276* was related to the occurrence, progression, and metastasis of HNSCC(Wang et al., 2021). Both *EGFR* and *CD276* were found highly expressed in Cluster 1, indicating that HNSCC of Cluster 1 were more likely to correlate with worse biological behavior, poorer clinical result, and insensitivity to ICIs-based immunotherapy. However, EGFR antibodies or CD276 blockade could be considered for HNSCC of Cluster 1.

To further investigate internal influencing factors between the two clusters, we screened DEGs and performed WGCNA to find the module closely related to clustering and ImmuneScores. Successively, we constructed a PPI network, we finally obtained 16 genes (*PDCD1*, *CD28*, *CD247*, *CD3D*, *CD3E*, *CD3G*, *CD8A*, *CD8B*, *HLA-DPB1*, *HLA-DQA2*, *HLA-DQB2*, *GRAP2*, *TRAT1*, *SKAP1*, *ZAP70,* and *ITK*). These 16 hub genes were all upregulated in Cluster 2. PDCD1 was a receptor of immunosuppression usually expressed in activated T cells. CD28 played an essential role in T cells proliferation, and survival, and provided the second signal for T cell activation (Esensten et al., 2016). CD247, CD3D, CD3E, and CD3G participated in constituting T-cell receptor-CD3 complex (TCR-CD3) to recognize antigens and deliver the first signal for T cell activation (Kuhns et al., 2006). CD8A and CD8B acted as co-receptors for TCR (Rudolph et al., 2006). MHC II molecule was encoded by *HLA-DPB1*, *HLA-DQA2,* and *HLA-DQB2* and played an important role in antigens binding and cross-presentation. Proteins encoded by *ZAP70* and *ITK* belonged to the tyrosine kinase family, which were critical for signal transduction in T cells (Berg, 2007; Au-Yeung et al., 2018). And proteins encoded by *GRAP2*, *TRAT1*, and *SKAP1* also played an important role in signal transduction (Raab et al., 2010). It was, therefore, suggested that most of these DEGs were highly relevant to immune response and might contribute to different immune profiles between the two clusters.

There were several limitations in our study. We performed the study based on TCGA and GEO databases without verification using a clinical dataset. Only correlation analysis on phenotype level was conducted. There lacked a demonstration on the protein level. Last but not least, there was a lack of mechanistic study.

In conclusion, our study divided HNSCC into two clusters based on *IGF2BP2* and *YTHDF1*, which provided a simple and feasible tool to identify HNSCC with different immune profiles and helped estimate sensitivity to ICIs-based immunotherapy. We preliminarily explore the possible mechanisms, combined with the previous research works about IGF2BP2 and YTHDF1, and we speculated that they might hamper the expression of specific genes, which were related to antigen recognition, signal transduction, proliferation, and activation of effector T cells. Meanwhile, they might increase the stability of tumorigenic genes, such as *EGFR* and *CD276*, and excessively activated downstream signal pathways. The joint effect led to different biologic behavior and immune profiles in the two clusters. The existing study had found knocking down *YTHDF1* could enhance the therapeutic efficiency of ICIs in mice (Han et al., 2019). It would be important to determine whether suppressing the expressions of *IGF2BP2* and *YTHDF1* could modulate the sensitivity to ICI-based immunotherapy. Specific molecular mechanisms awaited further exploration.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent from the patients/participants or patients/participants' legal guardian/ next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Conceptualization: QW, SL, and YZ; methodology: QW, SL, and YZ; software and data analysis: QW, JL, and SL; validation: QW, SL, JL, and YZ; writing, review and editing: QW, SL, JL, and YZ; and supervision: YZ. All authors have read and agreed to the published version of the manuscript.

FUNDING

This study was supported by a grant from the Scientific Research Project of Hubei Provincial Health and Family Planning Commission (Grant No. WJ 2019H060), a grant from the Leading Discipline Construction Project of Oncology of Zhongnan Hospital of Wuhan University, and a grant from the Science, Technology, and Innovation Seed Fund of Zhongnan Hospital of Wuhan University (Grant no. znpy2018123).

ACKNOWLEDGMENTS

The authors would like to acknowledge the Scientific Research Project of Hubei Provincial Health and Family Planning Commission for their support in their study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.903634/full#supplementary-material

REFERENCES

Argiris, A., Karamouzis, M. V., Raben, D., and Ferris, R. L. (2008). Head and Neck Cancer. *Lancet* 371 (9625), 1695–1709. doi:10.1016/s0140-6736(08)60728-x

Au-Yeung, B. B., Shah, N. H., Shen, L., and Weiss, A. (2018). ZAP-70 in Signaling, Biology, and Disease. *Annu. Rev. Immunol.* 36, 127–156. doi:10.1146/annurev-immunol-042617-053335

Bader, G. D., and Hogue, C. W. (2003). An Automated Method for Finding Molecular Complexes in Large Protein Interaction Networks. *BMC Bioinformatics* 4, 2. doi:10.1186/1471-2105-4-2

Berg, L. J. (2007). Signalling through TEC Kinases Regulates Conventional versus Innate CD8+ T-Cell Development. *Nat. Rev. Immunol.* 7 (6), 479–485. doi:10.1038/nri2091

Chow, L. Q. M. (2020). Head and Neck Cancer. *N. Engl. J. Med.* 382 (1), 60–72. doi:10.1056/nejmra1715715

Citri, A., and Yarden, Y. (2006). EGF-ERBB Signalling: towards the Systems Level. *Nat. Rev. Mol. Cel Biol.* 7 (7), 505–516. doi:10.1038/nrm1962

Dixit, D., Prager, B. C., Gimple, R. C., Poh, H. X., Wang, Y., Wu, Q., et al. (2021). The RNA m6A Reader YTHDF2 Maintains Oncogene Expression and Is a Targetable Dependency in Glioblastoma Stem Cells. *Cancer Discov.* 11 (2), 480–499. doi:10.1158/2159-8290.cd-20-0331

Esensten, J. H., Helou, Y. A., Chopra, G., Weiss, A., and Bluestone, J. A. (2016). CD28 Costimulation: From Mechanism to Therapy. *Immunity* 44 (5), 973–988. doi:10.1016/j.immuni.2016.04.020

Han, D., Liu, J., Chen, C., Dong, L., Liu, Y., Chang, R., et al. (2019). Anti-tumour Immunity Controlled through mRNA m6A Methylation and YTHDF1 in Dendritic Cells. *Nature* 566 (7743), 270–274. doi:10.1038/s41586-019-0916-x

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). GSVA: Gene Set Variation Analysis for Microarray and RNA-Seq Data. *BMC Bioinformatics* 14, 7. doi:10.1186/1471-2105-14-7

Hegde, P. S., and Chen, D. S. (2020). Top 10 Challenges in Cancer Immunotherapy. *Immunity* 52 (1), 17–35. doi:10.1016/j.immuni.2019.12.011

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

Kuhns, M. S., Davis, M. M., and Garcia, K. C. (2006). Deconstructing the Form and Function of the TCR/CD3 Complex. *Immunity* 24 (2), 133–139. doi:10.1016/j.immuni.2006.01.006

Langfelder, P., and Horvath, S. (2008). WGCNA: An R Package for Weighted Correlation Network Analysis. *BMC Bioinformatics* 9, 559. doi:10.1186/1471-2105-9-559

Li, B., Zhu, L., Lu, C., Wang, C., Wang, H., Jin, H., et al. (2021). circNDUFB2 Inhibits Non-small Cell Lung Cancer Progression via Destabilizing IGF2BPs and Activating Anti-tumor Immunity. *Nat. Commun.* 12 (1), 295. doi:10.1038/s41467-020-20527-z

Li T, T., Hu, P.-S., Zuo, Z., Lin, J.-F., Li, X., Wu, Q.-N., et al. (2019). METTL3 Facilitates Tumor Progression via an m6A-IGF2BP2-dependent Mechanism in Colorectal Carcinoma. *Mol. Cancer* 18 (1), 112. doi:10.1186/s12943-019-1038-7

Li Y, Y., Xiao, J., Bai, J., Tian, Y., Qu, Y., Chen, X., et al. (2019). Molecular Characterization and Clinical Relevance of m6A Regulators across 33 Cancer Types. *Mol. Cancer* 18 (1), 137. doi:10.1186/s12943-019-1066-3

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

Mootha, V. K., Lindgren, C. M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., et al. (2003). PGC-1α-responsive Genes Involved in Oxidative Phosphorylation Are Coordinately Downregulated in Human Diabetes. *Nat. Genet.* 34 (3), 267–273. doi:10.1038/ng1180

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Nicholson, R. I., Gee, J. M., and Harper, M. E. (2001). EGFR and Cancer Prognosis. *Eur. J. Cancer* 37 (Suppl. 4), S9–S15. doi:10.1016/s0959-8049(01)00231-3

Niu, Y., Lin, Z., Wan, A., Chen, H., Liang, H., Sun, L., et al. (2019). RNA N6-Methyladenosine Demethylase FTO Promotes Breast Tumor Progression through Inhibiting BNIP3. *Mol. Cancer* 18 (1), 46. doi:10.1186/s12943-019-1004-4

Nombela, P., Miguel-López, B., and Blanco, S. (2021). The Role of m6A, m5C and Ψ RNA Modifications in Cancer: Novel Therapeutic Opportunities. *Mol. Cancer* 20 (1), 18. doi:10.1186/s12943-020-01263-w

Raab, M., Wang, H., Lu, Y., Smith, X., Wu, Z., Strebhardt, K., et al. (2010). T Cell Receptor "Inside-Out" Pathway via Signaling Module SKAP1-RapL Regulates T Cell Motility and Interactions in Lymph Nodes. *Immunity* 32 (4), 541–556. doi:10.1016/j.immuni.2010.03.007

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

Rudolph, M. G., Stanfield, R. L., and Wilson, I. A. (2006). How TCRs Bind MHCs, Peptides, and Coreceptors. *Annu. Rev. Immunol.* 24, 419–466. doi:10.1146/annurev.immunol.23.021704.115658

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics. *CA Cancer J. Clin.* 71, 7–33.

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

Visvanathan, A., Patil, V., Arora, A., Hegde, A. S., Arivazhagan, A., Santosh, V., et al. (2018). Essential Role of METTL3-Mediated m6A 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 N6-Methyladenosine (m6A)-Forming Enzyme METTL3 Controls Myeloid Differentiation of normal Hematopoietic and Leukemia Cells. *Nat. Med.* 23 (11), 1369–1376. doi:10.1038/nm.4416

Wang, L., Hui, H., Agrawal, K., Kang, Y., Li, N., Tang, R., et al. (2020). m6 A RNA Methyltransferases METTL3/14 Regulate Immune Responses to Anti-PD-1 therapyA RNA Methyltransferases METTL3/14 Regulate Immune Responses to Anti-PD-1 Therapy. *EMBO J.* 39 (20), e104514. doi:10.15252/embj.2020104514

Wang, C., Li, Y., Jia, L., Kim, J. k., Li, J., Deng, P., et al. (2021). CD276 Expression Enables Squamous Cell Carcinoma Stem Cells to Evade Immune Surveillance. *Cell Stem Cell* 28 (9), 1597–1613. doi:10.1016/j.stem.2021.04.011

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: a Class Discovery Tool with Confidence Assessments and Item Tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

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

Zaccara, S., Ries, R. J., and Jaffrey, S. R. (2019). Reading, Writing and Erasing mRNA Methylation. *Nat. Rev. Mol. Cel Biol.* 20 (10), 608–624. doi:10.1038/s41580-019-0168-5

Zhang, B., and Horvath, S. (2005). A General Framework for Weighted Gene Co-expression Network Analysis. *Stat. Appl. Genet. Mol. Biol.* 4, 17. doi:10.2202/1544-6115.1128

Zhong, L., Liao, D., Zhang, M., Zeng, C., Li, X., Zhang, R., et al. (2019). YTHDF2 Suppresses Cell Proliferation and Growth via Destabilizing the EGFR mRNA in Hepatocellular Carcinoma. *Cancer Lett.* 442, 252–261. doi:10.1016/j.canlet.2018.11.006

Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A. H., Tanaseichuk, O., et al. (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

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.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 08 June 2022
doi: 10.3389/fgene.2022.907754

Risk Model and Immune Signature of m7G-Related lncRNA Based on Lung Adenocarcinoma

Chuanhao Zhang[1,2], Dong Zhou[2], Zhe Wang[2], Zaishuang Ju[2], Jiabei He[2], Genghao Zhao[1,2] and Ruoyu Wang[2]**

[1]Graduate School of Dalian Medical University, Dalian, China, [2]Departement of Medical Oncology, Affiliated Zhongshan Hospital of Dalian University, Dalian, China

Lung cancer is a major cause of cancer-related deaths globally, with a dismal prognosis. N7-methylguanosine (m7G) is essential for the transcriptional phenotypic modification of messenger RNA (mRNA) and long noncoding RNA (lncRNA). However, research on m7G-related lncRNAs involved in lung adenocarcinoma (LUAD) regulation is still limited. Herein, we aim to establish a prognostic model of m7G-related lncRNAs and investigate their immune properties. Eight prognostic m7G-related lncRNAs were identified using univariate Cox analysis. Six m7G-related lncRNAs were identified using LASSO-Cox regression analysis to construct risk models, and all LUAD patients in The Cancer Genome Atlas (TCGA) cohort was divided into low-risk and high-risk subgroups. The accuracy of the model was verified by Kaplan-Meier analysis, time-dependent receiver operating characteristic, principal component analysis, independent prognostic analysis, nomogram, and calibration curve. Further studies were conducted on the gene set enrichment and disease ontology enrichment analyses. The gene set enrichment analysis (GSEA) revealed that the high-risk group enriched for cancer proliferation pathways, and the enrichment analysis of disease ontology (DO) revealed that lung disease was enriched, rationally explaining the superiority of the risk model. Finally, we found that the low-risk group had higher immune infiltration and checkpoint expression. It can be speculated that the low-risk group has a better effect on immunotherapy. Susceptibility to antitumor drugs in different risk subgroups was assessed, and it found that the high-risk group showed high sensitivity to first-line treatment drugs for non-small cell lung cancer. In conclusion, a risk model based on 6 m7G-related lncRNAs can not only predict the overall survival (OS) rate of LUAD patients but also guide individualized treatment for these patients.

Keywords: lung adenocarcinoma (AC), m7G, lncRNA, model, immune signature, treatment

OPEN ACCESS

Edited by:
Yan Gong,
Wuhan University, China

Reviewed by:
Xiancai Ma,
Guangzhou Laboratory, China
Xuefeng Li,
Guangzhou Medical University, China

****Correspondence:***
Zhe Wang
wangzhe@dlu.edu.cn
Ruoyu Wang
wangruoyu@dlu.edu.cn

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: *30 March 2022*
Accepted: *13 May 2022*
Published: *08 June 2022*

Citation:
Zhang C, Zhou D, Wang Z, Ju Z, He J, Zhao G and Wang R (2022) Risk Model and Immune Signature of m7G-Related lncRNA Based on Lung Adenocarcinoma. Front. Genet. 13:907754. doi: 10.3389/fgene.2022.907754

INTRODUCTION

Lung cancer is the second most common type of cancer worldwide and the leading cause of cancer mortality, accounting for approximately 11.4% of diagnosed cancers and 18.0% of deaths (Sung et al., 2021). Currently, the 5-years survival rate for lung cancer is still very low, only 10–20% in most countries (Allemani et al., 2018). Lung adenocarcinoma (LUAD) is the most commonly diagnosed subtype of lung cancer, accounting for approximately 40% of all cases

(Travis et al., 2015). With the development of surgery, radiotherapy, chemotherapy, targeted therapy, and immunotherapy, the 5-years survival rate of lung cancer has improved, but the performance remains unsatisfactory. There is an urgent need to develop a convenient and fast prognostic model that can accurately judge patient prognosis and guide individualized treatment, which could be very useful for both patients and clinicians.

In humans, the methyltransferase like 1 (METTL1)/WD repeat domain 4 (WDR4) complex catalyzes N7-methylguanosine, one of the most common tRNA modifications in the tRNA variable loop (Alexandrov et al., 2005; Lin et al., 2018). METTL1 is an m7G catalytic enzyme and WDR4 is important in the methyltransferase complex stabilization (Alexandrov et al., 2002). Recently, it was found that METTL1 and WDR4 were significantly up-regulated in lung cancer tissues and played an oncogenic role in lung cancer via mediating m7G tRNA modification and modulated the translation of mRNAs, especially METTL1-mediated m7G tRNA modification and m7G codon usage promoted mRNA translation and lung cancer progression (Ma et al., 2021). This suggests that METTL1 and WDR4 may play a significant role in tumor progression. Therefore, screening m7G-related genes is essential.

Long non-coding RNAs (lncRNAs) are defined as non-coding RNAs of more than 200 nucleotides in length. They are not generally considered to encode proteins but are involved in the regulation of different levels (epigenetic regulation, transcriptional regulation, and post-transcriptional regulation) of genes encoding proteins in the form of RNA (Juhling et al., 2009; Spizzo et al., 2012). Several lncRNAs, including ferroptosis-related lncRNAs (Chen et al., 2022), pyroptosis-related lncRNAs (Xu et al., 2022), and autophagy-related lncRNAs (Luo et al., 2022), have recently been implicated in prognosis in cancer patients, while m7G-related lncRNAs have rarely been reported.

Herein, we identified 6 prognostic risk models of m7G-related lncRNAs and the correlation between the risk model and immune characteristics. As expected, our model well predicted survival in LUAD patients and showed greater efficacy in terms of immune cell invasion and immune checkpoint expression.

MATERIALS AND METHODS

Data Set

RNA sequencing data and associated clinical characteristics of 594 LUAD patients were extracted from The Cancer Genome Atlas (TCGA) database, including 59 normal tissues and 535 LUAD tissues. Forty m7G-related genes were obtained from the gene set enrichment analysis (GSEA) website (http://www.gsea-msigdb.org/gsea/login.jsp) and published articles. Patients lacking clinical information were deleted from subsequent analyses.

Selection of m7G-Related lncRNAs

LncRNAs were screened from 594 patients with LUAD using Strawberry Perl (version 5.30.0). A total of 2093 m7G-related lncRNAs were identified using the limma R package with the following criteria: Pearson correlation coefficient >0.4 and $p < 0.001$. A total of 990 differentially expressed lncRNAs (DELs) were identified in normal lung tissues and LUAD tissues with the following criteria: $\log_2$ fold change (FC) > 1 and false discovery rate <0.05.

Development and Validation of m7G-Related lncRNA Prognostic Model

To rigorously screen out prognostic lncRNAs, the *p*-value was set to 0.01 and univariate Cox analysis was used to identify prognostic lncRNAs. Next, the TCGA cohort was randomly divided into a training and a validation group, each accounting for 50%. Based on these prognostic lncRNAs, Lasso-Cox regression analysis was used to select genes to minimize the risk of overfitting and a risk prediction model was constructed. The risk score was calculated using the following formula:

$$risk\ score = \sum_{i=1}^{n} (coef_i * expr_i) \quad (1)$$

where $coef_i$ represents the coefficients of each lncRNA and $expr_i$ represents the expression level of each lncRNA. Based on the median value of the risk score, patients were divided into low-risk and high-risk groups. Survival curves were drawn between low-risk and high-risk groups using the survival and survminer packages of the R software. The stability of the risk score was performed using the validation group. Clinical information (including age, gender, and stage) of TCGA-LUAD patients was extracted and combined with the risk score for univariate and multivariate Cox regression analysis to evaluate whether the risk score is an independent prognostic factor for overall survival (OS), and compared predictive results of different factors using receiver operating characteristic (ROC) curve analysis.

Nomogram and Calibration

The rms R package was utilized to construct nomograms. Calibration curves were used to quantify the agreement between the predicted and the actual results for 1-, 3-, and 5-years survival rates.

Gene Set Enrichment Analysis

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was performed, and significantly enriched pathways in different risk subgroups were identified using GSEA software ($p < 0.05$ and FDR <0.25).

Assessment of Immune Cell Infiltration and Immune Checkpoints

All TCGA tumor immune cell infiltration files were downloaded from TIMER 2.0 and the correlation between the explored immune infiltrating cells and the risk score was analyzed using limma, scales, ggplot2 and ggtext R packages. Additionally, immune cell infiltration, immune-related signaling pathways,

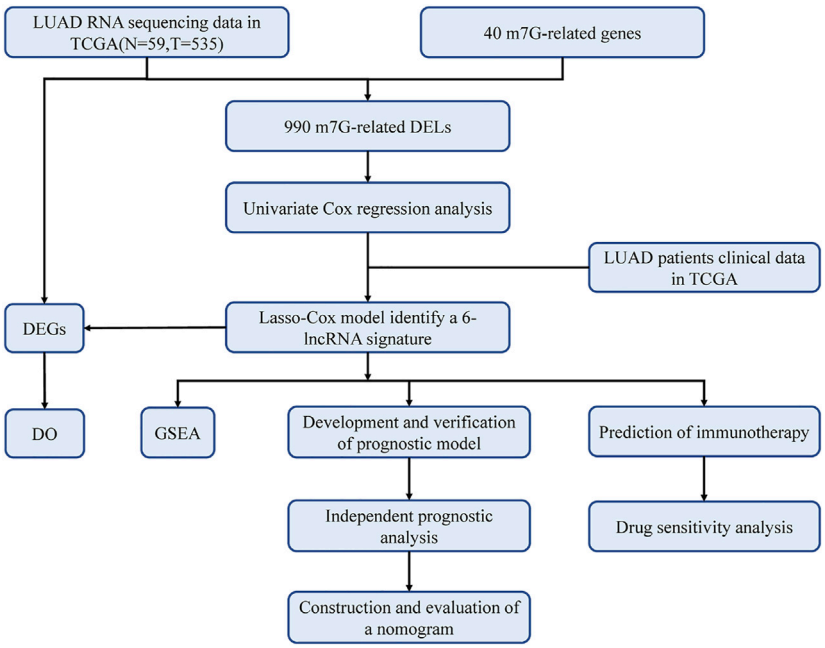


FIGURE 1 | Workflow of this study. The TCGA database was utilized to screen 990 differentially expressed lncRNAs (DELs) in LUAD, which were analyzed with LASSO-COX regression to develop a prognostic model for m7G-related lncRNAs. The prognostic model had been validated in multiple ways and proved to be stable and reliable. Therefore, based on this model, we also performed disease ontology enrichment analysis (DO), gene set enrichment analysis (GSEA), immune-related analysis and drug sensitivity analysis to determine the potential function of prognostic signatures.

tumor microenvironment (TME) scores, and immune checkpoints were compared between low-risk and high-risk groups using the ggpubr package.

Prediction of Drug Susceptibility

The pRRophetic R package was utilized to predict the half-maximal inhibitory concentration (IC_{50}) value of cancer drugs in different risk subgroups, which represents the effectiveness of a substance in inhibiting a specific biological or biochemical process.

Statistical Analysis

All statistical analyses were performed using R software (version 4.0.4). The Wilcoxon signed-rank test was used to investigate differences in the composition of immune infiltrating cells. The correlation between m7G-related genes and m7G-related lncRNAs was investigated using Spearman correlation analysis. Kaplan-Meier analysis was used to estimate survival curves. p values <0.05 (*), 0.01 (**), and 0.001 (***) were considered statistically significant.

RESULTS

Workflow of Study

The study flowchart is shown in **Figure 1**. The precise procedure is as follows: First, we obtained RNA sequencing from the TCGA database for 594 lung adenocarcinoma patients, as well as 40 m7G-associated genes from the GSEA database and relevant literature. Furthermore, a 6-lncRNA prognostic model was developed, and its stability was validated using multiple techniques. Finally, GSEA and DO enrichment analysis validated the superiority of the model, while immunological correlation analysis and drug sensitivity analysis extended on the idea of clinical treatment.

Identification of Differentially Expressed m7G-Related lncRNAs

Data for 594 LUAD samples were obtained from the TCGA database, and 14,056 lncRNAs and 19,573 mRNAs were detected. Forty m7G-related genes were obtained from published articles and the GSEA website (Letoquart et al., 2014; Trotman and Schoenberg, 2019; Galloway et al., 2021; Ma et al., 2021). The co-expression network between m7G-related genes and lncRNAs is shown in **Figure 2A**. A total of 990 DELs were screened from 59 normal tissues and 535 LUAD tissues ($|\text{Log}_2 FC| > 1$ and $p < 0.05$). Of these, 903 lncRNAs were up-regulated and 87 were down-regulated (**Figure 2B**).

Development and Validation of Prognostic Gene Models

Patients from the TCGA-LUAD database were randomly split into two groups: a training set and a validation set. To strictly screen prognosis-related DELs, the p-value was set to 0.01, and performed univariate Cox regression analysis was performed on the training group. Eight prognosis-related lncRNAs met the conditions, including AC092718.3, LINC01352, AP000695.1, AC018647.1, AL355472.3, AC026355.2, SALRNA1 and AL157895.1 (**Figure 3A**). These prognosis-related lncRNAs are shown in **Figure 3B**.

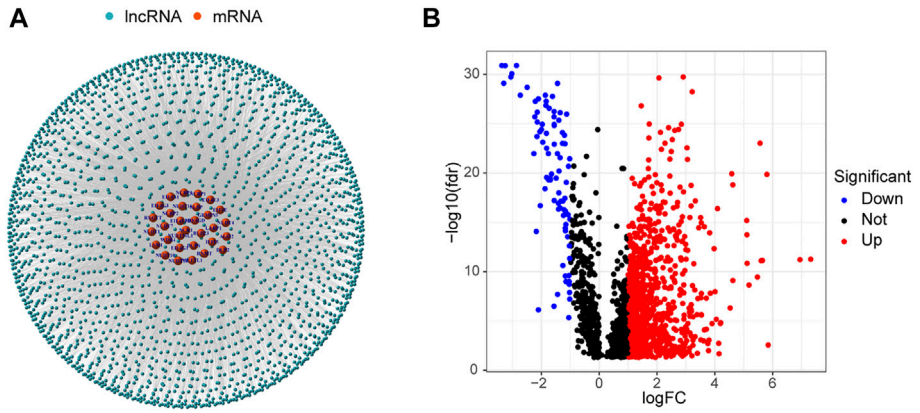


FIGURE 2 | Identification of m7G-related lncRNAs in LUAD patients. **(A)** Co-expression network of m7G-related lncRNAs and mRNAs. **(B)** 990 differentially expressed lncRNAs.

Furthermore, these lncRNAs were positively regulated by corresponding genes in the Sankey diagram (**Figure 3E**). LASSO regression analysis was then performed on these prognosis-associated lncRNAs. Cross-validation was also performed to obtain the best λ value from the smallest partial likelihood bias (**Figures 3C,D**), to further identify lncRNAs significantly associated with prognosis in LUAD patients. Moreover, multivariate Cox regression analysis was used to screen six prognosis-related lncRNAs and calculate the respective coefficients of these lncRNAs. Finally, six candidates, including LINC01352, AP000695.1, AC018647.1, AL355472.3, AC026355.2, SALRNA1, were selected to construct a risk model. The risk score was calculated using the following formula: LINC01352*(-1.42486) + AP000695.1*(0.37854) + AC018647.1*(-2.19905) + AL355472.3*(1.05547) + AC026355.2*(-0.38520) + SALRNA1*(-1.39428).

The median score was calculated based on the above formula, and the TCGA-LUAD cohort, training group, and validation group were classified into low-risk and high-risk subgroups, and the principal component analysis, risk score distribution, and survival status distribution were visualized, respectively (**Figures 4A–C**). The results revealed that the sample distribution of the two risk groups was reasonable. Kaplan-Meier survival analysis showed that the OS was shorter in the high-risk group than in the low-risk group (**Figures 4D–F**).

Independent Prognostic Value of Risk Models

Univariate and multivariate Cox regression analyses were performed on the TCGA-LUAD cohort to evaluate the accuracy of the risk model and determine whether risk score could serve as an independent prognostic factor for patient survival. Univariate Cox regression analysis showed that both the risk score and the stage were significantly related to the prognosis of the patient (**Figure 5A**). After controlling for other confounding factors, multivariate analysis revealed that risk score and stage were independent prognostic factors (**Figure 5B**). To expand the applicability of the risk model, the stage was divided into two subgroups: early-stage (Stage I and Stage II) and late-stage (Stage III and Stage IV). The survival curves are shown in **Figures 5C,D**. Patients with advanced Stage had a very poor prognosis, which is completely consistent with the clinical data.

A time-dependent ROC curve was generated in the TCGA-LUAD cohort, and the area under the curve (AUC) reached 0.705, 0.686, and 0.723 at 1, 3, and 5 years, respectively (**Figure 4I**). In addition, ROC curves confirmed that the risk signature had better prognostic accuracy compared with other clinicopathological features (**Figure 5F**). Time-dependent ROC curves also showed excellent predictive power in both the training and validation sets (**Figures 4G,H**).

Construction of Nomogram

Based on the TCGA-LUAD cohort, risk scores and clinical factors were integrated to create a nomogram (**Figure 5E**) to improve the predictive power of survival in LUAD patients. Calibration plots for 1-, 3- and 5-years OS revealed good agreement between nomogram prediction and actual observations (**Figure 5G**).

GSEA and DO

GSEA software was used to explore KEGG pathways in the entire collection to investigate differences in signaling pathways in different risk subgroups. It was found that pathways related to cancer proliferation, such as cell cycle, DNA replication, mismatch repair, proteasome, homologous recombination, etc., were enriched in the high-risk groups. In addition, the low-risk group was mainly enriched in pathways such as autoimmune thyroid disease, asthma, primary bile acid biosynthesis, arachidonic acid metabolism, and alpha linolenic acid metabolism (**Figure 6A**). The majority of enriched pathways in the high-risk group were closely related to radiotherapy (Azzam et al., 2012; Haro et al., 2012). Hence, we speculate that radiotherapy may have unexpected effects on

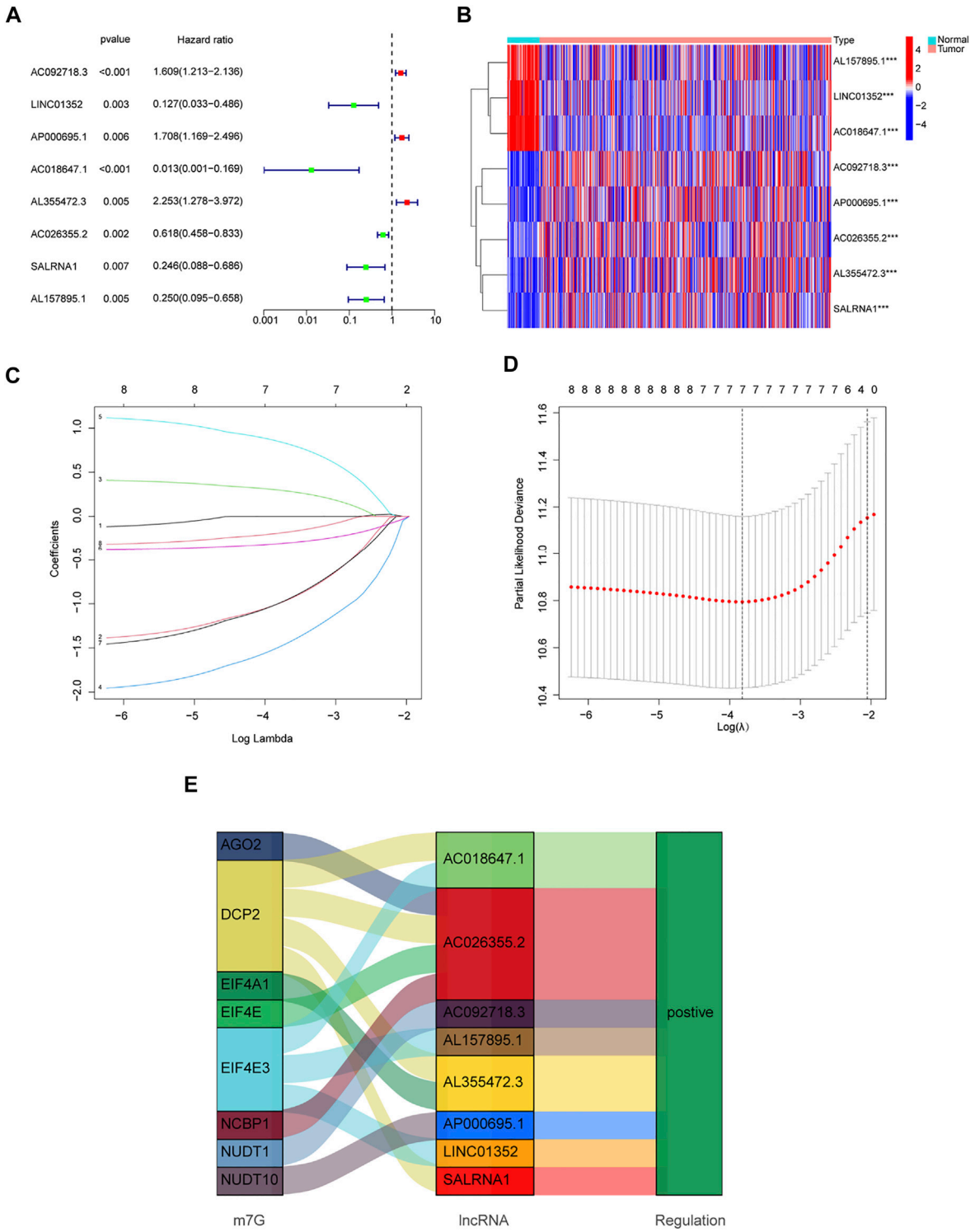


FIGURE 3 | Development of a new prognostic model for m7G-related lncRNAs in LUAD. **(A)** 8 prognosis-related lncRNAs screened by univariate Cox regression analysis ($p < 0.01$). **(B)** Differential expression of prognosis-related lncRNAs in lung normal tissues and adenocarcinoma tissues. **(C)** LASSO coefficient distribution of 8 m7G-related lncRNAs. **(D)** The tuning parameter (λ) in the LASSO model is chosen by the minimum criterion. **(E)** The Sankey diagram depicts the detailed connections between eight prognosis-related lncRNAs and m7G-related genes. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

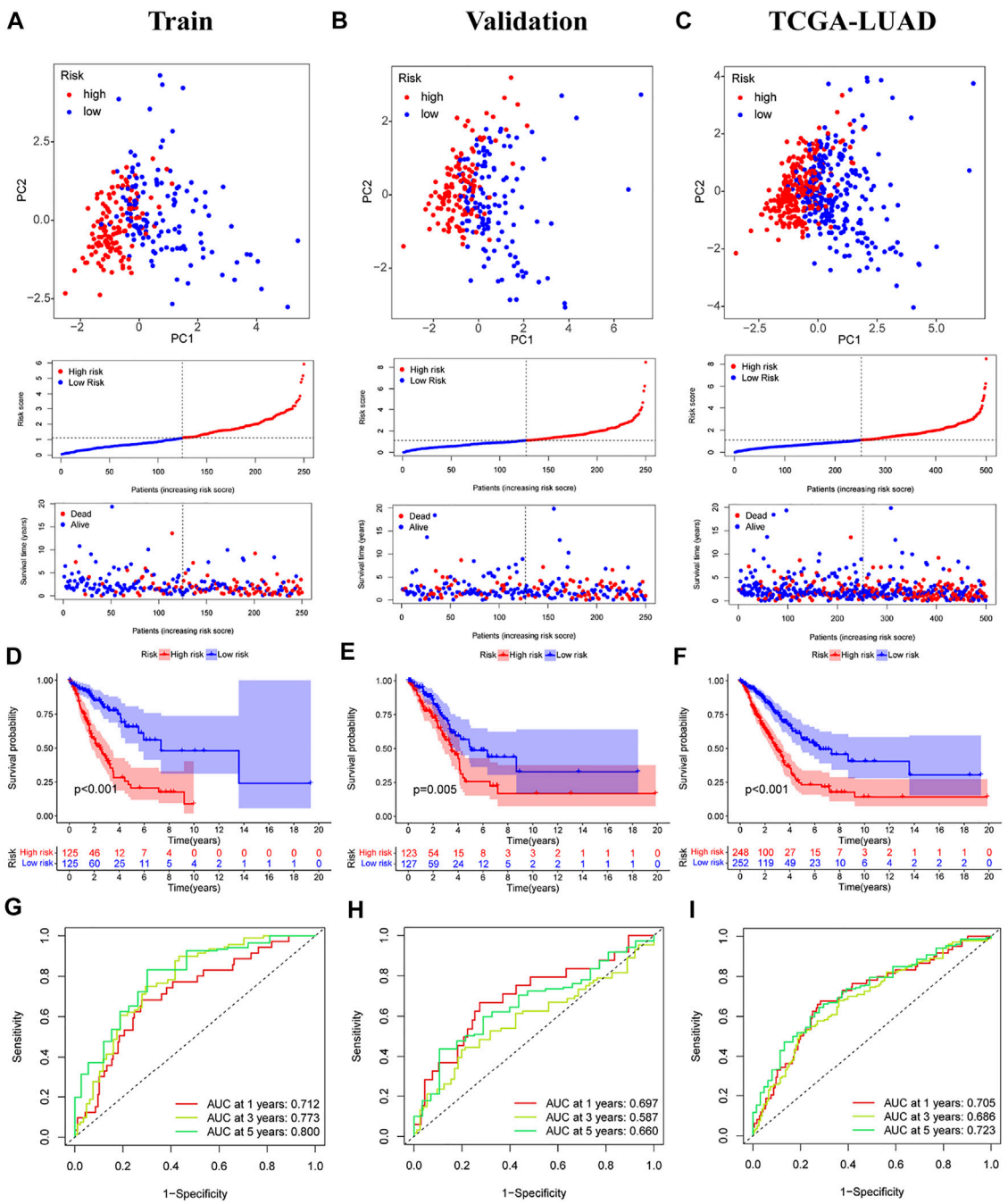


FIGURE 4 | Validation of prognostic models for six m7G-related lncRNAs. **(A–C)** Principal component analysis, risk score distribution, and survival status distribution for training, validation, and TCGA-LUAD. **(D–F)** Kaplan-Meier curves of training group, validation group and TCGA-LUAD at different risk groups. **(G–I)** ROC curves for 1 year, 3 years and 5 years.

patients in the high-risk group, providing foundations for future research directions. Immune-related pathways were enriched in the low-risk group, implying that the low-risk group may be closely related to immune characteristics. Disease differences of differentially expressed genes (DEGs) between the two risk subgroups were further investigated. First, DEGs ($|\text{Log}_2\ \text{FC}| > 1$ and $p < 0.05$) between the two risk subgroups were screened, followed by enrichment analysis of disease ontology (DO). DEGs were enriched in lung diseases, adenoma, coronary artery disease, and myocardial infarction (Yu et al., 2012; Yu et al., 2015). This confirms once again that our risk model is very superior. Surprisingly, DEGs are also involved in coronary artery disease and myocardial infarction (**Figures 6B,C**).

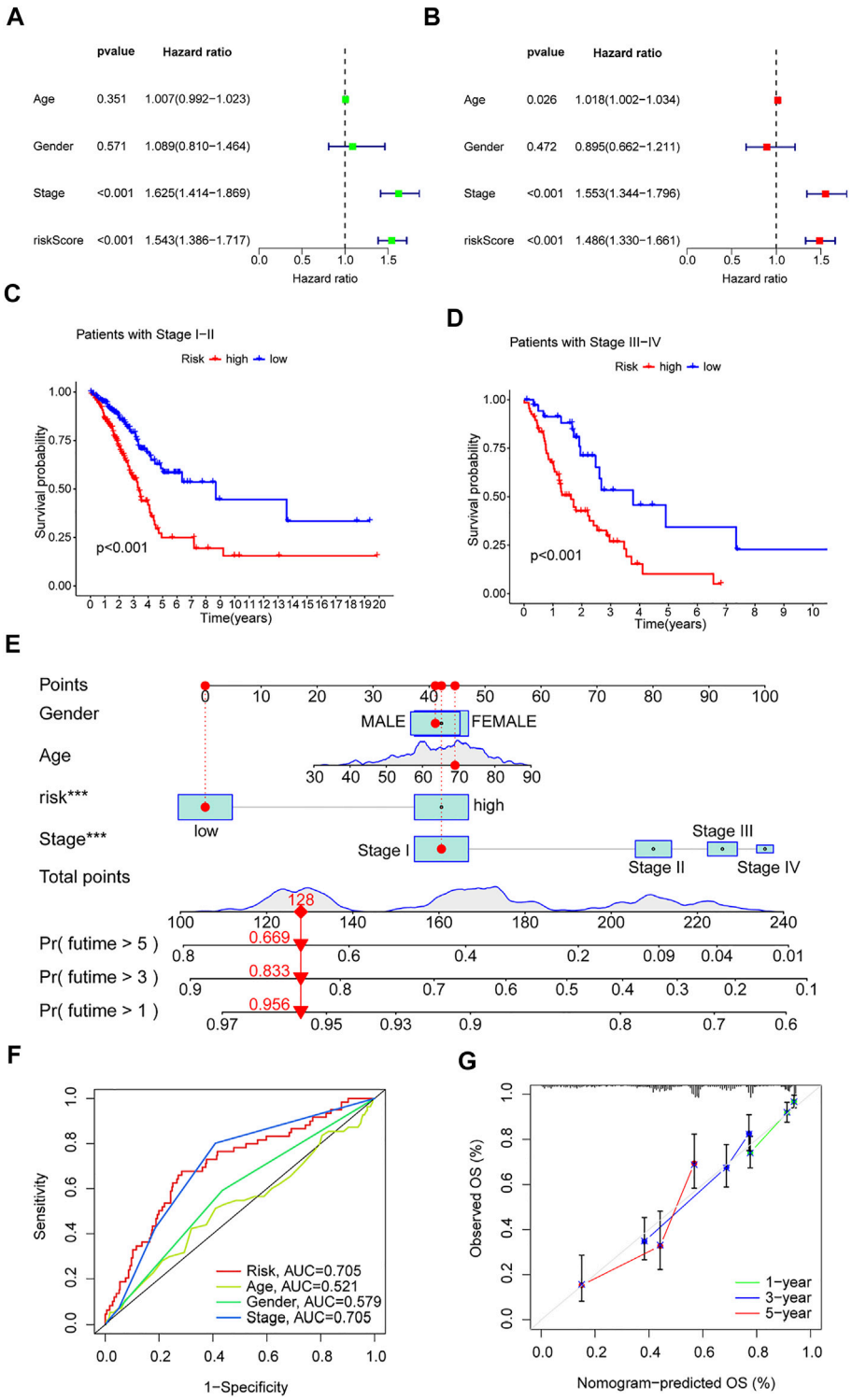


FIGURE 5 | Clinical value of risk characteristics in TCGA-LUAD. **(A)** Univariate Cox regression analysis of risk scores and clinical factors. **(B)** Multivariate Cox regression analysis of risk scores and clinical factors. **(C,D)** Pathological stage was stratified between low- and high-risk groups in the entire collection. **(E)** Nomogram combining gender, age, stage, and risk score predicts 1-, 3-, and 5-years overall survival. **(F)** Clinicopathological features and the predictive accuracy of risk models. **(G)** Calibration curves test the agreement between actual and predicted results at 1, 3, and 5 years. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Relationship Between Risk Model and Immune Characteristics

Because GSEA revealed that the low-risk group was enriched in immune-related pathways, we hypothesized that the m7G-related lncRNA-based risk model was strongly tied to immunity. Therefore, the relationship between the risk model and the immunological signature was investigated further. The relationship between immune cells and the risk score is shown in the bubble diagram. The majority of immune cells were negatively correlated with the risk score, especially hematopoietic stem cells of XCELL, tumor-related fibroblasts, stroma score, granulocyte-monocyte progenitor cells, and activated mast cells of CIBERSORT-ABS, resting memory CD4 + T cells, M2 macrophages, and Treg cells of QUANTISEQ and myeloid dendritic cells and endothelial cells of MCPCOUNTER (**Figure 7A**). Moreover, single sample gene set enrichment analysis (ssGSEA) was used to examine the enrichment fraction of 16 different types of immune cells as well as the activity of 13 different immune-related pathways. Interestingly, the low-risk group had more immune cell infiltration, particularly activated dendritic cells, B cells, immature dendritic cells, mast cells, neutrophils, T helper cells, and tumor-infiltrating lymphocytes (**Figure 7C**). The activity of type 2 interferon signaling pathway and human leukocyte antigen was higher in the low-risk group than in the high-risk group (**Figure 7D**). Differential analysis was used to detect differences in the tumor microenvironment between the high-risk and low-risk groups, and the results showed that the low-risk group had higher immune, stromal, and estimate scores (**Figure 7B**). It is possible to conclude that the low-risk group had greater immune cell infiltration and lower tumor purity. Furthermore, most immune checkpoints were highly expressed in low-risk groups (**Figure 7E**). Therefore, low-risk patients may benefit more from immune checkpoint inhibitor therapy in our risk model.

Clinical Application of Risk Model

Differences in drug sensitivity of different risk subgroups were analyzed to investigate the clinical application value of the risk model. Results showed that docetaxel, paclitaxel, etoposide, gemcitabine, erlotinib, and crizotinib had good effects on patients in high-risk groups (**Figures 8A–E**). Patients in low-risk groups were

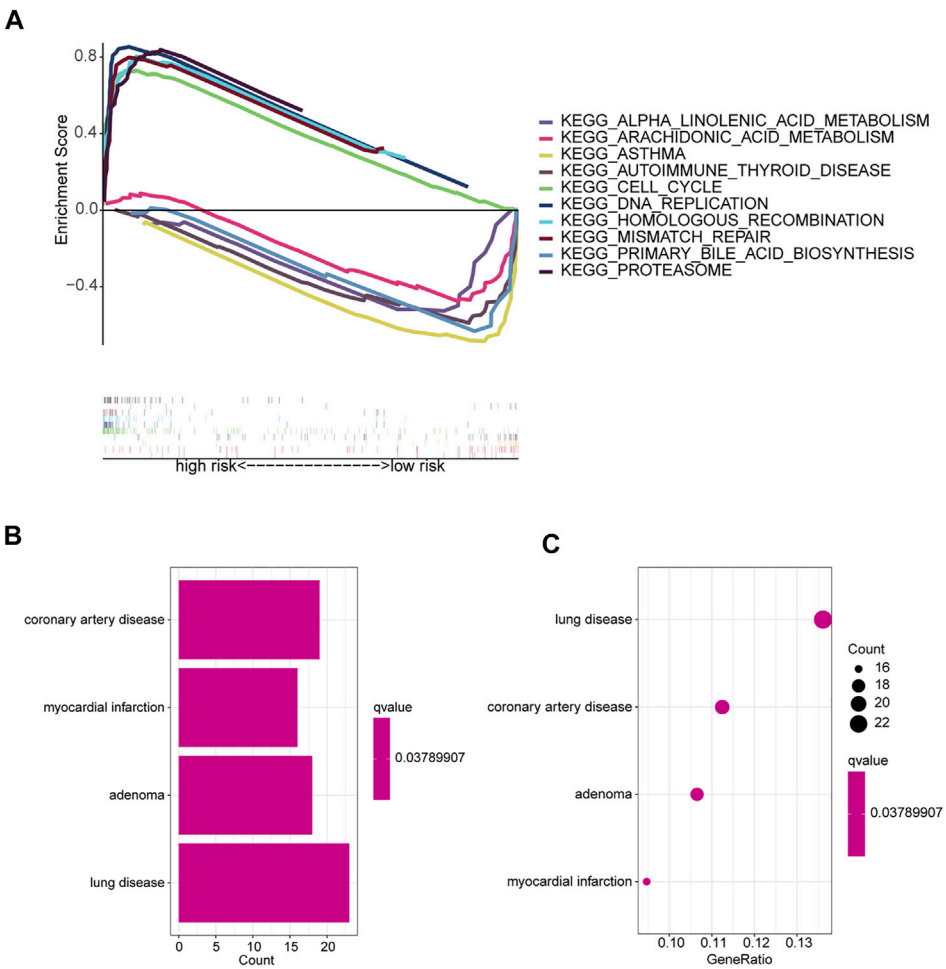


FIGURE 6 | Enrichment analysis of different risk subgroups. **(A)** Five pathways were significantly enriched in each of the high-risk and low-risk groups. **(B,C)** DO enrichment analysis of DEGs based on different risk subgroups.

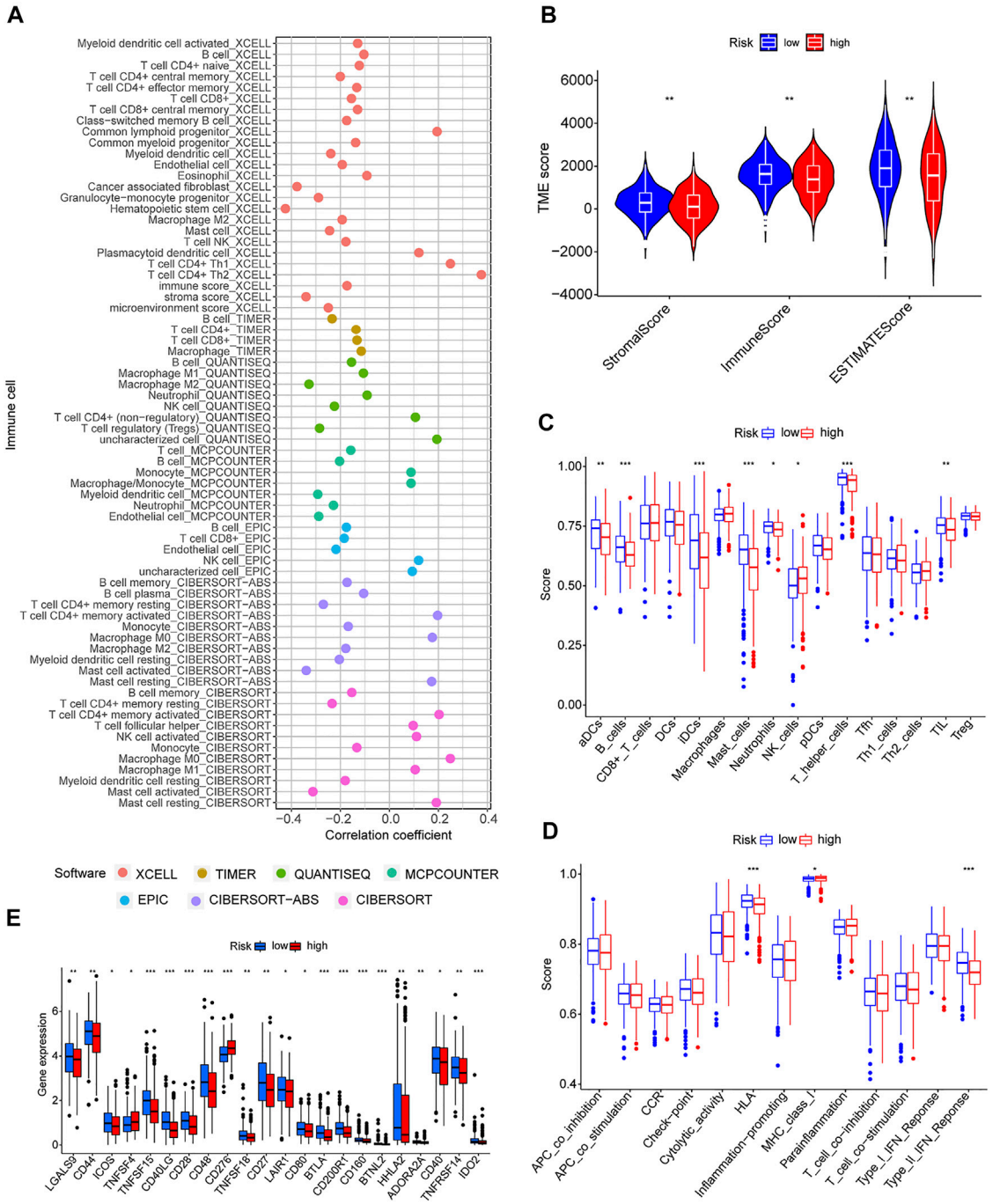


FIGURE 7 | Immune signatures of different risk groups. **(A)** Correlation between risk scores and immune cells. **(B)** Comparison of immune-related scores between low-risk and high-risk groups. **(C,D)** Enrichment scores for 16 immune cells and 13 immune-related pathways. **(E)** Differences in the expression of 22 checkpoints in different risk groups. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

more susceptible to drugs such as CDK4/6 inhibitors (PD.0332,991) and PI3K inhibitors (GDC0941); however, these drugs are currently used only for scientific research and may be promising in the future (**Figures 8F,G**). Reviewing the GSEA and DO enrichment analysis, it was found that the high-risk group in the TCGA-LUAD cohort had pathway enrichment such as cell cycle and DNA replication. The sensitive medications in the high-risk group are all first-line anti-tumor drugs for non-small cell lung cancer, among which chemotherapy drugs include docetaxel, paclitaxel, etoposide and gemcitabine, and their anti-tumor mechanisms are mainly directed against cell cycle and DNA replication. Erlotinib and crizotinib are two targeted medications, with erlotinib acting as an Epidermal

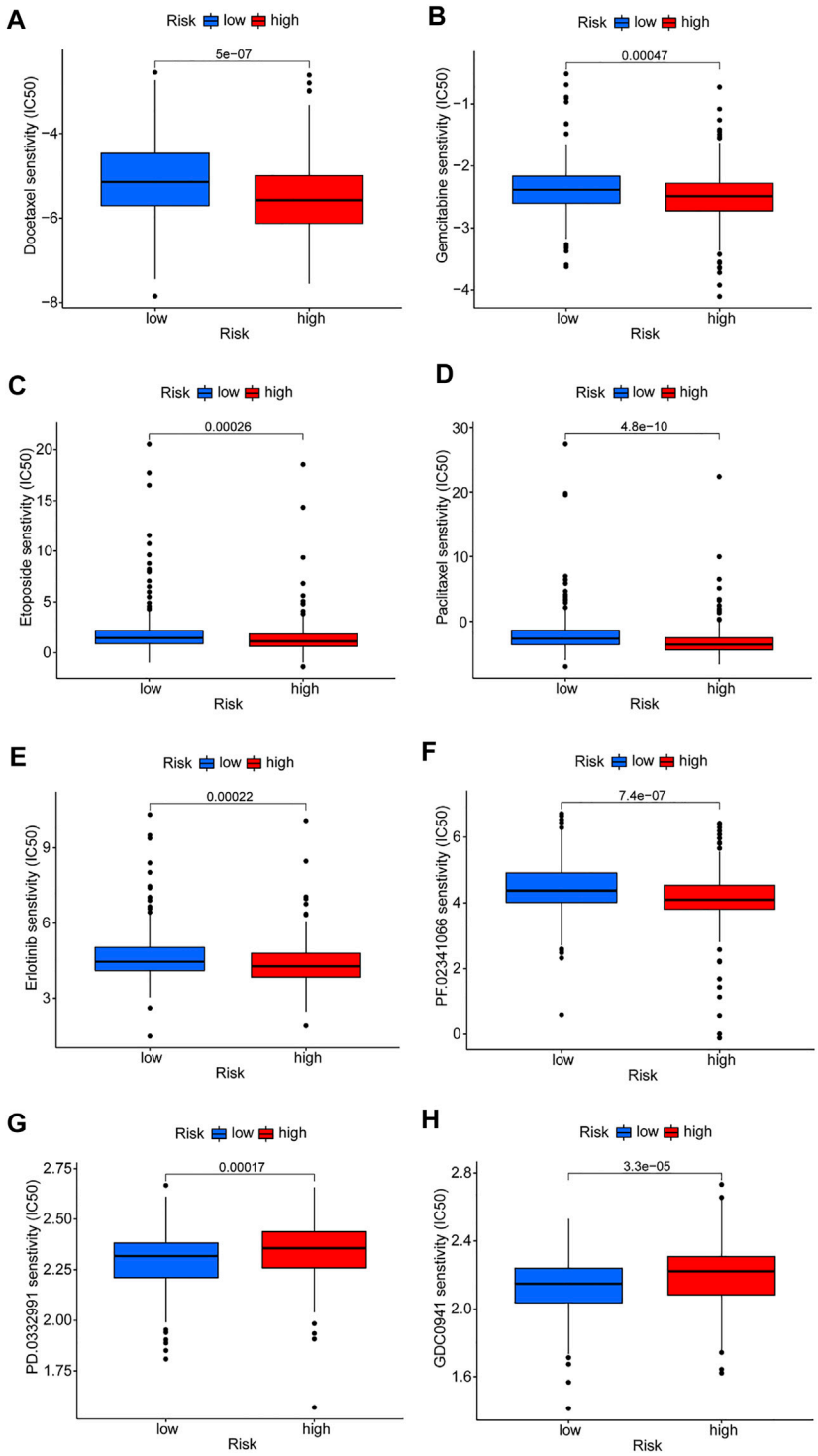


FIGURE 8 | Prediction of drug susceptibility in different risk groups. **(A–F)** Sensitive drugs in high-risk groups. **(G,H)** Sensitive drugs in low-risk groups.

Growth Factor Receptor (EGFR) inhibitor and crizotinib acting as an Anaplastic lymphoma kinase (ALK) inhibitor. Both EGFR and ALK targets are crucial for cancer proliferation. Presumably this is why the high-risk group is susceptible to the six antitumor drugs. The sensitive medications in the low-risk group have not been utilized in clinical practice. Fortunately, we discovered that the low-risk group had stronger immune infiltration and immune checkpoint expression, and it is expected that immunotherapy will be effective.

DISCUSSION

Numerous studies have recently revealed that m7G-related genes are closely linked to the development of cancer (Dai et al., 2021; Ma et al., 2021). A new class of lncRNAs has gradually become a research hotspot in various cancer fields in recent years. Some studies have found that abnormal expression of lncRNAs is associated with the occurrence and progression of LUAD and some lncRNAs may be highly correlated with prognosis (Spizzo et al., 2012; Cao et al., 2022; Xia et al., 2022). However, studies on m7G-related lncRNA predicting LUAD survival are scanty. The present study attempted to build a prognostic model of lncRNA in LUAD patients to test its clinical utility, and systematically explored the differences of risk models in immune cell infiltration, immune checkpoints, and drug sensitivity.

Forty publicly reported m7G-related genes were collected. First, lncRNAs that were differentially expressed in normal and LUAD tissues were explored. Univariate Cox regression was used to analyze the DELs, and 8 prognostic-related lncRNAs were screened out. Lasso-Cox regression analysis was then performed on these lncRNAs, and 6 prognosis-related lncRNAs (LINC01352, AP000695.1, AC018647.1, AL355472.3, AC026355.2, and SALRNA1) were finally identified and a risk prognosis model was constructed. The Sankey diagram showed that AC018647.1, AL355472.3, and SALRNA1 are related to DCP2. DCP2 is a decapping enzyme that plays a significant role in the regulation of the cell cycle and proliferation (Mugridge and Gross, 2018). DCP2 was found to promote lung cancer proliferation (Zhang et al., 2021). Our data also showed that AC026355.2 was highly correlated with four mRNAs (ACO2, DCP2, EIF4E, and NCBP1). Although AC026355.2 is rarely reported, we speculate that it plays a significant role in tumor development, but its precise role requires further investigation. EIF4E3 can promote translation, mRNA export, proliferation, and oncogenic transformation, and its related lncRNA LINC01352 was found to affect the growth and metastasis of hepatoma cells (Osborne et al., 2013). Bioinformatics analysis showed that AP000695.1 is closely related to immunity, and its related gene NUDT10, could be a potential immunotherapy target for LUAD in addition to promoting cell proliferation, inhibiting apoptosis, and causing tumor suppressor gene loss (Jin et al., 2020; Chen et al., 2021).

GSEA was performed on patients in both risk subgroups to reveal differences in biological function. Immune-related pathway enrichment was discovered in the low-risk group but not in the high-risk group. Immune cell bubble plots showed that low-risk groups had higher levels of immune infiltration. It has been reported that the high immune infiltration state tends to have a better immunotherapeutic effect (Guo et al., 2022; Luo et al., 2022). Interestingly, immune scores and immune checkpoint expression levels were also higher in the low-risk group, which is consistent with the results of Yu et al. (2021). Furthermore, susceptibility to antitumor drugs in different risk subgroups was assessed, and it found that the high-risk group showed high sensitivity to first-line treatment drugs for non-small cell lung cancer (including docetaxel, paclitaxel, etoposide, gemcitabine, erlotinib, and crizotinib) (Schiller et al., 2002; Zhou et al., 2011; Liang et al., 2017; Wu et al., 2018a; Wu et al., 2018b). Collectively, these results suggest that patients in the low-risk group will respond better to immunotherapy, while those in the high-risk group will respond better to chemotherapy and targeted drugs, which has important implications for individualized tumor therapy.

Although we verified the stability of the risk model from multiple aspects, there are still some limitations. First, the model was not externally validated because other databases lacked lncRNA information; thus, it was only be validated internally by TCGA. Further studies with a large sample size are required to draw definitive conclusions. Future studies will further explore the six lncRNAs.

CONCLUSION

In summary, this study conducted a comprehensive bioinformatics analysis and developed a risk model for six m7G-related lncRNAs, which not only accurately predicts patient survival but also reflects the immune characteristics of LUAD patients. This may provide important clues for the development of clinical individualized treatments and promote the progress of immunotherapy.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

AUTHOR CONTRIBUTIONS

CZ, ZW, and RW contributed to conception and design of the study. CZ and DZ organized the database. ZW performed the statistical analysis. CZ and DZ wrote the first draft of the manuscript. ZJ, GZ, and JH wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study was supported by Science and technology innovation project of Dalian City (2020JJ27SN080) and Key R&D Program of Liaoning province, China (2021JH1/10400051).

ACKNOWLEDGMENTS

We sincerely acknowledge The Cancer Genome Atlas (TCGA) for providing transcriptomic and clinicopathological data.

REFERENCES

Alexandrov, A., Martzen, M. R., and Phizicky, E. M. (2002). Two Proteins that Form a Complex Are Required for 7-Methylguanosine Modification of Yeast tRNA. *RNA* 8 (10), 1253–1266. doi:10.1017/s1355838202024019

Alexandrov, A., Grayhack, E. J., and Phizicky, E. M. (2005). tRNA m^7G Methyltransferase Trm8p/Trm82p: Evidence Linking Activity to a Growth Phenotype and Implicating Trm82p in Maintaining Levels of Active Trm8p. *RNA* 11 (5), 821–830. doi:10.1261/rna.2030705

Allemani, C., Matsuda, T., Di Carlo, V., Harewood, R., Matz, M., Nikšić, M., et al. (2018). Global Surveillance of Trends in Cancer Survival 2000-14 (CONCORD-3): Analysis of Individual Records for 37 513 025 Patients Diagnosed with One of 18 Cancers from 322 Population-Based Registries in 71 Countries. *Lancet* 391 (10125), 1023–1075. doi:10.1016/S0140-6736(17)33326-3

Azzam, E. I., Jay-Gerin, J.-P., and Pain, D. (2012). Ionizing Radiation-Induced Metabolic Oxidative Stress and Prolonged Cell Injury. *Cancer Lett.* 327 (1-2), 48–60. doi:10.1016/j.canlet.2011.12.012

Cao, Y., Zhang, H., Tang, J., and Wang, R. (2022). Long Non-coding RNA FAM230B Is a Novel Prognostic and Diagnostic Biomarker for Lung Adenocarcinoma. *Bioengineered* 13 (3), 7919–7925. doi:10.1080/21655979.2022.2034568

Chen, D., Zhang, R., Xie, A., Yuan, J., Zhang, J., Huang, Y., et al. (2021). Clinical Correlations and Prognostic Value of Nudix Hydroxylase 10 in Patients with Gastric Cancer. *Bioengineered* 12 (2), 9779–9789. doi:10.1080/21655979.2021.1995104

Chen, W., Chen, Y., Liu, L., Wu, Y., Fu, P., Cao, Y., et al. (2022). Comprehensive Analysis of Immune Infiltrates of Ferroptosis-Related Long Noncoding RNA and Prediction of Colon Cancer Patient Prognoses. *J. Immunol. Res.* 2022, 1–16. doi:10.1155/2022/9480628

Dai, Z., Liu, H., Liao, J., Huang, C., Ren, X., Zhu, W., et al. (2021). N(7)-Methylguanosine tRNA Modification Enhances Oncogenic mRNA Translation and Promotes Intrahepatic Cholangiocarcinoma Progression. *Mol. Cell* 81 (16), 3339–3355.e8. doi:10.1016/j.molcel.2021.07.003

Galloway, A., Kaskar, A., Ditsova, D., Atrih, A., Yoshikawa, H., Gomez-Moreira, C., et al. (2021). Upregulation of RNA Cap Methyltransferase RNMT Drives Ribosome Biogenesis during T Cell Activation. *Nucleic Acids Res.* 49 (12), 6722–6738. doi:10.1093/nar/gkab465

Guo, C., Liu, Z., Yu, Y., Liu, S., Ma, K., Ge, X., et al. (2022). Integrated Analysis of Multi-Omics Alteration, Immune Profile, and Pharmacological Landscape of Pyroptosis-Derived lncRNA Pairs in Gastric Cancer. *Front. Cell Dev. Biol.* 10, 816153. doi:10.3389/fcell.2022.816153

Haro, K. J., Scott, A. C., and Scheinberg, D. A. (2012). Mechanisms of Resistance to High and Low Linear Energy Transfer Radiation in Myeloid Leukemia Cells. *Blood* 120 (10), 2087–2097. doi:10.1182/blood-2012-01-404509

Jin, D., Song, Y., Chen, Y., and Zhang, P. (2020). Identification of a Seven-lncRNA Immune Risk Signature and Construction of a Predictive Nomogram for Lung Adenocarcinoma. *BioMed Res. Int.* 2020, 1–17. doi:10.1155/2020/7929132

Juhling, F., Morl, M., Hartmann, R. K., Sprinzl, M., Stadler, P. F., and Putz, J. (2009). tRNAdb 2009: Compilation of tRNA Sequences and tRNA Genes. *Nucleic Acids Res.* 37 (Database issue), D159–D162. doi:10.1093/nar/gkn772

Létoquart, J., Huvelle, E., Wacheul, L., Bourgeois, G., Zorbas, C., Graille, M., et al. (2014). Structural and Functional Studies of Bud23-Trm112 Reveal 18S rRNA N 7 -G1575 Methylation Occurs on Late 40S Precursor Ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* 111 (51), E5518–E5526. doi:10.1073/pnas.1413089111

Liang, J., Bi, N., Wu, S., Chen, M., Lv, C., Zhao, L., et al. (2017). Etoposide and Cisplatin versus Paclitaxel and Carboplatin with Concurrent Thoracic Radiotherapy in Unresectable Stage III Non-small Cell Lung Cancer: a Multicenter Randomized Phase III Trial. *Ann. Oncol.* 28 (4), 777–783. doi:10.1093/annonc/mdx009

Lin, S., Liu, Q., Lelyveld, V. S., Choe, J., Szostak, J. W., and Gregory, R. I. (2018). Mettl1/Wdr4-Mediated m(7)G tRNA Methylome Is Required for Normal mRNA Translation and Embryonic Stem Cell Self-Renewal and Differentiation. *Mol. Cell* 71 (2), 244–255. doi:10.1016/j.molcel.2018.06.001

Luo, Z., Nong, B., Ma, Y., and Fang, D. (2022). Autophagy Related Long Non-coding RNA and Breast Cancer Prognosis Analysis and Prognostic Risk Model Establishment. *Ann. Transl. Med.* 10 (2), 58. doi:10.21037/atm-21-6251

Ma, J., Han, H., Huang, Y., Yang, C., Zheng, S., Cai, T., et al. (2021). METTL1/WDR4-mediated m(7)G tRNA Modifications and m7G Codon Usage Promote mRNA Translation and Lung Cancer Progression. *Mol. Ther.* 29 (12), 3422–3435. doi:10.1016/j.ymthe.2021.08.005

Mugridge, J. S., and Gross, J. D. (2018). Decapping Enzymes STOP "cancer" Ribosomes in Their Tracks. *EMBO J.* 37 (23), e100801. doi:10.15252/embj.2018100801

Osborne, M. J., Volpon, L., Kornblatt, J. A., Culjkovic-Kraljacic, B., Baguet, A., and Borden, K. L. B. (2013). eIF4E3 Acts as a Tumor Suppressor by Utilizing an Atypical Mode of Methyl-7-Guanosine Cap Recognition. *Proc. Natl. Acad. Sci. U.S.A.* 110 (10), 3877–3882. doi:10.1073/pnas.1216862110

Schiller, J. H., Harrington, D., Belani, C. P., Langer, C., Sandler, A., Krook, J., et al. (2002). Comparison of Four Chemotherapy Regimens for Advanced Non-small-cell Lung Cancer. *N. Engl. J. Med.* 346 (2), 92–98. doi:10.1056/NEJMoa011954

Spizzo, R., Almeida, M. I., Colombatti, A., and Calin, G. A. (2012). Long Non-Coding RNAs and Cancer: a New Frontier of Translational Research? *Oncogene* 31 (43), 4577–4587. doi:10.1038/onc.2011.621

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

Travis, W. D., Brambilla, E., Nicholson, A. G., Yatabe, Y., Austin, J. H. M., Beasley, M. B., et al. (2015). The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *J. Thorac. Oncol.* 10 (9), 1243–1260. doi:10.1097/JTO.0000000000000630

Trotman, J. B., and Schoenberg, D. R. (2019). A Recap of RNA Recapping. *WIREs RNA* 10 (1), e1504. doi:10.1002/wrna.1504

Wu, Y.-L., Lu, S., Lu, Y., Zhou, J., Shi, Y.-k., Sriuranpong, V., et al. (2018a). Results of PROFILE 1029, a Phase III Comparison of First-Line Crizotinib versus Chemotherapy in East Asian Patients with ALK-Positive Advanced Non-small Cell Lung Cancer. *J. Thorac. Oncol.* 13 (10), 1539–1548. doi:10.1016/j.jtho.2018.06.012

Wu, Y.-L., Yang, J. C.-H., Kim, D.-W., Lu, S., Zhou, J., Seto, T., et al. (2018b). Phase II Study of Crizotinib in East Asian Patients with ROS1-Positive Advanced Non-small-cell Lung Cancer. *Jco* 36 (14), 1405–1411. doi:10.1200/JCO.2017.75.5587

Xia, M., Zhu, W., Tao, C., Lu, Y., and Gao, F. (2022). LncRNA LASTR Promote Lung Cancer Progression through the miR-137/TGFA/PI3K/AKT axis through Integration Analysis. *J. Cancer* 13 (4), 1086–1096. doi:10.7150/jca.66067

Xu, T., Gu, H., Zhang, C., Zhang, W., Liang, X., and Cheng, X. (2022). A Novel Risk Model Identified Based on Pyroptosis-Related lncRNA Predicts Overall Survival and Associates with the Immune Landscape of GC Patients. *Front. Genet.* 13, 843538. doi:10.3389/fgene.2022.843538

Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS A J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118

Yu, G., Wang, L.-G., Yan, G.-R., and He, Q.-Y. (2015). DOSE: an R/Bioconductor Package for Disease Ontology Semantic and Enrichment Analysis. *Bioinformatics* 31 (4), 608–609. doi:10.1093/bioinformatics/btu684

Yu, X., Dong, P., Yan, Y., Liu, F., Wang, H., Lv, Y., et al. (2021). Identification of N6-Methyladenosine-Associated Long Non-coding RNAs for Immunotherapeutic Response and Prognosis in Patients with Pancreatic Cancer. *Front. Cell Dev. Biol.* 9, 748442. doi:10.3389/fcell.2021.748442

Zhang, Q., Yan, Y.-F., Lv, Q., Li, Y.-J., Wang, R.-R., Sun, G.-B., et al. (2021). miR-4293 Upregulates lncRNA WFDC21P by Suppressing mRNA-Decapping Enzyme 2 to Promote Lung Carcinoma Proliferation. *Cell Death Dis.* 12 (8), 735. doi:10.1038/s41419-021-04021-y

Zhou, C., Wu, Y.-L., Chen, G., Feng, J., Liu, X.-Q., Wang, C., et al. (2011). Erlotinib versus Chemotherapy as First-Line Treatment for Patients with Advanced EGFR Mutation-Positive Non-small-cell Lung Cancer (OPTIMAL, CTONG-

0802): a Multicentre, Open-Label, Randomised, Phase 3 Study. *Lancet Oncol.* 12 (8), 735–742. doi:10.1016/s1470-2045(11)70184-x

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

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.

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 08 June 2022
doi: 10.3389/fgene.2022.907392

A Novel Necroptosis-Associated lncRNAs Signature for Prognosis of Head and Neck Squamous Cell Carcinoma

Jing Huang[1†], Rong Lu[2†], Dongta Zhong[3], Youliang Weng[4] and Lianming Liao[5*]*

[1]Department of Pharmacy, Fujian Medical University Cancer Hospital and Fujian Cancer Hospital, Fuzhou, China, [2]Department of Laboratory Medicine, The First Affiliated Hospital of Xiamen University, Xiamen Key Laboratory of Genetic Testing, School of Medicine, Xiamen University, Xiamen, China, [3]Department of Medical Oncology, Union Hospital of Fujian Medical University, Fuzhou, China, [4]Department of Radiation Oncology, Fujian Medical University Cancer Hospital and Fujian Cancer Hospital, Fuzhou 350014, China, [5]Center of Laboratory Medicine, Union Hospital of Fujian Medical University, Fuzhou, China

OPEN ACCESS

Edited by:
Yan Gong,
Wuhan University, China

Reviewed by:
Zhe Wang,
Affiliated Zhongshan Hospital of Dalian University, China
Jing Ke,
Wuhan University, China
Qiuji Wu,
Wuhan University, China

****Correspondence:***
Lianming Liao
liao772002@aliyun.com
Youliang Weng
wyl7788@sina.com

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: *29 March 2022*
Accepted: *04 May 2022*
Published: *08 June 2022*

Citation:
Huang J, Lu R, Zhong D, Weng Y and Liao L (2022) A Novel Necroptosis-Associated lncRNAs Signature for Prognosis of Head and Neck Squamous Cell Carcinoma. Front. Genet. 13:907392. doi: 10.3389/fgene.2022.907392

Purpose: The prognosis of head and neck squamous cell carcinoma (HNSCC) is poor. Necroptosis is a novel programmed form of necrotic cell death. The prognostic value of necroptosis-associated lncRNAs expression in HNSCC has not been explored.

Methods: We downloaded mRNA expression data of HNSCC patients from TCGA databases. Prognostic lncRNAs were identified by univariate Cox regression. LASSO was used to establish a model with necroptosis-related lncRNAs. Kaplan-Meier analysis and ROC were applied to verify the model. Finally, functional studies including gene set enrichment analyses, immune microenvironment analysis, and anti-tumor compound IC50 prediction were performed.

Results: We identified 1,117 necroptosis-related lncRNAs. The Cox regression showed 55 lncRNAs were associated with patient survival ($p < 0.05$). The risk model of 24- lncRNAs signature categorized patients into high and low risk groups. The patients in the low-risk group survived longer than the high-risk group ($p < 0.001$). Validation assays including ROC curve, nomogram and correction curves confirmed the prediction capability of the 24-lncRNA risk mode. Functional studies showed the two patient groups had distinct immunity conditions and IC50.

Conclusion: The 24-lncRNA model has potential to guide treatment of HNSCC. Future clinical studies are needed to verify the model.

Keywords: risk score, prognosis, squamous cell carcinoma, necroptosis, immune, tumor

INTRODUCTION

Head and neck squamous cell carcinomas (HNSCCs) arise from squamous cells in the oral cavity, pharynx and larynx. The most common risk factors for HNSCC include alcohol drinking, smoking and HPV infection (Fakhry et al., 2008). Athough HNSCC can be treated with surgery, radiotherapy and chemotherapy, patients with HNSCCs still suffer from poor survival. To improve patient survival, novel therapeutic targets and effective prognostic tools are needed.

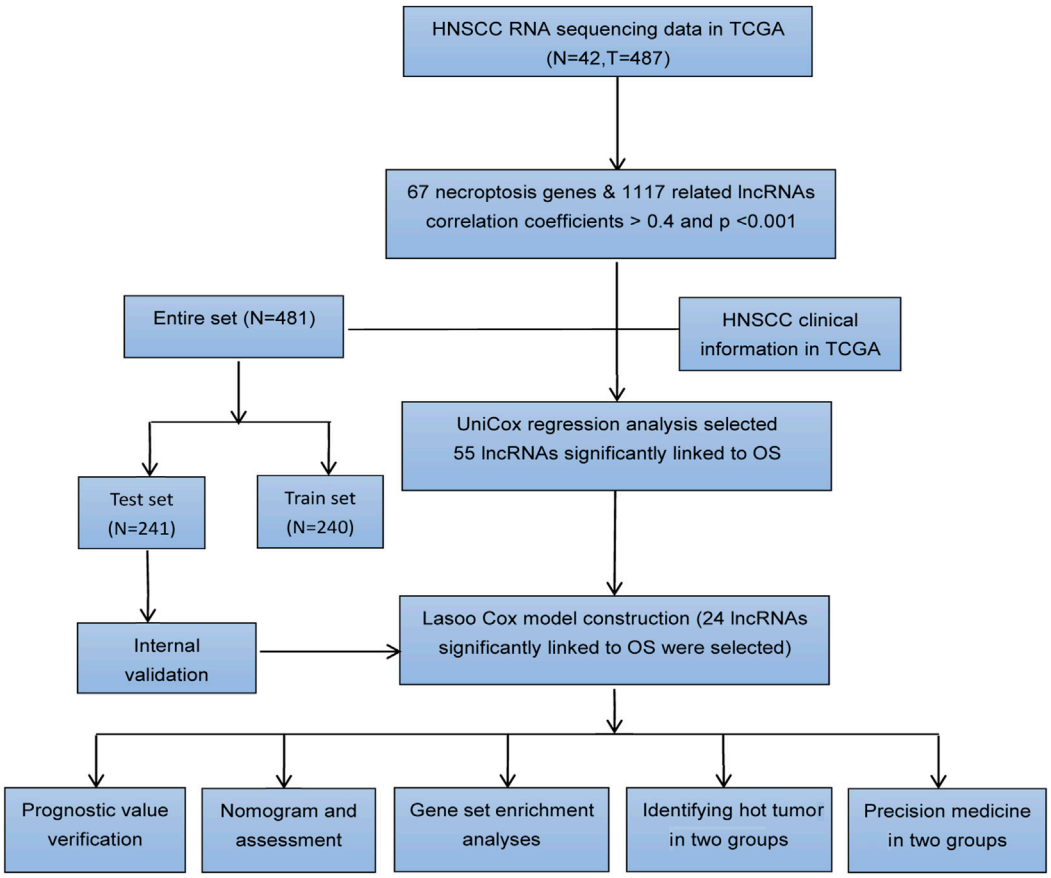


FIGURE 1 | Workflow diagram of data analysis.

Necroptosis is another mode of regulated cell death mimicking apoptosis and necrosis. Necroptosis is associated with a range of pathological conditions and diseases, including cancer. It is mediated by Fas, TNF, LPS, and death receptors (Vanden Berghe et al., 2014). Binding of ligands and receptors activates RIP3, which phosphorylates MLKL (Sun et al., 1999). Phosphorylated MLKL then translocates to and ruptures cellular membranes, leading to cell swelling and release of intracellular components (Dondelinger et al., 2014; Hildebrand et al., 2014; Wang et al., 2014).

A plethora of evidence shows necroptosis of tumor cells is often associated with tumor aggressiveness and metastasis. RIP3, a molecular marker of necroptosis, is an independent factor associated with survival in breast cancer (Koo et al., 2015). RIP3 expression was also decreased in colorectal cancer and was an independent prognostic factor of survival (Feng et al., 2015). In acute myeloid leukemia, RIP3 expression was reduced in most samples and overexpression of RIP3 in DA1-3b leukemia cells induced necroptosis (Nugues et al., 2014). Li et al. reported that necroptosis was associated with survival of HNSCC patients (Li et al., 2020).

Long non-coding RNAs (lncRNAs) regulate gene expression and are involved in tumorogenesis (Kumar and Goyal, 2017; Peng et al., 2017). Specially, Jiang et al. reported dysregulation of lncRNAs was involved in HNSCC(Jiang et al., 2019). Although necroptosis plays an important role in patient survival of a variety of tumors, the role of necroptosis-related lncRNAs in HNSCC has not been reported.

We thus explored the potential roles of different necroptosis-related lncRNAs on the survival of HNSCC patients. We developed a novel risk-score model with necroptosis-related lncRNAs according to their expression levels. The results might further our understanding of necroptosis in HNSCC.

MATERIALS AND METHODS

TCGA Data Acquisition

The Cancer Genome Atlas (TCGA) (https://portal.gdc.cancer.gov/repository) has transcriptomic data of more than 20,000 cancer and normal samples. In the present study RNA sequencing (RNA-seq) data of tumor tissues of 487 HNSCC patients and 42 matched normal tissues was downloaded from TCGA database. Our study was conducted by reviewing public database and ethical approval was not required.

Identification of Necroptosis-Related lncRNAs

The expression data of 67 necroptosis-associated genes was used for analysis (**Supplementary Table S1**). Correlation

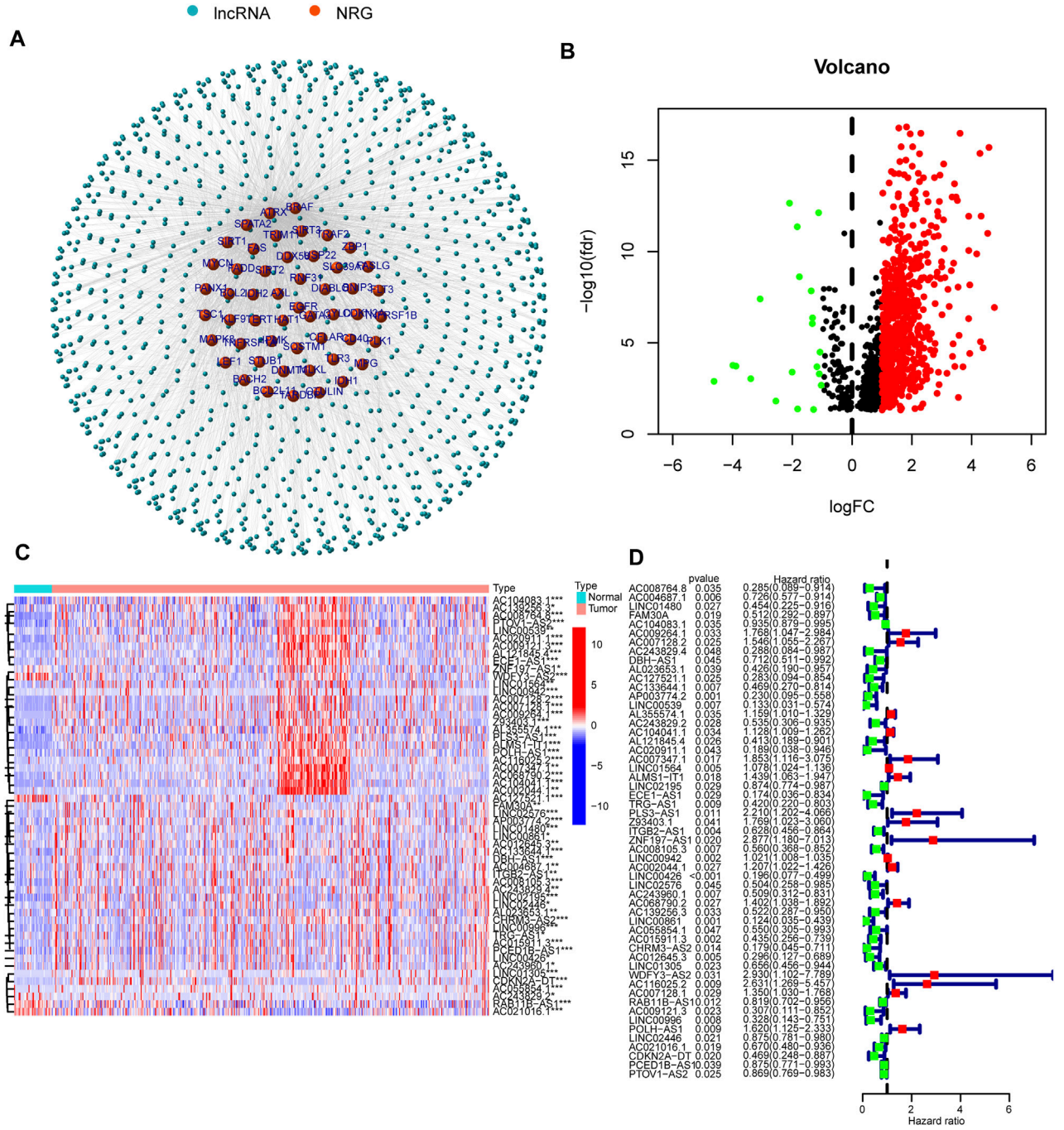


FIGURE 2 | Necroptosis-related lncRNA prognostic signature identified in HNSCC. **(A)** Correlation network of necroptosis-related genes and necroptosis-related lncRNAs. **(B)** The volcano plot of 717 lncRNAs. Red dots represented upregulated lncRNAs and black dots represented down-regulated lncRNAs. **(C,D)** The expression and univariate Cox regression of 55 prognostic lncRNAs. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

analysis was performed among 67 necroptosis-related genes and differentially expressed lncRNAs in the combined matrices. 1,117 lncRNAs with Pearson correlation coefficients >0.4 and $p < 0.001$ were identified to be necroptosis-related lncRNAs.

Establishment and Validation of the Risk Signature

The clinical data of HNSCC patients was downloaded from TCGA data portal. The univariate Cox proportional hazard regression analysis was used to screen prognostic genes. Least absolute

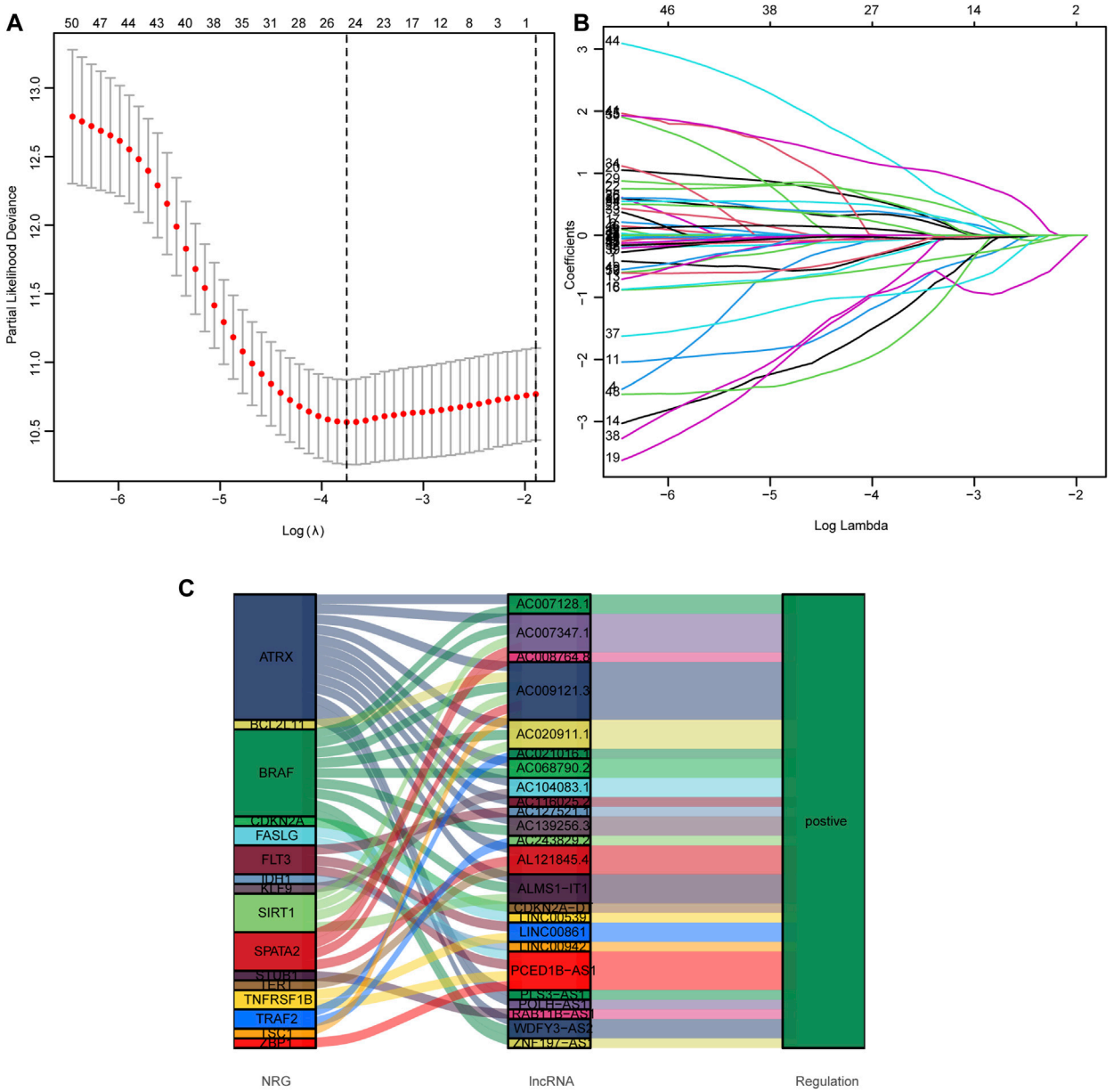


FIGURE 3 | Construction of prognostic signature in HNSCC. **(A)** LASSO with 10-fold cross-validation. **(B)** Coefficient profile plots. **(C)** The Sankey diagram shows the connection degree between the 24 prognostic lncRNAs and the necroptosis-related genes.

shrinkage and selection operator (Lasso) regression was conducted with 10-fold cross-validation and a *p* value of 0.05. After identification of the prognostic lncRNAs, the risk scores were determined as follow (X: coefficients, Y: expression level of lncRNAs):

$$\text{risk score} = \sum_{i}^{n} Xi * Yi$$

HNSCC patients were allocated into either low- or high-risk groups according to the median risk score (Meng et al., 2019; Hong et al., 2020). The Chi-square test was used to determine the prognostic significance value of the risk model, and overall survival (OS) time was compared between the two groups *via* Kaplan-Meier analysis. The "survival", "survminer" and "timeROC" R packages were used to plot the 1-, 3-, and 5-years receiver operating characteristics (ROC) curves. The risk scores were also evaluated as an independent risk factor with other clinical parameters by Cox regression with rms R package. Then a nomogram for prediction of the 1-, 3-, and 5-years OS was set up using risk score and clinical parameters. The ROC, calibration curves and Hosmer-Lemeshow test of the nomogram were assessed in the validation set.

GSEA

To explore the biological pathways that might be responsible for poor patient survival, we employed R (Bioconductor package

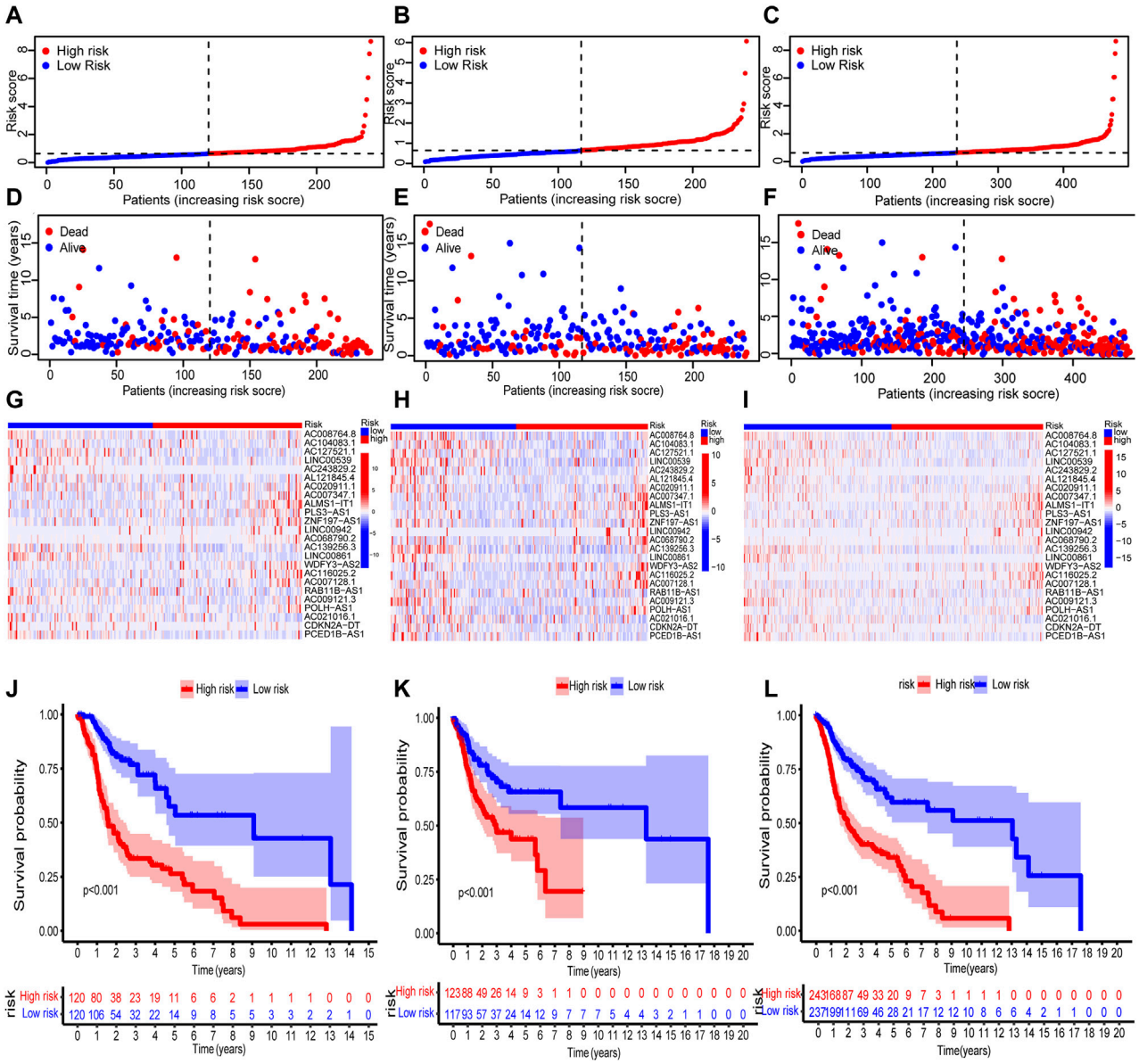


FIGURE 4 | The prognosis analysis of the three sets (training set: A, D, G, J; test set: B, E, H, K; entire set: C, F, I, L). **(A–C)** Risk scores of the high-risk and low-risk groups in the respective three sets. **(D–F)** Comparison of survival between the two groups in the respective three sets. **(G–I)** The heat maps of 24 lncRNAs. **(J–L)** Kaplan–Meier survival curves of patients in the two groups in the respective three sets. **(M)** Kaplan–Meier survival curves of OS stratified by clinicopathologic parameters between the two groups.

gsea) to perform gene set enrichment analyses (GSEA). Potential biological mechanisms of the prognostic model were also explored. KEGG gene sets in the GSEA database were downloaded. We chose gene sets with a FDR value < 0.05 and a FDR <0.25.

The Investigation of the TME and Immune Checkpoints

CIBERSORT, EPIC, MCPcounter, QUANTISEQ, TIMER, and XCELL were used to evaluate cells in the tumor microenvironment (TME) (http://timer.cistrome.org/). ggplot2, ggtext, limma, and scales R packages and Wilcoxon signed-rank test were performed to analyze cell types in TME (Hong et al., 2020). TME scores and immune checkpoint were compared between the two groups with ggpubr R package.

Prediction of Clinical Treatment Response

To predict therapeutic response, the R package pRRophetic was utilized to measure the half-maximal inhibitory concentrations (IC50) of each HNSCC sample on Genomics of Drug Sensitivity in Cancer (GDSC) (https://www.cancerrxgene.org/) (Geeleher et al., 2014).

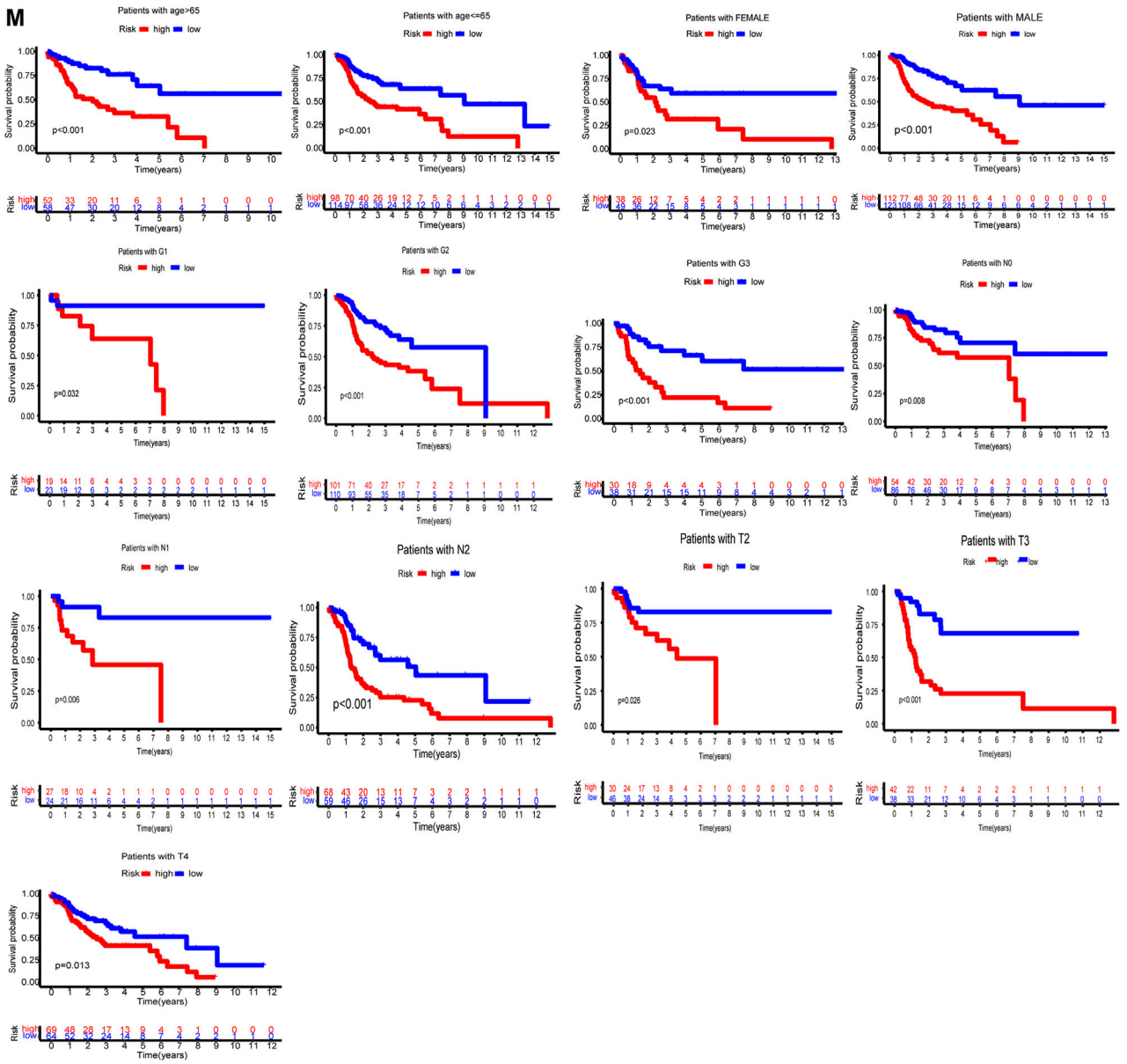


FIGURE 4 | Continued.

RESULTS

Extraction of lncRNAs

Figure 1 shows the flow diagram depicting the present study. We compared the expression levels of 67 necroptosis-associated genes (**Supplementary Table S1**) between 42 healthy samples and 487 HNSCC cancer samples from the TCGA data and identified 14,086 lncRNAs. Among these lncRNAs, 1,117 lncRNAs met the criteria (Pearson correlation coefficients >0.4 and p values <0.001) (**Figure 2A**). We identified 717 differentially expressed necroptosis-related lncRNAs (|Log2FC| > 1 and p < 0.05) (**Figure 2B**); 697 were upregulated and 20 were downregulated. Univariate Cox regression showed 55 lncRNAs were significantly correlated with OS (p < 0.05 for all) (**Figures 2C,D**).

Risk Model Construction and Verification

To avoid overfitting and to quantify the impact of necroptosis-related lncRNAs on the prognosis of each HNSCC patient, we constructed a 24 lncRNAs prognostic signature by LASSO regression analysis (**Figures 3A,B**). All 24 lncRNAs positively regulated necroptosis genes in the Sankey diagram (**Figure 3C**).

We established the following formula to calculate the risk score of every HNSCC patient.

Risk score = AC008764.8×(-0.2393)+AC104083.1×(-0.0743)+AC127521.1×(-0.9739)+LINC00539×(-

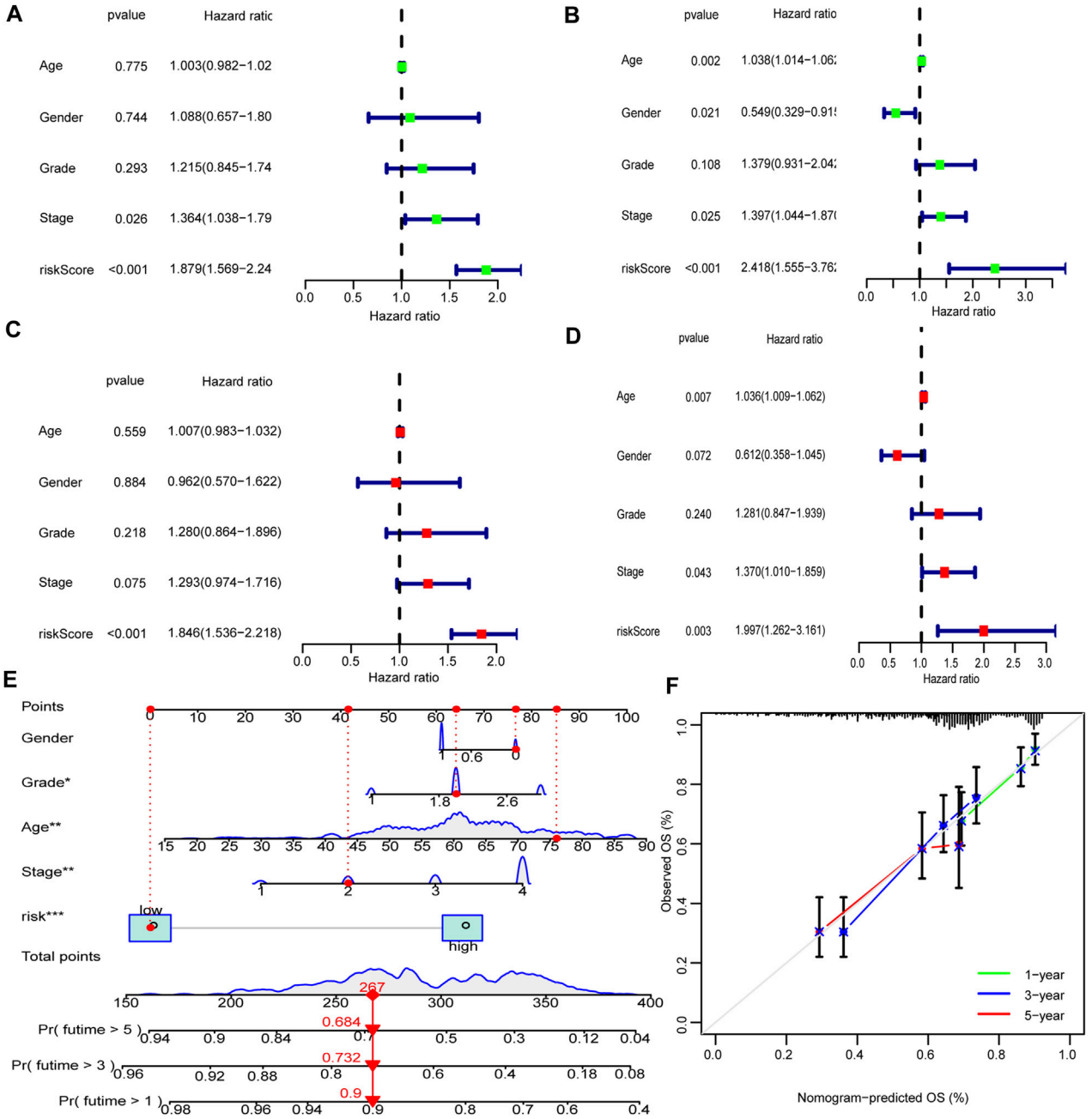


FIGURE 5 | Nomogram of the model. **(A,B)** Tumor stage and risk score were risk factors in the training set. **(C,D)** Only risk score was risk factor in the test sets. **(E)** The nomogram that integrated the risk score, age, and tumor stage to predict OS. **(F)** The calibration curves for OS. **(G–I)** ROC curves of the three sets (training set: A, C, G; test set: B, D, H; entire set: I).

1.2625)+AC243829.2×(-0.4868)+AL121845.4×(-0.2809)+AC020911.1×(-0.6943)+AC007347.1×(0.4674)+ALMS1-IT1×(0.6387)+PLS3-AS1×(0.3155)+ZNF197-AS1×(0.4949)+LINC00942×(0.0017)+AC068790.2×(0.3482)+AC139256.3×(-0.9441)+LINC00861×(-0.8428)+WDFY3-AS2×(1.1448)+AC116025.2×(1.0996)+AC007128.1×(0.0713)+RAB11B-AS1×(-0.0588)+AC009121.3× (-1.6935)+POLH-AS1×(0.4292)+AC021016.1×(-0.0412)+CDKN2A-DT×(-0.2129)+PCED1B-AS1×(0.1493) (Meng et al., 2019).

In the training set, test set and entire set, the distribution of risk scores and survival times were compared between the high-risk group and the low-risk group (**Figures 4A–C**). More patients died in the high-risk group (**Figures 4D–F**). The heat maps of 24 lncRNAs are shown in **Figures 4G–I**. Survival curves show the high-risk groups of the three sets had poor prognoses (**Figures**

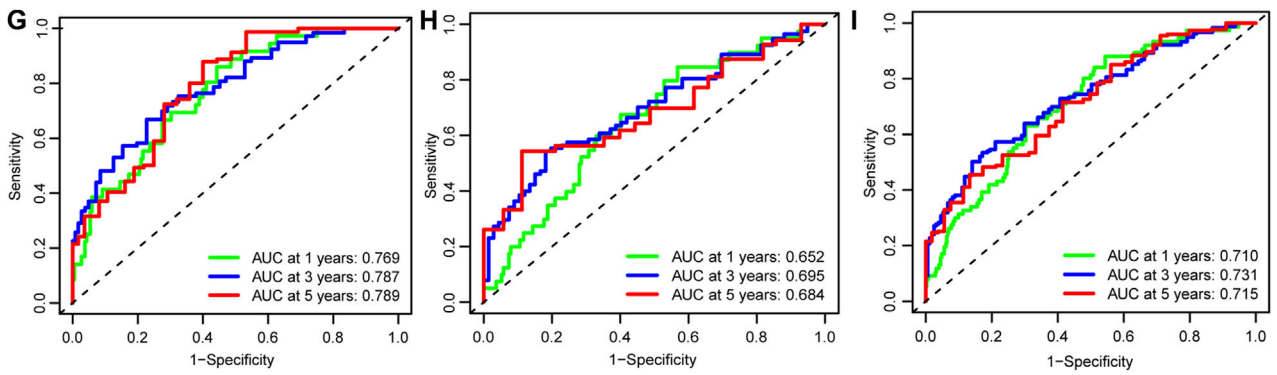


FIGURE 5 | Continued.

4J–L). Besides, some typical clinicopathologic parameters were identified to be prognostic factors (**Figure 4M**).

Nomogram

In both the training set and the test set, risk score, age and tumor stage were identified to be independent prognostic factors. The hazard ratios (HR) of these factors are shown in **Figures 5A–D**). A nomogram was established to predict the 1-, 3-, and 5-years OS (**Figure 5E**). The predicted survival showed close agreement with observed actual survival (**Figure 5F**).

Assessment of the Risk Model

The ROC curves showed the sensitivity and specificity of the model were high (**Figures 5G–I**).

GSEA

GSEA results showed that nine of top ten pathways were involved in carcinogenesis. For example, pentose and glucuronate interconversions, aldarate metabolism, and starch and sucrose metabolism were significantly enriched in the high-risk group. On the other hand, eight pathways enriched in the low-risk group were related to immunity ($p < 0.05$; FDR <0.25; |NES| > 1.9), such as T cell receptor signaling pathway and natural killer (NK) cell-mediated cytotoxicity (**Figure 6A**; **Supplementary Table S2**). Therefore, the low-risk group had a favorable TME. On the contrary, the high-risk group had an unfavorable TME.

Cold and Hot Tumors

Single sample GSEA (ssGSEA) was performed to calculate numbers for different types of immune cells. Tumors of the low-risk group were infiltrated by more immune cells as exhibited in the heatmap ($p < 0.05$ for all) (**Figure 6B**; **Supplementary Table S3**). Correlations between risk scores and activities of immune cell types are shown in **Figure 6C**. All of the 13 immune-related pathways had higher activity in the low-risk group (**Figure 6D**).

ESTIMATE was used to generate immune scores and stromal scores. **Figure 6E** shows both immune scores and stromal scores (microenvironment) were higher in the low-risk group. Besides, the immune checkpoint expression was lower in the low-risk group (**Figure 6F**).

Finally, we found IC50 of the anti-tumor compounds, such as AKT inhibitors, JNK inhibitor and sunitinib, was usually lower in the low-risk group (**Figure 6G**).

DISCUSSION

The human genome produces a large amount of RNA transcripts that do not encode for proteins (Djebali et al., 2012). lncRNAs are among those transcripts. They are usually longer than 200 nucleotides and have many functions, including regulating cancer development (Huarte, 2015; Marchese et al., 2017; Mattick, 2018).

Necroptosis is mediated by RIP1 and RIP3 (Chan and Baehrecke, 2012; Pasparakis and Vandenabeele, 2015). RIP1 phosphorylates RIP3, which phosphorylates MLKL. In necrosomes phosphorylation of MLKL leads to MLKL oligomerization. Oligomerized MLKL causes cell death by breaking down cell membranes (Sun et al., 2012; Guicciardi et al., 2013). Although necroptosis may cause cancer cell death, cell death may inhibit immune response (Pasparakis and Vandenabeele, 2015; Wang et al., 2017). Necroptosis may elicit necrosis-associated inflammation. Inflammation could contribute to progression of cancer and promote resistance to anticancer treatments. In addition necroptosis may also fail to elicit strong immunogenic reactions.

Massively parallel RNA sequencing has identified large amounts of novel lncRNAs. However, functional annotation of lncRNAs is lagging behind. In the present study, we explored the prognostic values of necroptosis-related lncRNAs in HNSCC. We found that several necroptosis-related lncRNAs were closely related to HNSCC prognosis. More specifically, AC007347.1, ALMS1-IT1, PLS3-AS1, ZNF197-AS1, AC068790.2, WDFY3-AS2, AC116025.2, POLH-AS1, and PCED1B-AS1 were risk factors. On the other hand, AC008764.8, AC127521.1, LINC00539, AC243829.2, AL121845.4, AC020911.1, AC139256.3, LINC00861, AC009121.3 and CDKN2A-DT were protective factors for HNSCC patients.

Further analysis showed that AC007347.1, ALMS1-IT1, PLS3-AS1, ZNF197-AS1, AC068790.2, WDFY3-AS2,

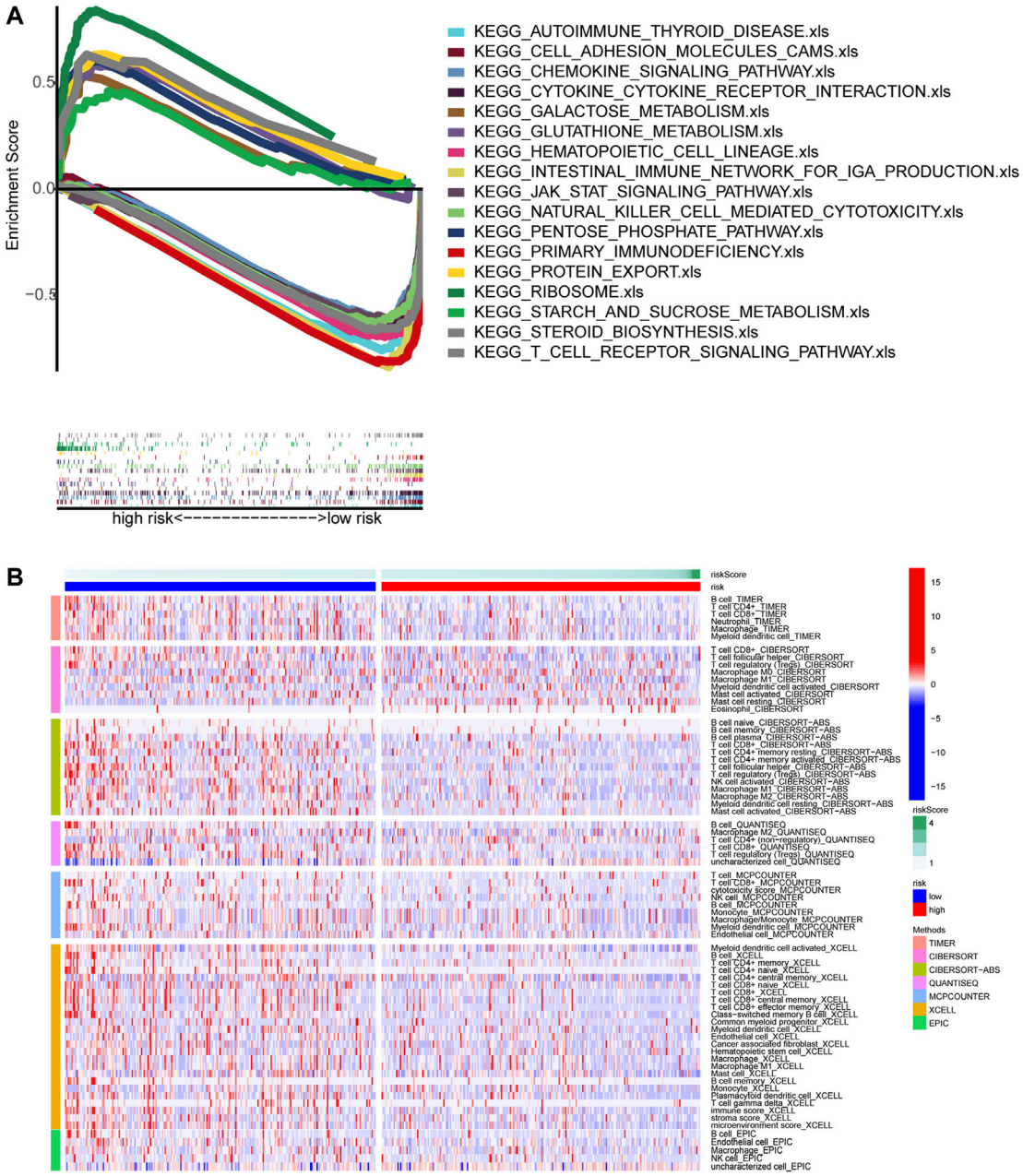


FIGURE 6 | TME and immunotherapy. **(A)** Top 10 pathways identified by GSEA **(B)**The heat maps of immune cells. **(C)** Risk scores were corrected with immune cells. **(D)** ssGSEA scores of immune functions. **(E)** Comparison of immune-related scores between the two groups. **(F)** The expression of 34 checkpoints. **(G)**Twelve targeted and immunotherapeutic drugs with different IC50 between the low-risk group (green) and the high-risk group (red).

AC116025.2, POLH-AS1, and PCED1B-AS1 are positive regulators of BRAF, SIRT1, FLT3, FASLG, TRAF2, ATRX, TERT, SPATA2, and TNFRSF1B. BRAF is a proto-oncogene that encodes for the B-Raf protein, a kinase of the RAF protein family (Rebocho and Marais, 2013). The Ras/Raf/MAPK pathway regulates cell growth, differentiation, cell motility and apoptosis (Rebocho and Marais, 2013; Schettini et al., 2018). Abnormal activation of the pathway is responsible for many tumors (Bouchè et al., 2021).

SIRT1 is a member of the HDAC family. Aberrant SIRT1 expression has been found in many tumors (Bradbury et al., 2005; Hida et al., 2007; Stünkel et al., 2007; Chen et al., 2014). ATRX is a member of the SWI-SNF protein family (Stayton et al., 1994; Picketts et al., 1996; Argentaro et al., 2007). SWI-SNF proteins are involved in DNA recombination and repair (Picketts et al., 1996), which are crucial for both development and cancer (Watson et al., 2015). SPATA2 is a TNF receptor modulator. TNF-α pathway modulates immune responses (Swann et al., 2008). TNF-α and

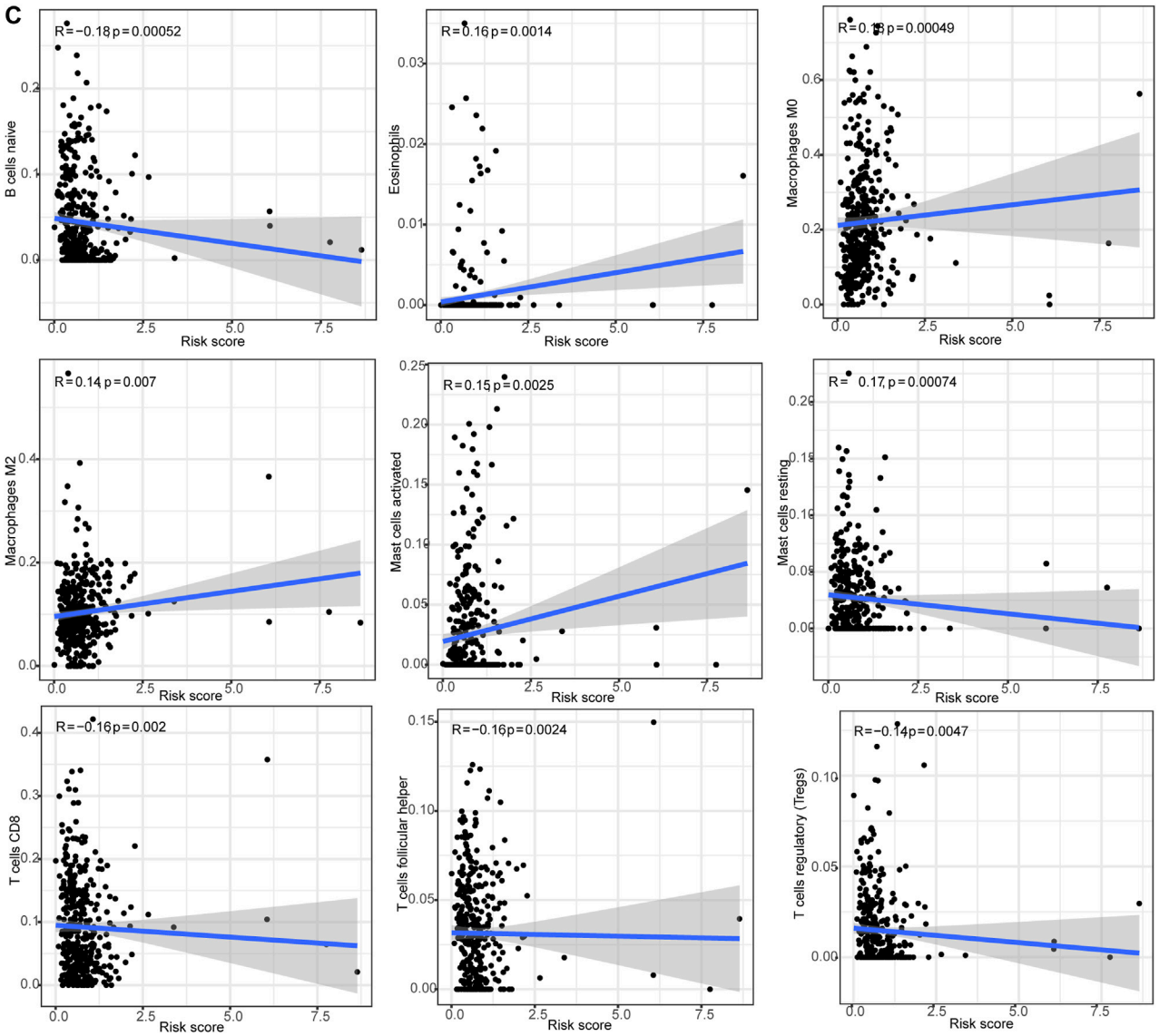


FIGURE 6 | Continued.

IL-1β induced SPATA2 expression in ovarian cancer cells and that increased SPATA2 expression was associated with poor prognosis of ovarian cancer patients (Wieser et al., 2019). Our study suggested SPATA2 expression is also associated with poor prognosis of HNSCC patients. ZBP1 is expressed in many tissues (Fu et al., 1999; Rothenburg et al., 2002) and is a interferon stimulated gene (Fu et al., 1999; Kuriakose and Kanneganti, 2018). ZBP1 expression in tumors is elevated. ZBP1 deletion blocks tumor necroptosis during tumor development and inhibits tumor metastasis (Baik et al., 2021). TNF-α is a pro-inflammatory cytokine mainly secreted by macrophages. There are two receptors for TNF-α, i.e., TNFRSF1A and TNFRSF1B. Although TNF can kill tumor cells, it also contribute to tumorigenesis (Aggarwal, 2003).

On the other hand, AC008764.8, AC127521.1, LINC00539, AC243829.2, AL121845.4, AC020911.1, AC139256.3, LINC00861, AC009121.3 and CDKN2A-DT were protective factors for HNSCC patients. Further analysis showed these lncRNAs were positive regulators of $p16^{INK4a}$, SPATA2, FLT3, FASLG, TRAF2, ATRX, TERT, BRAF, SIRT1, TNFRSF1B, and BCL2L11. $p16^{INK4a}$ is a tumor suppressor protein encoded by CDKN2A (Witcher and Emerson, 2009). $p16^{INK4a}$ is a negative regulator of cell cycle (Serrano et al., 1993). CDKN2A also encodes for another tumor suppressor protein, which interacts with p53 (Pomerantz et al., 1998). Inactivation of $p16^{INK4a}$ has been observed in various cancers *via* various mechanisms (Zhao et al., 2016). FLT3 is a receptor tyrosine kinase that is expressed in hematopoietic cells. Activation of FLT3 leads to autophosphorylation and mediates

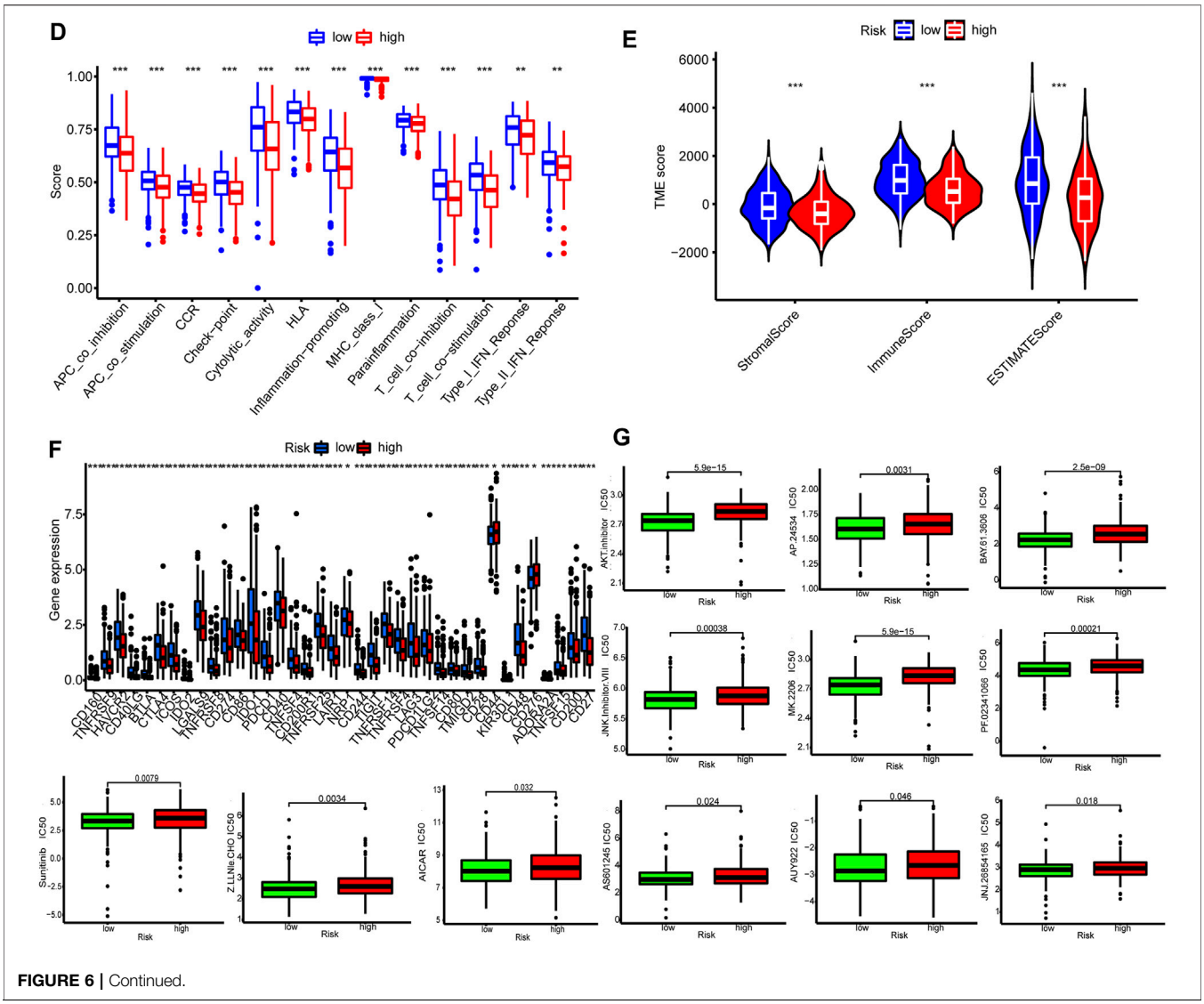


FIGURE 6 | Continued.

proliferation and differentiation of hematopoietic progenitor cells. However its role in tumorogenesis has not been reported. FASLG is a tumor suppressor and a member of the tumor necrosis factor superfamily (Magerus et al., 2021). FASLG/FAS signaling could induce apoptosis in various cancers (Liu et al., 2009; Kadam and Abhang, 2016; Magerus et al., 2021). TRAFs are intracellular adaptor signaling molecules of immune cells (Rothe et al., 1995; Ye et al., 2002; Park, 2018). TRAF2 promotes p53-dependent apoptosis by activating the JNK signaling cascade in cancer cells (Tsuchida et al., 2020). BCL2L11 is a member of BCL-2 family and regulates function of mitochondria (Concannon et al., 2010; Kilbride et al., 2010). BCL2L11 deletion/downregulation is found in many neoplasms and contribute to acquired drug resistance (Zhang et al., 2016).

By our model, we found pathways such as TNF, RAF and BCL-2 and FASLG/FAS are closely related to HNSCC. Although the protective lncRNAs are positive regulators of several tumor suppressors, they are also associated with several oncogenes. We propose that the prognostic value of a specific lncRNA is determined by the net effect of its multiple target genes.

Tumors have been described as "hot" or "cold" according to infiltration degree by T cells rushing to fight the cancerous cells. Hot tumors typically respond well to immunotherapy treatment using checkpoint inhibitors. Checkpoint inhibitors block signalling through checkpoint receptors to prevent the loss of T cell response to tumors. In contrast, nonimmunogenic "cold" tumors have not yet been infiltrated with T cells. The lack of T cells makes it difficult to provoke an immune response with immunotherapy drugs. IN addition, the microenvironment surrounding cold tumors contains myeloid-derived suppressor cells and T regulatory cells, which are known to dampen the immune response. In our model the patients in the high-risk group were more likely to have cold tumors, which may partially explain why the patients in the high-risk group had poor prognosis.

There were some limitations of our model. As a retrospective study, inherent biases might affect the model. We had performed internal validation by the test set, but we did not perform external validation.

In conclusion, we established a novel necroptosis-associated lncRNA signature for the prognosis of HNSCC. The established signatures suggest that lncRNAs might be associated with responses to targeted therapy and immunotherapy of HNSCC. The potential of this signature in predicting patient survival and treatment responses need to be validated in future tests.

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

JH, RL, YW, and LL contributed to conception and design of the study. JH organized the database. YW performed the statistical analysis. JH wrote the first draft of the manuscript. RL, DZ, YW, and LL wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This study was funded by the Fujian Provincial Health Technology Project (Grant number:2021CXA029), the Norman Bethune Medical Science Research Fund (Grant number: B19006CS), the Natural Science Foundation of Fujian Province (Grant number: 2019J01462) and Startup Fund for Scientific Research, Fujian Medical University (Grant number: 2018QH1226).

ACKNOWLEDGMENTS

The results shown here are in whole based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.907392/full#supplementary-material

REFERENCES

Aggarwal, B. B. (2003). Signalling Pathways of the TNF Superfamily: a Double-Edged Sword. *Nat. Rev. Immunol.* 3 (9), 745–756. doi:10.1038/nri1184

Argentaro, A., Yang, J.-C., Chapman, L., Kowalczyk, M. S., Gibbons, R. J., Higgs, D. R., et al. (2007). Structural Consequences of Disease-Causing Mutations in the ATRX-DNMT3-DNMT3L (ADD) Domain of the Chromatin-Associated Protein ATRX. *Proc. Natl. Acad. Sci. U.S.A.* 104 (29), 11939–11944. doi:10.1073/pnas.0704057104

Baik, J. Y., Liu, Z., Jiao, D., Kwon, H.-J., Yan, J., Kadigamuwa, C., et al. (2021). ZBP1 Not RIPK1 Mediates Tumor Necroptosis in Breast Cancer. *Nat. Commun.* 12 (1), 2666. doi:10.1038/s41467-021-23004-3

Berghe, T. V., Linkermann, A., Jouan-Lanhouet, S., Walczak, H., and Vandenabeele, P. (2014). Regulated Necrosis: the Expanding Network of Non-apoptotic Cell Death Pathways. *Nat. Rev. Mol. Cell Biol.* 15 (2), 135–147. doi:10.1038/nrm3737

Bouchè, V., Aldegheri, G., Donofrio, C. A., Fioravanti, A., Roberts-Thomson, S., Fox, S. B., et al. (2021). BRAF Signaling Inhibition in Glioblastoma: Which Clinical Perspectives? *Front. Oncol.* 11, 772052. doi:10.3389/fonc.2021.772052

Bradbury, C. A., Khanim, F. L., Hayden, R., Bunce, C. M., White, D. A., Drayson, M. T., et al. (2005). Histone Deacetylases in Acute Myeloid Leukaemia Show a Distinctive Pattern of Expression that Changes Selectively in Response to Deacetylase Inhibitors. *Leukemia* 19 (10), 1751–1759. doi:10.1038/sj.leu.2403910

Chan, F. K.-M., and Baehrecke, E. H. (2012). RIP3 Finds Partners in Crime. *Cell* 148 (1-2), 17–18. doi:10.1016/j.cell.2011.12.020

Chen, X., Hokka, D., Maniwa, Y., Ohbayashi, C., Itoh, T., and Hayashi, Y. (2014). Sirt1 Is a Tumor Promoter in Lung Adenocarcinoma. *Oncol. Lett.* 8 (1), 387–393. doi:10.3892/ol.2014.2057

Concannon, C. G., Tuffy, L. P., Weisová, P., Bonner, H. P., Dávila, D., Bonner, C., et al. (2010). AMP Kinase-Mediated Activation of the BH3-Only Protein Bim Couples Energy Depletion to Stress-Induced Apoptosis. *J. Cell Biol.* 189 (1), 83–94. doi:10.1083/jcb.200909166

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., et al. (2012). Landscape of Transcription in Human Cells. *Nature* 489 (7414), 101–108. doi:10.1038/nature11233

Dondelinger, Y., Declercq, W., Montessuit, S., Roelandt, R., Goncalves, A., Bruggeman, I., et al. (2014). MLKL Compromises Plasma Membrane Integrity by Binding to Phosphatidylinositol Phosphates. *Cell Rep.* 7 (4), 971–981. doi:10.1016/j.celrep.2014.04.026

Fakhry, C., Westra, W. H., Li, S., Cmelak, A., Ridge, J. A., Pinto, H., et al. (2008). Improved Survival of Patients with Human Papillomavirus-Positive Head and Neck Squamous Cell Carcinoma in a Prospective Clinical Trial. *JNCI J. Natl. Cancer Inst.* 100 (4), 261–269. doi:10.1093/jnci/djn011

Feng, X., Song, Q., Yu, A., Tang, H., Peng, Z., and Wang, X. (2015). Receptor-interacting Protein Kinase 3 Is a Predictor of Survival and Plays a Tumor Suppressive Role in Colorectal Cancer. *neo* 62 (4), 592–601. doi:10.4149/neo_2015_071

Fu, Y., Comella, N., Tognazzi, K., Brown, L. F., Dvorak, H. F., and Kocher, O. (1999). Cloning of DLM-1, a Novel Gene that Is Up-Regulated in Activated Macrophages, Using RNA Differential Display. *Gene* 240 (1), 157–163. doi:10.1016/s0378-1119(99)00419-9

Geeleher, P., Cox, N. J., and Huang, R. (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

Guicciardi, M. E., Malhi, H., Mott, J. L., and Gores, G. J. (2013). Apoptosis and Necrosis in the Liver. *Compr. Physiol.* 3 (2), 977–1010. doi:10.1002/cphy.c120020

Hida, Y., Kubo, Y., Murao, K., and Arase, S. (2007). Strong Expression of a Longevity-Related Protein, SIRT1, in Bowen's Disease. *Arch. Dermatol Res.* 299 (2), 103–106. doi:10.1007/s00403-006-0725-6

Hildebrand, J. M., Tanzer, M. C., Lucet, I. S., Young, S. N., Spall, S. K., Sharma, P., et al. (2014). Activation of the Pseudokinase MLKL Unleashes the Four-Helix Bundle Domain to Induce Membrane Localization and Necroptotic Cell Death. *Proc. Natl. Acad. Sci. U.S.A.* 111 (42), 15072–15077. doi:10.1073/pnas.1408987111

Hong, W., Liang, L., Gu, Y., Qi, Z., Qiu, H., Yang, X., et al. (2020). Immune-Related lncRNA to Construct Novel Signature and Predict the Immune Landscape of Human Hepatocellular Carcinoma. *Mol. Ther. - Nucleic Acids* 22, 937–947. doi:10.1016/j.omtn.2020.10.002

Huarte, M. (2015). The Emerging Role of lncRNAs in Cancer. *Nat. Med.* 21 (11), 1253–1261. doi:10.1038/nm.3981

Jiang, Y., Cao, W., Wu, K., Qin, X., Wang, X., Li, Y., et al. (2019). LncRNA LINC00460 Promotes EMT in Head and Neck Squamous Cell Carcinoma by Facilitating Peroxiredoxin-1 into the Nucleus. *J. Exp. Clin. Cancer Res.* 38 (1), 365. doi:10.1186/s13046-019-1364-z

Kadam, C. Y., and Abhang, S. A. (2016). Apoptosis Markers in Breast Cancer Therapy. *Adv. Clin. Chem.* 74, 143–193. doi:10.1016/bs.acc.2015.12.003

Kilbride, S. M., Farrelly, A. M., Bonner, C., Ward, M. W., Nyhan, K. C., Concannon, C. G., et al. (2010). AMP-activated Protein Kinase Mediates Apoptosis in Response to Bioenergetic Stress through Activation of the Pro-apoptotic Bcl-2 Homology Domain-3-Only Protein BMF. *J. Biol. Chem.* 285 (46), 36199–36206. doi:10.1074/jbc.M110.138107

Koo, G.-B., Morgan, M. J., Lee, D.-G., Kim, W.-J., Yoon, J.-H., Koo, J. S., et al. (2015). Methylation-dependent Loss of RIP3 Expression in Cancer Represses Programmed Necrosis in Response to Chemotherapeutics. *Cell Res.* 25 (6), 707–725. doi:10.1038/cr.2015.56

Kumar, M. M., and Goyal, R. (2017). LncRNA as a Therapeutic Target for Angiogenesis. *Ctmc* 17 (15), 1750–1757. doi:10.2174/1568026617666161116144744

Kuriakose, T., and Kanneganti, T.-D. (2018). ZBP1: Innate Sensor Regulating Cell Death and Inflammation. *Trends Immunol.* 39 (2), 123–134. doi:10.1016/j.it.2017.11.002

Li, J., Huang, S., Zeng, L., Li, K., Yang, L., Gao, S., et al. (2020). Necroptosis in Head and Neck Squamous Cell Carcinoma: Characterization of Clinicopathological Relevance and *In Vitro* Cell Model. *Cell Death Dis.* 11 (5), 391. doi:10.1038/s41419-020-2538-5

Liu, Y., Wen, Q.-J., Yin, Y., Lu, X.-T., Pu, S.-H., Tian, H.-P., et al. (2009). FASLG Polymorphism Is Associated with Cancer Risk. *Eur. J. Cancer* 45 (14), 2574–2578. doi:10.1016/j.ejca.2009.04.001

Magerus, A., Bercher-Brayer, C., and Rieux-Laucat, F. (2021). The Genetic Landscape of the FAS Pathway Deficiencies. *Biomed. J.* 44 (4), 388–399. doi:10.1016/j.bj.2021.06.005

Marchese, F. P., Raimondi, I., and Huarte, M. (2017). The Multidimensional Mechanisms of Long Noncoding RNA Function. *Genome Biol.* 18 (1), 206. doi:10.1186/s13059-017-1348-2

Mattick, J. S. (2018). The State of Long Non-coding RNA Biology. *ncRNA* 4 (3), 17. doi:10.3390/ncrna4030017

Meng, T., Huang, R., Zeng, Z., Huang, Z., Yin, H., Jiao, C., et al. (2019). Identification of Prognostic and Metastatic Alternative Splicing Signatures in Kidney Renal Clear Cell Carcinoma. *Front. Bioeng. Biotechnol.* 7, 270. doi:10.3389/fbioe.2019.00270

Nugues, A.-L., El Bouazzati, H., Hétuin, D., Berthon, C., Loyens, A., Bertrand, E., et al. (2014). RIP3 Is Downregulated in Human Myeloid Leukemia Cells and Modulates Apoptosis and Caspase-Mediated p65/RelA Cleavage. *Cell Death Dis.* 5 (8), e1384. doi:10.1038/cddis.2014.347

Park, H. H. (2018). Structure of TRAF Family: Current Understanding of Receptor Recognition. *Front. Immunol.* 9, 1999. doi:10.3389/fimmu.2018.01999

Pasparakis, M., and Vandenabeele, P. (2015). Necroptosis and its Role in Inflammation. *Nature* 517 (7534), 311–320. doi:10.1038/nature14191

Peng, W.-X., Koirala, P., and Mo, Y.-Y. (2017). LncRNA-mediated Regulation of Cell Signaling in Cancer. *Oncogene* 36 (41), 5661–5667. doi:10.1038/onc.2017.184

Picketts, D., Higgs, D. R., Bachoo, S., Blake, D. J., Quarrell, O. W., and Gibbons, R. J. (1996). ATRX Encodes a Novel Member of the SNF2 Family of Proteins: Mutations Point to a Common Mechanism Underlying the ATR-X Syndrome. *Hum. Mol. Genet.* 5 (12), 1899–1907. doi:10.1093/hmg/5.12.1899

Pomerantz, J., Schreiber-Agus, N., Liégeois, N. J., Silverman, A., Alland, L., Chin, L., et al. (1998). The Ink4a Tumor Suppressor Gene Product, p19Arf, Interacts with MDM2 and Neutralizes MDM2's Inhibition of P53. *Cell* 92 (6), 713–723. doi:10.1016/s0092-8674(00)81400-2

Rebocho, A. P., and Marais, R. (2013). ARAF Acts as a Scaffold to Stabilize BRAF: CRAF Heterodimers. *Oncogene* 32 (26), 3207–3212. doi:10.1038/onc.2012.330

Rothe, M., Sarma, V., Dixit, V. M., and Goeddel, D. V. (1995). TRAF2-Mediated Activation of NF-Kb by TNF Receptor 2 and CD40. *Science* 269 (5229), 1424–1427. doi:10.1126/science.7544915

Rothenburg, S., Schwartz, T., Koch-Nolte, F., and Haag, F. (2002). Complex Regulation of the Human Gene for the Z-DNA Binding Protein DLM-1. *Nucleic Acids Res.* 30 (4), 993–1000. doi:10.1093/nar/30.4.993

Schettini, F., De Santo, I., Rea, C. G., De Placido, P., Formisano, L., Giuliano, M., et al. (2018). CDK 4/6 Inhibitors as Single Agent in Advanced Solid Tumors. *Front. Oncol.* 8, 608. doi:10.3389/fonc.2018.00608

Serrano, M., Hannon, G. J., and Beach, D. (1993). A New Regulatory Motif in Cell-Cycle Control Causing Specific Inhibition of Cyclin D/CDK4. *Nature* 366 (6456), 704–707. doi:10.1038/366704a0

Stayton, C. L., Dabovic, B., Gulisano, M., Gecz, J., Broccoll, V., Glovanazzl, S., et al. (1994). Cloning and Characterization of a New Human Xq13 Gene, Encoding a Putative Helicase. *Hum. Mol. Genet.* 3 (11), 1957–1964. doi:10.1093/hmg/3.11.1957

Stünkel, W., Peh, B. K., Tan, Y. C., Nayagam, V. M., Wang, X., Salto-Tellez, M., et al. (2007). Function of the SIRT1 Protein Deacetylase in Cancer. *Biotechnol. J.* 2 (11), 1360–1368. doi:10.1002/biot.200700087

Sun, L., Wang, H., Wang, Z., He, S., Chen, S., Liao, D., et al. (2012). Mixed Lineage Kinase Domain-like Protein Mediates Necrosis Signaling Downstream of RIP3 Kinase. *Cell* 148 (1-2), 213–227. doi:10.1016/j.cell.2011.11.031

Sun, X., Lee, J., Navas, T., Baldwin, D. T., Stewart, T. A., and Dixit, V. M. (1999). RIP3, a Novel Apoptosis-Inducing Kinase. *J. Biol. Chem.* 274 (24), 16871–16875. doi:10.1074/jbc.274.24.16871

Swann, J. B., Vesely, M. D., Silva, A., Sharkey, J., Akira, S., Schreiber, R. D., et al. (2008). Demonstration of Inflammation-Induced Cancer and Cancer Immunoediting during Primary Tumorigenesis. *Proc. Natl. Acad. Sci. U.S.A.* 105 (2), 652–656. doi:10.1073/pnas.0708594105

Tsuchida, M., Yokosawa, T., Noguchi, T., Shimada, T., Yamada, M., Sekiguchi, Y., et al. (2020). Pro-apoptotic Functions of TRAF2 in P53-Mediated Apoptosis Induced by Cisplatin. *J. Toxicol. Sci.* 45 (4), 219–226. doi:10.2131/jts.45.219

Wang, H., Sun, L., Su, L., Rizo, J., Liu, L., Wang, L.-F., et al. (2014). Mixed Lineage Kinase Domain-like Protein MLKL Causes Necrotic Membrane Disruption upon Phosphorylation by RIP3. *Mol. Cell* 54 (1), 133–146. doi:10.1016/j.molcel.2014.03.003

Wang, T., Jin, Y., Yang, W., Zhang, L., Jin, X., Liu, X., et al. (2017). Necroptosis in Cancer: An Angel or a Demon? *Tumour Biol.* 39 (6), 101042831771153. doi:10.1177/1010428317711539

Watson, L. A., Goldberg, H., and Bérubé, N. G. (2015). Emerging Roles of ATRX in Cancer. *Epigenomics* 7 (8), 1365–1378. doi:10.2217/epi.15.82

Wieser, V., Tsibulak, I., Degasper, C., Welponer, H., Leitner, K., Parson, W., et al. (2019). Tumor Necrosis Factor Receptor Modulator Spermatogenesis-Associated Protein 2 Is a Novel Predictor of Outcome in Ovarian Cancer. *Cancer Sci.* 110 (3), 1117–1126. doi:10.1111/cas.13955

Witcher, M., and Emerson, B. M. (2009). Epigenetic Silencing of the p16INK4a Tumor Suppressor Is Associated with Loss of CTCF Binding and a Chromatin Boundary. *Mol. Cell* 34 (3), 271–284. doi:10.1016/j.molcel.2009.04.001

Ye, H., Arron, J. R., Lamothe, B., Cirilli, M., Kobayashi, T., Shevde, N. K., et al. (2002). Distinct Molecular Mechanism for Initiating TRAF6 Signalling. *Nature* 418 (6896), 443–447. doi:10.1038/nature00888

Zhang, H., Duan, J., Qu, Y., Deng, T., Liu, R., Zhang, L., et al. (2016). Onco-miR-24 Regulates Cell Growth and Apoptosis by Targeting BCL2L11 in Gastric Cancer. *Protein Cell* 7 (2), 141–151. doi:10.1007/s13238-015-0234-5

Zhao, R., Choi, B. Y., Lee, M.-H., Bode, A. M., and Dong, Z. (2016). Implications of Genetic and Epigenetic Alterations of CDKN2A (P16 INK4a) in Cancer. *EBioMedicine* 8, 30–39. doi:10.1016/j.ebiom.2016.04.017

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

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.

frontiers | Frontiers in Genetics

TYPE Review
PUBLISHED 10 August 2022
DOI 10.3389/fgene.2022.983564

OPEN ACCESS

EDITED BY
Zhongxiao Wang,
Biogen Idec (United States),
United States

REVIEWED BY
Yawen Guo,
Zhejiang University, China
Wei Wei,
Memorial Sloan Kettering Cancer Center, United States

*CORRESPONDENCE
Xinghua Liao,
xinghualiao@wust.edu.cn
Qi Zhang,
zhangqi06172@163.com
Zeming Liu,
6myt@163.com

†These authors have contributed equally to this work and share first authorship

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

RECEIVED 01 July 2022
ACCEPTED 18 July 2022
PUBLISHED 10 August 2022

CITATION
Zhou M, Dong M, Yang X, Gong J, Liao X, Zhang Q and Liu Z (2022), The emerging roles and mechanism of m6a in breast cancer progression.
Front. Genet. 13:983564.
doi: 10.3389/fgene.2022.983564

The emerging roles and mechanism of m6a in breast cancer progression

Mengying Zhou[1†], Menglu Dong[2†], Xue Yang[2†], Jun Gong[3], Xinghua Liao[1*], Qi Zhang[4*] Zeming Liu[4*]

[1]Institute of Biology and Medicine, College of Life and Health Sciences, Wuhan University of Science and Technology, Wuhan, China, [2]Department of Thyroid and Breast Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [3]Department of Biliary-Pancreatic Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China, [4]Department of Plastic and Cosmetic Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Breast cancer (BC) has continued to be the leading cause of cancer deaths in women, accompanied by highly molecular heterogeneity. N6-methyladenosine (m6A), a methylation that happens on adenosine N6, is the most abundant internal mRNA modification type in eukaryotic cells. Functionally, m6A methylation is a reversible modification process and is regulated by 3 enzymes with different functions, namely "writer", "reader", and "eraser". Abnormal m6A modifications trigger the expression, activation, or inhibition of key signaling molecules in critical signaling pathways and the regulatory factors acting on them in BC. These m6A-related enzymes can not only be used as markers for accurate diagnosis, prediction of prognosis, and risk model construction, but also as effective targets for BC treatment. Here, we have emphasized the roles of different types of m6A-related enzymes reported in BC proliferation, invasion, and metastasis, as well as immune regulation. The comprehensive and in-depth exploration of the molecular mechanisms related to m6A will benefit in finding effective potential targets and effective stratified management of BC.

KEYWORDS

breast cancer, M6A, methylation, gene expression, miRNA, metastasis

Abbreviations: AML, acute myeloid leukemia; ALKBH5, AlkB homolog 5; BC, breast cancer; BCBMs, breast cancer brain metastases ; CSC, cancer stem cell; CSCC, cervical squamous cell carcinoma; ceRNA, competitive endogenous RNA; FTO, fat mass and obesity-associated; GC, gastric cancer; HSCs, hematopoietic stem cells; HBXIP, hepatitis B X-interacting protein; METTL5, methyltransferase-like 5; m6A, N6-methyladenosine; OC, ovarian cancer; OS, overall survival; PD-L1, programmed cell death 1 ligand; PD-1, programmed cell death receptor 1; RBPs, RNA-binding proteins; SNPs, single nucleotide polymorphisms; TNBC, triple-negative breast cancer; TAMs, tumor-associated macrophages.

1 Introduction

Breast cancer (BC) has continued to be the leading cause of cancer deaths in women, and its incidence rates are still increasing globally (Sung et al., 2021). BC is a highly molecularly heterogeneous tumor type and is continually associated with headaches, such as recurrence, metastasis, and drug resistance (Yau et al., 2022) (Derose et al., 2011). These features cause certain bottlenecks in the diagnosis and treatment of BC. Therefore, exploring the molecular mechanism of the occurrence and development of BC, as well as the capability for the early diagnosis of BC, treatment monitoring, or the search for effective potential targets, is of great significance for the effective stratified management of BC and the development of new diagnosis and treatment methods (Ellis and Perou, 2013).

Epigenetic regulation, represented by N6-methyladenosine (m6A) modification, histone modification, DNA methylation, chromatin remodeling, and non-coding RNA (ncRNA) regulation, plays an overwhelming role in almost all biological behaviors, including cell differentiation and tissue development, and tumor progression (Wiener and Schwartz, 2021). M6A, a methylation that happens on adenosine N6, is the most abundant internal mRNA modification type in eukaryotic cells (Hu et al., 2022). There is a lot of evidence that m6A modification is an emerging important molecular modulation for tumors. In mammals, m6A modifications are known as a reversible dynamic process to influence different dimensions of RNA expression, including regulation of mRNA stability, splicing, translation efficiency, nuclear export, and degradation (Dierks et al., 2021). Functionally, this reversible m6A methylation modification is regulated by 3 enzymes with different functions, namely "writer", "reader", and "eraser" (Uddin et al., 2021). Among them, m6A can be installed by the methyltransferase complex, namely, writers, which include METTL3, METTL14, WTAP, RBM15/RBM15B, HAKAI, ZC3H13, and VIRMA/KIAA1429 (Huang et al., 2020). M6A eraser is able to remove methylation from m6A-modified RNA, mainly including FTO, ALKHB5, and ALKHB3 (Zhou et al., 2020). M6A reader is a selective RNA-binding protein that is responsible for catalyzing for recognizing m6A to activate downstream pathways, including YTH domain family 1-3 (YTHDF1, YTHDF1-2, YTHDF1), eIF3, IGF2BP1-3, hnRNPC, and hnRNPA2B1 (Dai et al., 2021) (Figure 1).

The enzymes involved in m6A modifications have been implicated in the regulation of gene expression and tumor evolution, including carcinogenesis, metastasis, and progression, especially in BC. For example, YTHDF1 overexpression is a not desirable signature for BC patients and is linked to lower immune infiltrate and poor clinical outcomes, while YTHDF1 inhibition promotes the

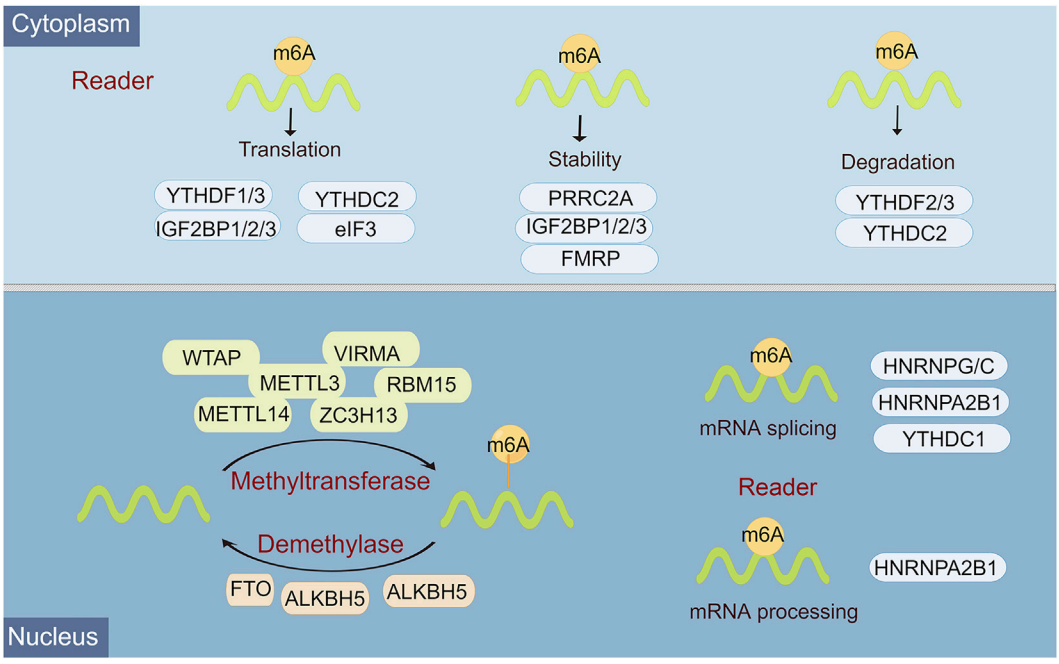


FIGURE 1
Overview of the classification and molecular mechanisms of m6A methylation. m6A RNA methylation is regulated by 3 different key enzymes, corresponding writers, erasers, and readers, which perform the functions of adding, deleting, or recognizing m6A, respectively. The consequences of m6A methylation lead to multiple processes in RNA metabolism and expression, including RNA splicing, miRNA processing, nuclear export, translation, stability, and RNA decay.

proliferation, migration, and invasion in BC cell lines (Li et al., 2022). The results of sequencing data mining showed that m6A and its target genes have corresponding changes at the gene and protein levels in tumors, endowing the potential to indicate the prognosis of BC. In a bioinformatics analysis, high expression of IGF2BP1, a key m6A regulator, was often associated with shorter overall survival (OS) in BC patients. This suggests that IGF2BP1 is an independent prognostic factor in BC (Zhong et al., 2021). Besides, Chen et al. confirmed that METTL3 methylation is involved in KRT7-mediated m6A-induced BC lung metastasis (Chen et al., 2021). Furthermore, the expression profile of m6A regulators in BC is also prominently related to tumor malignancy, tumor immune score, anti-tumor immune response, and therapeutic effect (He et al., 2021). Gong et al. reported that METTL14 and ZC3H13 were positively correlated with the abundance of $CD8^+$ T cells, neutrophils, macrophages, and dendritic cells (DCs) in BC.

Therefore, the systematic elucidation of the exact molecular mechanisms of m6A epigenetic regulation in BC progression is highly warranted. Here, we have reviewed and highlighted the roles of different types of m6A-related enzymes reported in BC proliferation, invasion, and metastasis, as well as immune regulation. Continued focus on the molecular mechanisms associated with m6A will benefit in finding effective potential targets and effective stratified management of BC.

2 The role of N6-methyladenosine modification in breast cancer progression

2.1 N6-methyladenosine writer in breast cancer progression

2.1.1 METTL3

At present, METTL3 is the most studied methylation-modified protein that plays a broad regulatory role in BC progression. METTL3 mainly participates in the biogenesis, decay, and translation control of mRNA through m6A modification (Lin et al., 2016). Both METTL3 and METTL14 have methyltransferase activity, and the methyltransferase complex formed by the two performs catalytic function during the m6A process (Yankova et al., 2021).

In TNBC, METTL3 is an important collaborator in regulating metastasis, and low expression of METTL3 is implicated in the poor prognosis of triple-negative breast cancer (TNBC) (Shi et al., 2020). As Shi et al. confirmed, this metastasis-suppressing function of METTL3 was achieved by suppressing the expression of COL3A1 and its m6A function (Shi et al., 2020). METTL3 could accelerate the protein levels of SOX2, CD133, and CD44 to maintain or promote BC cell stemness, which was triggered by the m6A modification of SOX2 mRNA by METTL3, ultimately leading to the alteration in enhanced BC invasion and migration capabilities (Xie et al., 2021). In BC lung metastasis cell lines, m6A and methyltransferase METTL3 expression was enhanced, while the expression level of demethylase FTO was reduced (Chen et al., 2021). M6A was capable of regulating lung metastasis in BC cells by regulating m6A/KRT7/KRT7-AS. The study by Wang et al. revealed a similar conclusion that METTL3 was a tumor promoter and its knockdown could inhibit tumor progression by reducing methylation levels (Wang H. et al., 2020). This mechanism of action was achieved through the Bcl-2 pathway targeted by METTL3. In addition, hepatitis B X-interacting protein (HBXIP) was identified to promote METTL3 expression by repressing miRNA let-7g, while METTL3 was simultaneously able to induce HBXIP expression (Cai et al., 2018). This mechanism caused a positive correlation between the expression of METTL3 and HBXIP in BC tissues and a positive feedback regulation phenomenon.

In the process of BC cell behavior, ncRNAs represented by miRNAs, lncRNAs, and circRNAs have been identified as very important direct regulators of METTL3. Therefore, interactions mediated by METTL3 and ncRNAs regulate the expression levels of post-transcriptionally regulated genes that determine tumor fate. METTL3 was capable of influencing the malignant behavior of BC EMT. Specifically, inhibition of METTL3 diminished the m6A modification of MALAT1, subsequently downregulated the MALAT1 expression to suppress EMT in BC by sponging miR-26b to reduce the expression of HMGA2 (Zhao et al., 2021). Fan et al. demonstrated that LINC00675 was a tumor protective factor, and its low expression was associated with higher tumor grade, lymphovascular invasion, and shorter survival (Fan and Wang, 2021). Furthermore, *in vitro* studies indicated that LINC00675 inhibited BC progression by suppressing miR-513b-5p in a METTL3-related m6A-dependent manner. From clinical, cellular, and tumor-bearing mouse levels, Xu et al. demonstrated that zinc finger protein 217 (ZNF217) silencing or miR-135 elevation inhibited BC cell migration, invasion, and EMT initiation (Xu et al., 2022). This was mediated by a mechanism that, ZNF217 could upregulate NANOG by reducing m6A levels through METTL3, thereby forming a miR-135/ZNF217/METTL3/NANOG axis.

LINC00958 was an overexpressed lncRNA that promoted the malignant progression of BC tumors (Rong et al., 2021). And, LINC00958 bound to miR-378a-3p to regulate YY1 expression, on the other hand, METTL3-mediated m6A modification promoted LINC00958 expression upregulation. In TNBC cell lines, the overexpressed METTL3 was an accelerator to suppress the proliferation and invasion (Ruan et al., 2021). Further validation showed that circMETTL3 served as a sponge for miR-34c-3p and exerted tumor-promoting functions by upregulating the expression of METTL3. METTL3-derived circRNAs contributed to the proliferation and invasion of BC cells, through the competitive endogenous RNA (ceRNA) effect

of miR-31–5p with upregulated CDK1 (Li et al., 2021). This function was also affected by the m6A modification mechanism of circMETTL3, which included METTL3.

2.1.2 METTL14

METTL14 is an important RNA methyltransferase that serves an essential and significant role in the growth of tumors by regulating RNA expression. METTL14 has been demonstrated to be a core component of the m6A methyltransferase complex and is implicated in the dynamic and reversible process of m6A modification (Zhu X. et al., 2021).

LINC00942 (LNC942) might function as an oncogene that promotes BC cell proliferation, and colony formation and inhibits apoptosis (Sun et al., 2020). In BC cells, LINC00942 increased the METTL14-mediated m6A methylation and its associated mRNA stability, as well as the CXCR4 and CYP1B1 expression of CXCR4 and CYP1B1, revealing a novel LNC942-METTL14-CXCR4/CYP1B1 regulatory axis. Zhao et al. demonstrated that silencing of lncRNA UCA1 suppressed DNA methylation of RNA methyltransferase METTL14 (Zhao et al., 2022). This event promoted m6A modification of miR-375, leading to reduced SOX12 expression and eventually restrained BC proliferation and invasion. METTL14 is also an m6A methyltransferase that is significantly elevated in BC tissues. When METTL14 was overexpressed or its activity was inhibited, the invasive ability of tumor cells becomes enhanced or weakened, accordingly (Yi et al., 2020). The abnormal expression of METTL14 reconstructed the miRNA expression profile of BC cells, and mainly regulated the cell adhesion and invasion ability by regulating the expression of Hsa-miR-146A-5p.

2.1.3 methyltransferase-like 5

Methyltransferase-like 5 (METTL5), can catalyze mA modification of 18S rRNA at adenosine 1832 (mA) in a critical position in the decoding center, possing the ability in regulating mRNA translation for impacting on cell growth (Sepich-Poore et al., 2022). In BC, METTL5 also exhibited a pattern of elevated expression and was required for the maintenance of BC cell lineage growth, reproduction, and S6K activation (Rong et al., 2020). The study by Rong et al. demonstrated that METTL5 was an 18S rRNA A1832-specific methyltransferase and was capable of regulating ribosome function via multiple models.

2.1.4 KIAA1429

KIAA1429, also known as VIRMA, is recognized as the largest m6A methyltransferase and is employed as a scaffold for the catalytic core component of the bridging m6A methyltransferase complex (Lan et al., 2019). KIAA1429 plays an instrumental function in m6A modification and has previously been found to be dysregulated in a variety of cancer types. KIAA1429 is considered to be involved in BC carcinogenesis and progression. Zhang et al. showed that KIAA1429 was a significant promoter of tumor invasion and metastasis *in vitro* and *in vivo*, and affected the course of BC in a non-m6A-regulated manner (Zhang et al., 2022). KIAA1429 failed to interfere with m6A levels of SMC1A mRNA, implying that m6A modifications did not affect the interplay between KIAA1429 and SMC1A mRNA. KIAA1429 directly bound to the 3′-UTR of SMC1A mRNA, leading to the stability enhancement of SMC1A mRNA. KIAA1429 showed an interesting expression pattern with high expression in tumor entities but low expression in nontumorous tissues (Qian et al., 2019). In terms of prognosis, the high expression of KIAA1429 was often associated with a lower OS. Mechanistically, KIAA1429 plays a carcinogenic role in BC progression by regulating CDK1 in an m6A-independent manner. The detailed mechanisms of m6A writer in regulating BC progression could be seen in Table 1.

2.2 N6-methyladenosine Eraser in breast cancer progression

2.2.1 Fat mass and obesity-associated

The fat mass and obesity-associated (FTO) gene is a well-known prominent factor in predicting obesity and is the first m6A eraser to be discovered in eukaryotic cells. FTO is responsible for controlling fatty acid transport, adipogenesis, fat metabolism, and obesity susceptibility. Single nucleotide polymorphisms (SNPs) of the FTO gene might be associated with various functions in different BC subtypes (Montazeri et al., 2022). It has been demonstrated that FTO expression is deregulated in a variety of tumors, including acute myeloid leukemia (AML), gastric cancer (GC), cervical squamous cell carcinoma (CSCC), ovarian cancer (OC), and BC (Deng et al., 2018).

As a key m6A demethylase, FTO is usually and aberrantly expressed up-regulated in BC tissues. High expression of FTO implies poor BC patient prognosis. Niu et al. determined that FTO remarkably contributed to BC cell proliferation and metastasis via the downregulation of tumor suppressor BNIP3, which involved FTO-mediated m6A demethylation in the 3′UTR of BNIP3 mRNA. Xu et al. demonstrated that in HER2-positive BC, the high FTO expression was linked to tumor progression, lymph node metastasis, TNM staging, and poor prognosis (Xu et al., 2020). *In vitro* experiments have similarly shown that FTO is a tumor-promoting factor that activates miR-181b-3p/ARL5B signaling leading to tumor migration.

2.2.2 AlkB homolog 5

AlkB homolog 5 (ALKBH5) is another key m6A demethylase for gene transcription, translation, and metabolism, and is

TABLE 1 The mechanisms of m6A writer in regulating BC progression.

Regulators	Expression pattern	Functions and mechanisms	Ref
METTL3	Low expression in TNBC	the low expression of METTL3-reduced m6A modification could promote TNBC metastasis by up-regulating COL3A1	(18)
METTL3	Upregulation in BC tissue, especially in T3-T4 or those accompanied with lymphatic metastasis	METTL3 promoted the stemness and malignant progression of BCa through mediating m6A modification on SOX2 mRNA	(19)
METTL3	Upregulation in BC tissue and cells	METTL3 knockdown could decrease the methylation level, reduce the proliferation, accelerate the apoptosis and inhibited the tumor growth by targeting Bcl-2	(20)
METTL3	Upregulation in BC tissue	HBXIP up-regulated METTL3 by suppressing let-7g, in which METTL3 increased HBXIP expression forming a positive feedback loop of HBXIP/let-7g/METTL3/HBXIP, leading to accelerated cell proliferation in BC	(21)
METTL3	Upregulation in BC tissue and cells	Silencing METTL3 down-regulated MALAT1 and HMGA2 by sponging miR-26b, and finally inhibited EMT, migration and invasion in BC	(22)
METTL3	–/–	METTL3 increased the m6A methylation of LINC00675, which enhanced the association between LINC00675 and miR-513b-5p	(23)
METTL3	–/–	MicroRNA-135 inhibited initiation of EMT in BC by targeting ZNF217 and promoting NANOG m6A modification	(24)
METTL3	Decreased in TNBC tissues and cell lines	circMETTL3 could act as a sponge for miR-34c-3p and inhibits cell proliferation, invasion, tumor growth and metastasis by up-regulating the expression of miR-34c-3p target gene METTL3	(26)
METTL3	–/–	circMETTL3 promotes BC progression through circMETTL3/miR-31–5p/CDK1 axis.	(27)
METTL14	Upregulation in BC cells and BC cohorts	LNC942 promoted METTL14-mediated m6A methylation in BC cell proliferation and progression	(29)
METTL14	Low expression in BC	LncRNA UCA1 promoted SOX12 expression by regulating m6A modification of miR-375 by METTL14 through DNA methylation	(30)
METTL14	Upregulation in BC tissue	METTL14 modulated m6A modification and hsa-miR-146a-5p expression, thereby promoting the migration and invasion of BC cells	(31)
METTL5	Elevated expression in BC tissue and cell lines	Ribosome 18S m6A methyltransferase METTL5 promotes translation initiation and BC cell growth, uncovering critical and conserved roles of METTL5 in the regulation of translation	(33)
KIAA1429	Overexpression in BC	KIAA1429/SMC1A/SNAIL axis in promoting EMT progress and metastasis in BC	(35)
KIAA1429	Highly expressed in BC tissues	KIAA1429 promoted BC progression and was correlated with pathogenesis by associating with CDK1 mRNA in an m6A-independent manner.	(36)

regarded as an effective biomarker for various diseases, especially cancers (Zhang et al., 2017). ALKBH3 preferentially acts on m6A in tRNA sites.

For instance, Wang et al. have previously reported that ALKBH5 could specifically regulate the function of AML leukemia stem cells without affecting normal hematopoietic stem cells (HSCs) (Wang J. et al., 2020). This highlighted the crucial role of ALKBH5 in maintaining cancer stem cell (CSC) renewal and cancer occurrence and development through the KDM4C-ALKBH5-AXL Signaling Axis. Hypoxia induces a series of stress changes to support the stable viability of cells, including changes in cytokine secretion profiles, and post-transcriptional and translational regulation. Reprogramming of the mA epitranscriptome is vital for the formation of the transcriptome and proteome in the setting of hypoxia (Wang et al., 2021). Similarly, Zhang et al. showed that hypoxia-induced an enhancement of NANOG mRNA and protein expression and breast CSC phenotype in a HIF- and ALKBH5-dependent manner, and that ALKBH5 deficiency was a debilitating factor for the hypoxia-induced BC CSC enrichment (Zhang et al., 2016).

2.3 N6-methyladenosine Reader in breast cancer progression

2.3.1 YTHDF1

YTHDF1-3 are the three major m6A binding proteins and the most widely studied and versatile m6A readers (Chen et al., 2019). YTHDF1-3, containing special YTH domains, possess the capabilities of targeting and recognizing m6A-modified RNAs and mediating their degradation. YTHDF1 is a typical and highly expressed m6A reader protein in BC tissues and cell lines, and its high expression status is thought to be indicative of tumor size, metastasis, poor prognosis, and chemotherapy resistance (Anita et al., 2020). YTHDF1 is involved in almost the whole process of tumor biological behavior, and plays an important role in regulating transcription, translation, protein synthesis,

angiogenesis, and EMT. Intriguingly, YTHDF1 as a target of tumor immune regulation has also attracted much attention.

Chen et al. showed that YTHDF1 promoted YTHDF1/FOXM1 to enhance FOXM1 expression, which in turn intensified the proliferation, invasion, and EMT phenotype of BC cells (Chen et al., 2022). Another 2022 similar study showed that a series of cascade reactions triggered by YTHDF1 were important molecular events in driving BC. Specifically, HIF1α expression could be induced and miR-16-5p levels were suppressed in a hypoxic microenvironment, resulting in upregulated YTHDF1 expression (Yao et al., 2022). Also, inhibition of YTHDF1 was able to promote the down-regulation of the glycolytic gene PKM2 to reduce BC glycolytic activity and lead to tumorigenicity and metastasis inhibition. These studies suggest that YTHDF1 can be used in a variety of regulatory pathways to modulate BC progression, and that depletion or targeted inhibition of YTHDF1 is a potentially efficient BC therapeutic strategy.

2.3.2 YTHDF2

YTHDF2 is an N-methyladenosine-binding protein and can modulate mRNA stability, thus impacting central nervous system responses, embryonic development, and tumor evolution. Tumor biology studies have shown that YTHDF2 can modulate m6A modification to regulate downstream signaling molecules to regulate tumor cell proliferation, invasion, and migration (Shen et al., 2021). For instance, SUMOylation of YTHDF2 promotes mRNA degradation and cancer progression by increasing its binding affinity to m6A-modified mRNA (Hou et al., 2021). Einstein et al. uncovered a mechanism associated with RNA-binding proteins (RBPs), that the suppression of YTHDF2 initiated proteotoxic cell death pattern in MYC-driven TNBC (Einstein et al., 2021). This work not only demonstrated that YTHDF2, aberrantly expressed RBP and its mode of interaction with RNA were essential for BC cell growth, but that targeting YTHDF2 and specific RBP possessed outstanding BC therapeutic potential. The important mechanism of post-translational protein modification of YTHDF2, fully confirmed that YTHDF2 has a delicate manipulation between the regulation of protein post-translational modification and RNA chemical modification.

2.3.3 YTHDF3

YTHDF3, in combination with YTHDF1 and YTHDF2, has a crucial effect in enhancing the synthesis of m6A-modified mRNAs in the cytoplasm (Shi et al., 2017). In TNBC subtypes, YTHDF3 expression was associated with poorer disease-free survival (DFS) and overall survival (OS) in patients (Lin et al., 2022). YTHDF3 could intensify the ZEB1 mRNA stability in an m6A-dependent manner, consequently leading to BC cell growth and EMT (Lin et al., 2022).

Intriguingly, the high expression level of YTHDF3 was also closely related to the prognosis of patients with breast cancer brain metastases (BCBMs). By enhancing the translation of m6A-enriched transcripts of ST6GALNAC5, GJA1, and EGFR, YTHDF3 promoted the communication between BC cells, endothelial cells, astrocytes, and tumor metastasis phenotypes represented by angiogenesis (Chang et al., 2020). Therefore, YTHDF3 could affect cascade steps in the BCBM, and then domesticate the evolution of BC cell changes in TME toward inducing brain metastatic polarity.

Totally, different binding proteins selectively recognize m6A-modified RNAs for achieving gene expression regulation. YTHDF1 is conducive to the enhanced mRNA translation, YTHDF2 is responsible for mRNA degradation, and YTHDF3 precipitates in the translation and degradation via the reciprocity actions with YTHDF1 and YTHDF2. This means that YTHDF3 can both collaborate with YTHDF1 to catalyze the translation of methylated RNAs, or directly engage with YTHDF2 to accelerate the decay of mRNAs. Thus, YTHDF1, YTHDF2, and YTHDF3 exert irreplaceable functions to foster BC progression and potentially even become robust therapeutic targets for prognostic stratification and effective treatment of BC.

2.3.4 IGF2BP1

As a post-transcriptional fine regulator, IGF2BP1 plays a role in remodeling tumor growth, chemotherapy resistance, and macroscopically, OS and recurrence of tumor patients. GF2BP1 potentiates tumor malignant progression in a variety of solid tumors and exhibits a poor prognostic indicative value (Glaß et al., 2021). The principal action of IGF2BP1 in oncogenic cells is to stabilize mRNA encoding oncogenic factors. In pan-cancer studies, the high expression and tumor-promoting characteristics of IGF2BP1 in specific tumors make it a promising therapeutic target, but it is also inhibitory in some tumors (Huang et al., 2018).

Zhu et al. identified a hypoxia-induced lncRNA KB-1980E6.3, that exhibited abnormal BC tissue upregulation and was associated with a poor prognosis (Zhu P. et al., 2021). LncRNA KB-1980E6.3 increased the stability of c-Myc mRNAs by binding to m6A reader IGF2BP1 and consequently maintained the stemness of BCSCs. Interrupting this mechanism was of the potential to provide a therapeutic strategy for hypoxic tumors. MIR210HG acts as an oncogenic lncRNA highly expressed in BC tissue, and could promote BC metastasis, by inhibiting its encoded miR-210 (Shi et al., 2022). Moreover, MYCN directly activated IGF2BP1, and both IGF2BP1 and ELAVL1 strengthened the MIR210HG stability, resulting in a MYCN/IGF2BP1/MIR210HG regulatory axis.

3 N6-methyladenosine modification in immune regulation

Immune cells, secreted factors, and the tumor immune microenvironment in which they are intertwined are indispensable key links in the anti-tumor response (Liu et al., 2021). Considerable evidence suggests that m6A is involved in processes that regulate innate and adaptive immune cells, which in turn have been assigned roles in anti-inflammatory, anti-infective, and anti-tumor immunity (Ma et al., 2021). There is also much literature based on the existing reported m6A regulators, mining the correlation and prognostic scoring model of m6A regulators and immune infiltrating cells, for providing a novel evaluation tool for the diagnosis, prognosis, and immune status of BC (Yuan et al., 2022).

The immunomodulatory role of METTL3 in BC has been frequently reported. Yin et al. found that knockdown of METTL3 in bone marrow cells triggered malignant tumor proliferation and metastasis and exhibited an elevated abundance of M1/M2-like tumor-associated macrophages (TAMs) and Treg infiltration (Yin et al., 2021). Mechanistic studies suggested that deletion of METTL3 disrupted YTHDF1-mediated SPRED2 translation, thereby enhancing NF-kB and STAT3 activation via the ERK pathway, leading to tumor progression. Meanwhile, as the therapeutic benefit of programmed cell death receptor 1 (PD-1) inhibitor was weakened in Mettl3−/− mice, METTL3 could be a potential target for tumor immunotherapy. In addition, it has been reported that the expression of programmed cell death 1 ligand (PD-L1) was positively linked to the expression of METTL3 and IGF2BP3 in BC tissues (Wan et al., 2022). Since METTL3-mediated m6A modification could enhance PD-L1 mRNA stability through the METTL3-IGF2BP3 axis, tumor immune cell infiltrations and $CD8^+$ T cell functions were enhanced forcefully when METTL3 or IGF2BP3 is inhibited. Ou et al. identified a specific C5aR1+ neutrophil subpopulation that potentiated BC cell glycolysis through ERK1/2-WTAP-ENO1 signaling, indicating that C5aR1+ neutrophils and the associated WTAP-ENO1 axis contribute to potential BC therapeutic target (Yin et al., 2021).

Another study reported that the co-expression network of YTHDF1 is critical in shaping immune responses, including antigen processing and presentation (Hu et al., 2021). YTHDF1 may act as a hopeful pan-cancer immune biomarker, as well as a novel promising marker for tumor immunotherapy. These results provide strong evidence that m6A modification is involved in the complex immune regulation of BC. M6A modifications can reshape antitumor immune responses by affecting immune cell state and function, and post-transcriptional regulation of specific cytokines and proteins.

4 Discussion

Based on these current developments, it is evident that m6A plays a dual role in shaping tumor progression. Specifically, m6A regulates the expression of its target genes to influence tumor progression, and whether the target genes act as tumor promoters or tumor suppressors determines the tumor-promoting or tumor-suppressing function of m6A. The presence of m6A modifications contributes to the promotion/suppression of various cellular functions, such as precursor mRNA splicing, nuclear translocation, stability, translation, and microRNA biogenesis, as the modification represented by tumor cells and immune cells, thus remodeling the BC progression.

In terms of BC diagnosis, m6A enzymes have also demonstrated good predictive efficacy. This is due to the fact that m6A enzymes show a characteristic pattern of differential expression in different BC subtypes and BC staging classifications. There have been more than 20 risk models based on these screened m6A regulators in BC research. For example, the overexpression of YTHDF1, YTHDF3, and KIAA1429 predicted a poor prognosis in terms of overall survival (OS), and the upregulation of YTHDF3 was an independent prognostic factor for OS in BR patients (Liu et al., 2019). There is even a strong performance of m6A enzymes in the treatment efficacy and recurrent metastasis of BC.

Aberrant expression of m6A regulators, are potential indicators for BC prediction, individually or synergistically. Other reported factor-based (such as ferroptosis, autophagy, lncRNA, m6A regulator-mediated immune Genes) models are also of huge value that have been confirmed in multiple studies. He et al. constructed an m6A regulator pattern, which could be effective for predicting malignancy, outcomes, and antitumor immune response (He et al., 2021). Moreover, the established models based on both m6A and other factors are attracting more and more attention. For instance, we previously constructed a risk signature based on 6 screened m6A-related lncRNAs, including Z68871.1, AL122010.1, OTUD6B-AS1, AC090948.3, AL138724.1, EGOT (Lv et al., 2021). This model could identify the prognosis and immune state in BC. Zhang et al. also adopt 21 m6A-related lncRNAs to establish a predictive model for predicting prognostic situations and BC subtypes with different immunogenicity (Zhang et al., 2021). Most of these models tend to build more accurate and extensive prediction models based on the inclusion of M6A-related indicators, thus providing rich information on tumor malignancy, tumor metastasis and recurrence, immune microenvironment, efficacy monitoring, and drug resistance evaluation.

Moreover, m6A-related models may have superior diagnostic value by associating with other models, including death (necroptosis, autophagy, and ferroptosis) genes, immune genes, glycosylation genes, and so on. However, since most of these are studies based on database excavation, they are retrospective studies. There are still relatively few prospective

studies that can actually be performed in a large sample of realistic cohorts, and need to be further corroborated in actual clinical practice to prove their credibility. Overall, as a single regulator included in the model can offer many disease information, more gene combination models potentially provide a more comprehensive for breast cancer diagnosis and therapeutic outcomes. We must first admit that the reported single or the risk models based on these screened m6A regulators, is not the replacement for traditional and classical methods such as serology, pathology, and imaging. To be more precise, the traditional multi-method routine is the gold standard for diagnosis and prediction, while the risk models based on these screened m6A regulators are potential multifactorial predictors, which can be used as a useful supplement to clinical routine evaluation methods.

In terms of m6A-targeted BC tumor therapy, the following points still deserve in-depth consideration. First, m6A can regulate tumor progression through multiple mechanisms, which has been confirmed in BC. How m6A modification dysregulation is regulated by affecting tumor stem cells, immune environment, tumor cell fate, and other multiple ways is an important theoretical basis for strengthening m6A as tumor therapy. However, the role of m6A is dual in tumors, and a comprehensive assessment of how to regulate m6A-related enzymes by inhibition or activation requires specific BC subtypes, tumor microenvironment, and other underlying diseases. Moreover, ncRNA is an important player involved in tumor regulation, and how m6A RNA modification affects ncRNA function deserves further exploration.

Secondly, the primary prerequisite for targeting m6A-modified enzymes is the resolution of the protein crystal complexes of these enzymes in order to mine and design docked high-affinity small molecules and antibodies by conformational relationships (Oerum et al., 2021). These inhibitors and antibodies, in turn, will be considered to provide a good pre-requisite for clinical targeting only after effective validation in cellular and animal experiments (You et al., 2022). Finally, numerous studies have explained the critical role of m6A recognition proteins in BC, but a considerable number of m6A regulators have not been fully validated, including RBM15/RBM15B, HAKAI, ZC3H13, ALKBH3, eIF3, hnRNPC, and hnRNPA2B1. Multi-omics information mining based on single-cell sequencing, proteomics, RNA-seq, and m6A methylation sequencing will provide comprehensive information mining for BC tumor sites and cellular models. This unreported regulator in BC is a research gap and therefore has considerable research value, including the relationship between expression abundance and prognosis, diagnostic and therapeutic potential. These can provide a more profound complement to the discovery of new molecules or previous regulatory networks. Thus, continued research is still needed to fully elucidate the role of m6A regulators in mRNA and ncRNA biology.

Finally, aberrant regulation of m6A regulatory proteins is involved in BC drug resistance and tumor immune response. Targeted m6A-based therapies will contribute to the oncological treatment of BC. However, it is worth mentioning that because the regulatory network of m6A modifications is complex and involves multiple signaling molecules and pathways, inhibition of a single molecule may lead to unintended responses. Therefore, for targeted therapies of m6A, or combination strategies with other tumor-targeting drugs, the optimal combination will lead to the best efficacy.

5 Conclusion

Overall, m6A is an important mechanism for epigenetic modifications that regulate BC progression. Abnormal m6A modifications trigger the expression, activation, or inhibition of key signaling molecules in critical signaling pathways and the regulatory factors acting on them in BC. M6A-related proteins and their targets show differentially expressed patterns in BC tissue and blood. These m6A-related genes can not only be used as markers for accurate diagnosis, prediction of prognosis, and risk model construction, but also as effective targets for BC treatment.

Author contributions

All authors contributed to the design of the study and the writing of the manuscript. MZ, MD, and XY designed the project and revised the manuscript. XL, QZ, and ZL performed the literature search and wrote the manuscript. JG helped revise the manuscript. All authors reviewed the manuscript and approved the final version.

Funding

This work was supported by the National Natural Science Foundation of China (81900518).

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

Anita, R., Paramasivam, A., Priyadharsini, J. V., and Chitra, S. (2020). The m6A readers YTHDF1 and YTHDF3 aberrations associated with metastasis and predict poor prognosis in breast cancer patients. *Am. J. Cancer Res.* 10, 2546–2554.

Cai, X., Wang, X., Cao, C., Gao, Y., Zhang, S., Yang, Z., et al. (2018). HBXIP-elevated methyltransferase METTL3 promotes the progression of breast cancer via inhibiting tumor suppressor let-7g. *Cancer Lett.* 415, 11–19. doi:10.1016/j.canlet.2017.11.018

Chang, G., Shi, L., Ye, Y., Shi, H., Zeng, L., Tiwary, S., et al. (2020). YTHDF3 induces the translation of m6A-enriched gene transcripts to promote breast cancer brain metastasis. *Cancer Cell* 38, 857–871. e7. doi:10.1016/j.ccell.2020.10.004

Chen, F., Chen, Z., Guan, T., Zhou, Y., Ge, L., Zhang, H., et al. (2021). N^6-methyladenosine regulates mRNA stability and translation efficiency of KRT7 to promote breast cancer lung metastasis. *Cancer Res.* 81, 2847–2860. doi:10.1158/0008-5472.CAN-20-3779

Chen, H., Yu, Y., Yang, M., Huang, H., Ma, S., Hu, J., et al. (2022). YTHDF1 promotes breast cancer progression by facilitating FOXM1 translation in an m6A-dependent manner. *Cell Biosci.* 12, 19. doi:10.1186/s13578-022-00759-w

Chen, X.-Y., Zhang, J., and Zhu, J.-S. (2019). The role of m6A RNA methylation in human cancer. *Mol. Cancer* 18, 103. doi:10.1186/s12943-019-1033-z

Dai, X.-Y., Shi, L., Li, Z., Yang, H.-Y., Wei, J.-F., and Ding, Q. (2021). Main N6-methyladenosine readers: YTH family proteins in cancers. *Front. Oncol.* 11, 635329. doi:10.3389/fonc.2021.635329

Deng, X., Su, R., Stanford, S., and Chen, J. (2018). Critical enzymatic functions of FTO in obesity and cancer. *Front. Endocrinol.* 9, 396. doi:10.3389/fendo.2018.00396

Derose, Y. S., Wang, G., Lin, Y. C., Bernard, P. S., Buys, S. S., Ebbert, M. T. W., et al. (2011). Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nat. Med.* 17, 1514–1520. doi:10.1038/nm.2454

Dierks, D., Garcia-Campos, M. A., Uzonyi, A., Safra, M., Edelheit, S., Rossi, A., et al. (2021). Multiplexed profiling facilitates robust m6A quantification at site, gene and sample resolution. *Nat. Methods* 18, 1060–1067. doi:10.1038/s41592-021-01242-z

Einstein, J. M., Perelis, M., Chaim, I. A., Meena, J. K., Nussbacher, J. K., Tankka, A. T., et al. (2021). Inhibition of YTHDF2 triggers proteotoxic cell death in MYC-driven breast cancer. *Mol. Cell* 81, 3048–3064. e9. doi:10.1016/j.molcel.2021.06.014

Ellis, M. J., and Perou, C. M. (2013). The genomic landscape of breast cancer as a therapeutic roadmap. *Cancer Discov.* 3, 27–34. doi:10.1158/2159-8290.CD-12-0462

Fan, S., and Wang, L. (2021). N 6 -Methyladenosine-regulated LINC00675 suppress the proliferation, migration and invasion of breast cancer cells via inhibiting miR-513b-5p. *Bioengineered* 12, 10690–10702. doi:10.1080/21655979.2021.2001905

Glaß, M., Misiak, D., Bley, N., Müller, S., Hagemann, S., Busch, B., et al. (2021). IGF2BP1, a conserved regulator of RNA turnover in cancer. *Front. Mol. Biosci.* 8, 632219. doi:10.3389/fmolb.2021.632219

He, X., Tan, L., Ni, J., and Shen, G. (2021). Expression pattern of m6A regulators is significantly correlated with malignancy and antitumor immune response of breast cancer. *Cancer Gene Ther.* 28, 188–196. doi:10.1038/s41417-020-00208-1

Hou, G., Zhao, X., Li, L., Yang, Q., Liu, X., Huang, C., et al. (2021). SUMOylation of YTHDF2 promotes mRNA degradation and cancer progression by increasing its binding affinity with m6A-modified mRNAs. *Nucleic Acids Res.* 49, 2859–2877. doi:10.1093/nar/gkab065

Hu, J., Qiu, D., Yu, A., Hu, J., Deng, H., Li, H., et al. (2021). YTHDF1 is a potential pan-cancer biomarker for prognosis and immunotherapy. *Front. Oncol.* 11, 607224. doi:10.3389/fonc.2021.607224

Hu, L., Liu, S., Peng, Y., Ge, R., Su, R., Senevirathne, C., et al. (2022). m6A RNA modifications are measured at single-base resolution across the mammalian transcriptome. *Nat. Biotechnol.* [Epub ahead of print]. doi:10.1038/s41587-022-01243-z

Huang, H., Weng, H., and Chen, J. (2020). The biogenesis and precise control of RNA m6A methylation. *Trends Genet.* 36, 44–52. doi:10.1016/j.tig.2019.10.011

Huang, X., Zhang, H., Guo, X., Zhu, Z., Cai, H., and Kong, X. (2018). Insulin-like growth factor 2 mRNA-binding protein 1 (IGF2BP1) in cancer. *J. Hematol. Oncol.* 11, 88. doi:10.1186/s13045-018-0628-y

Lan, T., Li, H., Zhang, D., Xu, L., Liu, H., Hao, X., et al. (2019). KIAA1429 contributes to liver cancer progression through N6-methyladenosine-dependent post-transcriptional modification of GATA3. *Mol. Cancer* 18, 186. doi:10.1186/s12943-019-1106-z

Li, C., Zhang, C., Zhang, G., Chen, B., Li, X., Li, K., et al. (2022). YTHDF1 amplification is correlated with worse outcome and lower immune cell infiltrations in breast cancer. *Cancer Biomark.* 1, 1–16. doi:10.3233/CBM-203103

Li, Z., Yang, H., Dai, X., Zhang, X., Huang, Y., Shi, L., et al. (2021). CircMETTL3, upregulated in a m6A-dependent manner, promotes breast cancer progression. *Int. J. Biol. Sci.* 17, 1178–1190. doi:10.7150/ijbs.57783

Lin, S., Choe, J., Du, P., Triboulet, R., and Gregory, R. I. (2016). The m6A methyltransferase METTL3 promotes translation in human cancer cells. *Mol. Cell* 62, 335–345. doi:10.1016/j.molcel.2016.03.021

Lin, Y., Jin, X., Nie, Q., Chen, M., Guo, W., Chen, L., et al. (2022). YTHDF3 facilitates triple-negative breast cancer progression and metastasis by stabilizing ZEB1 mRNA in an m6A-dependent manner. *Ann. Transl. Med.* 10, 83. doi:10.21037/atm-21-6857

Liu, C., Yang, Z., Li, R., Wu, Y., Chi, M., Gao, S., et al. (2021). Potential roles of N6-methyladenosine (m6A) in immune cells. *J. Transl. Med.* 19, 251. doi:10.1186/s12967-021-02918-y

Liu, L., Liu, X., Dong, Z., Li, J., Yu, Y., Chen, X., et al. (2019). N6-methyladenosine-related genomic targets are altered in breast cancer tissue and associated with poor survival. *J. Cancer* 10, 5447–5459. doi:10.7150/jca.35053

Lv, W., Wang, Y., Zhao, C., Tan, Y., Xiong, M., Yi, Y., et al. (2021). Identification and validation of m6A-related lncRNA signature as potential predictive biomarkers in breast cancer. *Front. Oncol.* 11, 745719. doi:10.3389/fonc.2021.745719

Ma, Z., Gao, X., Shuai, Y., Xing, X., and Ji, J. (2021). The m6A epitranscriptome opens a new charter in immune system logic. *Epigenetics* 16, 819–837. doi:10.1080/15592294.2020.1827722

Montazeri, F., Hatami, H., Fathi, S., Hasanpour Ardekanizadeh, N., Bourbour, F., Rastgoo, S., et al. (2022). FTO genotype was associated with breast cancer in HER2 negative patients. *Clin. Nutr. ESPEN* 49, 495–498. doi:10.1016/j.clnesp.2022.02.122

Oerum, S., Meynier, V., Catala, M., and Tisne, C. (2021). A comprehensive review of m6A/m6Am RNA methyltransferase structures. *Nucleic Acids Res.* 49, 7239–7255. doi:10.1093/nar/gkab378

Qian, J. Y., Gao, J., Sun, X., Cao, M. Da, Shi, L., Xia, T. S., et al. (2019). KIAA1429 acts as an oncogenic factor in breast cancer by regulating CDK1 in an N6-methyladenosine-independent manner. *Oncogene* 38, 6123–6141. doi:10.1038/s41388-019-0861-z

Rong, B., Zhang, Q., Wan, J., Xing, S., Dai, R., Li, Y., et al. (2020). Ribosome 18S m6A methyltransferase METTL5 promotes translation initiation and breast cancer cell growth. *Cell Rep.* 33, 108544. doi:10.1016/j.celrep.2020.108544

Rong, D., Dong, Q., Qu, H., Deng, X., Gao, F., Li, Q., et al. (2021). m6A-induced LINC00958 promotes breast cancer tumorigenesis via the miR-378a-3p/YY1 axis. *Cell Death Discov.* 7, 27. doi:10.1038/s41420-020-00382-z

Ruan, H., Gu, W., Xia, W., Gong, Y., Zhou, X., Chen, W., et al. (2021). METTL3 is suppressed by circular RNA circMETTL3/miR-34c-3p signaling and limits the tumor growth and metastasis in triple negative breast cancer. *Front. Oncol.* 11, 778132. doi:10.3389/fonc.2021.778132

Sepich-Poore, C., Zheng, Z., Schmitt, E., Wen, K., Zhang, Z. S., Cui, X.-L., et al. (2022). The METTL5-TRMT112 N6-methyladenosine methyltransferase complex regulates mRNA translation via 18S rRNA methylation. *J. Biol. Chem.* 298, 101590. doi:10.1016/j.jbc.2022.101590

Shen, X., Zhao, K., Xu, L., Cheng, G., Zhu, J., Gan, L., et al. (2021). YTHDF2 inhibits gastric cancer cell growth by regulating FOXC2 signaling pathway. *Front. Genet.* 11, 592042. doi:10.3389/fgene.2020.592042

Shi, H., Wang, X., Lu, Z., Zhao, B. S., Ma, H., Hsu, P. J., et al. (2017). YTHDF3 facilitates translation and decay of N 6-methyladenosine-modified RNA. *Cell Res.* 27, 315–328. doi:10.1038/cr.2017.15

Shi, W., Tang, Y., Lu, J., Zhuang, Y., and Wang, J. (2022). MIR210HG promotes breast cancer progression by IGF2BP1 mediated m6A modification. *Cell Biosci.* 12, 38. doi:10.1186/s13578-022-00772-z

Shi, Y., Zheng, C., Jin, Y., Bao, B., Wang, D., Hou, K., et al. (2020). Reduced expression of METTL3 promotes metastasis of triple-negative breast cancer by m6A methylation-mediated COL3A1 up-regulation. *Front. Oncol.* 10, 1126. doi:10.3389/fonc.2020.01126

Sun, T., Wu, Z., Wang, X., Wang, Y., Hu, X., Qin, W., et al. (2020). LNC942 promoting METTL14-mediated m6A methylation in breast cancer cell proliferation and progression. *Oncogene* 39, 5358–5372. doi:10.1038/s41388-020-1338-9

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

Uddin, M. B., Wang, Z., and Yang, C. (2021). The m6A RNA methylation regulates oncogenic signaling pathways driving cell malignant transformation and carcinogenesis. *Mol. Cancer* 20, 61. doi:10.1186/s12943-021-01356-0

Wan, W., Ao, X., Chen, Q., Yu, Y., Ao, L., Xing, W., et al. (2022). METTL3/IGF2BP3 axis inhibits tumor immune surveillance by upregulating N6-methyladenosine modification of PD-L1 mRNA in breast cancer. *Mol. Cancer* 21, 60. doi:10.1186/s12943-021-01447-y

Wang, H., Xu, B., and Shi, J. (2020a). N6-methyladenosine METTL3 promotes the breast cancer progression via targeting Bcl-2. *Gene* 722, 144076. doi:10.1016/j.gene.2019.144076

Wang, J., Li, Y., Wang, P., Han, G., Zhang, T., Chang, J., et al. (2020b). Leukemogenic chromatin alterations promote AML leukemia stem cells via a KDM4C-ALKBH5-AXL signaling Axis. *Cell Stem Cell* 27, 81–97. e8. doi:10.1016/j.stem.2020.04.001

Wang, Y. J., Yang, B., Lai, Q., Shi, J. F., Peng, J. Y., Zhang, Y., et al. (2021). Reprogramming of m6A epitranscriptome is crucial for shaping of transcriptome and proteome in response to hypoxia. *RNA Biol.* 18, 131–143. doi:10.1080/15476286.2020.1804697

Wiener, D., and Schwartz, S. (2021). The epitranscriptome beyond m6A. *Nat. Rev. Genet.* 22, 119–131. doi:10.1038/s41576-020-00295-8

Xie, J., Ba, J., Zhang, M., Wan, Y., Jin, Z., and Yao, Y. (2021). The m6A methyltransferase METTL3 promotes the stemness and malignant progression of breast cancer by mediating m6A modification on SOX2. *J. B.U.ON.* 26, 444–449.

Xu, L., Zhang, J., Ma, Y., Yuan, Y., Yu, H., Wang, J., et al. (2022). MicroRNA-135 inhibits initiation of epithelial-mesenchymal transition in breast cancer by targeting ZNF217 and promoting m6A modification of NANOG. *Oncogene* 41, 1742–1751. doi:10.1038/s41388-022-02211-2

Xu, Y., Ye, S., Zhang, N., Zheng, S., Liu, H., Zhou, K., et al. (2020). The FTO/miR-181b-3p/ARL5B signaling pathway regulates cell migration and invasion in breast cancer. *Cancer Commun.* 40, 484–500. doi:10.1002/cac2.12075

Yankova, E., Blackaby, W., Albertella, M., Rak, J., De Braekeleer, E., Tsagkogeorga, G., et al. (2021). Small-molecule inhibition of METTL3 as a strategy against myeloid leukaemia. *Nature* 593, 597–601. doi:10.1038/s41586-021-03536-w

Yao, X., Li, W., Li, L., Li, M., Zhao, Y., Fang, D., et al. (2022). YTHDF1 upregulation mediates hypoxia-dependent breast cancer growth and metastasis through regulating PKM2 to affect glycolysis. *Cell Death Dis.* 13, 258. doi:10.1038/s41419-022-04711-1

Yau, C., Osdoit, M., van der Noordaa, M., Shad, S., Wei, J., de Croze, D., et al. (2022). Residual cancer burden after neoadjuvant chemotherapy and long-term survival outcomes in breast cancer: a multicentre pooled analysis of 5161 patients. *Lancet. Oncol.* 23, 149–160. doi:10.1016/S1470-2045(21)00589-1

Yi, D., Wang, R., Shi, X., Xu, L., Yilihamu, Y., and Sang, J. (2020). METTL14 promotes the migration and invasion of breast cancer cells by modulating N6–methyladenosine and hsa–miR–146a–5p expression. *Oncol. Rep.* 43, 1375–1386. doi:10.3892/or.2020.7515

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, 1394. doi:10.1038/s41467-021-21514-8

You, Y., Fu, Y., Huang, M., Shen, D., Zhao, B., Liu, H., et al. (2022). Recent advances of m6A demethylases inhibitors and their biological functions in human diseases. *Int. J. Mol. Sci.* 23, 5815. doi:10.3390/ijms23105815

Yuan, B., Liu, W., Huo, M., Zhang, J., Yang, Y., Gao, T., et al. (2022). The hub ten gene-based risk score system using RNA m6A methylation regulator features and tumor immune microenvironment in breast cancer. *Breast Cancer* 29, 645–658. doi:10.1007/s12282-022-01341-5

Zhang, C., Samanta, D., Lu, H., Bullen, J. W., Zhang, H., Chen, I., et al. (2016). Hypoxia induces the breast cancer stem cell phenotype by HIF-dependent and ALKBH5-mediated m^{6}A-demethylation of NANOG mRNA. *Proc. Natl. Acad. Sci. U. S. A.* 113, E2047–E2056. doi:10.1073/pnas.1602883113

Zhang, J., Shan, B., Lin, L., Dong, J., Sun, Q., Zhou, Q., et al. (2021). Corrigendum: Dissecting the role of N6-methylandenosine-related long non-coding RNAs signature in prognosis and immune microenvironment of breast cancer. *Front. Cell Dev. Biol.* 9, 812770. doi:10.3389/fcell.2021.812770

Zhang, S., Zhao, B. S., Zhou, A., Lin, K., Zheng, S., Lu, Z., et al. (2017). m6A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. *Cancer Cell* 31, 591–606. e6. doi:10.1016/j.ccell.2017.02.013

Zhang, X., Dai, X. Y., Qian, J. Y., Xu, F., Wang, Z. W., Xia, T., et al. (2022). SMC1A regulated by KIAA1429 in m6A-independent manner promotes EMT progress in breast cancer. *Mol. Ther. Nucleic Acids* 27, 133–146. doi:10.1016/j.omtn.2021.08.009

Zhao, C., Ling, X., Xia, Y., Yan, B., and Guan, Q. (2022). LncRNA UCA1 promotes SOX12 expression in breast cancer by regulating m6A modification of miR-375 by METTL14 through DNA methylation. *Cancer Gene Ther.* 29, 1043–1055. doi:10.1038/s41417-021-00390-w

Zhao, C., Ling, X., Xia, Y., Yan, B., and Guan, Q. (2021). The m6A methyltransferase METTL3 controls epithelial-mesenchymal transition, migration and invasion of breast cancer through the MALAT1/miR-26b/HMGA2 axis. *Cancer Cell Int.* 21, 441. doi:10.1186/s12935-021-02113-5

Zhong, S., Lin, Z., Chen, H., Mao, L., Feng, J., and Zhou, S. (2021). The m 6 A-related gene signature for predicting the prognosis of breast cancer. *PeerJ* 9, e11561. doi:10.7717/peerj.11561

Zhou, Z., Lv, J., Yu, H., Han, J., Yang, X., Feng, D., et al. (2020). Mechanism of RNA modification N6-methyladenosine in human cancer. *Mol. Cancer* 19, 104. doi:10.1186/s12943-020-01216-3

Zhu, P., He, F., Hou, Y., Tu, G., Li, Q., Jin, T., et al. (2021a). A novel hypoxic long noncoding RNA KB-1980E6.3 maintains breast cancer stem cell stemness via interacting with IGF2BP1 to facilitate c-Myc mRNA stability. *Oncogene* 40, 1609–1627. doi:10.1038/s41388-020-01638-9

Zhu, X., Feng, J., Zheng, M., Yang, Z., Zhao, L., Zhang, W., et al. (2021b). Metal–protein nanoparticles facilitate anti-VSV and H1N1 viruses through the coordinative actions on innate immune responses and METTL14. *Macromol. Biosci.* 21, 2000382. doi:10.1002/mabi.202000382

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 10 August 2022
DOI 10.3389/fgene.2022.935601

OPEN ACCESS

EDITED BY
Wei Li,
Xiangya Hospital, Central South University, China

REVIEWED BY
Shixiong Zhang,
Xidian University, China
Hongpan Zhang,
Affiliated Hospital of North Sichuan Medical College, China

*CORRESPONDENCE
Conghua Xie,
chxie_65@whu.edu.cn
Yan Gong,
yan.gong@whu.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 04 May 2022
ACCEPTED 15 July 2022
PUBLISHED 10 August 2022

CITATION
Zeng Z, Zhang J, Li J, Li Y, Huang Z, Han L, Xie C and Gong Y (2022), SETD2 regulates gene transcription patterns and is associated with radiosensitivity in lung adenocarcinoma. *Front. Genet.* 13:935601. doi: 10.3389/fgene.2022.935601

SETD2 regulates gene transcription patterns and is associated with radiosensitivity in lung adenocarcinoma

Zihang Zeng[1†], Jianguo Zhang[1†], Jiali Li[1], Yangyi Li[1], Zhengrong Huang[1,2], Linzhi Han[1], Conghua Xie[1,3,4*] and Yan Gong[2,5*]

[1]Department of Radiation and Medical Oncology, Zhongnan Hospital of Wuhan University, Wuhan, China, [2]Department of Biological Repositories, Zhongnan Hospital of Wuhan University, Wuhan, China, [3]Hubei Key Laboratory of Tumor Biological Behaviors, Zhongnan Hospital of Wuhan University, Wuhan, China, [4]Hubei Cancer Clinical Study Center, Zhongnan Hospital of Wuhan University, Wuhan, China, [5]Tumor Precision Diagnosis and Treatment Technology and Translational Medicine, Hubei Engineering Research Center, Zhongnan Hospital of Wuhan University, Wuhan, China

Lung adenocarcinoma (LUAD) has high morbidity and mortality worldwide, and its prognosis remains unsatisfactory. Identification of epigenetic biomarkers associated with radiosensitivity is beneficial for precision medicine in LUAD patients. SETD2 is important in repairing DNA double-strand breaks and maintaining chromatin integrity. Our studies established a comprehensive analysis pipeline, which identified SETD2 as a radiosensitivity signature. Multi-omics analysis revealed enhanced chromatin accessibility and gene transcription by SETD2. In both LUAD bulk RNA sequencing (RNA-seq) and single-cell RNA sequencing (scRNA-seq), we found that SETD2-associated positive transcription patterns were associated with DNA damage responses. SETD2 knockdown significantly upregulated tumor cell apoptosis, attenuated proliferation and migration of LUAD tumor cells, and enhanced radiosensitivity *in vitro*. Moreover, SETD2 was a favorably prognostic factor whose effects were antagonized by the m6A-related genes RBM15 and YTHDF3 in LUAD. In brief, SETD2 was a promising epigenetic biomarker in LUAD patients.

KEYWORDS

lung adenocarcinoma, radiosensitivity, SETD2, DNA damage response, multi-omics, prognosis, epigenetic

1 Introduction

Lung cancer is one of the main causes of cancer-related deaths worldwide (Ferlay et al., 2019; Sung et al., 2021). Approximately 85% of lung cancers are non-small cell lung cancers (NSCLC), of which around half are lung adenocarcinomas (LUAD) (Behrend et al., 2021). The prognosis of lung cancer is still unsatisfactory (Goldstraw et al., 2016). Radiotherapy has clear benefits for patients unsuitable for surgery, and is widely used in the radical and palliative treatment of LUAD patients (Ettinger et al., 2021). Radioresistance is a major cause of lesion

recurrence and metastasis (Pollom et al., 2016). In the era of precision medicine, research is gradually shifting from the population level to the individual level. Radiosensitivity is not only determined by tumor histology, but also affected by gene pathways (Césaire et al., 2022).

With the development of sequencing technology, a large amount of omics sequencing data have been documented (Hanash et al., 2002). Multi-omics data provided unparalleled dimensions of information and reflected the inherent features of individuals (Xu et al., 2021). Previous studies have successfully developed a genome-based model for adjusting radiotherapy dose (GARD) (Scott et al., 2017). Therefore, the identification of potential biomarkers is conducive to the development of precision radiotherapy.

Epigenetic regulation affects tumor heterogeneity, and is involved in radiosensitivity (Peng et al., 2021). On the post-transcriptional level, N6-methyladenosine (m6A) modification transferase METTL3 increases radioresistance *via* promoting the stability of target RNAs in multiple cancers (Huang et al., 2021). For histone modifications, histone deacetylase (HDAC) inhibitors demonstrate radiosensitization of various cancers in preclinical studies *via* targeting DNA damage responses (DDR) (Shirbhate et al., 2020). Moreover, Xue et al. (2015); Wu et al. (2020) found upregulated DNA methyl-transferase DNMT3B in radioresistant nasopharyngeal and prostate cancer cells. However, the impact of epigenetics on radiosensitivity is still not well understood, and the identification of novel epigenetic markers has a substantial clinical interest.

SETD2 is the sole transferase of histone H3 trimethylation on lysine 36 (H3K36me3) in humans. SETD2 is involved in DNA repair and maintaining chromatin integrity (Carvalho et al., 2014; Pfister et al., 2014). SETD2 is necessary to recruit DDR factors 53BP1 and RAD51 (Carvalho et al., 2014). In the Mayo cohort, renal carcinoma patients without H3K36me3 had worse cancer-specific survival (Ho et al., 2016). Moreover, SETD2 mutation promotes MLL-AF9-induced leukemia progression and chemoresistance (Mar et al., 2017). Our previous studies found that SETD2 knockdown triggers DNA double-strand breaks (DSB) and activates the cGAS-STING pathway (Zeng et al., Forthcoming 2022). On the other hand, SETD2-mediated H3K36me3 guides m6A modifications on nascent RNA transcripts (Kumari and Muthusamy, 2020). However, studies on the roles of SETD2 in LUAD are still lacking. Therapeutical values and possible mechanisms of SETD2 remain to be investigated in LUAD.

Here, we utilized comprehensive omics-data analysis to determine that SETD2 was a key radiosensitivity-related signature. Our results indicated that SETD2 enhanced chromatin opening and transcription, especially in the DDR-related pathways. *In vitro* experiments indicated that SETD2 knockdown upregulated tumor cell apoptosis, attenuated proliferation and migration of LUAD cells, and enhanced their radiosensitivity. Furthermore, SETD2 was a prognostic protection factor whose effect interacted with m6A-related genes. Our finding suggested SETD2 as a potential epigenetic marker in LUAD patients.

2 Materials and methods

2.1 Collection and processing of omics data

A total of 11 datasets were included in this study (Supplementary Table S1). The survival fraction at 2 Gy (SF2) was a common index to describe cellular radiosensitivity. In this study, the determination of SF2 was based on previous studies, which reported colony formation with irradiation (Supplementary Table S2) (Torres-Roca et al., 2005; Eschrich et al., 2009; Gao et al., 2012; Oleinick et al., 2016; Zhong et al., 2016). For Microarray data, GEO datasets were standardized by the default method. When multiple probes corresponded to the same gene, the maximum value of the probes was selected. GSE20549 contained 42 samples of H460 and H1299 cells at six time points (0, 2, 4, 8, 12, and 24 h) after 2 Gy ionizing radiation (IR). We collected 10 NSCLC cell samples from GSE32036 and 16 samples from GSE57083 (Byers et al., 2013), which were normalized and integrated with Z-score. GSE5949 contained 59 pan-cancer cell samples (Reinhold et al., 2010). We collected SETD2 expression data for survival analysis in GSE50081 (Der et al., 2014) and GSE3141 (Bild et al., 2006). For RNA-seq data, we collected RNA-seq (standardized by RPKM) of the 16 HepG2 cell samples treated with shSETD2 from GSE121949 (Huang et al., 2019). The clinical data of The Cancer Genome Atlas (TCGA) cohorts were downloaded by R TCGAbiolinks package (Colaprico et al., 2016). Xena was used to obtain omics data of TCGA (Goldman et al., 2020). The RNA-seq data were normalized by log2 (TPM +1). ChIP-seq data was annotated by R ChIPseeker package (Yu et al., 2015). H3K36me3 ChIP-seq of HepG2 cells with/without shSETD2 was obtained from GSE110318 (Huang et al., 2019). The lung cancer ChIP-seq data (H3K36me3, H3K27me3, H3K9me3, H3K27ac, H3K4me3, and H3K4me1) were downloaded from Roadmap (ID: EN96) (Kundaje et al., 2015). Moreover, we collected scRNA-seq for 43,704 cells from tumor tissues of 11 LUAD patients in GSE131907 (Kim et al., 2020). ATAC-seq data of TCGA was obtained from NCI GDC (https://gdc.cancer.gov/about-data/publications/ATACseq-AWG). The R ChIPseeker package was also used for annotation (Yu et al., 2015). To compare the expression of SETD2 in different NSCLC cell lines, we collected SETD2 RNA-seq in the Cancer Cell Line Encyclopedia (CCLE) database (Barretina et al., 2012).

2.2 Cell culture and radiation

The Type Culture Center of the Chinese Academy of Sciences (Shanghai, China) provided the LUAD A549 and H1299 cells, cultivated in RPMI-1640 media (HyClone, United States) containing 10% fetal bovine serum. The cells were grown in a standard tissue culture incubator at 37°C, with 95% humidity and 5% CO_2. Radiation was conducted using a small animal radiation research platform (6 Gy, PXI X-RAD 225Cx, Gulmay, CT, United States).

2.3 Cell transfection

Small interfering RNAs (siRNAs) and negative control (NC) were transfected at 20 nM *via* jetPRIME® transfection reagent. SETD2 siRNA (siSETD2) 2 sequences were as follows: sense, CCUUCAGGCUCAGAGUUAATT, and anti-sense, UUAACU CUGAGCCUGAAGGTT; siSETD2 3 sequences were as follows: sense, CCGGAAACCUGACUGCAAATT, and anti-sense, UUU GCAGUCAGGUUUCCGGTT.

2.4 RNA isolation and quantitative real-time PCR

Using the TRIzol reagent, total RNA was isolated from cells (Vazyme, China). We used HiScript® Q RT SuperMix (Vazyme, China) to transcribe RNA and ChamQTM SYBR® qPCR Master Mix (Vazyme, China) for qRT-PCR. The relative mRNA levels were calculated with the $2^{-\Delta\Delta Ct}$ method. All experiments were performed in triplicates.

2.5 Wound healing and colony formation assays

For wound healing assays, we seeded the transfected cells into 6-well plates. A straight line was scratched with a pipette tip. The migration rate was calculated using the following formula: wound closure rate (%) = (area of initial scratch—the area of final imaged cell-free area)/area of initial scratch * 100. For colony formation assays, we subjected the transfected cells to radiotherapy, and they were seeded into 6-well plates at 1,000 cells/well 48 h later. After 2 weeks, the medium was aspirated, and 4% paraformaldehyde was added and fixed for 30 min. Then after PBS washing, they were stained with 0/5% crystal violet for 30 min and finally washed with water, dried, and photographed.

2.6 Flow cytometry for cell apoptosis

After 48 h, the treated cells were collected and washed twice with PBS. We suspended the cells in binding buffer with Annexin V-FITC staining solution and propidium iodide (PI) solution on ice. The samples were detected by flow cytometry (Beckman, China).

2.7 Immunohistochemistry from the human protein atlas database

We collected SETD2 immunohistochemistry images from the HPA database (https://www.proteinatlas.org/) (Karlsson and Zhang, 2021), including available 5 LUAD and 3 normal lung tissues. All images were made of antibody HPA042451. SETD2 staining score was calculated as intensity times quantity. The intensity score consisted of 0 (Negative), 1 (Weak), 2 (Moderate), and 3 (Strong). The quantity score consisted of 0 (None), 1 (<25% cells), 2 (25–75% cells), and 3 (>75% cells).

2.8 Analysis of single-cell RNA sequencing

The Seurat workflow was adopted to analyze scRNA-seq data (Satija et al., 2015). Cells with less than 200 genes (min.features = 200) and genes with less than 3 cells (min.cells = 3) were screened out. Only the cells with less than 15% of mitochondrial genes were retained. A total of 2,000 hypervariable genes were selected with the vst method. An Elbow diagram was drawn to select the best number of principal components. The resolution parameter was set as 0.5. Uniform manifold approximation and projection were used to visualize single-cell atlas (McInnes et al., 2018), which was realized by Seurat DimPlot and FeaturePlot functions. Cell types were identified using marker genes from the previous study (Lambrechts et al., 2018). Specifically, tumor cell markers were EPCAM and KRT19; T/NK cell markers were NKG7, CD3E, CD3G, and CD3D; B cell markers were CD79A and CD79B; myeloid cell marker was LYZ; mast cell markers were TPSB2 and TPSAB1; fibroblast markers were COL1A1 and COL1A2; endothelial cell marker was CLDN5; normal epithelial cell marker was CAPS. GSVA was used to calculate the gene set enrichment score of individual cells (Hanzelmann et al., 2013).

2.9 Principal component analysis

PCA was a classic linearly dimensionality reduction algorithm. We used the R FactoMineR package to perform PCA (Lê et al., 2008). The first principal component was considered to be the vector with the largest variance. In this study, since gene clusters contained a large number of genes, we used the first principal component as the eigenvalue to characterize the gene clusters.

2.10 Short time-series expression miner analysis

Short time-series expression miner analysis was an algorithm to cluster, compare and visualize time-course gene expression (Ernst and Bar-Joseph, 2006). Genes with similar time expression patterns were grouped into the same clusters. We extracted eigenvalue of time-course clusters using PCA. Next, we performed Spearman's correlation between the eigenvalue of gene clusters and SF2 to recognize SF2-related clusters.

2.11 Random forest

The RF was the ensemble methods with multiple decision trees. We used R randomForest packages to train RF model for SF2 fitting (Breiman, 2001). The mean-square error was used to calculate importance of genes.

2.12 Single-gene liner quadratic model

The Linear-Quadrac (LQ) model proposed by Kellerer and Rossi was a classical model widely used in the field of radiotherapy (Kellerer and Rossi, 1978). The LQ model estimated the survival fraction (SF) of cells exposed to radiation:

$$SF = e^{-\alpha \times D - \beta \times D^2}$$

Which e represents the natural logarithm, α and β represent radiation-specific parameters describing the radiosensitivity of tumor cells, D is the radiation dose.

However, the equation for fitting SF using gene expression was still unclear. We next analyzed a gene expression data set exposed to different doses (0, 2, 5, 6, and 7 Gy) of radiation (GSE102971, n = 100) (Park et al., 2017). Compared with quadratic equation and cubic equation, linear equation (dose-gene expression) has the smallest Akaike information criterion in the analysis of each gene, suggesting that gene expression was linearly related to radiation dose (Supplementary Figure S1). In order to establish a simulation model of gene expression and SF, we constructed a single-gene linear-quadratic (SGLQ) model inspired by the LQ model:

$$SF = e^{\alpha \times Ei - \beta \times Ei^2}$$

Here, α is the linear radiosensitivity parameter of a single gene i, and β represents the quadratic radiosensitivity parameter of gene i. Ei is the expression value of gene i. The SGLQ contributed to modeling the relationship between gene expression and SF in the era of omics.

2.13 The 4-omics system biological network

Gene regulations were the complex systems biology networks. Analysis of nodes in the network helped to identify key genes. The 4-omics system biological networks consisted of mutation, copy number alteration (CNA), mRNA co-expression and protein interaction sub-networks. The gene interaction of protein interaction sub-network was formed by STRING database (Szklarczyk et al., 2017). The mutation and CNA sub-networks were constructed by HotNet diffusion-oriented subnetworks (HotNet2) algorithm, which was based on random walk with restart (Leiserson et al., 2015). The HotNet2 included not only the topology of gene interaction networks from STRING, but also the heat values. Here, mutation frequency and copy number were set as heat values of the mutation and CNA sub-networks. Finally, HotNet2 identified sub-networks with close topology structure and high overall thermal diffusivity. The mRNA co-expression sub-network was formed by weight gene co-expression analysis (WGCNA) (Langfelder and Horvath, 2008). In this study, we used multi-omics data in TCGA pan-cancer cohorts to build the 4-omics system biological networks. Network analysis and visualization were realized by Cytoscape (Shannon et al., 2003). The innovation of this network was the inclusion of multi-omics data.

2.14 The maSigPro algorithm

Differential expression analysis of time-course transcriptome was performed by maSigPro using a 2-step regression strategy (Conesa et al., 2006). Since GSE121949 contained gene expression data at 4 time points (0, 1, 3, and 6 h) (Huang et al., 2019), we constructed the cubic equation in the maSigPro algorithm to identify treatment group related genes.

2.15 Binding and expression target analysis

BETA was a tool to integrate ChIP-seq and gene differential expression list from transcriptome (Wang et al., 2013). In this study, we explored the transcriptional activation or inhibition of H3K36me3, and identified the motif of H3K36me3 and its collaborators by combining H3K36me3 ChIP-seq (GSE110318) and RNA-seq (GSE121949) *via* BETA (Huang et al., 2019). BETA was realized by cistrome (http://cistrome.org/ap/root) (Liu et al., 2011).

2.16 Enrichment analysis

Gene set enrichment analysis (GSEA) was used to identify GO terms that were activated or inhibited in a predefined list of gene differential expression *via* permutation test (Ashburner et al., 2000). Over-representation analysis (ORA) was performed to identify GO terms associated with a predefined gene set *via* a hypergeometric test. GSEA and ORA were realized by R clusterProfiler packages (Yu et al., 2012).

2.17 Weighted gene co-expression network analysis

WGCNA clustered genes into different modules according to expression similarity through kmeans clustering and dynamic branch cutting (Langfelder and Horvath, 2008). In

this study, we correlated the eigenvalues of the modules with the SETD2 expression to identify SETD2-related modules.

2.18 Statistical analysis

Most statistical analysis was analyzed in R software 4.1.0. The basic statistical analysis was performed by the R stats package. Cox proportional hazards regression was realized by the R survival package. When the survival curve crossed, landmark analysis was used to assess the prognostic value of SETD2 at different time periods (Der et al., 2014). The landmark analysis was realized by R jskm package (https://rdrr.io/cran/jskm/). Gene expression plots for TCGA data were implemented by GEPIA (Tang et al., 2017a) and TIMER (Li et al., 2017). p values less than 0.05 were considered statistically significant. All the p values were two-sided.

3 Results

3.1 Comprehensive analysis suggested a critical role of SETD2 in radiosensitivity

3.1.1 Identification of radiosensitivity related transcriptome patterns

Gene expression induced by temporal changes in radiation may be related to radiotherapy response and sensitivity. We collected 42 NSCLC cell samples from GSE20549 (Clough and Barrett, 2016) with six time points (0, 2, 4, 8, 12, and 24 h) after 2 Gy IR. The ANOVA identified 3,337 genes variously expressed at different time points ($p < 0.05$). We next explored IR time-dependent gene patterns using Short Time-series Expression Miner method (Ernst and Bar-Joseph, 2006). A total of 11 time-course gene clusters reached statistical significance (false discovery rate, FDR q < 0.01, Supplementary Figure S2).

To determine the radiosensitivity-related time-course clusters, we collected 26 untreated NSCLC cell samples from GSE32036 and GSE57083 (Byers et al., 2013), whose SF2 were provided by colony formation assays from previous studies (Supplementary Table S2, see Methods). There were 5 first principal component of clusters correlated with SF2 (cor >0.1, Figures 1A,B), which were considered as SF2-related clusters. Gene ontology (GO) enrichment analysis suggested that the 832 genes in these 5 clusters were linked to cell cycle, DDR and histone methylation (all, FDR $p < 0.05$, Figure 1C).

3.1.2 SETD2 was a radiosensitivity signature at the single-gene scale

Since the above analysis was performed at the gene cluster level, we subsequently determined the SF2-related signatures at the single-gene level. Here we used the RF (Breiman, 2001) and SGLQ model to quantitate the importance of a single gene on SF2 in GSE32036 and GSE57083. A total of 289 (34.7%) genes did not perform well both in the RF and SGLQ models (Figures 1D,E), suggesting that these genes were unlikely to be SF2-related signatures.

In the remaining 543 (65.3%) genes, we next identified the hub ones. We constructed the "4-level network" (see Methods), containing the mutation, CNA, RNA, and protein subnetworks from TCGA pan-cancer data (Figure 1F). The 3 subnetworks were identified in the 4-level biological network, representing histone methylation, cell cycle, and DNA damage checkpoint (Supplementary Figure S3). Considering the topological structure of networks, the 280 genes were selected as irradiation-related genes (Figure 1G) with a high degree, betweenness, and closeness centrality (all, > median value). The significant enrichment of histone H3K36 methylation, cell cycle, DDR were observed in these 280 genes (Figure 1H), containing ATM, ERCC4, H2AFX, DTX3L, CCNA2, RAD9A, POLE3, BRSK1, CLOCK, CNOT3, CNOT4, CNOT6, BRD4, CAMK2A, EZH2, RB1, ACTR1A, AURKB, RBX1, NSD1.

We next validated the relationship between H3K36 methylation regulatory genes (SETD2, SETD3, NSD1, PAXIP1, BRD4, IWS1, SETMAR, SMYD2, ASH1L) and SF2 in the dataset GSE5949 (pan-cancer cell lines, $n = 59$) (Reinhold et al., 2010). SETD2 was the only gene linked to SF2 *via* Spearman correlation analysis ($p = 0.04$, Figure 1I).

3.2 SETD2 enhanced transcription and chromosomal accessibility

SETD2 was the main methyltransferase that specifically trimethylated "Lys-36" of histone H3 in mammals. The H3K36me3 signals decreased in the whole genome after SETD2 knockdown in GSE110318 (Huang et al., 2019) (Figure 2A), which affected both H3K36me3 coverage and average peak signals (Figure 2B).

GSE121949 (Huang et al., 2019) provided RNA-seq of HepG2 cells with or without shSETD2 at 4 time points (0, 1, 3, and 6 h, Figure 2C). We performed gene differential expression analysis between the shSETD2 and control groups using the maSigPro algorithm (Conesa et al., 2006) (Figure 2D). Enrichment analysis showed inhibited transcriptiona in the shSETD2 groups (Figure 2E). Compare with the control group, more genes were downregulated in the shSETD2 group at all time points (Figure 2F). Furthermore, binding and expression target analysis (BETA) (Wang et al., 2013) of integrating H3K36me3 ChIP-seq (GSE110318) and RNA-seq (GSE121949) demonstrated that H3K36me3 enhanced transcription (Figure 2G). Motif analysis of H3K36me3 ChIP-seq suggested that H3K36me3 regulated transcription factors (NFIC, ELF2, ELF4, EHF, Figure 2H).

Open chromatin facilitates transcription. Next, we investigated the relation between SETD2/H3K36me3 and chromosomal accessibility. We collected ChIP-seq data for the six histone modifications (H3K36me3, H3K27me3, H3K9me3, H3K27ac,

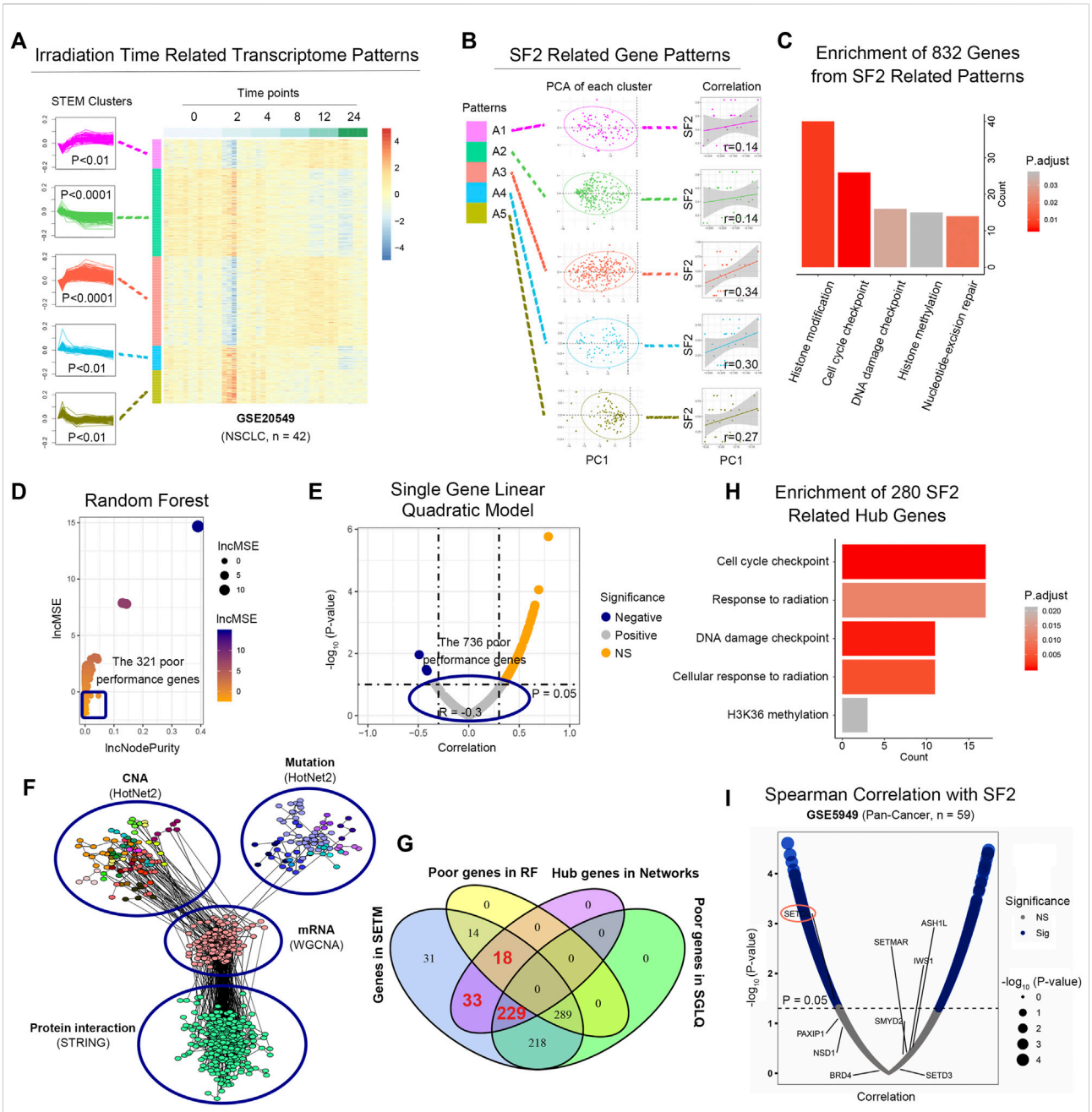


FIGURE 1
SETD2 functions as a radiosensitivity signature. **(A)** Radiosensitivity-related time-course clusters by the short time-series expression miner method. **(B)** Correlation between SF2 and the first principal component (PC1) of each time-course cluster. **(C)** GO enrichment analysis of the 832 genes in 5 SF2-related clusters. **(D)** RF analysis of single gene on SF2 in GSE32036 and GSE57083. **(E)** SGLQ analysis of single gene on SF2 in GSE32036 and GSE57083. **(F)** The 4-omics biological networks included mutation, CNA, mRNA co-expression, and protein interaction sub-networks. **(G)** Venn diagram for screening key genes. **(H)** GO enrichment analysis of the 280 hub genes. **(I)** Correlation between gene expression and SF2.

H3K4me3, and H3K4me1) of lung cancer sample from Roadmap (Sample ID: EN96) (Kundaje et al., 2015). Correlation analysis and PCA showed that H3K36me3 had the similar patterns to open-chromosome-related histone modifications (H3K27ac, H3K4me3), but was distant from closed histone modifications (H3K27me3, H3K9me3, Figures 3A,B). Figure 3C showed a specific example of peak distributions in chromosome 17: 1-6850845. Furthermore, ATAC-seq of TCGA cohorts demonstrated that expression and promoter methylation of SETD2 were associated with chromosomal accessibility in TCGA LUAD (Figures 3D,E). The combined analysis

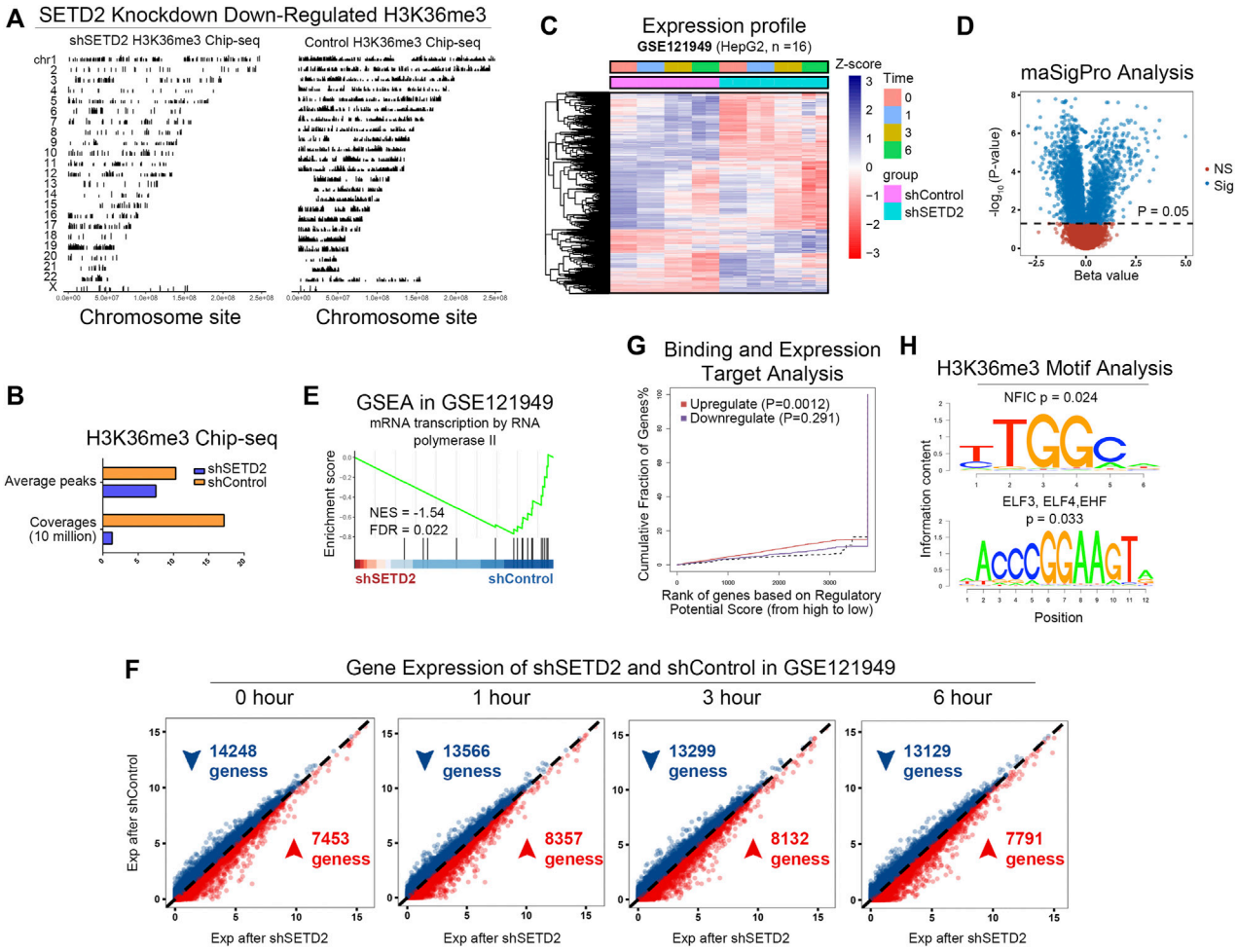


FIGURE 2
SETD2 and H3K36me3 enhanced transcription. **(A)** The H3K36me3 signals of shSETD2 vs. control in GSE110318. **(B)** The quantitative H3K36me3 signals. **(C)** Expression profile of GSE121949. **(D)** Gene different expression analysis between shSETD2 and control groups using maSigPro algorithm. **(E)** GSEA of shSETD2 vs. control in GSE121949. **(F)** Gene expression of shSETD2 vs. control in GSE121949. **(G)** Binding and expression target analysis (BETA) of H3K36me3 Chip-seq (GSE110318) and RNA-seq (GSE121949). **(H)** Motif analysis of H3K36me3 Chip-seq.

of ATAC-seq and RNA-seq showed that the expression of SETD2 was positively correlated with the openness of the promoter region of the extensive genes in TCGA NSCLC (n = 76, Figures 3F,G). These results indicated that SETD2 enhanced transcription and chromosomal accessibility.

3.3 SETD2 positively regulated transcriptional patterns associated with DNA damage responses

The next question was whether the enhanced transcription by SETD2 was gene-specific. We analyzed SETD2-related co-expression genes in bulk RNA-seq. In the TCGA LUAD cohort, WGCNA (Langfelder and Horvath, 2008) identified 35 co-expression modules (Figure 4A, Supplementary Figures S4, S5). Correlation analysis of module eigenvalues with SETD2 expression revealed the 4 SETD2 positive correlation modules (Turquoise, Green, Midnightgreen, and Blue modules, Figure 4B). GO enrichment analysis showed that genes of the above 4 modules enriched in DDR, DNA repair, RNA splicing, and histone modification signals (Figure 4C).

In the TCGA lung squamous cell carcinoma (LUSC) cohort, we repeated the WGCNA (Figure 4D, Supplementary Figures S6, S7). Similarly, we identified 50 modules, the 5 of which were positively correlated with SETD2 (Turquoise, Brown, Blue, Darkgreen, and Royalblue Figure 4E). DNA damage repair, cell cycle, RNA splicing, and histone modification signals were enriched in genes of these 5 modules (Figure 4F).

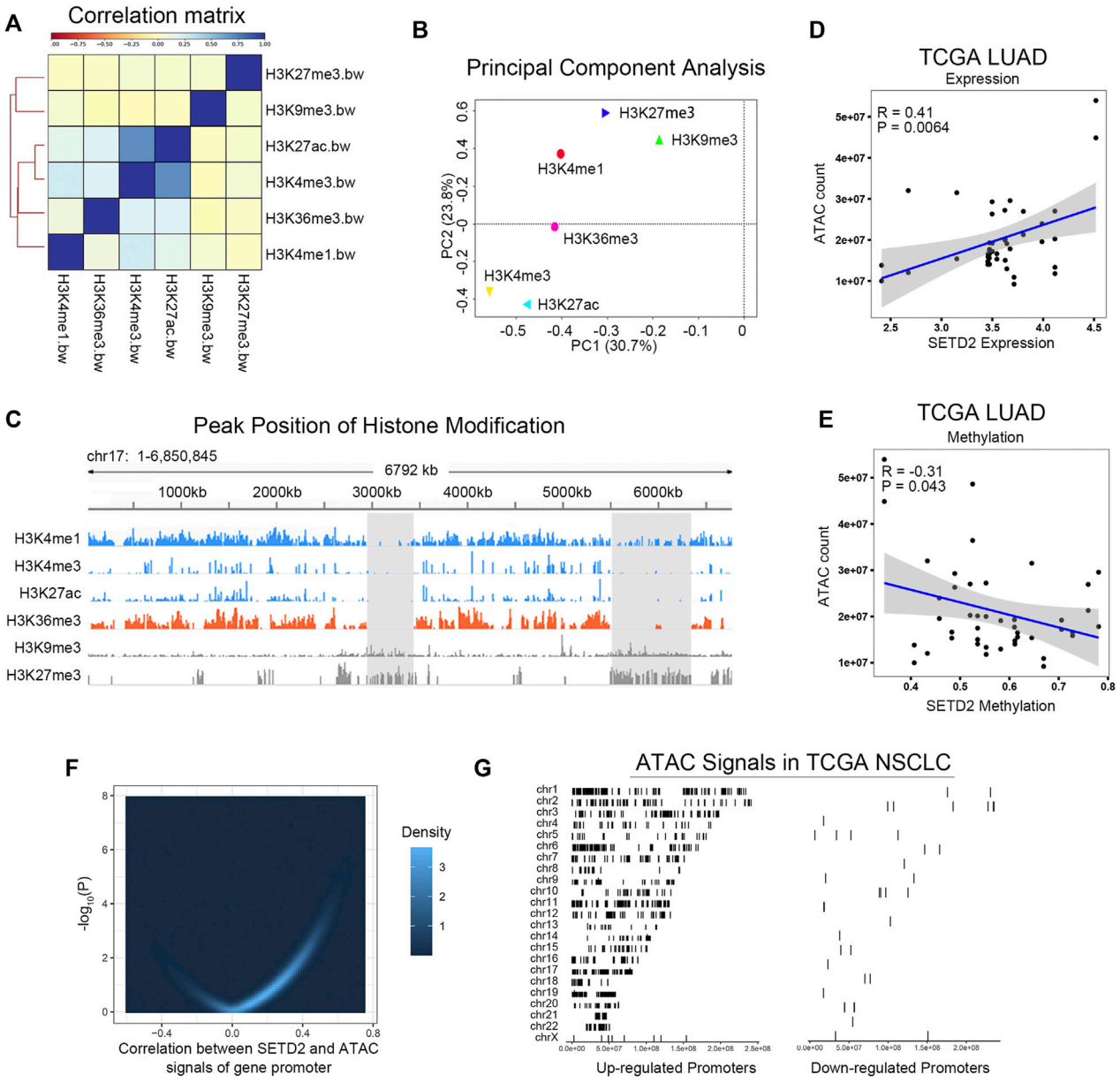


FIGURE 3
SETD2 and H3K36me3 enhanced chromosomal accessibility. **(A)** Correlation analysis of H3K36me3 and other histone modifications. **(B)** PCA of H3K36me3 and other histone modifications. **(C)** Peak distributions in chromosome 17: 1-6850845 of H3K36me3 and other histone modifications. **(D,E)** Correlation between SETD2 and ATAC signals in TCGA LUAD. **(F)** Correlation between SETD2 expression and the accessibility of the whole gene promoters. **(G)** The left side shows the location of the ATAC-seq signal of genes whose promoter region accessibility is positively correlated with SETD2 on the chromosome; The right side shows the position of the ATAC-seq signal of the negatively related gene on the chromosome.

3.4 Single-cell analysis validated the co-expression patterns of SETD2 in lung adenocarcinoma

We next investigated 43,704 single-cell RNA profiles of the 11 primary LUAD patients in scRNA-seq GSE131907 (Kim et al., 2020). Through Seurat workflows, we identified the 8 cell types: tumor cells, fibroblasts, endothelial cells, epithelial cells, T/NK cells, B cells, myeloid cells, and mast cells (Figure 5A). SETD2 was widely distributed in different cell types (Figure 5B, Supplementary Figure S8). Overall, The SETD2 positive rate in tumor cells was lower than that in normal epithelial cells (15.2 vs. 20.2%, $p = 0.2$, Figure 5C), but higher than that in immune cells, including T/NK cells (10.7%,

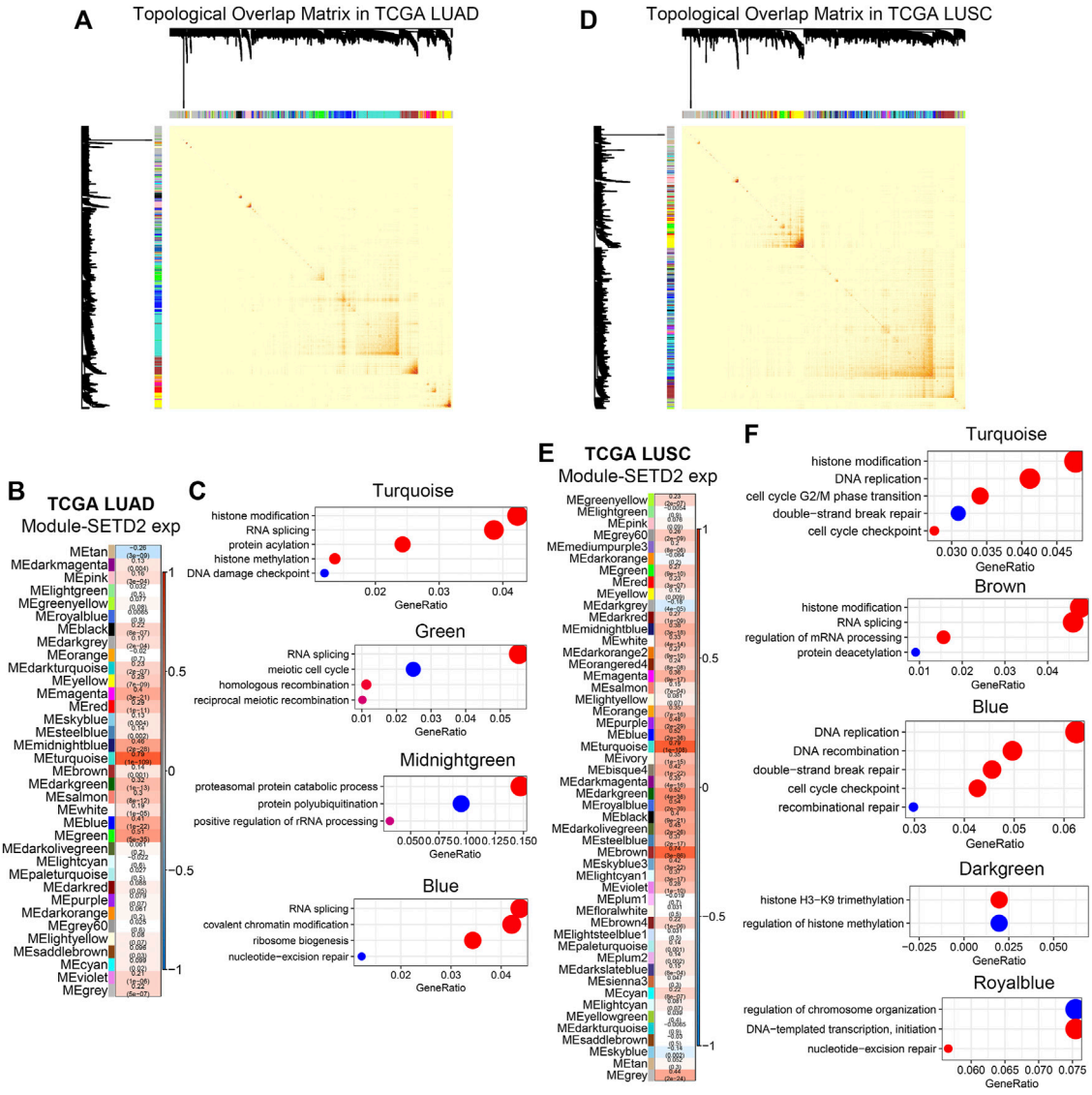


FIGURE 4
SETD2 regulated transcriptional patterns in bulk RNA-seq. **(A)** Visualization of topological overlap matrix in TCGA LUAD. **(B)** WGCNA revealed the gene clusters related to SETD2 in TCGA LUAD. **(C)** GO enrichment analysis of SETD2-related clusters in TCGA LUAD. **(D)** Visualization of topological overlap matrix in TCGA LUSC. **(E)** WGCNA revealed the gene clusters related to SETD2 in TCGA LUSC. **(F)** GO enrichment analysis of SETD2-related clusters in TCGA LUSC.

$p < 0.0001$), B cells (8.4%, $p < 0.0001$), and myeloid cells (12.1%, $p < 0.0001$). Moreover, in log2 (TPM+1) normalized profiles, we found that SETD2 was positively associated with the expression of more genes (Figure 5D), especially in tumor cells (positive rate: 78.7%) and myeloid cells (positive rate: 73.7%), while the opposite was observed in mast cells (positive rate: 46.4%) and epithelial cells (positive rate: 44.1%). This finding was consistent with Section 3.2. Next, we compared the gene expression of SETD2-positive and negative tumor cells. Genes highly expressed in SETD2-positive tumor cells were enriched in DDR, RNA splicing, and histone modification signals (Figure 5E).

3.5 Knockdown of SETD2 upregulated apoptosis, attenuated proliferation and migration of tumor cells, and enhanced the radiosensitivity in lung adenocarcinoma

SETD2 was downregulated in TCGA LUAD and LUSC tissues than normal ones (Supplementary Figure S9). In immunohistochemistry of 5 LUAD and 3 normal lung tissues from the HPA database (Karlsson and Zhang, 2021), we compared SETD2 staining scores of tumor and alveolar cells. The results

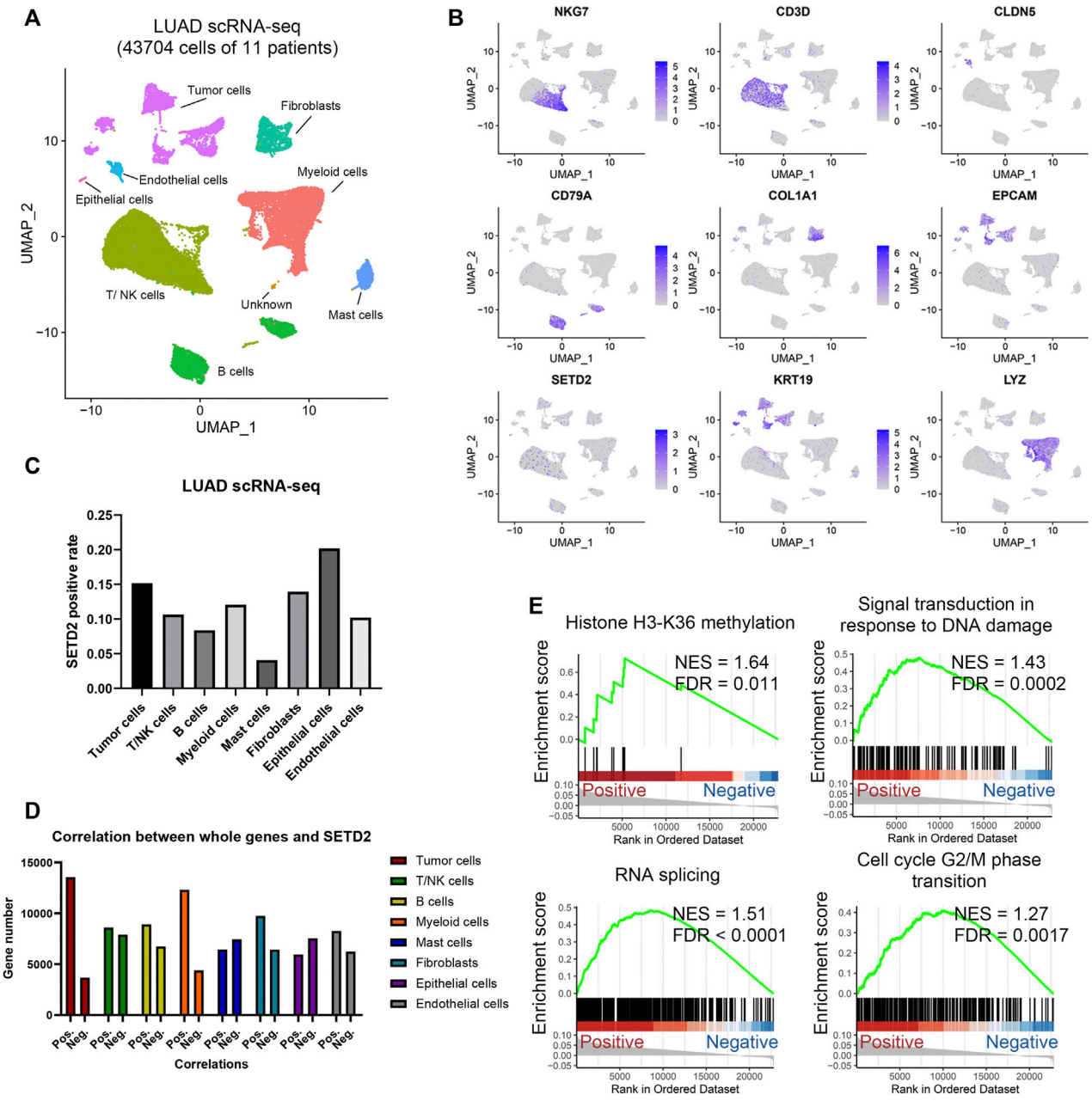


FIGURE 5
Analysis of SETD2 in LUAD scRNA-seq. **(A)** UMAP of the 43,704 cells in LUAD scRNA-seq. **(B)** Marker gene expressions in scRNA-seq. **(C)** Positive rate of SETD2 for each cell type. **(D)** The number of genes positively or negatively correlated with SETD2 expression. **(E)** GSEA of positive SETD2 vs. negative cells in scRNA-seq.

showed that SETD2 staining scores were low but not significantly different between tumor and alveolar cells, and that the staining was mainly concentrated in the nuclear (Supplementary Figures S10A,B).

We next investigated the effects of SETD2 on tumor malignant behaviors and radiosensitivity *in vitro*. According to our preliminary studies (Zeng et al., Forthcoming 2022), the expression levels of SETD2 in H1975 and A549 cells were high. In this study, we collected RNA-seq for NSCLC cells in the CCLE dataset. SETD2 expression remained higher in H1975 and A549 cells than H1299, PC9, and H460 cells (Supplementary Figure S11).

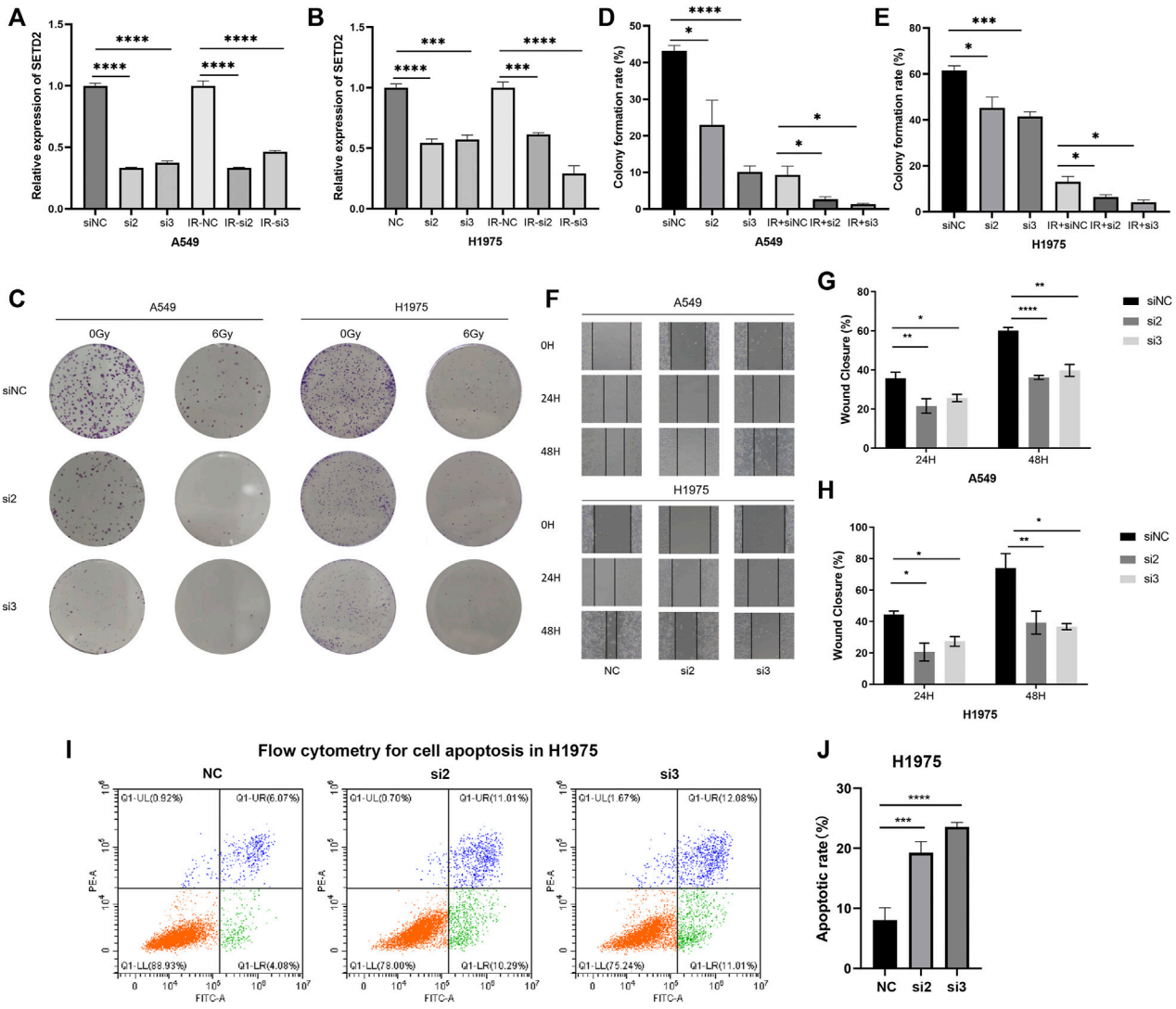


FIGURE 6
Flow cytometry, wound healing assays and clone formation experiments treated with siSETD2 or radiation. **(A)** The qRT-PCR of SETD2 in A549 cells. **(B)** The qRT-PCR of SETD2 in H1975 cells. **(C)** The six well plate image of clone formation experiment. **(D)** Clone formation rate of A549 cells. **(E)** Clone formation rate of H1975 cells. **(F)** Wound healing assay image. **(G)** Wound closure rate of A549 cells. **(H)** Wound closure rate of H1975 cells. **(I)** Flow cytometry for cell apoptosis in H1975. **(J)** Quantitative results of apoptosis.

Therefore, we cultured LUAD A549 and H1975 cells and divided them into six groups: negative control, siNC; siSETD2-2, si2; siSETD2-3, si3; negative control plus 6 Gy IR, IR-NC; siSETD2-2 plus 6 Gy IR, IR-s2; siSETD2-3 plus 6 Gy IR, IR-s3. With or without IR, siSETD2 showed high knockdown efficiency in A549 and H1975 (Figures 6A,B). Colony formation assays indicated attenuated tumor proliferation after SETD2 knockdown (Figures 6C–E). Cell proliferation was diminished after 6 Gy IR, and cells were more sensitive to radiation upon siSETD2 treatment. Moreover, SETD2 knockdown decreased cell migration (Figures 6F–H). Due to the severe killing of tumor cells by IR after transfection with siSETD2, we did not perform wound healing assays in the IR groups. Next, we performed flow cytometry for cell apoptosis in H1975 cells. SETD2 knockdown significantly upregulated LUAD cell apoptosis (Figures 6I,J).

3.6 SETD2 interacted with N6-methyladenosine-related genes RBM15 & YTHDF3 statistically and was associated with a favorable prognosis

Previous studies reported the possible association between SETD2 and m6A (Kumari and Muthusamy, 2020). We

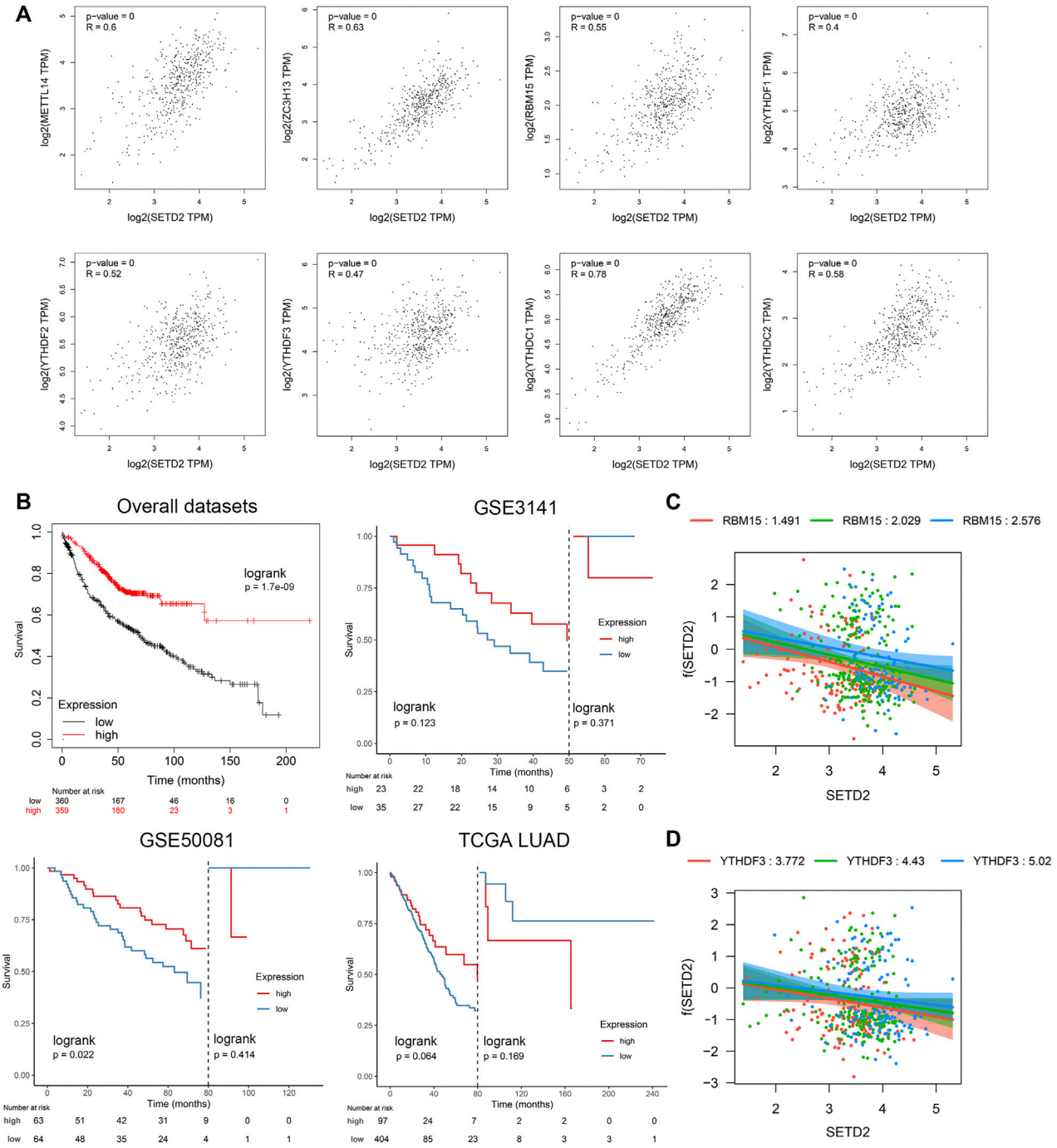


FIGURE 7
SETD2 was linked to favorable prognosis whose effect was negatively affected by the interaction of m6A-related genes RBM15 and YTHDF3. **(A)** Correlations between SETD2 and m6A-related genes. **(B)** Kaplan-Meier survival curves of SETD2 in the Kmplot database. **(C)** Visualization of interaction effect of RBM15 on Cox regression coefficient of SETD2. **(D)** Visualization of interaction effect of YTHDF3 on Cox regression coefficient of SETD2.

comprehensively analyzed m6A-related genes, including 8 "writer" genes, 9 "reader" genes, and 2 "eraser" genes (Gu et al., 2020) in TCGA LUAD dataset. SETD2 was positively related to "writer" and "reader" genes ($r > 0.4$ & $p < 0.01$): METTL14, ZC3H13, RBM15, YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2 (Figure 7A, Supplementary Figure

S12). SETD2 was linked to a favorable prognosis in multiple LUAD datasets (Figure 7B). However, the landmark analysis showed that patients with high SETD2 expression changed from favorable prognosis to unfavorable prognosis after more than 50–80 months. We next investigated the interaction effects of m6A-related genes with SETD2 on prognosis. Using multivariate Cox regression with interaction terms (Survival ~ SETD2 + SETD2 *m6A gene + age + gender + stage), we identified that RBM15 (interaction term HR = 1.15, p = 0.02) and YTHDF3 (interaction term HR = 1.06, p = 0.18) interacted with SETD2 (Supplementary Tables S3, S4). The protective effects of SETD2 on prognosis were enhanced with the reduction of RBM15 or YTHDF3 (Figures 7C,D). The prognostic effects of SETD2 may be explained with low microsatellite instability and frequency of mutations (Liu et al., 2015; Zeng et al., Forthcoming 2022).

4 Discussion

Our work demonstrated that SETD2 as a radiosensitivity signature positively regulated DDR-related transcriptional patterns. Possibly due to inhibition of DDR, SETD2 knockdown upregulated the radiosensitivity of LUAD cells. Clinically, SETD2 was a promising epigenetic biomarker for prognosis and radiotherapy in LUAD.

Previous studies identified radiosensitivity-related genes based on regression. Torres-Roca et al. (2015) identified 10 genes associated with SF2 by linear regression models (Eschrich et al., 2009). Tang et al. (2017b) identified 65 radiosensitivity-related genes by logistic regression models in soft tissue sarcoma. In this study, we implemented a novel gene identification pipeline. Through pattern analysis of radiation time-associated transcriptomes, machine learning algorithms, and 4-omics networks, we successfully identified SETD2 as a key gene for radiosensitivity, which was validated in other omics datasets and cell experiments. Our novel pipeline can provide a case for other gene identification related studies.

Poly (ADP-ribose) polymerase (PARP) inhibitors targeted DNA damage repair, induced further DNA damage, and had a synthetic lethal effect in DNA repair-deficient tumors (Slade, 2020). PARP inhibitors improved progression-free survival in recurrent ovarian cancer patients with BRCA1/2 mutation and platinum-sensitive by 13.6 months (Pujade-Lauraine et al., 2017). PARP inhibitors also had a higher response rate in BRCA mutant triple-negative breast cancer (Pahuja et al., 2014). In addition to BRCA, other DNA repair-related genes also affected PARP inhibitor responses, such as RAD51 (Liu et al., 2017). Due to the important roles of SETD2 in homologous recombination repair (Skucha et al., 2019), the application of PARP inhibitor in SETD2-deficient tumors may achieve favorable curative effects, which needed to be confirmed by further studies.

SETD2 also played important roles in m6A RNA modification (Huang et al., 2019), which was related to prognosis and radioresistance. Li et al. (2020) found that low expression of FTO and METTL14 and high expression of METTL3, HNRNPA2B1, and YTHDF3 were related to the poor prognosis of osteosarcoma. The m6A "writer" METTL3 was demonstrated to promote radioresistance in pancreatic cancer (Taketo et al., 2018), hypopharyngeal squamous cell carcinoma (Wu et al., 2021), and glioma stem-like cells (Visvanathan et al., 2018). Radiosensitization caused by SETD2 knockdown may be related to both DDR and m6A.

LUAD and LUSC were highly heterogeneous for radiotherapy. Previous studies found that LUSC shrunk faster than LUAD after stereotactic body radiotherapy (Miyakawa et al., 2013). However, the local control rates of LUAD were not inferior to LUSC (Miyakawa et al., 2013; Hörner-Rieber et al., 2017; McAleese et al., 2019; Katagiri and Jingu, 2021). Moreover, LUSC was more likely to relapse locally, while LUAD was more likely to metastasize after radiation (McAleese et al., 2019; Katagiri and Jingu, 2021). For the overall survival of LUAD and LUSC after radiotherapy, there was some seemingly contradictory evidence (Nakayama et al., 1997; Holgersson et al., 2011), possibly due to the heterogeneity of the included populations. Despite the controversy, based on current evidence, the radiocurability of LUAD was not inferior to that of LUSC.

In the bulk transcriptome, SETD2 was associated with a favorable prognosis. However, it was unclear whether the favorable prognosis of SETD2 in LUAD was determined by the tumor or microenvironmental cells. Previous studies found that SETD2 histological staining scores of tumor cells were linked to good survival in gastric cancer (Chen et al., 2018) and nonmetastatic clear-cell renal cell carcinoma (Liu et al., 2015), possibly because lack of SETD2 increased microsatellite instability and frequency of spontaneous mutations (Liu et al., 2015; Zeng et al., Forthcoming 2022). Moreover, we collected 7 LUAD single-cell-derived metastasis-associated genes (PRSS3, GPI, CCL20, KRT18, TCN1, SLCO1B3, and GNPNAT1) from a previous study (He et al., 2021). However, in GSE131907, the results of GSVA analysis showed no significant difference in the scores of 7 metastasis-associated genes between SETD2-positive and negative tumor cells (Supplementary Figure S13). We expected further clinical studies to investigate the prognostic implications of SETD2 in cancer.

There were still some open issues. One focus was to identify genomic signatures associated with radioresistance and explore their mechanisms. Although acquired small deletion mutations were suggested as possible causes of radioresistance (Kocakavuk et al., 2021), further studies were required to investigate the complexity of tumor heterogeneity. The landmark analysis showed that the protective effects of SETD2 were reversed at more than 50–80 months, however the mechanism was not clear. Moreover, SETD2 also plays important roles in RNA splicing (Bhattacharya et al., 2021). Other mechanisms by which SETD2 affected therapeutic effectiveness also remained to be

investigated. In addition, preclinical studies investigating the toxicity and efficacy of interventions targeting SETD2 were lacking.

In this study, we mainly analyzed various genomics data supplemented by a limited amount of *in vitro* cell data. This study had some limits: 1) Since SETD2 expression was not high in various cells, we expected further studies to confirm the implications of SETD2 over-expression in LUAD; 2) The role of SETD2 regulated m6A in radiosensitivity remains unclear; 3) More clinical evidence was needed to use SETD2 as a prognosis and radiotherapy marker.

5 Conclusion

Our comprehensive analysis pipeline demonstrated that SETD2 was a key radiosensitivity signature. SETD2 enhanced chromatin accessibility and gene transcription which focused on DDR, DNA damage repair, and histone modification. Knockdown of SETD2 attenuated the proliferation and migration of LUAD cells, and enhanced cell apoptosis and radiosensitivity *in vitro*. Furthermore, SETD2 was a positively prognostic factor whose effects were negatively affected by the interaction of m6A-related genes RBM15 and YTHDF3.

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

ZZ, CX, and YG designed the experiments and wrote the paper. ZZ, JZ, JL, and YL conducted the experiments. ZZ, JZ, ZH, and LH analyzed the data. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (81800429 and 81972852), Key Research & Development Project of Hubei Province (2020BCA069), Health Commission of Hubei Province Medical Leading Talent Project, Young and Middle-Aged Medical Backbone Talents of Wuhan (WHQG201902), Application Foundation Frontier Project of Wuhan (2020020601012221), Translational Medicine and Interdisciplinary Research Joint Fund of Zhongnan Hospital of Wuhan University (ZNJC201922 and ZNJC202007), and Medical Science and Technology Innovation Platform Construction Support Project of Zhongnan Hospital of Wuhan University (PTXM2022009).

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.935601/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Akaike information criterion of linear equation, quadratic equation and cubic equation: Radiation dose ~ gene expression. The slope of the red line is 1 and the intercept is 0.

SUPPLEMENTARY FIGURE S2
Time-dependent gene clusters *via* Short Time-series Expression Miner. The blue line shows the changes of genes in the cluster over time. The black line is the eigenvalue of the cluster. The P value of the first 11 clusters is less than 0.01.

SUPPLEMENTARY FIGURE S3
The subnetworks in the 4-level biological network. **(A)** Subnetwork structures which were identified by MCODE app in Cytoscape (Shannon et al., 2003); **(B)** GO enrichment analysis of genes in the corresponding left subnetworks.

SUPPLEMENTARY FIGURE S4
Selection of soft threshold in scale-free co-expression networks in TCGA LUAD.

SUPPLEMENTARY FIGURE S5
Co-expression modules in TCGA LUAD.

SUPPLEMENTARY FIGURE S6
Selection of soft threshold in scale-free co-expression networks in TCGA LUSC.

SUPPLEMENTARY FIGURE S7
Co-expression modules in TCGA LUSC.

SUPPLEMENTARY FIGURE S8
Positive proportion of SETD2 in different cells of the 11 patients in scRNA-seq GSE131907.

SUPPLEMENTARY FIGURE S9

Expression of SETD2 in TCGA pan-cancer datasets.

SUPPLEMENTARY FIGURE S10
Immunohistochemistry of SETD2 from the HPA database. **(A)** SETD2 staining scores of LUAD tumor cells and alveolar cells. **(B)** The immunohistochemistry images of 5 lung adenocarcinoma and 3 normal lung tissues.

SUPPLEMENTARY FIGURE S11
SETD2 expression of NSCLC cell lines from the CCLE database.

SUPPLEMENTARY FIGURE S12
Correlations between SETD2 and m6A-related genes.

SUPPLEMENTARY FIGURE S13
GSVA analysis of 7 LUAD single-cell-derived metastasis-associated genes (PRSS3, GPI, CCL20, KRT18, TCN1, SLCO1B3, and GNPNAT1) between SETD2-positive and negative tumor cells in GSE131907.

SUPPLEMENTARY TABLE S1
The basic information of included datasets.

SUPPLEMENTARY TABLE S2
SF2 values from previous laboratory studies.

SUPPLEMENTARY TABLE S3
Multivariate Cox regression with RBM15 * SETD2 interaction.

SUPPLEMENTARY TABLE S4
Multivariate Cox regression with YTHDF3 * SETD2 interaction.

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. *Nat. Genet.* 25 (1), 25–29. doi:10.1038/75556

Barretina, J., Caponigro, G., Stransky, N., Venkatesan, K., Margolin, A. A., Kim, S., et al. (2012). The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* 483 (7391), 603–607. doi:10.1038/nature11003

Behrend, S. J., Giotopoulou, G. A., Spella, M., and Stathopoulos, G. T. (2021). A role for club cells in smoking-associated lung adenocarcinoma. *Eur. Respir. Rev.* 30 (162), 210122. doi:10.1183/16000617.0122-2021

Bhattacharya, S., Levy, M. J., Zhang, N., Li, H., Florens, L., Washburn, M. P., et al. (2021). The methyltransferase SETD2 couples transcription and splicing by engaging mRNA processing factors through its SHI domain. *Nat. Commun.* 12 (1), 1443. doi:10.1038/s41467-021-21663-w

Bild, A. H., Yao, G., Chang, J. T., Wang, Q., Potti, A., Chasse, D., et al. (2006). Oncogenic pathway signatures in human cancers as a guide to targeted therapies. *Nature* 439 (7074), 353–357. doi:10.1038/nature04296

Breiman, L. (2001). Random forests. *Mach. Learn.* 45 (1), 5–32. doi:10.1023/A:1010933404324

Byers, L. A., Diao, L., Wang, J., Saintigny, P., Girard, L., Peyton, M., et al. (2013). An epithelial-mesenchymal transition gene signature predicts resistance to EGFR and PI3K inhibitors and identifies Axl as a therapeutic target for overcoming EGFR inhibitor resistance. *Clin. Cancer Res.* 19 (1), 279–290. doi:10.1158/1078-0432.ccr-12-1558

Carvalho, S., Vítor, A. C., Sridhara, S. C., Martins, F. B., Raposo, A. C., Desterro, J. M. P., et al. (2014). SETD2 is required for DNA double-strand break repair and activation of the p53-mediated checkpoint. *eLife* 3, e02482. doi:10.7554/eLife.02482

Césaire, M., Montanari, J., Curcio, H., Lerouge, D., Gervais, R., and Demontrond, P. (2022). Radioresistance of non-small cell lung cancers and therapeutic perspectives. *Cancers* 14 (12), 2829. doi:10.3390/cancers14122829

Chen, Z., Raghoonundun, C., Chen, W., Zhang, Y., Tang, W., Fan, X., et al. (2018). SETD2 indicates favourable prognosis in gastric cancer and suppresses cancer cell proliferation, migration, and invasion. *Biochem. Biophys. Res. Commun.* 498 (3), 579–585. doi:10.1016/j.bbrc.2018.03.022

Clough, E., and Barrett, T. (2016). The gene expression omnibus database. *Methods Mol. Biol.* 1418, 93–110. doi:10.1007/978-1-4939-3578-9_5

Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAbiolinks: an R/bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44 (8), e71. doi:10.1093/nar/gkv1507

Conesa, A., Nueda, M. J., Ferrer, A., and Talón, M. (2006). maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. *Bioinformatics* 22 (9), 1096–1102. doi:10.1093/bioinformatics/btl056

Der, S. D., Sykes, J., Pintilie, M., Zhu, C. Q., Strumpf, D., Liu, N., et al. (2014). Validation of a histology-independent prognostic gene signature for early-stage, non-small-cell lung cancer including stage IA patients. *J. Thorac. Oncol.* 9 (1), 59–64. doi:10.1097/jto.0000000000000042

Ernst, J., and Bar-Joseph, Z. (2006). STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinforma.* 7 (1), 191. doi:10.1186/1471-2105-7-191

Eschrich, S., Zhang, H., Zhao, H., Boulware, D., Lee, J. H., Bloom, G., et al. (2009). Systems biology modeling of the radiation sensitivity network: a biomarker discovery platform. *Int. J. Radiat. Oncol. Biol. Phys.* 75 (2), 497–505. doi:10.1016/j.ijrobp.2009.05.056

Ettinger, D. S., Wood, D. E., Aisner, D. L., Akerley, W., Bauman, J. R., Bharat, A., et al. (2021). NCCN guidelines insights: Non-small cell lung cancer, version 2.2021. *J. Natl. Compr. Canc. Netw.* 19 (3), 254–266. doi:10.6004/jnccn.2021.0013

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., PinerosM., et al. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* 144 (8), 1941–1953. doi:10.1002/ijc.31937

Gao, Z., Zhuang, L., and Chen, Y. (2012). Effect and mechanism of gefitinib inhibition on non-small cell lung cancer radiosensitivity of HCC827 and H358 cell lines. *Zhongguo Fei Ai Za Zhi* 15 (6), 324–331. doi:10.3779/j.issn.1009-3419.2012.06.02

Goldman, M. J., Craft, B., Hastie, M., Repečka, K., McDade, F., Kamath, A., et al. (2020). Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* 38 (6), 675–678. doi:10.1038/s41587-020-0546-8

Goldstraw, P., Chansky, K., Crowley, J., Rami-Porta, R., Asamura, H., Eberhardt, W. E. E., et al. (2016). The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (eighth) edition of the TNM classification for lung cancer. *J. Thorac. Oncol.* 11 (1), 39–51. doi:10.1016/j.jtho.2015.09.009

Gu, C., Shi, X., Dai, C., Shen, F., Rocco, G., Chen, J., et al. (2020). RNA m^{6}A modification in cancers: molecular mechanisms and potential clinical applications. *Innovation.* 1 (3), 100066. doi:10.1016/j.xinn.2020.100066

Hanash, S. M., Bobek, M. P., Rickman, D. S., Williams, T., Rouillard, J. M., Kuick, R., et al. (2002). Integrating cancer genomics and proteomics in the post-genome era. *Proteomics* 2 (1), 69–75. doi:10.1002/1615-9861(200201)2:1<69::aid-prot69>3.0.co;2-8

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

He, J., Zhang, W., Li, F., and Yu, Y. (2021). Development of metastasis-associated seven gene signature for predicting lung adenocarcinoma prognosis using single-cell RNA sequencing data. *Math. Biosci. Eng.* 18 (5), 5959–5977. doi:10.3934/mbe.2021298

Ho, T. H., Kapur, P., Joseph, R. W., Serie, D. J., Eckel-Passow, J. E., Tong, P., et al. (2016). Loss of histone H3 lysine 36 trimethylation is associated with an increased risk of renal cell carcinoma-specific death. *Mod. Pathol.* 29 (1), 34–42. doi:10.1038/modpathol.2015.123

Holgersson, G., Bergström, S., Bergqvist, M., Nyman, J., Hoye, E., Helsing, M., et al. (2011). Swedish lung cancer radiation study group: predictive value of histology for radiotherapy response in patients with non-small cell lung cancer. *Eur. J. Cancer* 47 (16), 2415–2421. doi:10.1016/j.ejca.2011.06.011

Hörner-Rieber, J., Bernhardt, D., Dern, J., König, L., Adeberg, S., Paul, A., et al. (2017). Histology of non-small cell lung cancer predicts the response to stereotactic body radiotherapy. *Radiother. Oncol.* 125 (2), 317–324. doi:10.1016/j.radonc.2017.08.029

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, W., Chen, T.-Q., Fang, K., Zeng, Z.-C., Ye, H., Chen, Y.-Q., et al. (2021). N6-methyladenosine methyltransferases: functions, regulation, and clinical potential. *J. Hematol. Oncol.* 14 (1), 117. doi:10.1186/s13045-021-01129-8

Karlsson, M., Zhang, C., Mear, L., Zhong, W., Digre, A., Katona, B., et al. (2021). A single-cell type transcriptomics map of human tissues. *Sci. Adv.* 7 (31), eabh2169. doi:10.1126/sciadv.abh2169

Katagiri, Y., Jingu, K., Yamamoto, T., Matsushita, H., Umezawa, R., Ishikawa, Y., et al. (2021). Differences in patterns of recurrence of squamous cell carcinoma and adenocarcinoma after radiotherapy for stage III non-small cell lung cancer. *Jpn. J. Radiol.* 39 (6), 611–617. doi:10.1007/s11604-021-01091-y

Kellerer, A. M., and Rossi, H. H. (1978). A generalized formulation of dual radiation action. *Radiat. Res.* 75 (3), 471. doi:10.2307/3574835

Kim, N., Kim, H. K., Lee, K., Hong, Y., Cho, J. H., Choi, J. W., et al. (2020). Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. *Nat. Commun.* 11 (1), 2285. doi:10.1038/s41467-020-16164-1

Kocakavuk, E., Anderson, K. J., Varn, F. S., Johnson, K. C., Amin, S. B., Sulman, E. P., et al. (2021). Radiotherapy is associated with a deletion signature that contributes to poor outcomes in patients with cancer. *Nat. Genet.* 53 (7), 1088–1096. doi:10.1038/s41588-021-00874-3

Kumari, S., and Muthusamy, S. (2020). SETD2 as a regulator of N6-methyladenosine RNA methylation and modifiers in cancer. *Eur. J. Cancer Prev.* 29 (6), 556–564. doi:10.1097/cej.0000000000000587

Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature* 518 (7539), 317–330. doi:10.1038/nature14248

Lambrechts, D., Wauters, E., Boeckx, B., Aibar, S., Nittner, D., Burton, O., et al. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. *Nat. Med.* 24 (8), 1277–1289. doi:10.1038/s41591-018-0096-5

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 9, 559. doi:10.1186/1471-2105-9-559

Lê, S., Josse, J., and Husson, F. (2008). FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* 25, 1. doi:10.18637/jss.v025.i01

Leiserson, M. D., Vandin, F., Wu, H. T., Dobson, J. R., Eldridge, J. V., Thomas, J. L., et al. (2015). Pan-cancer network analysis identifies combinations of rare somatic mutations across pathways and protein complexes. *Nat. Genet.* 47 (2), 106–114. doi:10.1038/ng.3168

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77 (21), e108–e110. doi:10.1158/0008-5472.can-17-0307

Li, J., Rao, B., Yang, J., Liu, L., Huang, M., Liu, X., et al. (2020). Dysregulated m6A-related regulators are associated with tumor metastasis and poor prognosis in osteosarcoma. *Front. Oncol.* 10, 769. doi:10.3389/fonc.2020.00769

Liu, T., Ortiz, J. A., Taing, L., Meyer, C. A., Lee, B., Zhang, Y., et al. (2011). Cistrome: an integrative platform for transcriptional regulation studies. *Genome Biol.* 12 (8), R83. doi:10.1186/gb-2011-12-8-r83

Liu, W., Fu, Q., An, H., Chang, Y., Zhang, W., Zhu, Y., et al. (2015). Decreased expression of SETD2 predicts unfavorable prognosis in patients with nonmetastatic clear-cell renal cell carcinoma. *Med. (Baltimore)* 94 (45), e2004. doi:10.1097/md.0000000000002004

Liu, Y., Burness, M. L., Martin-Trevino, R., Guy, J., Bai, S., Harouaka, R., et al. (2017). RAD51 mediates resistance of cancer stem cells to PARP inhibition in triple-negative breast cancer. *Clin. Cancer Res.* 23 (2), 514–522. doi:10.1158/1078-0432.CCR-15-1348

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

McAleese, J., Taylor, A., Walls, G. M., and Hanna, G. G. (2019). Differential relapse patterns for non-small cell lung cancer subtypes adenocarcinoma and squamous cell carcinoma: implications for radiation oncology. *Clin. Oncol.* 31 (10), 711–719. doi:10.1016/j.clon.2019.07.008

McInnes, L., Healy, J., Saul, N., and Grossberger, L. (2018). UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* 3, 861. doi:10.21105/joss.00861

Miyakawa, A., Shibamoto, Y., Kosaki, K., and Hashizume, C. (2013). Early response and local control of stage I non-small-cell lung cancer after stereotactic radiotherapy: difference by histology. *Cancer Sci.* 104 (1), 130–134. doi:10.1111/cas.12048

Nakayama, Y., Hayakawa, K., Mitsuhashi, N., Saito, Y., and Niibe, H. (1997). Long-term survivors of non-small cell lung cancer after radiation therapy: the significance of histological type. *Anticancer Res.* 17 (4), 2769–2773.

Oleinick, N. L., Biswas, T., Patel, R., Tao, M., Patel, R., Weeks, L., et al. (2016). Radiosensitization of non-small-cell lung cancer cells and xenografts by the interactive effects of pemetrexed and methoxyamine. *Radiother. Oncol.* 121 (2), 335–341. doi:10.1016/j.radonc.2016.10.007

Pahuja, S., Beumer, J. H., Appleman, L. J., Tawbi, H. A.-H., Stoller, R. G., Lee, J. J., et al. (2014). Outcome of BRCA 1/2-mutated (BRCA+) and triple-negative, BRCA wild type (BRCA-wt) breast cancer patients in a phase I study of single-agent veliparib (V). *J. Clin. Oncol.* 32 (26), 135. doi:10.1200/jco.2014.32.26_suppl.135

Park, J. G., Paul, S., Briones, N., Zeng, J., Gillis, K., Wallstrom, G., et al. (2017). Developing human radiation biodosimetry models: testing cross-species conversion approaches using an *Ex Vivo* model system. *Radiat. Res.* 187 (6), 708–721. doi:10.1667/rr14655.1

Peng, Q., Weng, K., Li, S., Xu, R., Wang, Y., Wu, Y., et al. (2021). A perspective of epigenetic regulation in radiotherapy. *Front. Cell Dev. Biol.* 9, 624312. doi:10.3389/fcell.2021.624312

Pfister, S. X., Ahrabi, S., Zalmas, L.-P., Sarkar, S., Aymard, F., Bachrati, C. Z., et al. (2014). SETD2-dependent histone H3K36 trimethylation is required for homologous recombination repair and genome stability. *Cell Rep.* 7 (6), 2006–2018. doi:10.1016/j.celrep.2014.05.026

Pollom, E. L., Qian, Y., Durkee, B. Y., von Eyben, R., Maxim, P. G., Shultz, D. B., et al. (2016). Hypofractionated intensity-modulated radiotherapy for patients with non-small-cell lung cancer. *Clin. Lung Cancer* 17 (6), 588–594. doi:10.1016/j.cllc.2016.05.024

Pujade-Lauraine, E., Ledermann, J. A., Selle, F., Gebski, V., Penson, R. T., Oza, A. M., et al. (2017). Olaparib tablets as maintenance therapy in patients with platinum-sensitive, relapsed ovarian cancer and a BRCA1/2 mutation (SOLO2/ENGOT-Ov21): a double-blind, randomised, placebo-controlled, phase 3 trial. *Lancet. Oncol.* 18 (9), 1274–1284. doi:10.1016/s1470-2045(17)30469-2

Reinhold, W. C., Reimers, M. A., Lorenzi, P., Ho, J., Shankavaram, U. T., Ziegler, M. S., et al. (2010). Multifactorial regulation of E-cadherin expression: an integrative study. *Mol. Cancer Ther.* 9 (1), 1–16. doi:10.1158/1535-7163.mct-09-0321

Satija, R., Farrell, J. A., Gennert, D., Schier, A. F., and Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* 33 (5), 495–502. doi:10.1038/nbt.3192

Scott, J. G., Berglund, A., Schell, M. J., Mihaylov, I., Fulp, W. J., Yue, B., et al. (2017). A genome-based model for adjusting radiotherapy dose (GARD): a retrospective, cohort-based study. *Lancet. Oncol.* 18 (2), 202–211. doi:10.1016/S1470-2045(16)30648-9

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi:10.1101/gr.1239303

Shirbhate, E., Patel, P., Patel, V. K., Veerasamy, R., Sharma, P. C., Rajak, H., et al. (2020). The combination of histone deacetylase inhibitors and radiotherapy: a promising novel approach for cancer treatment. *Future Oncol.* 16 (30), 2457–2469. doi:10.2217/fon-2020-0385

Skucha, A., Ebner, J., and Grebien, F. (2019). Roles of SETD2 in leukemia-transcription, DNA-damage, and beyond. *Int. J. Mol. Sci.* 20 (5), E1029. doi:10.3390/ijms20051029

Slade, D. (2020). PARP and PARG inhibitors in cancer treatment. *Genes Dev.* 34 (5-6), 360–394. doi:10.1101/gad.334516.119

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 (3), 209–249. doi:10.3322/caac.21660

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45 (D1), D362–D368. doi:10.1093/nar/gkw937

Taketo, K., Konno, M., Asai, A., Koseki, J., Toratani, M., Satoh, T., et al. (2018). The epitranscriptome m^6A writer METTL3 promotes chemo- and radioresistance in pancreatic cancer cells. *Int. J. Oncol.* 52 (2), 621–629. doi:10.3892/ijo.2017.4219

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., Zhang, Z., et al. (2017a). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45 (W1), W98–W102. doi:10.1093/nar/gkx247

Tang, Z., Zeng, Q., Li, Y., Zhang, X., Suto, M. J., Xu, B., et al. (2017b). Predicting radiotherapy response for patients with soft tissue sarcoma by developing a molecular signature. *Oncol. Rep.* 38 (5), 2814–2824. doi:10.3892/or.2017.5999

Torres-Roca, J. F., Eschrich, S., Zhao, H., Bloom, G., Sung, J., McCarthy, S., et al. (2005). Prediction of radiation sensitivity using a gene expression classifier. *Cancer Res.* 65 (16), 7169–7176. doi:10.1158/0008-5472.can-05-0656

Visvanathan, A., Patil, V., Arora, A., Hegde, A. S., Arivazhagan, A., Santosh, V., et al. (2018). Essential role of METTL3-mediated m(6)A modification in glioma stem-like cells maintenance and radioresistance. *Oncogene* 37 (4), 522–533. doi:10.1038/onc.2017.351

Wang, S., Sun, H., Ma, J., Zang, C., Wang, C., Wang, J., et al. (2013). Target analysis by integration of transcriptome and ChIP-seq data with BETA. *Nat. Protoc.* 8 (12), 2502–2515. doi:10.1038/nprot.2013.150

Wu, C., Guo, E., Ming, J., Sun, W., Nie, X., Sun, L., et al. (2020). Radiation-induced DNMT3B promotes radioresistance in nasopharyngeal carcinoma through methylation of p53 and p21. *Mol. Ther. Oncolytics* 17, 306–319. doi:10.1016/j.omto.2020.04.007

Wu, P., Fang, X., Liu, Y., Tang, Y., Wang, W., Li, X., et al. (2021). N6-methyladenosine modification of circCUX1 confers radioresistance of hypopharyngeal squamous cell carcinoma through caspase1 pathway. *Cell Death Dis.* 12 (4), 298. doi:10.1038/s41419-021-03558-2

Xu, Y., Su, G. H., Ma, D., Xiao, Y., Shao, Z. M., Jiang, Y. Z., et al. (2021). Technological advances in cancer immunity: from immunogenomics to single-cell analysis and artificial intelligence. *Signal Transduct. Target. Ther.* 6 (1), 312. doi:10.1038/s41392-021-00729-7

Xue, G., Ren, Z., Chen, Y., Zhu, J., Du, Y., Pan, D., et al. (2015). A feedback regulation between miR-145 and DNA methyltransferase 3b in prostate cancer cell and their responses to irradiation. *Cancer Lett.* 361 (1), 121–127. doi:10.1016/j.canlet.2015.02.046

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

Yu, G., Wang, L. G., and He, Q. Y. (2015). ChIPseeker: an R/bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* 31 (14), 2382–2383. doi:10.1093/bioinformatics/btv145

Zeng, Z., Gao, Y., Li, J., Zhang, J., Li, Y., He, F., et al. (Forthcoming 2022). SETD2 mediates immunotherapy and radiotherapy efficacy via regulating DNA damage responses and genomic stability in lung adenocarcinoma. *Genes Dis.* doi:10.1016/j.gendis.2022.02.016

Zhong, X., Luo, G., Zhou, X., Luo, W., Wu, X., Zhong, R., et al. (2016). Rad51 in regulating the radiosensitivity of non-small cell lung cancer with different epidermal growth factor receptor mutation status. *Thorac. Cancer* 7 (1), 50–60. doi:10.1111/1759-7714.12274

frontiers | Frontiers in Genetics

ORIGINAL RESEARCH
published: 25 August 2022
doi: 10.3389/fgene.2022.907219

CircPUM1 Knockdown Confers Radiosensitivity in Oral Squamous Cell Carcinoma by Regulating the miR-580/STAT3 Pathway

Linghui Jia[1,2†], *Pengcheng Huang*[1,3†], *Tingting Lin*[4], *Chunyan Lin*[4], *Xiaofen Ding*[4], *Liping Lin*[4], *Lifeng Zhu*[4*] and *Zhilian Zhou*[4*]

[1]*Fujian Key Laboratory of Oral Diseases, Fujian Provincial Engineering Research Center of Oral Biomaterial of Stomatological Key laboratory of Fujian College and University, School and Hospital of Stomatology, Fujian Medical University, Fuzhou, China,* [2]*Department of Prosthodontics, School and Hospital of Stomatology, Fujian Medical University, Fuzhou, China,* [3]*Department of Orthodontics, School and Hospital of Stomatology, Fujian Medical University, Fuzhou, China,* [4]*Affiliated Sanming First Hospital, Fujian Medical University, Sanming, China*

OPEN ACCESS

Edited by:
Yan Gong,
Wuhan University, China

Reviewed by:
Zhe Wang,
Affiliated Zhongshan Hospital of Dalian University, China
Fang Tang,
Wuhan University, China

***Correspondence:**
Zhilian Zhou
zhouzl_0608@163.com
Lifeng Zhu
zhulfsm@163.com

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

Specialty section:
This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

Received: *29 March 2022*
Accepted: *20 June 2022*
Published: *25 August 2022*

Citation:
Jia L, Huang P, Lin T, Lin C, Ding X, Lin L, Zhu L and Zhou Z (2022) CircPUM1 Knockdown Confers Radiosensitivity in Oral Squamous Cell Carcinoma by Regulating the miR-580/STAT3 Pathway. Front. Genet. 13:907219. doi: 10.3389/fgene.2022.907219

Background: CircPUM1 acts as an oncogene in a variety of tumors, and there is no related research on oral squamous cell carcinoma. This study aimed to evaluate the clinical significance of CircPUM1 in oral squamous cell carcinoma radiotherapy.

Methods: Radio-resistant cell lines were established by increasing the X-ray dose. Analysis of CircPUM1 expression in oral squamous cell carcinoma was carried out using bioinformatics tools. Cell proliferation was analyzed with CCK-8 and colony formation. Protein and gene expressions were detected by Western blotting and qPCR. RNA interference inhibits endogenous gene expression. A luciferase reporter system and immunoprecipitation were used to validate the target of CircPUM1.

Result: CircPUM1 was highly expressed in OSCC. The higher the expression level of CircPUM1 in OSCC, the worse the clinical features and prognosis. Knockdown of CircPUM1 enhances the sensitivity of OSCC cells to X-rays, and expression of exogenous CircPUM1 makes OSCC cells acquire radiation resistance. The absence of CircPUM1 blocked the cells in the G0/G1 phase and triggered apoptosis. The prediction of mir-580-binding site, luciferase reporter system, and immunoprecipitation confirmed that mir-580 is the binding site of CircPUM1. In addition, STAT3 was predicted and confirmed as the binding site of mir-580. Overexpression of STAT3 partially attenuated the radiosensitivity of OSCC cells to knockdown of CircPUM1.

Conclusion: CircPUM1 has the oncogene expression profile in oral squamous cell carcinoma; patients with high expression of CircPUM1 have less benefit from radiotherapy and need more frequent follow-up. In addition, CircPUM1 may be a potential therapeutic target for oral squamous cell carcinoma. The CircPUM1/mir-580/STAT3 axis has a certain effect on the radiosensitivity of OSCC. These results suggest that patients with low expression of CircPUM1 may gain more benefits.

Keywords: oral squamous cell carcinoma, radiosensitivity, CircPUM1, mir-580, STAT3

INTRODUCTION

Oral squamous cell carcinoma is one of the most common malignant tumors in oral and maxillofacial regions (Müller, 2017; Kennedy, 2018). Oral squamous cell carcinoma has poor clinical prognosis and low survival rate (Al-Sarraf, 2007). Currently, the combined treatment of surgery plus radiotherapy and chemotherapy can significantly improve the overall survival rate of patients with oral squamous cell carcinoma (Aslam et al., 2012). Radiotherapy, one of the curative treatments for oral squamous cell carcinoma, can be used with surgery, chemotherapy, or alone (Pentenero et al., 2005). Radiotherapy, an important part of comprehensive sequential therapy for advanced oral squamous cell carcinoma, can improve the local control rate (Berdugo et al., 2019). However, improving the effect of radiotherapy for oral squamous cell carcinoma is the main problem that we are facing currently.

Studies have shown that circular RNA (circRNA) is a kind of non-coding RNA, which has high stability due to its closed circular molecule (Zheng et al., 2019). Studies have shown that circRNA can inhibit its expression by acting as a sponge molecule of microRNA (miRNA), thus regulating the biological processes of tumor cell proliferation, apoptosis, migration, and so on (Lee et al., 2021). However, there are relatively few studies on circRNAs in oral squamous cell carcinoma. Circular RNA pum1 (circular RNA pum1) is highly expressed in lung cancer, ovarian cancer, colon cancer, and other tumors and can promote the occurrence and development of tumors (Jeck et al., 2013; Salzman et al., 2013; Rybak-Wolf et al., 2015). CircPUM1 (circ_0000043) is highly expressed in endometrial adenocarcinoma (Deng et al., 2020), ovarian adenocarcinoma (Guan et al., 2019), and lung adenocarcinoma (Chen et al., 2019) and plays a role as an oncogene, but it has not been reported in oral squamous cell carcinoma. Circinteractome predicted that CircPUM1 targeted mir-580 (the highest predicted score was 99). starBase predicted that mir-580 targeted Stat3. STAT3 has been reported to be involved in radiosensitivity of oral squamous cell carcinoma (Yu et al., 2020).

However, it is not known whether CircPUM1 affects the malignant biological behavior of oral squamous cells by targeting the expression of mir-580/Stat3. Therefore, this study mainly explored the expression of CircPUM1 and mir-580/STAT3 in oral squamous cell carcinoma and analyzed whether CircPUM1 regulates radiosensitivity through mir-580/STAT3. Therefore, we believe that our findings will provide a theoretical basis for improving the radiosensitivity of OSCC.

MATERIALS AND METHODS

Patients and Sample

A total of 50 OSCC and matched normal tissues were collected. All samples were independently confirmed by two pathologists. Immediately after operation, the specimens were frozen in liquid nitrogen and then stored in the refrigerator at − 80 C for future use. Patients who received any neoadjuvant therapy and had a history of cancer were excluded from the study. Prior to the start of the study, informed consent was obtained from the participants, and the study was conducted in strict accordance with the Helsinki declaration.

Cell Culture

Human oral squamous cell carcinoma cell lines Cal-27 and hsc3 were cultured in DMEM medium containing 10% fetal bovine serum, 100 units/ml penicillin–streptomycin, and 100% glucose at 37°C and 5% CO_2.

Irradiation of Cells and Development of Radio-Resistant Cell Lines

The cells were irradiated using the Faxitron cabinet X-ray system 43855d (Faxitron X-ray Company, IL, USA). Mcf-7r and mda-mb-231r were established on their parent cell lines. In short, the initial dose was 2 Gy, then the dose increased by 0.5 per week, and the total radiation dose reached 57 Gy in 12 weeks. After that, 5 Gy per week was used for further maintenance.

Cell Transfection

The shRNA was synthesized and subcloned into a plko.1-trc-puro plasmid. The sequence of shRNA is shown in **Table 1**. We used the Lipofectamine 2000 system to transfect the plasmid. In short, 8 × 105 cells were seeded in a six-well dish for 24 h, and then Liposome 2000 was mixed with 5 μg plasmid. Using Opti-MEM, the mixture was successively incubated at room temperature for 30 min, then the mixture was added to each well, supplemented with Opti-MEM to 1 ml, and incubated for 6 h. After 6 h, DMEM containing 20% FBS was added to each well and 2 ml medium and then cultured for 48 h for mRNA extraction and protein collection. According to the results predicted by Circinteractome, in order to overexpress CircPUM1, the whole coding sequence 5 'GACUUUUUGACUACAAUUCUC AA3′ of CircPUM1 was subcloned into the pZsG vector.

RNA Extraction and Quantitative Real-Time Polymerase Chain Reaction (QRT-PCR) Analysis

According to the manufacturer's protocol, TRIzol reagent (Invitrogen, Carlsbad, CA) was used to extract total RNA. The first-strand cDNA was synthesized by a reverse transcription kit (Takara, Dalian, China). GAPDH is used as an internal control. PCR primers for CircPUM1 and 18S RNA were found in starBase prediction. The relative expression of CircPUM1 was expressed with the $2^{-\Delta\Delta CT}$ method. All samples are in triplicate.

Cell Proliferation Assay

Cell proliferation test: Cal-27 and HSC3 cells were seeded onto a 96-well plate (Müller, 2017) and cultured for 24, 48, and 72 h. The incubation time was 37 h. μ CCK8 was injected into each well and incubated at 37°C. The absorbance at 480 nm was measured using the Rayto-6000 system (Rayto, China) after 2 h of storage of C and normalized to DMEM medium as control.

Flow Cytometry Analysis

In cell cycle analysis, the cells were harvested after 6 h of starvation, fixed overnight with cold ethanol, and then

TABLE 1 | Patient information and clinicopathological characteristics of 58 patients with OSCC *P < 0.05 or **P < 0.01 was considered significant (chi-square test between 2 groups).

Patients	Low circPUM1 (*n* = 29)	High circPUM1 (*n* = 29)	*P*-value
Age(years)			0.11556
≥65	16	17	
<65	13	12	
Gender			0.06496
Female	14	17	
Male	15	12	
Remote metastasis			0.00049**
No	9	21	
Yes	20	8	
Lymph node metastasis			0.00025**
No	7	19	
Yes	22	10	
TNM satge			<0.01**
I	3	6	
II	5	11	
III	8	8	
IV	13	4	
Differentiation			0.07813
Poor	12	11	
Moderate	5	4	
Well	13	14	

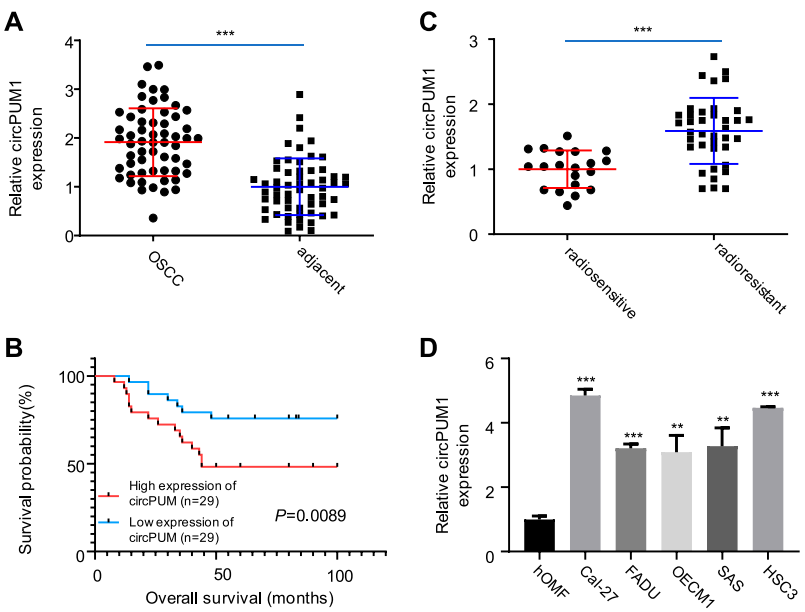


FIGURE 1 | Expression of CircPUM1 in OSCC carcinoma and adjacent tissues **(A)**; overall survival rate of low-expression group and high-expression group of CircPUM1 **(B)**; expression of CircPUM1 in OSCC tissues with different radiation doses **(C)**; expression of CircPUM1 in different OSCC cells **(D)**. $p < 0.01$ (**) and $p < 0.001$ (***).

incubated in the dark with propidium iodide and ribonuclease (BD, USA) for 15 min. In apoptosis analysis, after 6 h of starvation, the cells were washed twice with cold PBS, stained with FITC-binding Annexin V for 20 min, and stained with propidium iodide for 15 min. The stained cells were detected by flow cytometry (FACSAria III, BD, USA) and analyzed by FlowJo vx 0.7 software.

Western Blot Analysis

Protein was extracted with RIPA lysis buffer containing protease inhibitor (Roche). Quantitative analysis of protein was conducted using BCA ™ protein analysis kit (Pierce, Appleton, Wisconsin, USA). After that, protein (30 μg/sample) was transferred to a poly (vinylidene fluoride two fluoroethylene) membrane by 10% alkyl sulfate-polyacrylamide gel electrophoresis.

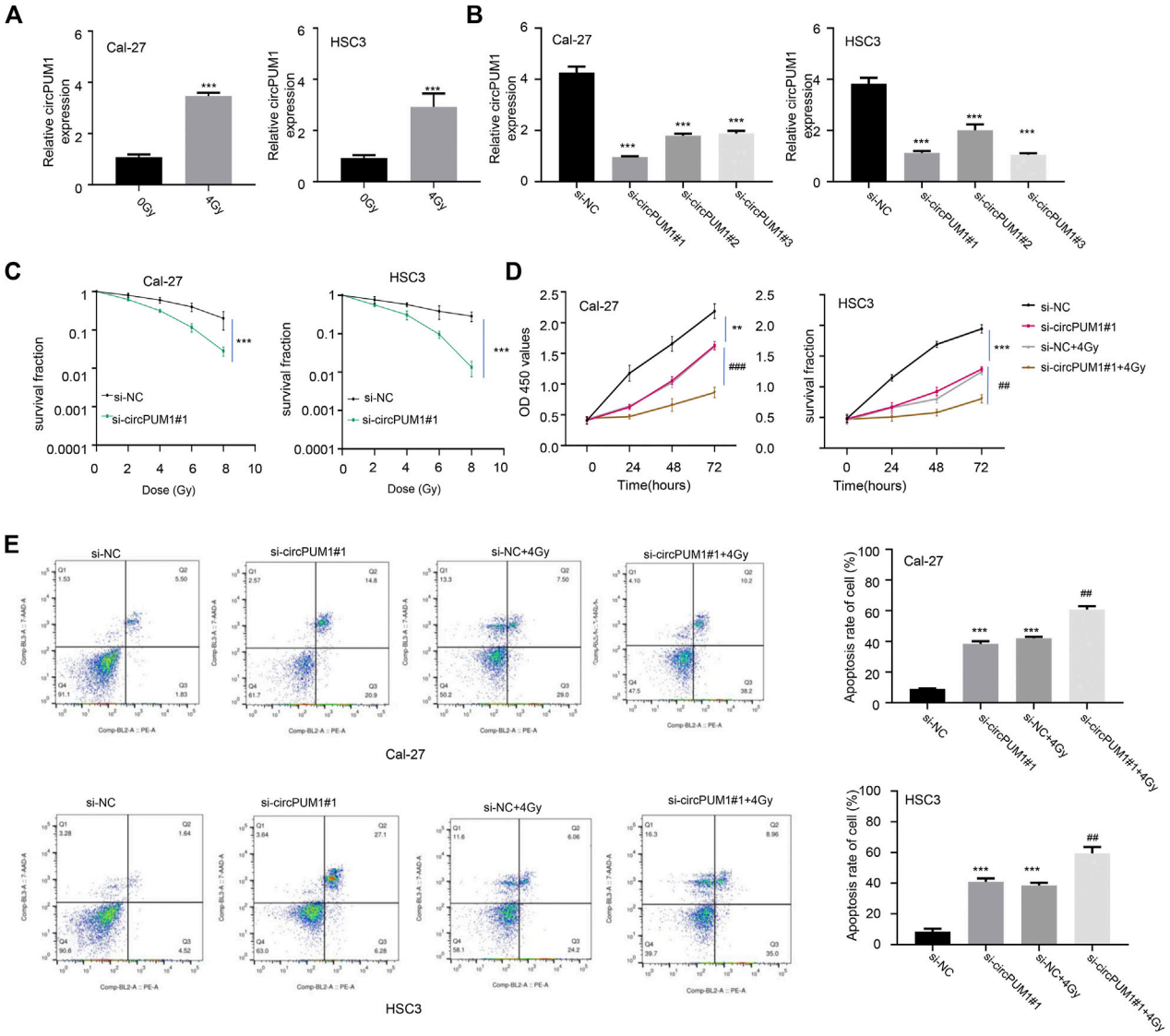


FIGURE 2 | QRT-PCR was used to detect the expression of CircPUM1 in Cal-27 and hsc3 cells**(A)**; expression of CircPUM1 in different groups of OSCC cells **(B)**; colony survival fraction of Cal-27 and hsc3 cells in different groups under different radiation doses**(C)**; optical absorption values of Cal-27 and hsc3 cells in different groups at 450 nm at different time periods were measured**(D)**; apoptosis was observed in different groups of Cal-27 and hsc3 cells**(E)**. $p < 0.01$ (**) and $p < 0.001$ (***).

The antibody was prepared in 5% blocking buffer with a dilution of 1:1,000, incubated with the membrane at 4 °C overnight, washed twice with TBST, then cultured with a secondary antibody (1:2000), and labeled with horseradish peroxidase for 2 h at room temperature. Immuno Western chemiluminescence HRP substrate (Millipore) was used to cover the film surface. Finally, the signal was captured, and the concentration of the band was quantified by Image Lab ™ software (Bio-Rad Laboratories, Hercules, CA, USA).

Luciferase Report Analysis

Methods: circPUM1-WT and circPUM1-Mut were subcloned into pGL3 vectors (Promega, Madison, WI) to construct plasmids. The plasmids were further transfected with designated mimics or siRNA for 48 h. The PCR products were cloned into the polyclonal sites of the recombinant pGL3 expression vector. Finally, the luciferase activity was detected by Dual Luciferase Report Analysis kit (Promega, Madison, Wisconsin, USA).

Statistical Analysis

All statistical analyses were performed using GraphPad Prism version 8.0 (GraphPad Software, La Jolla, California). The significant differences between groups were estimated using Student's t-test. A p-value less than 0.05 was considered statistically significant. The results are reported as average ± standard deviation. All experiments were carried out in triplicate.

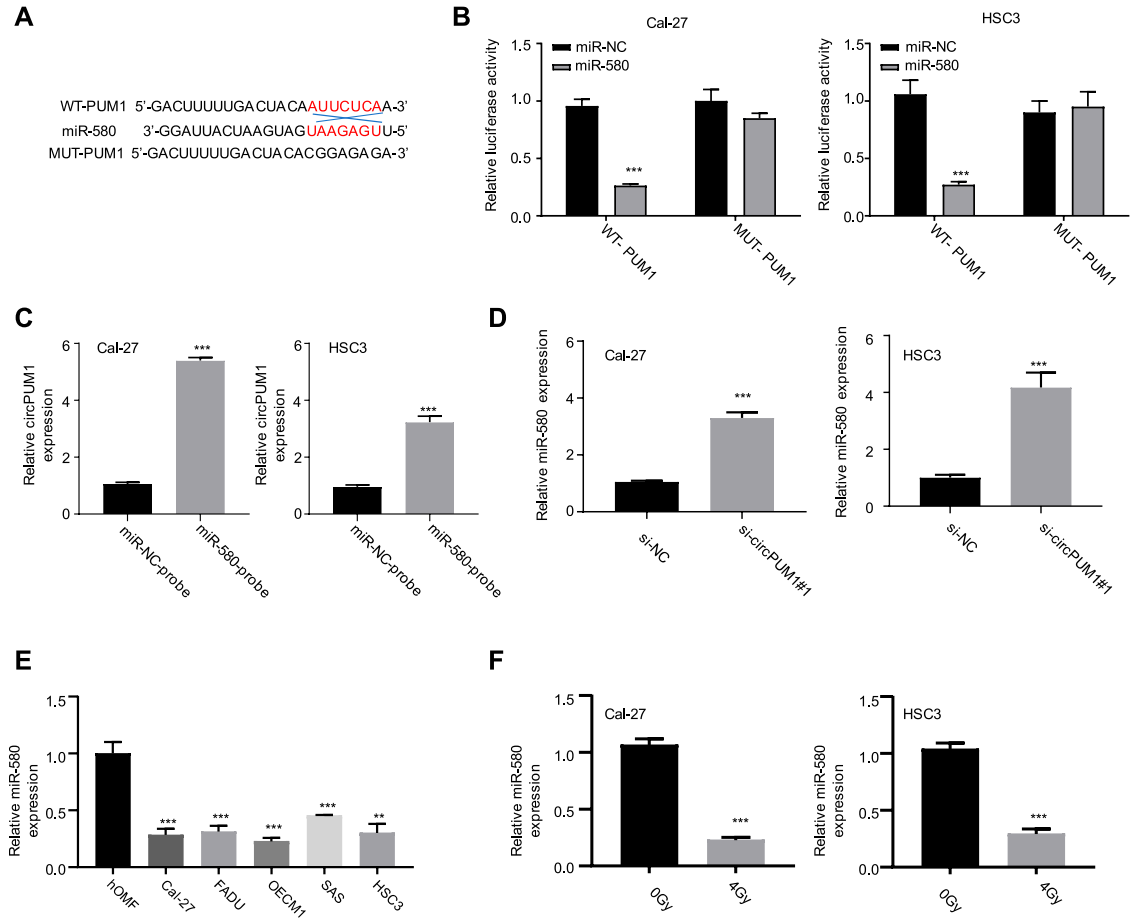


FIGURE 3 | Binding site of CircPUM1 and mir-580 **(A)**; effect of mir-580 on the activity of CircPUM1 **(B,C)**; relationship between knockdown of CircPUM1 and mir-580 expression **(D)**; expression level of mir-580 in the OSCC cell line **(E)**; expression of mir-580 in OSCC cells after 4 Gy radiation **(F)**. $p < 0.01$ (**) and $p < 0.001$ (***).

RESULTS

CircPUM1 Was Highly Expressed in OSCC

First, we analyzed the expression of CircPUM1 in patients. QRT-PCR was used to detect the expression level of CircPUM1 in 58 pairs of OSCC cancer tissues and corresponding adjacent tissues. It was found that the expression level of CircPUM1 in OSCC was significantly increased ($p < 0.001$, **Figure 1A**). At the same time, 58 OSCC patients were divided into two groups: low-expression group (n = 29) and high-expression group (n = 29) according to the cut-off value of median expression of CircPUM1 in **Figure 1A**. Chi square test was used to analyze the relationship between the expression of circpum1 and the clinicopathological data of OSCC, $p < 0.05$, but not related to the patient's age, gender and tumor differentiation. The Kaplan–Meier survival curve was used to evaluate the overall survival rate of the two groups. It was found that the prognosis of high-expression group of CircPUM1 was poor, and the difference was statistically significant ($p = 0.0089$, **Figure 1B**). Subsequently, the expression levels of CircPUM1 in radiosensitive (n = 20) and radio-resistant (n = 38) were detected by QRT-PCR. The expression of CircPUM1 in radio-resistant OSCC was significantly increased ($p < 0.001$, **Figure 1C**). As mentioned previously, we successfully established the expression level of CircPUM1 in two radiation-resistant OSCC cell lines (Cal-27, FaDu, OECM1, SAS, and HSC3) and the human oral mucous fibroblasts (HMFs). The expression level of CircPUM1 in OSCC cell lines was higher than that in normal tissues ($p < 0.001$, **Figure 1D**).

Knockdown of CircPUM1 Enhances Radiosensitivity of OSCC Cells

We transfected the CircPUM1 siRNA and expression plasmid into Cal-27 and hsc3 cells ($p < 0.005$, **Figures 2A,B**). The cells were exposed to different doses of ionizing radiation (0 and 4 Gy), and the expression level of CircPUM1 increased after 4 Gy; the difference was statistically significant. QRT-PCR was used to detect the expression level of CircPUM1 in Cal-27 and hsc3 cells of different groups (Si-nc, si-CircPUM1 # 1, si-CircPUM1 # 2, and si-CircPUM1 # 3). Colony formation assay was used to detect the colony survival fraction of Cal-27 and hsc3 cells in different groups (Si-nc and si-CircPUM1 #

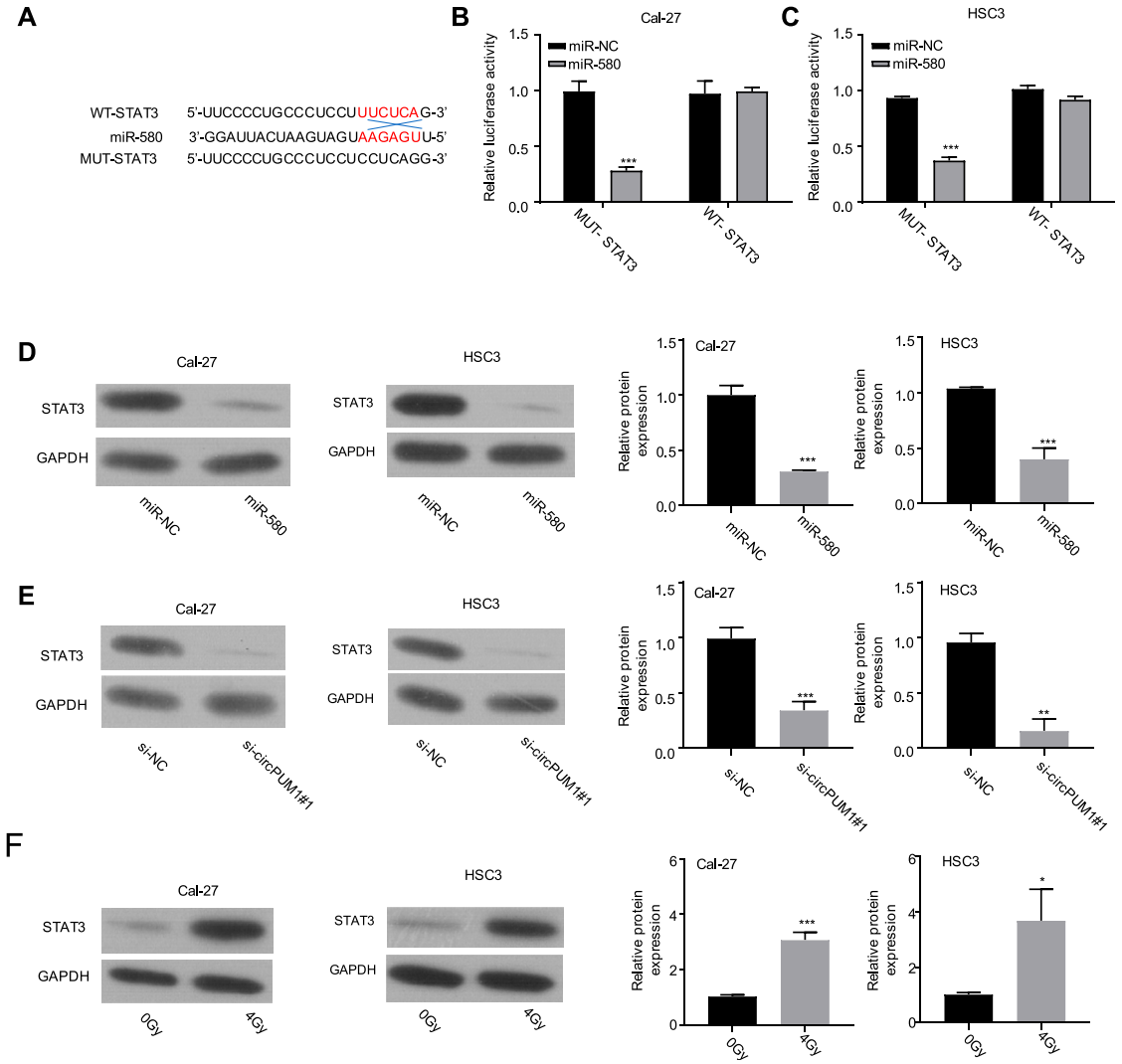


FIGURE 4 | Binding site of mir-580 and STAT3 **(A)**; luciferase reporter gene assay verified the targeting relationship between mir-580 and STAT3 **(B,C)**; protein expression level of STAT3 after mir-580 overexpression **(D)**; knockdown of CircPUM1 and STAT3 protein expression in the OSCC cell line **(E)**; expression of STAT3 in OSCC cells after 4 Gy irradiation **(F)**. $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***).

1) under different radiation doses (0, 2, 4, and 8 Gy). Compared with the Si NC group, the colony survival fraction of si-CircPUM1 # 1 cells gradually decreased with the gradual increase in radiation dose Gy ($p < 0.005$, **Figure 2C**). The results confirmed that radiosensitivity was a dose-dependent inhibitory effect. The loss of function of CircPUM1 induced more apoptosis when exposed to 4 Gy X-ray.

CCK8 was used to detect the light absorption values of Cal-27 and hsc3 cells in different groups (Si-nc, si-CircPUM1#1, Si-nc + 4Gy, and si-CircPUM1#1 + 4Gy) at the wavelength of 450 nm at 0h, 24h, 48h, and 72 h. Compared with the Si NC group, the light absorption values of si-CircPUM1#1 and Si-nc + 4Gy groups were lower at the wavelength of 450nm; compared with the Si NC + 4Gy group, the optical absorption value of si-CircPUM1 # 1 + 4Gy at 450 nm wavelength decreased, and the difference was statistically significant ($p < 0.01$, **Figure 2D**). The apoptosis level of Cal-27 and hsc3 cells in different groups (Si NC, si-CircPUM1 # 1, Si NC + 4Gy, and si-CircPUM1 # 1 + 4Gy) was detected by flow cytometry. Compared with the Si NC group, the apoptosis level of si-CircPUM1 # 1 and Si-nc + 4Gy groups was high; compared with the Si NC + 4Gy group, the apoptosis level of si-CircPUM1 # 1 + 4Gy group was significantly high ($p < 0.01$, FIG.2E). The results showed that the low expression of CircPUM1 could promote the radiosensitivity and apoptosis of OSCC tumor cells.

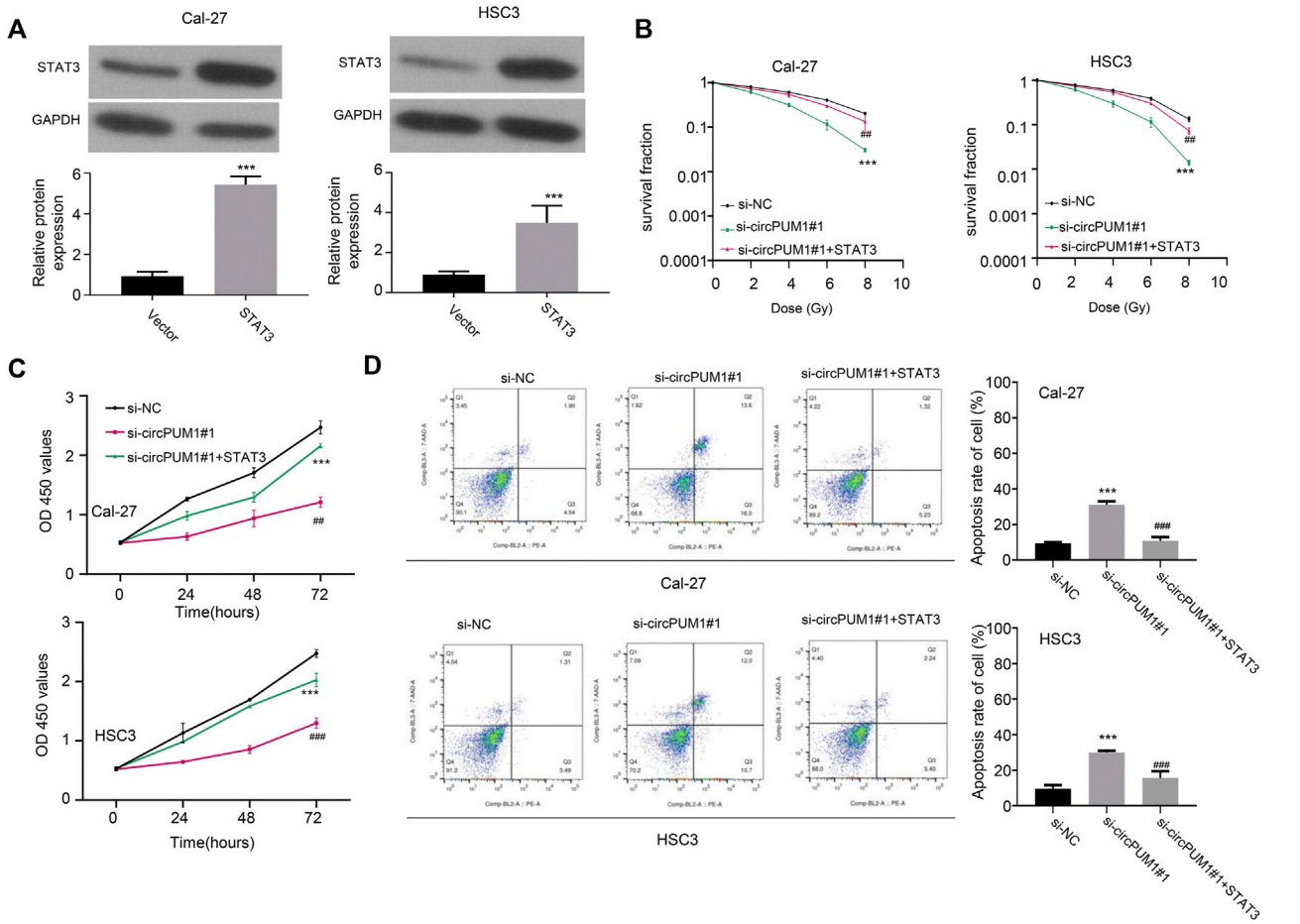


FIGURE 5 | Overexpression of STAT3 protein in OSCC cells **(A)**; grouping of Cal-27 and hsc3 cells at 4Gy: the optical absorption values of Si-nc, si-CircPUM1#1, and si-CircPUM1#1 + STAT3 at 0h, 24h, 48h, and 72 h **(B,C)** at 450 nm wavelength of different groups of cells under different time periods of radiation; knockdown of si-CircPUM1#1 and cotransfection of STAT3 plasmid induced apoptosis **(D)**. $p < 0.01$ (**) and $p < 0.001$ (***).

CircPUM1 Targets mir-580 and Inhibits Its Expression in OSCC Cells

Next, we want to know how CircPUM1 plays its biological function in radiation-resistant cells. Through the analysis of the Circinteractome online database, we found that there was a mir-580-binding site in CircPUM1. Luciferase reporter gene experiments were carried out in Cal-27 and hsc3 cells, respectively. The results showed that overexpression of mir-580 could inhibit the luciferase activity of wild-type CircPUM1 vector in Cal-27 and hsc3 cells, compared with mir-nc. After mutating the predicted mir-580-binding site, the inhibitory effect disappeared (**Figures 3A,B**); compared with the NC probe, the mir-580 probe enriched more CircPUM1 in Cal-27 and hsc3 cells (**Figure 3C**). QRT-PCR was used to detect the expression level of mir-580 in Cal-27 and hsc3 cells in different groups (Si-nc, si-CircPUM1 # 1). Knockdown of CircPUM1 increased the level of mir-580 in cells, and the difference was statistically significant ($p < 0.01$, **Figure 3D**). The results showed that the activity of CircPUM1 was negatively correlated with the expression of mir-580. QRT-PCR was used to detect the expression of mir-580 in OSCC cell lines (Cal-27, FaDu, OECM1, SAS, and HSC3) and the human oral mucous fibroblasts (HMFs). The expression of mir-580 in OSCC cell lines was low, and the difference was statistically significant ($p < 0.01$, **Figure 3E**). QRT-PCR was used to detect the expression level of mir-580 in Cal-27 and hsc3 cells in different groups (0gy, 4Gy). After 4Gy radiation, the expression level of mir-580 decreased, and the difference was statistically significant ($p < 0.001$,**Figure 3F**).

STAT3 Is a Downstream Target Gene of miR-580

After that, we analyze the target of CircPUM1 through starBase. Among the targets found, STAT3 scored higher, which predicted that mir-580 and STAT3 had binding sites, as shown in **Figure 4A**. Luciferase reporter gene experiments were carried out in Cal-27 and hsc3 cells to verify the targeting relationship. Compared with mir-nc, overexpression of mir-580 could inhibit the activity of luciferase in cells, and the inhibition disappeared after mutating the

predicted STAT3-binding site ($p < 0.001$, **Figures 4B,C**). Our samples showed that the mir-580 expression level was negatively correlated with STAT3 expression. WB detected the protein expression level of STAT3 after overexpression of mir-580 in Cal-27 and hsc3 cells. Compared with mir-nc, mir-580 overexpression downregulated the protein expression of STAT3, and the difference was statistically significant ($p < 0.001$, **Figure 4D**). WB was used to detect the protein expression level of STAT3 in Cal-27 and hsc3 cells of different groups (Si NC and si-CircPUM1 # 1). Knockdown of CircPUM1 downregulated the protein expression of STAT3, and the difference was statistically significant ($p < 0.01$, **Figure 4E**). WB method was used to detect the expression level of STAT3 in Cal-27 and hsc3 cells after different doses of ionizing radiation (0gy and 4Gy groups). The expression level of STAT3 increased after 4Gy radiation, and the difference was statistically significant ($p < 0.05$,**Figure 4F**).

Overexpression of STAT3 Can Partially Reverse the Effect of Knockdown of CircPUM1 on Radiosensitivity of OSCC Cells

The target of CircPUM1 is analyzed continuously by starBase. The predicted binding sites of CircPUM1 and STAT3 are shown in **Figure 5A**, $p < 0.001$. In addition, studies have shown that upregulation of STAT3 can promote radiotherapy tolerance. Therefore, this study investigated whether STAT3 affects the stability of CircPUM1 protein by regulating the expression of STAT3 in OSCC cells. First, we found that STAT3 was elevated in OSCC tissues and radiation-resistant cells ($p < 0.001$, **Figures 5B,C**). As mentioned previously, CircPUM1 deficiency can make cells sensitive to radiation. Knockdown of si-CircPUM1 # 1 increased the level of apoptosis. After cotransfection with the STAT3 plasmid, the level of apoptosis was partially decreased ($p < 0.001$, **Figure 5D**). Therefore, high expression of STAT3 could partially inhibit the effect of knockdown of CircPUM1 on radiosensitivity of OSCC cells.

DISCUSSION

Recent studies have shown that lncRNA plays a key regulatory role in the pathogenesis, progression, and phenotype development of oral squamous cell carcinoma (Xu et al., 2020). Accumulating data suggest that lncRNA plays a role in oral squamous cell carcinoma (Liu et al., 2017). As mentioned previously, CircPUM1 plays a role in the treatment of various cancers. In our review of published articles, CircPUM1 may be closely associated with the prognosis of colon cancer (Ju et al., 2019), ovarian cancer (Zhang et al., 2019), and gastric cancer (Li et al., 2019a). Scholars have expounded on the different mechanisms of CircPUM1 in different cancers, including regulating the mir-524-5p axis of colon cancer, circc3p1/mir-21, tumor-suppressor gene *PTEN*, and NF-κB and PI3K/Akt pathways (Zhao et al., 2016; Su et al., 2020). However, the role of CircPUM1 in oral squamous cell carcinoma is rare.

CircRNAs are a group of non-coding RNAs with stable closed-loop structures that prevent them from being broken down by enzymes (Li et al., 2019b). Accumulating data suggest that circRNAs are closely related to the progression of NSCLC (Yu et al., 2019). For example, inhibition of circRNA VANGL1 can inhibit bladder cancer progression (Yang et al., 2020). Circ_000984 promotes cell proliferation and metastasis in NSCLC by regulating the Wnt/β-catenin pathway (Li et al., 2019c). Some researchers found that (Pang et al., 2020) the expression of circ_0072309 was downregulated in NSCLC tissues and cells, and the overexpression of circ_0072309 significantly prevented the proliferation, migration, and invasion of cells, which indicated that circ_0072309 played a tumor-suppressor role in NSCLC.

The circular PUM1 RNA (circPUM1, has_circ_0000043) is derived from exon backsplicing of the *PUM1* gene. Recent studies have shown that circPUM1 is highly expressed in endometrial cancer, lung adenocarcinoma, and ovarian cancer tissues (Chen et al., 2019; Guan et al., 2019; Zong et al., 2020). Furthermore, circPUM1 can inhibit tumor development through cavernous microRNAs. These studies suggest that circPUM1 may play an important role in the development of diseases such as tumors. Related studies have reported that circPUM1 can promote the proliferation, invasion, and migration of HCC *in vitro*, and studies have shown that circPUM1 can act as an oncogene of HCC (Zhang et al., 2021). In conclusion, circPUM1 may function as an oncogene in human cancers. In our data, we found that reduction of CircPUM1 induced apoptosis of OSCC cells and enhanced their radiosensitivity.

Aberrant expression of miR-580 in many tumors, such as glioma and breast cancer, has been investigated (Loberg et al., 2006). Also, studies have shown that the level of miR-580 in tumor tissue is significantly higher than that in adjacent normal tissue, which can promote the proliferation, invasion, and migration of HCC (Wang et al., 2021). Mir-580 was identified as a potential target of CircPUM1. Among the discovered targets, mir-580 and CircPUM1 had higher binding sites and predicted values. Furthermore, our data suggest that CircPUM1 can competitively adsorb mir-580 to enhance STAT3 expression in OSCC cells. After we blocked the function of STAT3, OSCC cells triggered the apoptotic pathway and restored their radiosensitivity. This phenotype can be obtained by exogenous mir-580 expression or STAT3 deletion.

Many studies have shown that constitutive STAT3 is activated in a variety of human tumors (Grandis et al., 2000). Evidence suggests that abnormal STAT3 signaling promotes the occurrence and development of human cancers by inhibiting apoptosis, inducing cell proliferation, angiogenesis, invasion, and metastasis (Melinda et al., 2000; Leong et al., 2003; Bollrath et al., 2009; Zhu et al., 2019; Chen et al., 2020), as well as inducing inflammation and immunosuppression (Leaman et al., 1996; Mohan et al., 2022; Tse et al., 2022). In OSCC studies, STAT3 is the most common signal transducer and activator of transcription. STAT3 plays a variety of biological effects in the degree of invasion, lymph node metastasis, and different clinical grades of oral squamous cell carcinoma. It is of great value for early diagnosis and can be used as an important biological indicator for judging prognosis.

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

Conceptualization: LJ and PH; investigation: LJ and TL; original draft preparation: PH, ZZ, LL, LZ, XD, and LL; review and editing: HC, ZZ, LL, and LZ; and funding acquisition: ZZ and LZ.

FUNDING

This research was funded by the Program of Startup Fund for Scientific Research, Fujian Medical University (No. 2018QH1165), the Natural Science Foundation of Fujian Province, (No. 2020J011274), the Scientific Fund of Sanming Science and Technology Bureau (No. 2019-S-9), the Natural Science Foundation of Fujian Province (No. 2020J0112 69), the Natural Science Foundation of Fujian Province, (No. 2021J011 388), and the Natural Science Foundation of Fujian Province (No. 2021J011 3 91).

REFERENCES

Al-Sarraf, M. D. (2007). The Role of Concurrent Chemoradiotherapy in Patients with Head and Neck Cancers: a Review [J]. *Gulf J. Oncol.* 7 (2), 8–16. PMID: 20084719.

Aslam, F., Atique, M., Aslam, M., Sarfraz, T., Ayaz, B. A. B., and Alamgir, W. (2012). Relation of Tumour Thickness with Lymph Node Metastasis in Oral Squamous Cell Carcinoma. *Pak. Armed Forces Med. J.* 62, 529–533.

Berdugo, J., Thompson, L. D. R., Purgina, B., Sturgis, C. D., Tuluc, M., Seethala, R., et al. (2019). Measuring Depth of Invasion in Early Squamous Cell Carcinoma of the Oral Tongue: Positive Deep Margin, Extratumoral Perineural Invasion, and Other Challenges. *Head Neck Pathol* 13, 154–161. doi:10.1007/s12105-018-0925-3

Bollrath, J., Phesse, T. J., von Burstin, V. A., Putoczki, T., Bennecke, M., Bateman, T., et al. (2009). gp130-Mediated Stat3 Activation in Enterocytes Regulates Cell Survival and Cell-Cycle Progression during Colitis-Associated Tumorigenesis. *Cancer Cell.* 15, 91–102. doi:10.1016/j.ccr.2009.01.002

Chen, J., Xu, S., Chen, S., Zong, Z., Han, X., Zhao, Y., et al. (2019). CircPUM1 Promotes the Malignant Behavior of Lung Adenocarcinoma by Regulating miR-326. *Biochem. Biophysical Res. Commun.* 508 (3), 844–849. Epub 2018 Dec 7. PMID: 30528736. doi:10.1016/j.bbrc.2018.11.176

Chen, T., Yu, Q., Xin, L., and Guo, L. (2020). Retracted : Circular RNA circC3P1 Restrains Kidney Cancer Cell Activity by Regulating miR-21/PTEN axis and Inactivating PI3K/AKT and NF- K B Pathways. *J. Cell. Physiol.* 235 (4), 4001–4010. doi:10.1002/jcp.29296

Deng, L., Chen, Q., Xie, J., Wei, W., and Hui, H. (2020). circPUM1 Promotes Polycystic Ovary Syndrome Progression by Sponging to miR-760. *Gene* 754, 144903. Epub 2020 Jun 12. PMID: 32540374. doi:10.1016/j.gene.2020.144903

Grandis, J. R., Drenning, S. D., Zeng, Q., Watkins, S. C., Melhem, M. F., Endo, S., et al. (2000). Constitutive Activation of Stat3 Signaling Abrogates Apoptosis in Squamous Cell Carcinogenesis *In Vivo*. *Proc. Natl. Acad. Sci. U.S.A.* 97 (8), 4227–4232. doi:10.1073/pnas.97.8.4227

Guan, X., Zong, Z.-h., Liu, Y., Chen, S., Wang, L.-l., and Zhao, Y. (2019). circPUM1 Promotes Tumorigenesis and Progression of Ovarian Cancer by Sponging miR-615-5p and miR-6753-5p. *Mol. Ther. - Nucleic Acids* 18, 882–892. Epub 2019 Oct 23. PMID: 31751911; PMCID: PMC6881671. doi:10.1016/j.omtn.2019.09.032

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, 141–157. doi:10.1261/rna.035667.112

Ju, H. Q., Zhao, Q., Wang, F., Lan, P., Wang, Z., Zuo, Z-X., et al. (2019). A circRNA Signature Predicts Postoperative Recurrence in Stage II/III Colon Cancer [J]. *EMBO Mol. Med.* 11 (10), e10168. 201810168. doi:10.15252/emmm

Kennedy, R. A. (2018). WHO Is in and WHO Is Out of the Mouth, Salivary Glands, and Jaws Sections of the 4th Edition of the WHO Classification of Head and Neck Tumours. *Br. J. Oral Maxillofac. Surg.* 56, 90–95. doi:10.1016/j.bjoms.2017.12.009

Leaman, D. W., Leungs, X., Li, X., and Stark, G. R. (1996). Regulation of STAT-dependent Pathways by Growth Factors and Cytokines [J]. *FASEB J.* 10 (14), B1578–B1588. doi:10.1096/fasebj.10.14.9002549

Lee, Y.-J., Kwon, T.-G., Kim, J.-W., Lee, S.-T., Hong, S.-H., and Choi, S.-Y. (2021). Evaluation of Depth of Invasion and Tumor Thickness as a Prognostic Factor for Early-Stage Oral Squamous Cell Carcinoma: A Retrospective Study. *Diagnostics* 12 (1), 20. PMID: 35054187; PMCID: PMC8774458. doi:10.3390/diagnostics12010020

Leong, P. L., Andrews, G. A., Johnson, D. E., Dyer, K. F., Xi, S., Mai, J. C., et al. (2003). Targeted Inhibition of Stat3 with a Decoy Oligonucleotide Abrogates Head and Neck Cancer Cell Growth. *Proc. Natl. Acad. Sci. U.S.A.* 100 (7), 4138–4143. doi:10.1073/pnas.0534764100

Li, H., Jin, X., Liu, B., Zhang, P., Chen, W., and Li, Q. (2019). CircRNA CBL.11 Suppresses Cell Proliferation by Sponging miR-6778-5p in Colorectal Cancer. *BMC Cancer* 19 (1), 826. PMID: 31438886; PMCID: PMC6704711. doi:10.1186/s12885-019-6017-2

Li, X. Y., Liu, Y. R., Zhou, J. H., Li, W., Guo, H. H., and Ma, H. P. (2019). Enhanced Expression of Circular RNA Hsa_circ_000984 Promotes Cells Proliferation and Metastasis in Non-small Cell Lung Cancer by Modulating Wnt/β-Catenin Pathway. *Eur. Rev. Med. Pharmacol. Sci.* 23 (8), 3366–3374. PMID: 31081091. doi:10.26355/eurrev_201904_17700

Li, Y.-F., Zhang, J., and Yu, L. (2019). Circular RNAs Regulate Cancer Onset and Progression via Wnt/β-Catenin Signaling Pathway. *Yonsei Med. J.* 60 (12), 1117–1128. PMID: 31769242; PMCID: PMC6881706. doi:10.3349/ymj.2019.60.12.1117

Liu, Y. H., He, A. B., and Liao, X. H. (2017). Research Progress on Tumor-Related Circular RNA [J]. *China Cancer* 26 (11), 886–892. 2017,11,A009. doi:10.11735/j.issn,1004-0242

Loberg, R. D., Day, L. L., Harwood, J., Ying, C., St. John, L. N., Giles, R., et al. (2006). CCL2 Is a Potent Regulator of Prostate Cancer Cell Migration and Proliferation. *Neoplasia* 8 (7), 578–586. PMID: 16867220; PMCID: PMC1601934. doi:10.1593/neo.06280

Melinda, H., Carmen, P., Kevin, R., and Lin, J. (2000). Constitutive Activation of Stat 3 Oncogene Product in Human Ovarian Carcinoma Cells [J]. *Gyn Ecol. ic Onco logy* 79 (1), 67–73. doi:10.1006/gyno.2000.5931

Mohan, C. D., Yang, M. H., Rangappa, S., Chinnathambi, A., Alharbi, S. A., Alahmadi, T. A., et al. (2022). 3-Formylchromone Counteracts STAT3 Signaling Pathway by Elevating SHP-2 Expression in Hepatocellular Carcinoma. *Biology* 11, 29. doi:10.3390/biology11010029

Müller, S. (2017). Update from the 4th Edition of the World Health Organization of Head and Neck Tumours: Tumours of the Oral Cavity and Mobile Tongue. *Head. Neck Pathol.* 11, 33–40. doi:10.1007/s12105-017-0792-3

Pang, W., Huang, F., Zhang, X., Ye, M., Huang, Y., Huang, X., et al. (2020). Circular RNA Hsa_circ_0072309 Inhibits Non-small Cell Lung Cancer Progression by Sponging miR-580-3p. *Biosci. Rep.* 40 (5), BSR20194237. doi:10.1042/BSR20194237

Pentenero, M., Gandolfo, S., and Carrozzo, M. (2005). Importance of Tumor Thickness and Depth of Invasion in Nodal Involvement and Prognosis of Oral Squamous Cell Carcinoma: A Review of the Literature. *Head. Neck* 27, 1080–1091. doi:10.1002/hed.20275

Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., et al. (2015). Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. *Mol. Cell.* 58, 870–885. doi:10.1016/j.molcel.2015.03.027

Salzman, J., Chen, R. E., Olsen, M. N., Wang, P. L., and Brown, P. O. (2013). Cell-Type Specific Features of Circular RNA Expression. *PLoS Genet.* 9, e1003777. doi:10.1371/journal.pgen.1003777

Su, Y., Feng, W., Shi, J., Chen, L., Huang, J., and Lin, T. (2020). circRIP2 Accelerates Bladder Cancer Progression via miR-1305/Tgf-B2/smad3 Pathway. *Mol. Cancer* 19 (1), 23. Erratum in: Mol Cancer. 2021 Jan 2;20(1):1. PMID: 32019579; PMCID: PMC6998850. doi:10.1186/s12943-019-1129-5

Tse, J., Pierce, T., Carli, A. L. E., Alorro, M. G., Thiem, S., Marcusson, E. G., et al. (2022). Onco-miR-21 Promotes Stat3-dependent Gastric Cancer Progression. *Cancers* 14, 264. doi:10.3390/cancers14020264

Wang, X., Sheng, W., Xu, T., Xu, J., Gao, R., and Zhang, Z. (2021). CircRNA Hsa_circ_0110102 Inhibited Macrophage Activation and Hepatocellular Carcinoma Progression via miR-580-5p/PPARα/CCL2 Pathway. *Aging* 13 (8), 11969–11987. Epub 2021 Apr 23. PMID: 33891564; PMCID: PMC8109088. doi:10.18632/aging.202900

Xu, X., Zhang, J., Tian, Y., Gao, Y., Dong, X., Chen, W., et al. (2020). CircRNA Inhibits DNA Damage Repair by Interacting with Host Gene. *Mol. Cancer* 19, 128. doi:10.1186/s12943-020-01246-x

Yang, D., Qian, H., Fang, Z., Xu, A., Zhao, S., Liu, B., et al. (2020). Silencing Circular RNA VANGL1 Inhibits Progression of Bladder Cancer by Regulating miR-1184/IGFBP2 axis. *Cancer Med.* 9 (2), 700–710. Epub 2019 Nov 23. PMID: 31758655; PMCID: PMC6970048. doi:10.1002/cam4.2650

Yu, D., Ma, Y., Feng, C., Ma, Z., Guo, J., Chen, H., et al. (2020). PBX1 Increases the Radiosensitivity of Oesophageal Squamous Cancer by Targeting of STAT3. *Pathol. Oncol. Res.* 26 (4), 2161–2168. Epub 2020 Mar 13. PMID: 32170580. doi:10.1007/s12253-020-00803-5

Yu, T., Wang, Y., Fan, Y., Fang, N., Wang, T., Xu, T., et al. (2019). CircRNAs in Cancer Metabolism: a Review. *J. Hematol. Oncol.* 12 (1), 90. PMID: 31484561; PMCID: PMC6727394. doi:10.1186/s13045-019-0776-8

Zhang, Q., Zhang, C., Ma, J.-X., Ren, H., Sun, Y., and Xu, J.-Z. (2019). Circular RNA PIP5K1A Promotes Colon Cancer Development through Inhibiting miR-1273a. *Wjg* 25 (35), 5300–5309. PMID: 31558874; PMCID: PMC6761237. doi:10.3748/wjg.v25.i35.5300

Zhang, Y., Wang, D., Zhu, T., Yu, J., Wu, X., Lin, W., et al. (2021). CircPUM1 Promotes Hepatocellular Carcinoma Progression through the miR-1208/MAP3K2 axis. *J. Cell. Mol. Med.* 25 (1), 600–612. doi:10.1111/jcmm.15998

Zhao, L., Wu, P., Zhang, P. G., Xie, D. Z., Gao, G., Zhou, N. J., et al. (2016). Effect of Triptolide on Expression of Bax/Bcl-2 and Cleaved Caspase 3 in HCT116 Cells [J]. *World Chin. J. Dig.* 24, 3580–3586. doi:10.11569/wcjdv24.i24.3580

Zheng, Y. Y., Li, W., and Chen, Y. Q. (2019). Research Progress on the Correlation between circRNAs and Tumors [J]. *Chin. J. General Pract.* 17 (7), 1181–1185. doi:10.16766/j.cnki.issn.1674-4152.000896

Zhu, C. Y., Meng, F. Q., and Liu, J. (2019). MicroRNA-524-5p Suppresses Cell Proliferation and Promotes Cell Apoptosis in Gastric Cancer by Regulating CASP3 [J]. *Eur. Rev. Med. Pharmacol. Sci.* 23 (18), 7968–7977. doi:10.26355/eurrev_201909_19013

Zong, Z. H., Liu, Y., Chen, S., and Zhao, Y. (2020). Circ_PUM1 Promotes the Development of Endometrial Cancer by Targeting the miR-136/NOTCH3 Pathway. *J. Cell. Mol. Med.* 24 (7), 4127–4135. Epub 2020 Feb 19. PMID: 32073729; PMCID: PMC7171399. doi:10.1111/jcmm.15069

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.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 29 August 2022
DOI 10.3389/fgene.2022.970507

OPEN ACCESS

EDITED BY
Zhongxiao Wang,
Biogen Idec, United States

REVIEWED BY
Yajuan Zhou,
Hubei Cancer Hospital, China
Wanzun Lin,
Shanghai Cancer Center, Fudan University, China

*CORRESPONDENCE
Conghua Xie,
chxie_65@whu.edu.cn

†These authors contributed equally to this work

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

RECEIVED 16 June 2022
ACCEPTED 03 August 2022
PUBLISHED 29 August 2022

CITATION
Xu L, Huang Z, Zeng Z, Li J, Xie H and Xie C (2022), An integrative analysis of DNA methylation and gene expression to predict lung adenocarcinoma prognosis. *Front. Genet.* 13:970507. doi: 10.3389/fgene.2022.970507

An integrative analysis of DNA methylation and gene expression to predict lung adenocarcinoma prognosis

Liexi Xu[1†], Zhengrong Huang[1,2†], Zihang Zeng[1], Jiali Li[1], Hongxin Xie[1] and Conghua Xie[1,3*]

[1]Department of Radiation and Medical Oncology, Wuhan University of Zhongnan Hospital, Wuhan, China, [2]Tumor Precision Diagnosis and Treatment Technology and Translational Medicine, Hubei Engineering Research Center, Zhongnan Hospital of Wuhan University, Wuhan, China, [3]Hubei Key Laboratory of Tumor Biological Behaviors, Zhongnan Hospital of Wuhan University, Wuhan, China

Background: Abnormal DNA methylation of gene promoters is an important feature in lung adenocarcinoma (LUAD). However, the prognostic value of DNA methylation remains to be further explored. *Objectives.* We sought to explore DNA methylation characteristics and develop a quantifiable criterion related to DNA methylation to improve survival prediction for LUAD patients.

Methods: Illumina Human Methylation450K array data, level 3 RNA-seq data and corresponding clinical information were obtained from TCGA. Cox regression analysis and the Akaike information criterion were used to construct the best-prognosis methylation signature. Receiver operating characteristic curve analysis was used to validate the prognostic ability of the DNA methylation-related feature score. qPCR was used to measure the transcription levels of the identified genes upon methylation.

Results: We identified a set of DNA methylation features composed of 11 genes (*MYEOV*, *KCNU1*, *SLC27A6*, *NEUROD4*, *HMGB4*, *TACR3*, *GABRA5*, *TRPM8*, *NLRP13*, *EDN3* and *SLC34A1*). The feature score, calculated based on DNA methylation features, was independent of tumor recurrence and TNM stage in predicting overall survival. Of note, the combination of this feature score and TNM stage provided a better overall survival prediction than either of them individually. The transcription levels of all the hypermethylated genes were significantly increased after demethylation, and the expression levels of 3 hypomethylated proteins were significantly higher in tumor tissues than in normal tissues, as indicated by immunohistochemistry data from the Human Protein Atlas. Our results suggested that these identified genes with prognostic features were regulated by DNA methylation of their promoters.

Conclusion: Our studies demonstrated the potential application of DNA methylation markers in the prognosis of LUAD.

KEYWORDS

DNA methylation, integrative analysis, lung adenocarcinoma, prognosis, TCGA

1 Introduction

Lung adenocarcinoma (LUAD) is the most common histological subtype of lung cancer, accounting for approximately 50% of all lung cancer cases in most countries (Goldstraw et al., 2011; Bray et al., 2018). Previous studies have revealed that in addition to cigarette smoking, risk factors such as age, environmental pollution, occupational exposure, race, sex, and preexisting lung disease are also substantially involved in lung cancer. With the development and popularization of public databases in recent years, an increasing number of researchers have tried to identify prognostic biomarkers for LUAD by analyzing clinical characteristics and molecular information (Edge and Compton, 2010). The TNM staging system of the American Joint Commission on Cancer (AJCC) was reported to have great value in LUAD prognosis (Folkman, 1971). Liu et al. stated that conventional staging alone was not enough to predict prognosis and guide treatment decisions. They analyzed large cohorts from The Cancer Genome Atlas (TCGA) database and developed a 4-gene feature related to glycolysis (Kaishang et al., 2018). Su et al. identified an RNA sequencing network of 29 key lncRNAs, 72 mRNAs and 24 miRNAs as potential biomarkers to optimize the diagnosis and prognosis of LUAD patients by using the TCGA database (Li et al., 2017). The findings from these studies indicate that it is feasible to use different molecular markers and clinical features in public databases to establish practical models that have great application potential. Although the effectiveness of these prediction models has not been tested in clinical practice, it is necessary to continue to mine and improve the gene signatures related to the prognosis of LUAD.

Epigenetic disorders, especially abnormal DNA methylation in gene promoters, are a fundamental feature of human malignant tumors (Liu et al., 2019). As one of the most well-studied epigenetic modifications, DNA methylation mainly occurs at 5′-cytosine-phosphate-guanine-3’ (CpG) dinucleotides and is regulated by DNA methyltransferases and DNA demethylases (Yang et al., 2022). Methylation and cancer formation are associated in 2 main ways: one is the regulation of tumor suppressor gene expression by gene hypermethylation in the promoter, and the other is genome-wide hypomethylation, which plays an important role in the stability of the heterochromatin structure (Feinberg, 2007). In virtually every step of tumor progression, there is abnormal promoter methylation regulation (Sui et al., 2016). *APC*, *CDH13*, *MLH1* and *IRX1* have hypermethylation in promoter CpG islands (CGIs). The hypermethylation of the *APC* and *CDH13* genes in LUAD is associated with cancer cell adhesion, and the loss of *MLH1* and *IRX1* expression is associated with poor tumor survival (Goto et al., 2009; Küster et al., 2020). The hypomethylation of *LINE-1* and *ELF3* induces protein overexpression in LUAD. The overexpression of *ELF3* can stimulate the carcinogenic phenotype of LUAD cells and reduce the survival time of patients, suggesting that the hypomethylation of *LINE-1* is a prognostic marker of LUAD development and progression (Ikeda et al., 2013; Enfield et al., 2019).

TCGA has disclosed the clinical information of more than 10,000 patients and the molecular phenotype information of their tumor tissues. This information covers 33 different types of tumors and multiple data from different sources, including transcriptomic, methylomic and proteomic sources (Mazor et al., 2016). By integrating data from different sources, we can identify specific events in the carcinogenic process and identify potential biomarkers associated with patient survival.

In this study, we obtained the TCGA Illumina Human Methylation 450K microarray data, RNA-seq data, and clinical data of LUAD patients and performed an integrative analysis to identify a set of DNA methylation features for 11 genes. We performed area under the receiver operating characteristic curve (AUC-ROC) analysis to verify the ability of the identified DNA methylation feature to predict the survival of LUAD. In addition, we performed qPCR, and the results suggested that these identified genes with prognostic features were regulated by DNA methylation in their promoters.

2 Materials and methods

2.1 Data preparation

The steps of data acquisition and analysis, as well as methylation feature acquisition and verification, are shown in the flow chart (Figure 1). Illumina Human Methylation 450K array data were obtained from TCGA, and a total of 24,587 DMSs from 478 pretreated methylation arrays were screened using the camp and Minfi R software packages. After that, 478 samples (449 LUAD samples and 29 normal samples) were included after being filtered, inspected, and standardized with the ChAMP R package. Level 3 RNA-seq data from TCGA were normalized and log2 transformed by the edgeR package. For the preprocessing of clinical information corresponding to the sample, patients with nonsurvival status or survival time less than 1 month were excluded because of other disease-related deaths.

2.2 Differential methylation analysis and differential expression analysis

In total, 449 LUAD samples and 29 normal samples were subjected to differential methylation analysis with the ChAMP R package and the Minfi R package (Li et al., 2019). Principal component analysis was used to detect the sample quality. The ChAMP DMP function and the Minfi R package defined the methylation loci with an average methylation difference >0.2 and a false discovery rate <0.05 as differentially methylated sites (DMSs), and the final DMSs were obtained through the

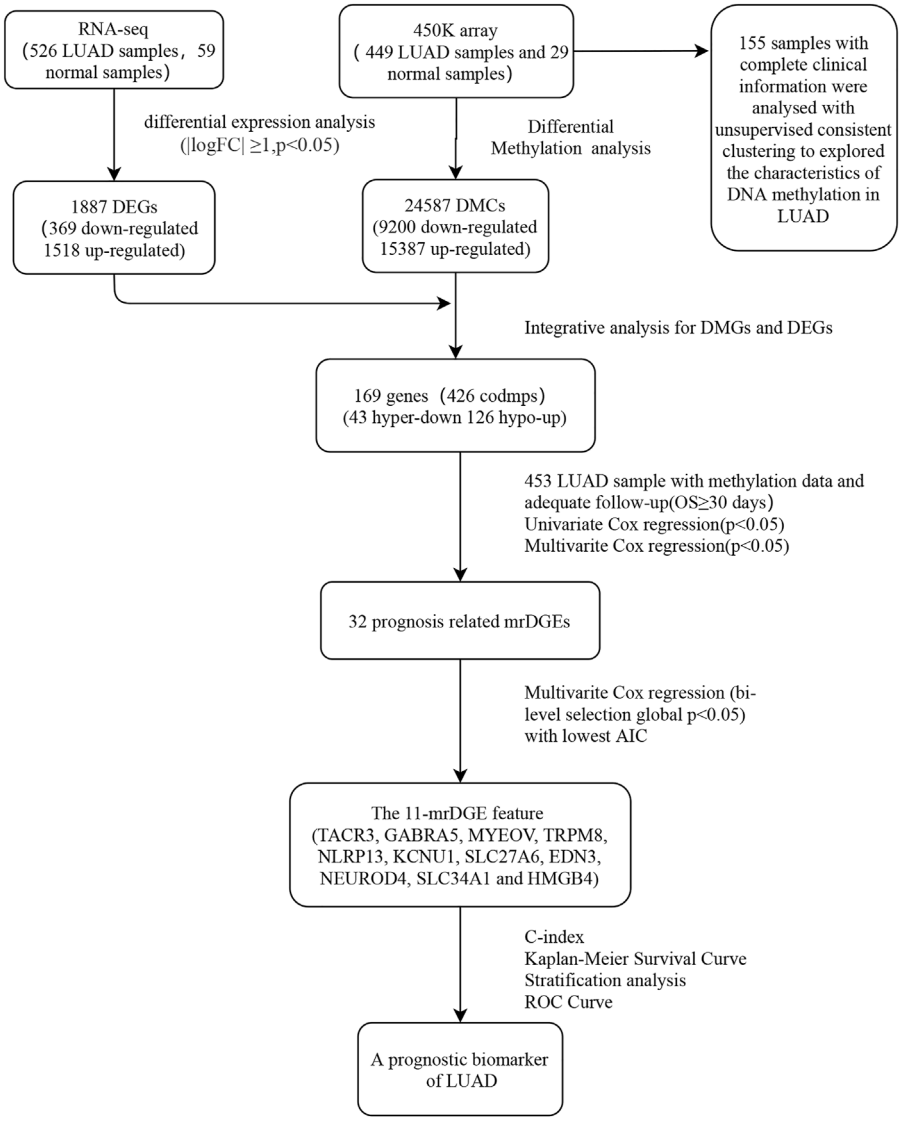


FIGURE 1
Flow chart for obtaining and verifying methylation feature.

intersection of the two (Li et al., 2018). Differentially expressed genes (DEGs) between the 526 LUAD samples and the 59 normal samples were analyzed with the Limma R package ($p < 0.05$ and $|\log2FC| \geq 1$). Metascape (https://metascape.org/gp/index.html#/main/step1) was used to analyze the pathway enrichment of the hyper-down and hypo-up methylated related differential expression genes (mrDEGs) groups.

2.3 Survival model construction process

A prognosis prediction model was established according to the DNA methylation β value of mrDEGs and matched prognostic data of patients. According to the methylation β value, univariate Cox regression analysis was used to screen mrDEGs ($p < 0.01$) that were significantly associated with overall survival (OS). Then, mrDEGs identified in the univariate Cox regression analysis were subjected to multivariate Cox regression analysis (Lian et al., 2019). At the same time, the Akaike information criterion (AIC) was used to screen out the genes with subtly individual but significantly synergistic effects to determine the most appropriate gene feature (Tozzi et al., 2020). A Kaplan–Meier (K-M) curve with a log rank test was used to validate the survival difference of patients (Zhao et al., 2020). Harrell's concordance index (C-index) and the corresponding 95% confidence intervals

TABLE 1 All primer sequences used in qPCR.

Primer name	5′ Sequence 3′
EDN3 Fp	ATTGCCACCTGGACATCATT
EDN3 Rp	GCAGGCCTTGTCATATCTCC
TACR3 Fp	TTCATCCAAACCGGCAAAGC
TACR3 Rp	AAACTTGGGTCTCTTGGCGT
SLC27A6 Fp	AAAAAGGGGGACACGGTG
SLC27A6 Rp	AGGAGGGAGTTGGAGCGA

(CIs) were calculated to determine the prognostic model's ability. These steps were performed by R with the survival and survcomp R packages (Schröder et al., 2011).

2.4 Consensus clustering analysis

We selected 155 LUAD samples with complete clinical information. For our standard, we considered the standard deviation of the β value in tumor samples to be greater than 0.2 and the average β value in normal tissues to be less than 0.05. With this approach, we selected 641 methylation probes according to the standard. According to the PAM algorithm and Euclidean distance, we then performed unsupervised consistent clustering on 641 probes of 155 samples. The Consensus Cluster Plus R package was used for the clustering analysis (Wilkerson and Hayes, 2010). The Kruskal–Wallis test was used to validate the significance of clinical features among clusters.

2.5 Validation experiments in cell lines

qPCR was used to verify the changes in gene transcription levels upon methylation. A549, PC9 and H1975 cells were purchased from the Canadian Standards Association (CSA). All cell lines were cultured in RPMI 1640 with 10% fetal bovine serum. All experimental cells were treated with 5-Aza-2′- deoxycytidine (5-aza, Aladdin) for 96 h at 1 μM. qPCR analyses of all cell lines were repeated at least 3 times (Christman, 2002). All primer sequences used in qPCR are listed in Attachment 1: Table 1.

2.6 Statistical analysis

The correlation between feature scores and clinical factors was analyzed by the chi square or Fisher exact test (Jung, 2014; Pandis, 2016). Multivariate Cox regression combined with hierarchical data analysis was used to evaluate the predictive power of the clinical features, TNM stage, and methylation feature score for prognosis. The forest map was drawn by Prism, and other statistical tests were performed by R using the corresponding R packages.

3 Results

3.1 Differential methylation and the identification of mrDMGs

Illumina Human Methylation 450K array data were obtained from TCGA. We screened a total of 24,587 DMSs from data from 478 pretreated methylation arrays using the camp and Minfi R software packages. We then divided DMSs into 15,387 hypermethylated and 9,200 hypomethylated sites and evaluated their distribution in the genome. Compared to 63% in the whole genome, hypermethylated sites increased significantly in the promoter, CGIs, and CGI promoters (71%, 96% and 99%, respectively, Figure 2A). At the same time, most DMSs on CGIs were hypermethylated (96%), and most DMSs on shelf CpG positions were hypomethylated (Figure 2B). When we detected the distribution of DMSs around the gene, we found that the hypermethylation of CpGs was higher near the transcription start site (TSS). For example, the proportions of hypermethylated CpGs in the 5′UTR, tss200 and first exon were 70%, 77%, and 76%, respectively (Figure 2C). Next, we located DMSs on the gene and obtained 5,900 differentially methylated genes (DMGs) (Figure 2D). Next, we identified 1,887 DEGs based on the RNA-seq data. Then, 406 mrDEGs were determined through the intersection of the DMGs and the identified DEGs (Figure 2E). Among them, 43 mrDEGs were in the hypermethylation downregulation group (hyper-down group), and 126 mrDEGs were in the hypomethylation upregulation group (hypo-up group) (Figure 2F).

3.2 mrDEGs involved in biological processes

The Metascape website was used to analyze the pathway enrichment of the hyper-down and hypo-up mrDEGs. In the hypo-up group, the genes showed a significant abundance in fatty acid degradation, cyclic adenosine monophosphate (cAMP)-mediated signaling, glycolysis/gluconeogenesis, etc. (Figure 3A). Cancer is usually accompanied by nutritional metabolic imbalances, such as abnormal glucose and lipid metabolism (Li and Liao, 2021). cAMP was the first second messenger to be discovered, and it plays key roles in physiological defects caused by metabolic disorders (Zhang et al., 2020; Chi et al., 2021). Interestingly, in the hyper-down group, there was also gene enrichment related to fatty acid degradation (Figure 3B). The effects of lipid metabolism disorders on cancer have attracted increasing attention in recent years (Karagiota et al., 2022). There is no doubt that to reprogram their metabolic state and ensure cell survival, tumor cells need

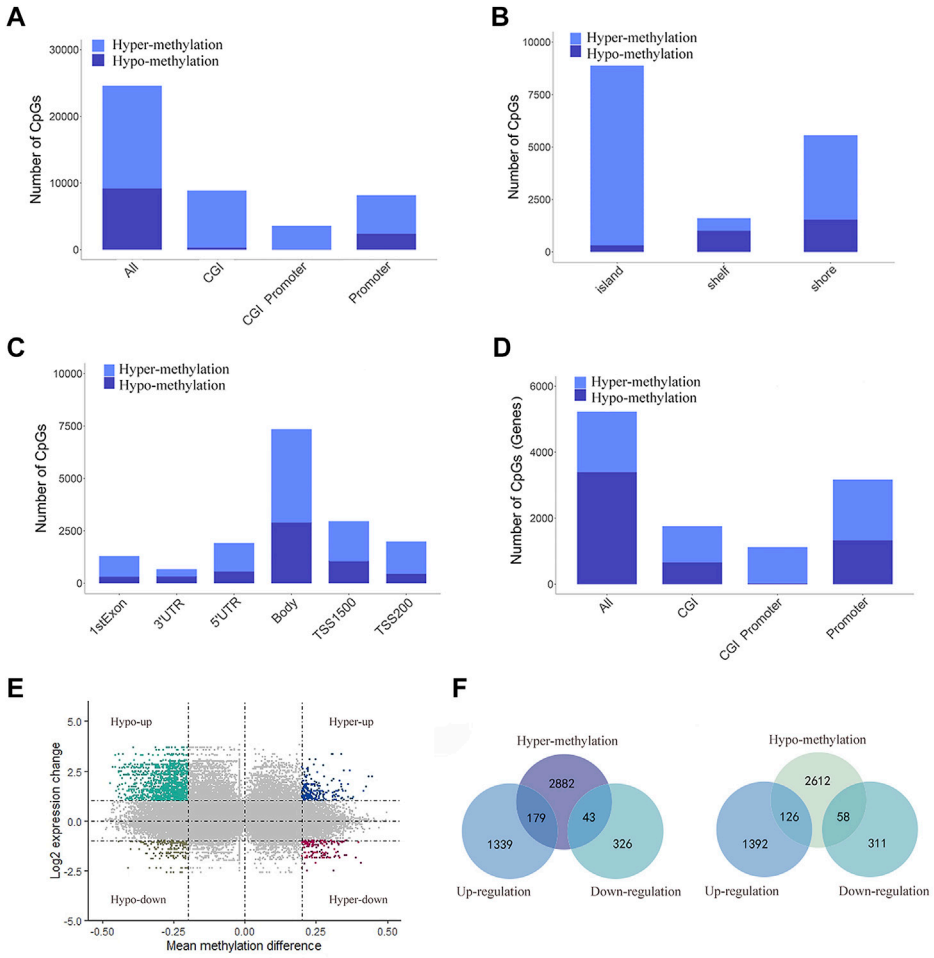


FIGURE 2
Distribution of DMSs and obtain mrDEGs of LUAD. **(A)** Distribution of DMSs across various genomic regions, including CpG islands (CGI), promoters, CGI promoter, and the whole genome (all). **(B)** Distribution of DMSs in various areas related to CGI distance, including CpG shelves, CpG shores and CpG islands. **(C)** Distribution of DMSs in gene location, including 3′ UTRs, gene bodies, first exons, 5′ UTRs, TSS200 and TSS1500. **(D)** Distribution of DMGs across various genomic regions. **(E)** Scatter plot shows mean methylation difference versus log2 expression change, and each point represents a pair of methylation site and gene. **(F)** Venn diagrams shows the intersection between DEGs and hypermethylated genes (left) and between DEGs and hypomethylated genes (right).

epigenetic modifications to regulate gene expression. For example, brother of the regulator of imprinted sites can mediate the Warburg effect and promote breast cancer by regulating the methylation of pyruvate kinase M1/2 (PKM) exons (Singh et al., 2017; Huo et al., 2021). Our results suggest that abnormal lipid metabolism in LUAD may be closely mediated by DNA methylation (Figure 3B).

3.3 Identification of prognostic mrDEGs

We established a prognosis prediction model according to the methylation data of mrDEGs and the matched prognostic data. We first analyzed 406 mrDEGs in 453 patients by univariate Cox regression and identified 32 mrDEGs related to prognosis ($p < 0.05$). Next, multivariate Cox regression analysis was performed on the 32 mrDEGs. The AIC, as the indicator for model fitness, determined the most suitable prognostic model. Finally, we identified 11 mrDEGs (*MYEOV*, *KCNU1*, *SLC27A6*, *NEUROD4*, *HMGB4*, *TACR3*, *GABRA5*, *TRPM8*, *NLRP13*, *EDN3* and *SLC34A1*) to be included in a DNA methylation feature prognostic model. There were 5 genes (*MYEOV*, *KCNU1*, *SLC27A6*, *NEUROD4* and *HMGB4*) with statistically nonsignificant *p* values in multivariate Cox regression analysis (Table 2). However, the AIC of this prognostic model was the lowest (AIC = 1733.8, p = 2e-05), indicating that this model was

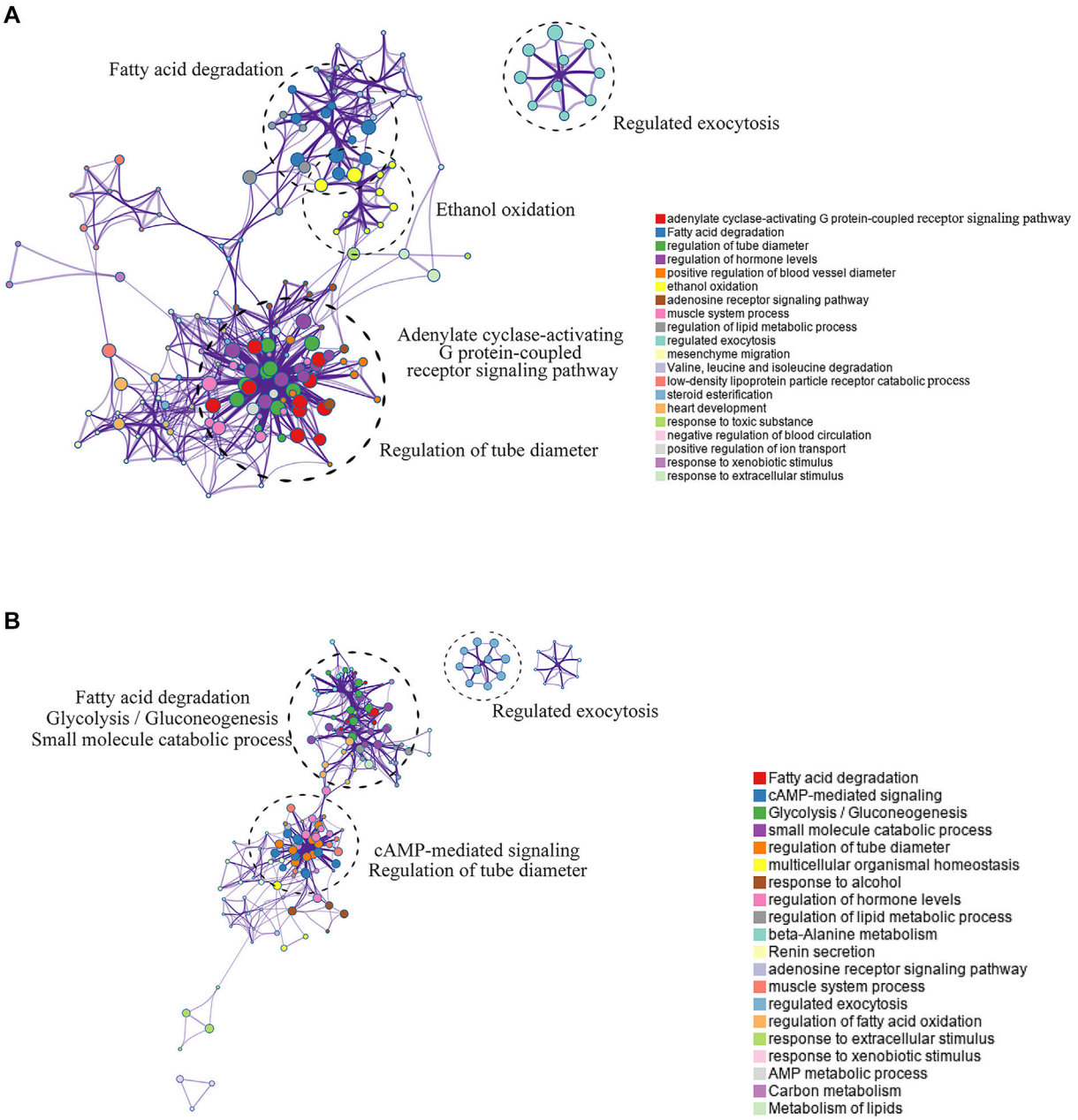


FIGURE 3
Pathway enrichment analysis of mrDEGs in LUAD. **(A)** The pathway enrichment analysis of the upregulated mrDEGs. Each node represents a gene group. The node size is proportional to the total number of genes in each gene set. The width of the line between nodes represents the proportion of genes shared among gene sets. **(B)** The pathway enrichment analysis of the downregulated mrDEGs.

the most suitable, and the overall effect of the model was significant. The C-index of the identified DNA methylation feature model was 0.666 (95% CI = 0.641–0.690), indicating great discrimination ability. The correlation between methylation level and gene expression of these 11 genes is shown in the appendix (Supplementary Figure.S1).

3.4 DNA methylation feature model for predicting the OS of LUAD patients

According to the correlation coefficients of the eleven mrDEGs obtained by multivariate Cox regression analysis, we established a feature score formula.

TABLE 2 Eleven mrDEGs identified as a DNA methylation signature prognostic model.

Gene symbol	Full name	Chr	Coefficient	*p* value
TACR3	Tachykinin receptor 3	4q24	2.513	0.049
GABRA5	Gamma-aminobutyric acid type A receptor subunit Alpha5	15q12	1.74	0.028
MYEOV	Myeloma overexpressed	11q13.3	−0.752	0.147
TRPM8	Transient receptor potential cation channel subfamily M member 8	2q37.1	−1.611	0.008
NLRP13	NLR family pyrin domain containing 13	19q13.43	−2.074	0.01
KCNU1	Potassium calcium-activated channel subfamily U member 1	8p11.23	−1.158	0.072
SLC27A6	Solute carrier family 27 member 6	5q23.3	−1.694	0.398
EDN3	Endothelin 3	20q13.32	−2.059	0.01
NEUROD4	Neuronal differentiation 4	12q13.2	−1.243	0.098
SLC34A1	Solute carrier family 34 member 1	5q35.3	1.626	0.004
HMGB4	High mobility group box 4	1p35.1	0.91	0.138

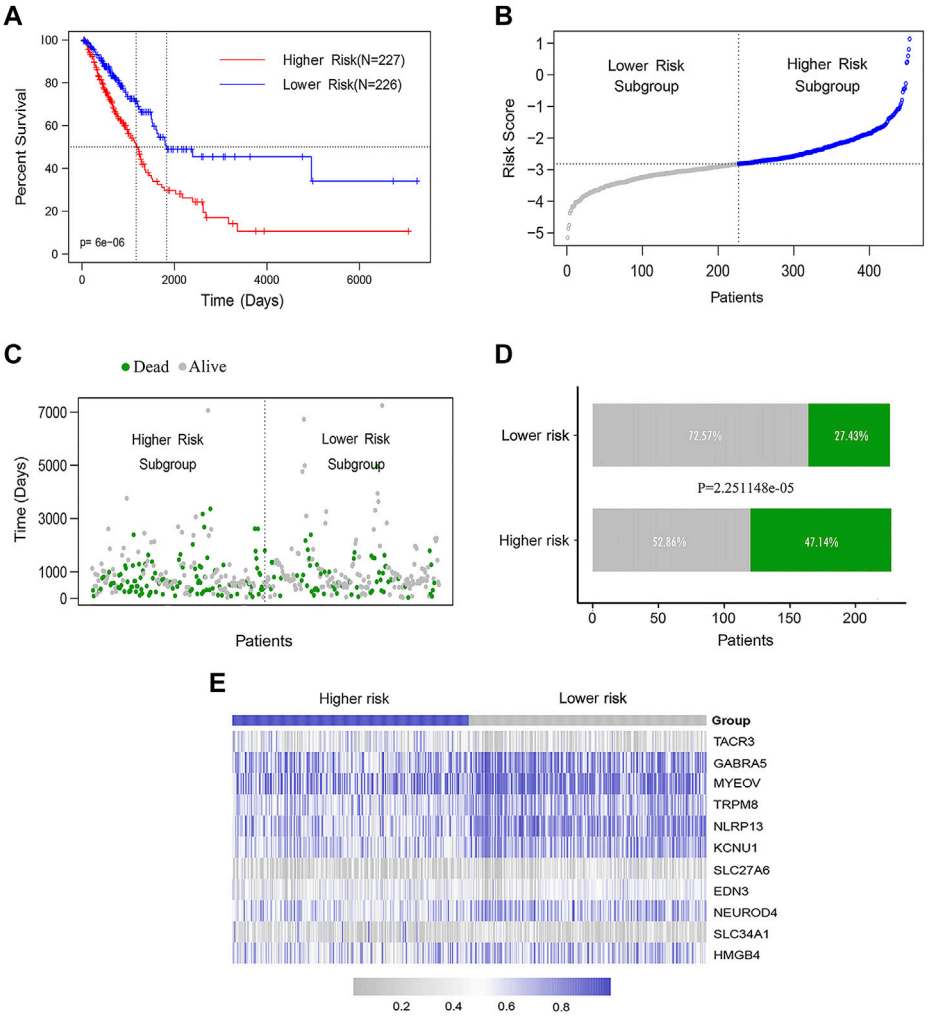


FIGURE 4
The DNA methylation feature for predicting OS prediction in LUAD patients. **(A)** K-M assessed OS based on the DNA methylation feature. The LUAD patients were divided into the lower-risk (n = 226) and higher-risk (n = 227) subgroups according to the median of the methylation scores. Log rank test was used between curves ($p < 0.0001$). **(B)** The distribution of feature scores for DNA methylation feature of patients. **(C–D)** The distribution of survival status of LUAD patients in the lower- and higher-risk groups (Chi-square test, $p < 0.0001$). **(E)** The methylation β value spectrum of 11 DNA methylation feature genes.

$$\begin{aligned}\text{feature score} = {} & (2.513 * methylation\,\beta\ value\ of\ TACR3) \\ & + (1.740 * methylation\,\beta\ value\ of\ GABRA5) \\ & + (-0.752 * methylation\,\beta\ value\ of\ MYEOV) \\ & + (-1.611 * methylation\,\beta\ value\ of\ TRPM8) \\ & + (-2.074 * methylation\,\beta\ value\ of\ NLRP13) \\ & + (-1.158 * methylation\,\beta\ value\ of\ KCNU1) \\ & + (-1.694 * methylation\,\beta\ value\ of\ SLC27A6) \\ & + (-2.059 * methylation\,\beta\ value\ of\ EDN3) \\ & + (-1.243 * methylation\,\beta\ value\ of\ NEUROD4) \\ & + (1.626 * methylation\,\beta\ value\ of\ SLC34A1) \\ & + (0.910 * methylation\,\beta\ value\ of\ HMGB4)\end{aligned}$$

The LUAD patients were ranked according to their calculated methylation-related feature scores and divided into higher-risk (n = 226) and lower-risk groups (n = 227) according to the median. The K-M curve showed that the median OS of the higher-risk group was significantly shorter than that of the lower-risk group (log rank test $p < 0.0001$) (Figure 4A). We also analyzed the distribution of the methylation feature scores, patient survival statuses and methylated β values in LUAD patients, as well as the methylation β value spectra of 11 DNA methylation feature genes (Figures 4B–E).

3.5 DNA methylation feature model with clinicopathological features

First, we attempted to validate the correlation between DNA methylation levels and clinicopathological features in LUAD. We then performed unsupervised consistent clustering of the 641 most variable DNA methylation probes in 155 samples (with complete clinical information) into 4 clusters: CGI Methylator Phenotype (CIMP) high, CIMP medium high, CIMP medium low and CIMP low (Figure 5A). The average methylation levels among the different clusters were significant ($p < 2.2e-16$) (Figure 5B). DNA methylation was significantly correlated with tumor subtype (X-squared = 31.457, $p = 2.073e-05$) (Figure 5C). Most of magnoid tumors were enriched in the CIMP low and CIMP medium low clusters, while the CIMP high group had more squamoid tumors Furthermore, DNA methylation showed a trend related to tumor recurrence; however, there was no statistical Figure 5A Significance ($p = 0.2864$) (Figure 5D).

Then, we analyzed the correlation between DNA methylation feature scores and clinicopathological features. The results showed that the methylation feature score was significantly correlated with tumor subtype (Table 3, $p = 0.032$). In addition, the DNA methylation feature score was associated with smoking history in LUAD patients (Table 3, $p = 0.004$). Previous studies have shown that smoking is associated with methylation levels. For example, hypomethylation at cg05575921 in the aryl hydrocarbon receptor repressor gene was strongly associated with the smoking behavior of an individual (Jamieson et al., 2020). Therefore, the hypothesis that prognostic signals are related to smoking is reasonable. To delve into the effects of DNA methylation and clinicopathological features on prognosis, 391 patients with complete clinicopathological features were analyzed in a Cox regression model. The forest map showed that the feature score (HR = 1.69, 95% CI = 1.17–2.45, $p = 0.006$), the TNM staging system, and tumor recurrence (HR = 2.58, 95% CI = 1.79–3.72, $p < 0.0001$) were independent prognostic factors for LUAD, while smoking history and tumor subtype were not (Figure 5E). DNA methylation is closely related to tumor immune microenvironment (Chiappinelli and Baylin, 2022; Li et al., 2022). We investigated the relationship between methylation feature scores and tumor immune infiltration. In the group with low methylation feature scores, we observed an increase in monocyte, dendritic cell (resting) and mast cell (resting) infiltration, as well as a decrease in macrophage (M0) cell infiltration. However, there is no changes of immune effector cells observed (Supplementary Figure S2).

3.6 Prognostic value of the DNA methylation feature score is independent of TNM stage and cancer recurrence

Since a high feature score, tumor recurrence and a high TNM stage were independent adverse prognostic factors for LUAD (Figure 5D), we performed a combined analysis between DNA methylation features and the other 2 influencing factors. We found that the prognosis of patients in the higher-feature score group was poorer, whether in the recurrence (log rank test, $p < 0.0009$) or nonrecurrence subgroup (log rank test, $p = 0.02$) (Figure 6A). In the combined analysis of TNM stages and DNA methylation features, we found that patients in the lower TNM stage (I and II) subgroups had a notably worse prognosis when they were also in the high-feature score subgroups (p = 2e-05) but not in the higher TNM stage (III and IV) subgroups ($p = 0.3$) (Figure 6B). We further classified patients at low TNM stages and found that the p value of the K-M curve in the stage I subgroup (p = 9e-04) was more significant than that in the stage II subgroup ($p = 0.01$) (Figures 6C,D). These results suggested that the DNA methylation feature score was more valuable in patients at a lower TNM stage. AUC-ROC analysis was used to evaluate the sensitivity and specificity of the prediction model (Figure 6E). The combination of this feature score and TNM stage was significantly superior to that of TNM stage alone (0.697 vs 0.658, $p = 0.0275$) or feature score alone (0.697 vs 0.603, $p = 0.0001$). These results suggested that the combination of the DNA methylation feature score and TNM stage might help to improve OS prediction in LUAD patients.

3.7 The expression of the eleven identified genes

The prognostic methylation signature consists of 11 genes. Three of them are hypermethylated in LUAD (*TACR3*, *EDN3*

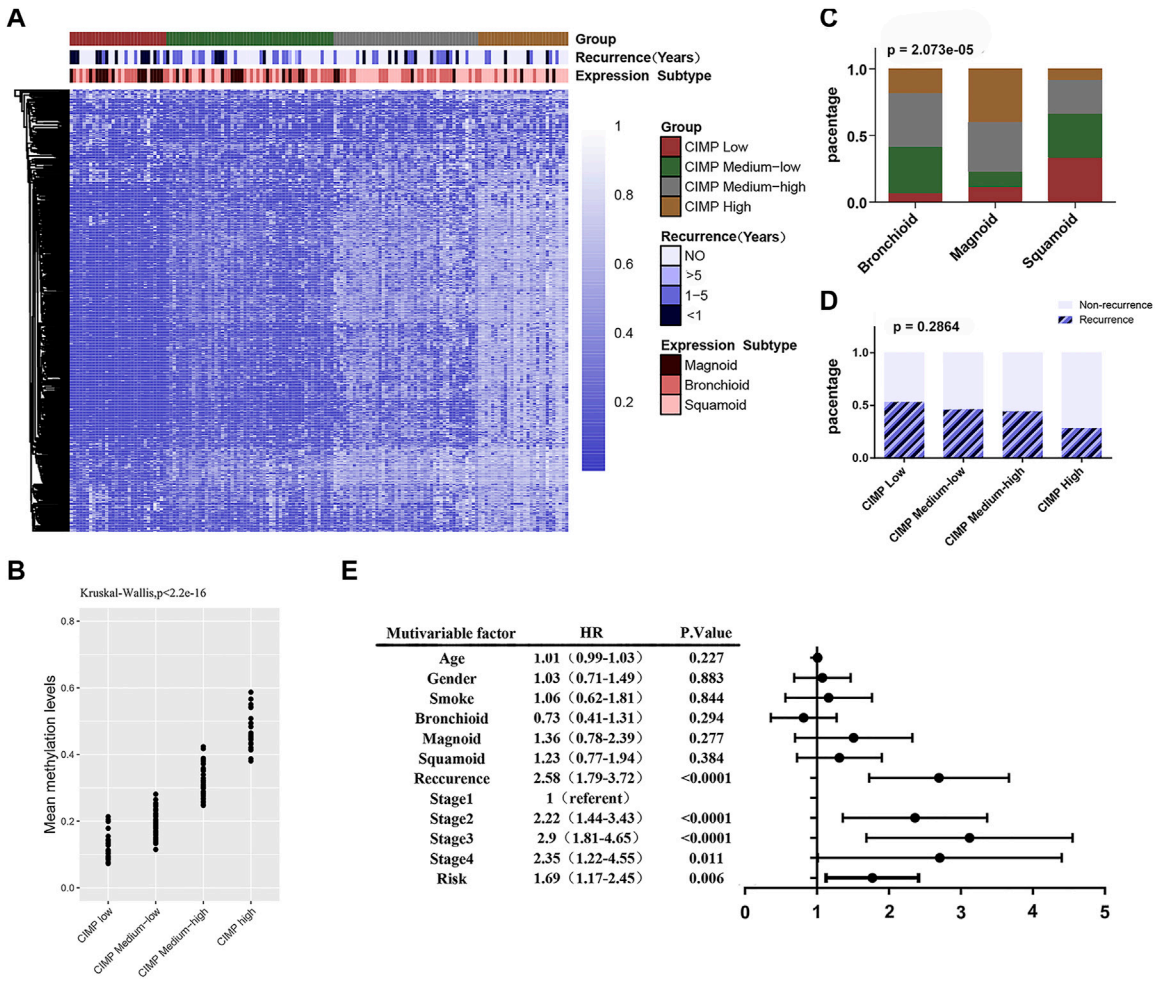


FIGURE 5
The DNA methylation feature with clinicopathological features. **(A)** Unsupervised cluster analysis of the methylation levels in LUAD. A total of 155 samples are presented in rows, and 641 CpG loci with the largest variation (mean methylation level β < 0.05 in normal samples and standard deviation σ > 0.20 in tumor samples) are listed. The 4 identified clusters are represented as CIMP high (n = 28), CIMP medium high (n = 45), CIMP medium low (n = 52) and CIMP low (*n* = 30). **(B)** Significant differences ($p < 0.0001$) in the methylation levels of the 4 clusters. **(C)** The sample distributions in terms of tumor subtype (Chi-square test, $p < 0.0001$). **(D)** The sample distributions in terms of recurrence (Chi-square test, $p = 0.2864$). **(E)** Forest map: multivariate Cox regression analysis was used to analyze the prognostic values of age, gender, smoking, tumor expression subtype, tumor recurrence, TNM stage, feature score and other clinicopathological features in 391 cases.

and *SLC27A6*). We selected the broad-spectrum demethylation drug 5-aza-2′-deoxycytidine (5-aza) to treat LUAD cells (A549, PC9 and H1975) and measured the mRNA levels of *TACR3*, *EDN3* and *SLC27A6* by qPCR 4 days after treatment. The results showed that, compared with the control group, the transcription levels of these 3 genes were significantly increased after treatment with 5-aza (Figures 7A–C), suggesting that the transcriptional regulation of *TACR3*, *EDN3* and *SLC27A6* was related to promoter methylation. The other 8 genes showed low methylation and high mRNA expression in LUAD cells (*MYEOV*, *NLRP13*, *SLC34A1*, *NEUROD4*, *HMGB4*, *KCNU1*, *GABRA5* and *TRPM8*). Since there is no broad-spectrum drug to improve DNA methylation, we used the Human Protein Atlas (HPA) (https://www.proteinatlas.org/) to verify the expression of the proteins encoded by these genes in LUAD and normal tissues. Five of them (*MYEOV*, *NLRP13*, *SLC34A1*, *NEUROD4* and *HMGB4*) had protein expression data in the database, and *MYEOV*, *NLRP13* and *SLC34A1* were highly expressed in tumor tissues (Figure 7D), while *NEUROD4* and *HMGB4* showed no significant difference (Figure 7E).

4 Discussion

DNA methylation, as one of the most studied epigenetic alterations related to tumor phenotype, is of great significance for

TABLE 3 The correlation between DNA methylation feature scores and clinicopathological feature.

	N	High	Low	P
Age (years)	453			
≥ 60	313	147 (46%)	166 (54%)	0.098
< 60	130	73 (56%)	57 (44%)	
Sex	453			
female	239	109 (46%)	130 (54%)	0.067
male	214	117 (55%)	97 (45%)	
Tumor location	440			
right	257	126 (49%)	131 (51%)	0.784
left	183	93 (51%)	90 (49%)	
T stage	453			
TX+T1+T2	398	200 (50%)	198 (50%)	0.787
T3+T4	55	26 (47%)	29 (53%)	
N stage	452			
N0	299	145 (48%)	154 (52%)	0.427
N1+N2+N3	153	81 (53%)	72 (47%)	
M stage	448			
M0	428	216 (50%)	212 (50%)	1
M1	20	10 (50%)	10 (50%)	
TNM stage	448			
I+II	350	175 (50%)	175 (50%)	0.949
III+IV	98	50 (51%)	48 (49%)	
Recurrence	416			
YES	162	85 (52%)	77 (48%)	0.39
NO	254	121 (48%)	133 (52%)	
Subtype	193			
Bronchioid	69	24 (35%)	45 (65%)	0.032
Magnoid	49	28 (57%)	21 (43%)	
Squamoid	75	39 (52%)	36 (48%)	
Smoke	430			
Non-smoker	63	20 (32%)	43 (68%)	0.004
Current smoker	367	192 (52%)	175 (48%)	

High and low groups were divided according to median of feature scores.

tumor research (Mazor et al., 2016; Biswas S Rao, 2017). Previous research showed that the overall DNA methylation pattern in tumor cell genomes is hypomethylation, while many CGIs associated with promoters showed focal hypermethylation (Hansen et al., 2011). Promoter hypermethylation was associated with tumor suppressor-related gene silencing, while the hypomethylation of the tumor cell genome could increase genomic instability (Feinberg and Vogelstein, 1983; Goelz et al., 1985; Pfeifer, 2018). Abnormal DNA methylation can be used not only as a target for tumor therapy but also as a biomarker for diagnosis and prognosis (Yamashita et al., 2018; Szejniuk et al., 2019). With the public information provided in TCGA, we conducted a comprehensive analysis and identified 11 methylation-related genes (*TACR3*, *SLC27A6*, *EDN3*, *TRPM8*, *MYEOV*, *NLRP13*, *KCNU1*, *NEUROD4*, *GABRA5*, *SLC34A1* and *HMGB4*) to predict the prognosis of LUAD. These genes not only were differentially methylated and expressed in LUAD tumor tissues from TCGA but also were related to the prognosis of patients. The survival curves showed that there was a significant difference in the survival curve between the higher-risk and lower-risk groups, especially in patients with early LUAD. Some of the 11 identified methylation-related genes have been shown to be abnormally expressed and important in cancer or other diseases. For example, *TRPM8* is a calcium permeability channel abnormally expressed in multiple malignant tumors. There is evidence that *TRPM8* plays a major role in promoting cell invasion and preventing replicative senescence (Yee, 2016). In the present study, *TRPM8* showed hypermethylation and low RNA expression in LUAD samples from TCGA, and the hypermethylation and low expression of *TRPM8* were associated with long survival. Previous studies have shown that decreased expression or inactivation of *EDN3* can inhibit the migration of cancer cells and improve survival (Wang et al., 2013; Kim et al., 2017). Our results revealed that *EDN3* was hypermethylated and expressed at low levels in LUAD, which was associated with longer OS. Another gene, *MYEOV*, is a region of cancer-associated genomic amplification. The amplification of this gene was reported to promote the progression of NSCLC pancreatic ductal adenocarcinoma and colorectal cancer (Lawlor et al., 2010; Fang et al., 2019). Subsequent mechanistic studies showed that the overexpression of *MYEOV* might be regulated by promoter hypomethylation (Liang et al., 2020). In accordance with the above results, our studies showed that *MYEOV* was hypermethylated and expressed at low levels in LUAD and that hypermethylation was positively correlated with survival time. As for other genes in the DNA methylation feature. TACR3 was found to be highly elevated in endometrial carcinoma. Although the role of *TAC1-TACR3* axis is not clear. Haixu et al. found that highly methylated *TAC1* promoted the development of endometrial carcinoma through the deregulation of *TAC* (Xu et al., 2018). Kyoichi Obata et al. found that *TACR3* protein showed significant and significant overexpression at the onset of bone matrix invasion in oral squamous cell carcinoma (Obata et al., 2016). *SLC27A6* is used as a predictor in the genetic analysis of colorectal cancer, prostate cancer, pancreatic cancer, and other tumors (Mohammed et al., 2019; Uhan et al., 2020; Verma et al.,

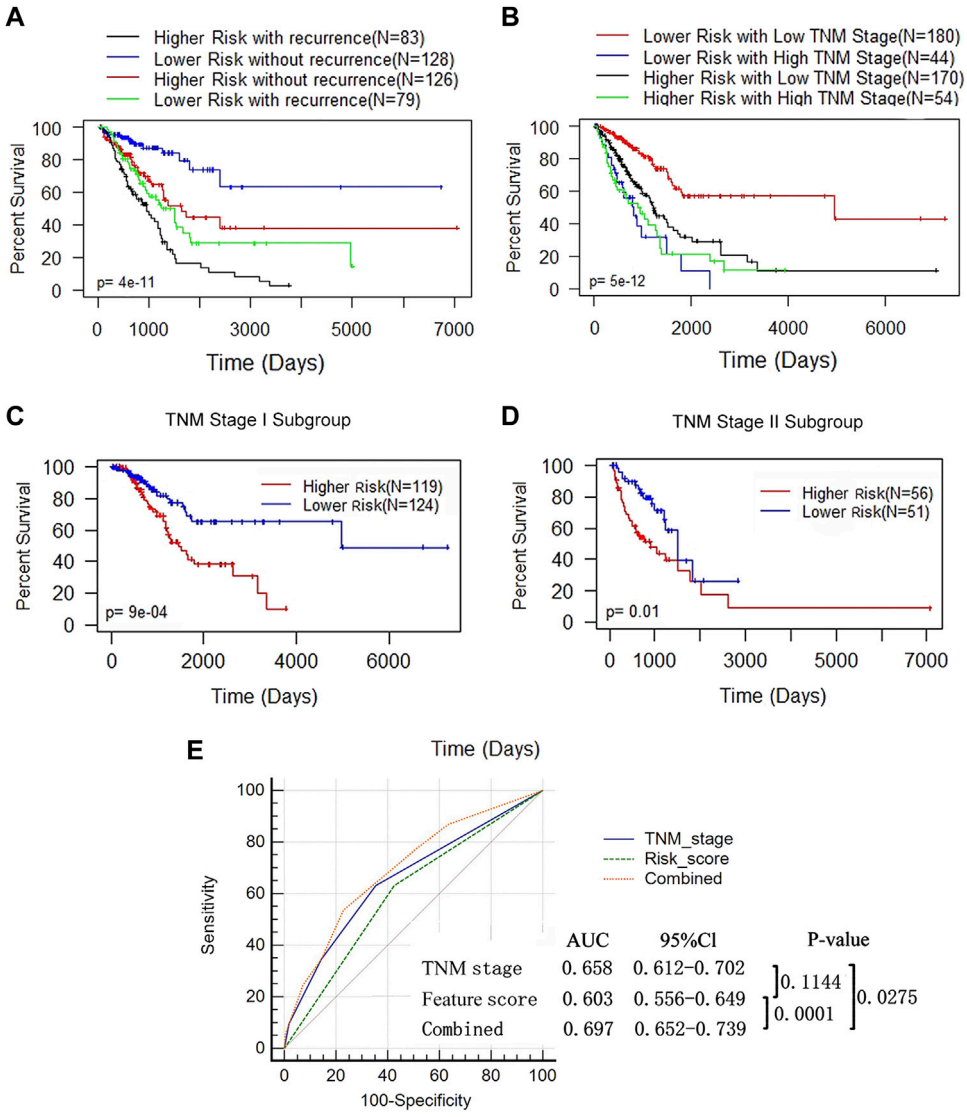


FIGURE 6
The prognostic value of DNA methylation feature score was not associated with tumor recurrence status and TNM stage. **(A)** K-M analysis of OS based on the DNA methylation feature and recurrence status. The LUAD patients were divided into the lower- and higher-risk subgroups according to the median of the methylation feature scores, with or without recurrence. Log-rank test ($p < 0.0001$). **(B)** K-M analysis of OS based on the DNA methylation feature and TNM stage. The LUAD patients were divided into the lower- and higher-risk subgroups according to the median of the methylation feature scores, and divided into the low (stage I+ II, n = 350) and high (stage III+IV, $n = 98$) stages according to the TNM stage (Log-rank test, $p < 0.0001$). **(C)** K-M curves for patients in the TNM stage I subgroup ($n = 350$). **(D)** K-M curves for patients in the TNM stage II subgroup ($n = 98$). **(E)** ROC analysis assessed the sensitivity and specificity of DNA methylation feature score, TNM stage and the combination of the 2 factors in predicting OS.

2020; Zhong et al., 2021). However, basic research on this topic is still very limited. As a neuron differentiation factor, *NEUROD4* has been reported overexpressed in neuroendocrine tumors. Studies have shown that the continuous expression of *NEUROD4* in neuronal cells may be related to the regeneration of neural cells, and its expression level gradually decreases with the maturation of neurons (Masserdotti et al., 2015; Cecil et al., 2016). *SLC27A1* is rarely studied in tumors, but in recent years, some articles have pointed out that *SLC27A1* is highly expressed in melanoma and breast cancer and enhances tumor invasion, migration, and growth (Kwaepila et al., 2006; Zhang et al., 2018). The expression of *GABRA5*, which encodes the α *5-GABAA* receptor, has a synthetic lethal role in *MYC*-driven medulloblastoma (Sengupta et al., 2014). *NLRP13*, *KCNU1*, and *HMGB4*, although not as studied in tumors compared to the other genes in the model, need further

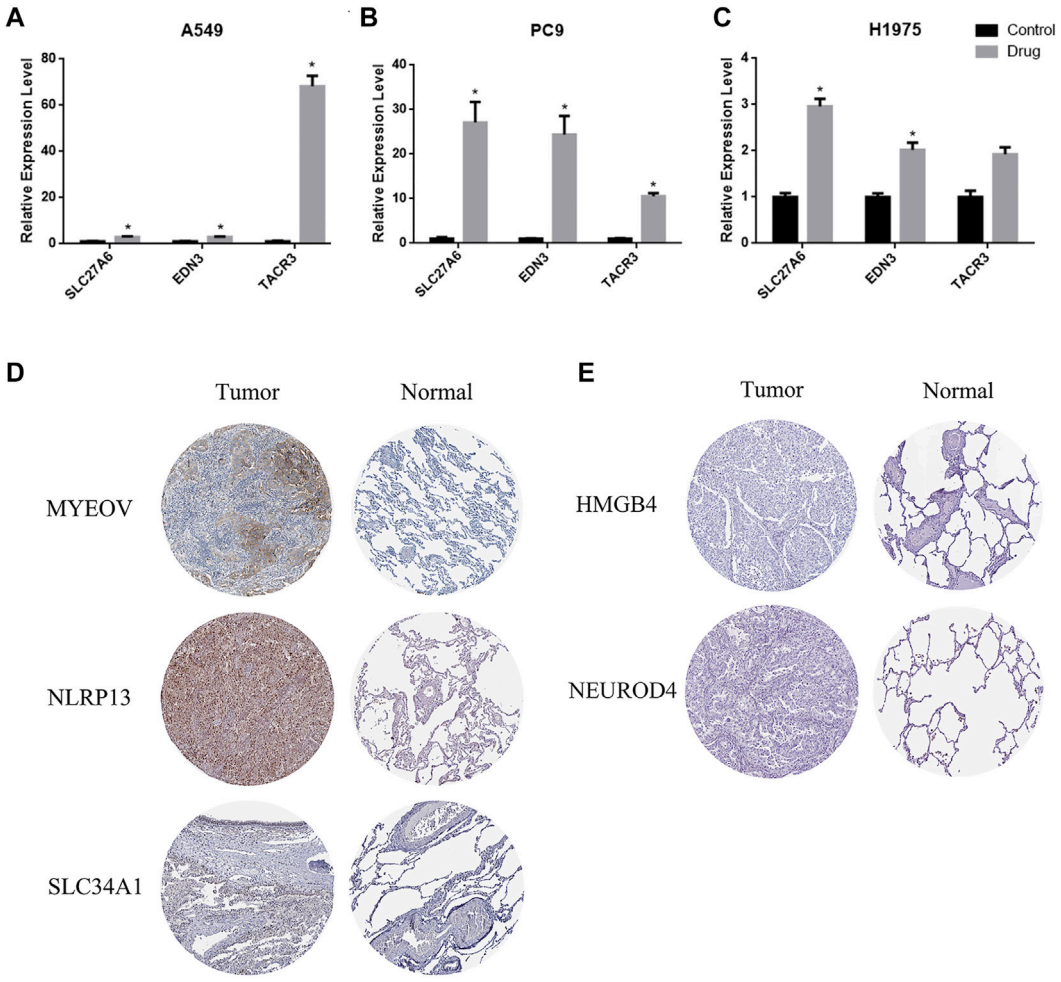


FIGURE 7
The expression of TACR3, EDN3 and SLC27A6 is related to promoter region methylation. **(A–C)** qPCR was used to detect the mRNA levels of TACR3, EDN3 and SLC27A6 in A549, PC9 and H1975 cells before and after the 5-Aza-2′-deoxycytidine treatment. **(D–E)** Immunohistochemistry images obtained from the HPA database demonstrated the protein expression of the 5 hypomethylated genes.

exploration. Many articles verify the effectiveness of their own prediction formulas by comparing them with TNM staging (Peng et al., 2020; Zhu et al., 2021; Qiu et al., 2022). In our research, the DNA methylation feature score is an independent predictor and is not associated with the TNM stage. ROC curve analysis showed that the combination of the DNA methylation feature score and TNM stage was better for prognosis than TNM stage alone, suggesting that the combination of the 2 might help to improve the prediction of OS in LUAD patients. In addition, we found that the transcription levels of *TACR3*, *EDN3* and *SLC27A6* in LUAD cells were significantly increased by treatment with broad-spectrum demethylating drugs. These results suggested that the low expression of these genes was related to promoter hypermethylation. At the same time, the expression of some hypomethylated genes (*MYEOV*, *NLRP13* and *SLC34A1*)in immunohistochemical sections of LUAD was significantly stronger than that in lung tissues. These results showed that the identified genes are worthy of further study as biomarkers of methylation in LUAD.

Because of its noninvasive and fast characteristics, detecting circulating tumor DNA (ctDNA) in blood to monitor epigenetic changes in tumor DNA has become a very promising technology. Although this technology is not sufficiently mature, blood testing based on a single DNA methylation biomarker has been approved (Frankell and Jamal-Hanjani, 2022). Our study shows that the methylation signals of these 11 genes may be used as candidate markers to detect ctDNA methylation in LUAD patients. This model can predict the prognosis of patients with low cost and high efficiency.

However, our research still has several limitations. First, DNA methylation biomarkers are not effective in predicting advanced LUAD. Considering the small sample size of the

advanced LUAD group, the results will have a certain deviation. Second, because it was difficult to obtain data with a sufficient sample size and consistent methylation detection platform, we did not use other datasets to verify the methylation formula. However, considering the large sample size of this study, this model was less likely to be an accidental feature of methylation noise but more likely to be a determinant of LUAD survival. Finally, our basic experiments were limited. We did not regulate the specific methylation site of genes. Further experimental studies on these genes will help to determine further their therapeutic potential.

5 Conclusion

In conclusion, we explored the characteristics of DNA methylation in LUAD. Furthermore, we confirmed a DNA methylation feature consisting of 11 genes. DNA methylation is associated with the survival of LUAD patients and can provide a better OS predictive ability when combined with TNM stage. Unfortunately, it was difficult for us to obtain a sufficient sample size and consistent methylation detection platform data to verify this methylation formula. However, our experiments indicated that the transcription of the hypermethylated genes was increased after demethylation with 5-aza, suggesting the validity of these results and indicating the potential value of these 11 genes in the study of LUAD prognosis.

Data availability statement

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

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

All authors have contributed significantly. LX, CX contributed to the conception of the study. LX, ZH and HX performed the experiment. LX, ZH, ZZ and CX contributed significantly to analysis and manuscript preparation. LX, ZH and JL performed the data analyses and wrote the manuscript. LX, ZZ, CX helped perform the analysis with constructive discussions. All authors approve of the version to be submitted finally.

Funding

This work was supported by National Natural Science Foundation of China (grant number 81972852), Key Research & Development Project of Hubei Province (grant number 2020BCA069), Health Commission of Hubei Province Medical Leading Talent Project, and Chinese Society of Clinical Oncology TopAlliance Tumor Immune Research Fund (grant number Y-JS2019-036).

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.2022.970507/full#supplementary-material

References

Biswas S Rao, C. M. (2017). Epigenetics in cancer: Fundamentals and beyond. *Pharmacol. Ther.* 173, 118–134. doi:10.1016/j.pharmthera.2017.02.011

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

Cecil, C. A., Walton, E., Smith, R. G., Viding, E., McCrory, E. J., Relton, C. L., et al. (2016). DNA methylation and substance-use risk: A prospective, genome-wide

study spanning gestation to adolescence. *Transl. Psychiatry* 6 (12), e976. doi:10.1038/tp.2016.247

Chi, Z., Kong, C. C., Wang, Z. Z., Wang, Z., Liu, G. L., Hu, Z., et al. (2021). The signaling pathways involved in metabolic regulation and stress responses of the yeast-like fungi Aureobasidium spp. *Biotechnol. Adv.* 55, 107898. doi:10.1016/j.biotechadv.2021.107898

Chiappinelli, K. B., and Baylin, S. B. (2022). Inhibiting DNA methylation improves antitumor immunity in ovarian cancer. *J. Clin. Invest.* 132 (14), e160186. doi:10.1172/JCI160186

Christman, J. K. (2002). 5-Azacytidine and 5-aza-2'-deoxycytidine as inhibitors of DNA methylation: Mechanistic studies and their implications for cancer therapy. *Oncogene* 21 (35), 5483–5495. doi:10.1038/sj.onc.1205699

Edge, S. B., and Compton, C. C. (2010). The American Joint committee on cancer: The 7th edition of the AJCC cancer staging manual and the future of TNM. *Ann. Surg. Oncol.* 17 (6), 1471–1474. doi:10.1245/s10434-010-0985-4

Enfield, K. S. S., Marshall, E. A., Anderson, C., Ng, K. W., Rahmati, S., Xu, Z., et al. (2019). Epithelial tumor suppressor ELF3 is a lineage-specific amplified oncogene in lung adenocarcinoma. *Nat. Commun.* 10 (1), 5438. doi:10.1038/s41467-019-13295-y

Fang, L., Wu, S., Zhu, X., Cai, J., Wu, J., He, Z., et al. (2019). MYEOV functions as an amplified competing endogenous RNA in promoting metastasis by activating TGF-β pathway in NSCLC. *Oncogene* 38 (6), 896–912. doi:10.1038/s41388-018-0484-9

Feinberg, A. P. (2007). Phenotypic plasticity and the epigenetics of human disease. *Nature* 447 (7143), 433–440. doi:10.1038/nature05919

Feinberg, A. P., and Vogelstein, B. (1983). Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* 301 (5895), 89–92. doi:10.1038/301089a0

Folkman, J. (1971). Tumor angiogenesis: Therapeutic implications. *N. Engl. J. Med.* 285 (21), 1182–1186. doi:10.1056/NEJM197111182852108

Frankell, A., and Jamal-Hanjani, M. (2022). When can we be confident of surgical cure with ctDNA? *Nat. Rev. Clin. Oncol.* doi:10.1038/s41571-022-00664-8

Goelz, S. E., Vogelstein, B., Hamilton, S. R., and Feinberg, A. P. (1985). Hypomethylation of DNA from benign and malignant human colon neoplasms. *Science* 228 (4696), 187–190. doi:10.1126/science.2579435

Goldstraw, P., Ball, D., Jett, J. R., Le Chevalier, T., Lim, E., Nicholson, A. G., et al. (2011). Non-small-cell lung cancer. *Lancet* 378 (9804), 1727–1740. doi:10.1016/S0140-6736(10)62101-0

Goto, Y., Shinjo, K., Kondo, Y., Shen, L., Toyota, M., Suzuki, H., et al. (2009). Epigenetic profiles distinguish malignant pleural mesothelioma from lung adenocarcinoma. *Cancer Res.* 69 (23), 9073–9082. doi:10.1158/0008-5472.Can-09-1595

Hansen, K. D., Timp, W., Bravo, H. C., Sabunciyan, S., Langmead, B., McDonald, O. G., et al. (2011). Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* 43 (8), 768–775. doi:10.1038/ng.865

Huo, M., Zhang, J., Huang, W., and Wang, Y. (2021). Interplay among metabolism, epigenetic modifications, and gene expression in cancer. *Front. Cell Dev. Biol.* 9, 793428. doi:10.3389/fcell.2021.793428

Ikeda, K., Shiraishi, K., Eguchi, A., Shibata, H., Yoshimoto, K., Mori, T., et al. (2013). Long interspersed nucleotide element 1 hypomethylation is associated with poor prognosis of lung adenocarcinoma. *Ann. Thorac. Surg.* 96 (5), 1790–1794. doi:10.1016/j.athoracsur.2013.06.035

Jamieson, E., Korologou-Linden, R., Wootton, R. E., Guyatt, A. L., Battram, T., Burrows, K., et al. (2020). Smoking, DNA methylation, and lung function: A mendelian randomization analysis to investigate causal pathways. *Am. J. Hum. Genet.* 106 (3), 315–326. doi:10.1016/j.ajhg.2020.01.015

Jung, S. H. (2014). Stratified Fisher's exact test and its sample size calculation. *Biom. J.* 56 (1), 129–140. doi:10.1002/bimj.201300048

Kaishang, Z., Xue, P., Shaozhong, Z., Yingying, F., Yan, Z., Chanjun, S., et al. (2018). Elevated expression of Twinfilin-1 is correlated with inferior prognosis of lung adenocarcinoma. *Life Sci.* 215, 159–169. doi:10.1016/j.lfs.2018.10.067

Karagiota, A., Chachami, G., and Paraskeva, E. (2022). Lipid metabolism in cancer: The role of acylglycerolphosphate acyltransferases (AGPATs). *Cancers (Basel)* 14 (1), 228. doi:10.3390/cancers14010228

Kim, I. S., Heilmann, S., Kansler, E. R., Zhang, Y., Zimmer, M., Ratnakumar, K., et al. (2017). Microenvironment-derived factors driving metastatic plasticity in melanoma. *Nat. Commun.* 8, 14343. doi:10.1038/ncomms14343

Küster, M. M., Schneider, M. A., Richter, A. M., Richtmann, S., Winter, H., Kriegsmann, M., et al. (2020). Epigenetic inactivation of the tumor suppressor IRX1 occurs frequently in lung adenocarcinoma and its silencing is associated with impaired prognosis. *Cancers* 12 (12), E3528. doi:10.3390/cancers12123528

Kwaepila, N., Burns, G., and Leong, A. S. (2006). Immunohistological localisation of human FAT1 (hFAT) protein in 326 breast cancers. Does this adhesion molecule have a role in pathogenesis? *Pathology* 38 (2), 125–131. doi:10.1080/00313020600559975

Lawlor, G., Doran, P. P., MacMathuna, P., and Murray, D. W. (2010). MYEOV (myeloma overexpressed gene) drives colon cancer cell migration and is regulated by PGE2. *J. Exp. Clin. Cancer Res.* 29 (1), 81. doi:10.1186/1756-9966-29-81

Li, C. H., and Liao, C. C. (2021). The metabolism reprogramming of microRNA let-7-mediated glycolysis contributes to autophagy and tumor progression. *Int. J. Mol. Sci.* 23 (1), 113. doi:10.3390/ijms23010113

Li, J., Liang, Y., Fan, J., Xu, C., Guan, B., Zhang, J., et al. (2022). DNA methylation subtypes guiding prognostic assessment and linking to responses the DNA methyltransferase inhibitor SGI-110 in urothelial carcinoma. *BMC Med.* 20 (1), 222. doi:10.1186/s12916-022-02426-w

Li, Y., Gu, J., Xu, F., Zhu, Q., Ge, D., and Lu, C. (2018). Transcriptomic and functional network features of lung squamous cell carcinoma through integrative analysis of GEO and TCGA data. *Sci. Rep.* 8 (1), 15834. doi:10.1038/s41598-018-34160-w

Li, Y., Sun, N., Lu, Z., Sun, S., Huang, J., Chen, Z., et al. (2017). Prognostic alternative mRNA splicing signature in non-small cell lung cancer. *Cancer Lett.* 393, 40–51. doi:10.1016/j.canlet.2017.02.016

Li, Z., Zhang, R., Yang, X., Zhang, D., Li, B., Zhang, D., et al. (2019). Analysis of gene expression and methylation datasets identified ADAMTS9, FKBP5, and PFKBF3 as biomarkers for osteoarthritis. *J. Cell. Physiol.* 234 (6), 8908–8917. doi:10.1002/jcp.27557

Lian, H., Han, Y. P., Zhang, Y. C., Zhao, Y., Yan, S., Li, Q. F., et al. (2019). Integrative analysis of gene expression and DNA methylation through one-class logistic regression machine learning identifies stemness features in medulloblastoma. *Mol. Oncol.* 13 (10), 2227–2245. doi:10.1002/1878-0261.12557

Liang, E., Lu, Y., Shi, Y., Zhou, Q., and Zhi, F. (2020). MYEOV increases HES1 expression and promotes pancreatic cancer progression by enhancing SOX9 transactivity. *Oncogene* 39 (41), 6437–6450. doi:10.1038/s41388-020-01443-4

Liu, C., Li, Y., Wei, M., Zhao, L., and Yu Y Li, G. (2019). Identification of a novel glycolysis-related gene signature that can predict the survival of patients with lung adenocarcinoma. *Cell Cycle* 18 (5), 568–579. doi:10.1080/15384101.2019.1578146

Masserdotti, G., Gillotin, S., Sutor, B., Drechsel, D., Irmler, M., Jørgensen, H. F., et al. (2015). Transcriptional mechanisms of proneural factors and REST in regulating neuronal reprogramming of astrocytes. *Cell Stem Cell* 17 (1), 74–88. doi:10.1016/j.stem.2015.05.014

Mazor, T., Pankov, A., and Song J S Costello, J. F. (2016). Intratumoral heterogeneity of the epigenome. *Cancer Cell* 29 (4), 440–451. doi:10.1016/j.ccell.2016.03.009

Mohammed, A., Janakiram, N. B., Suen, C., Stratton, N., Lightfoot, S., Singh, A., et al. (2019). Targeting cholecystokinin-2 receptor for pancreatic cancer chemoprevention. *Mol. Carcinog.* 58 (10), 1908–1918. doi:10.1002/mc.23084

Obata, K., Shimo, T., Okui, T., Matsumoto, K., Takada, H., Takabatake, K., et al. (2016). Tachykinin receptor 3 distribution in human oral squamous cell carcinoma. *Anticancer Res.* 36 (12), 6335–6341. doi:10.21873/anticanres.11230

Pandis, N. (2016). The chi-square test. *Am. J. Orthod. Dentofac. Orthop.* 150 (5), 898–899. doi:10.1016/j.ajodo.2016.08.009

Peng, Y., Wu, Q., Wang, L., Wang, H., and Yin, F. (2020). A DNA methylation signature to improve survival prediction of gastric cancer. *Clin. Epigenetics* 12 (1), 15. doi:10.1186/s13148-020-0807-x

Pfeifer, G. P. (2018). Defining driver DNA methylation changes in human cancer. *Int. J. Mol. Sci.* 19 (4), E1166. doi:10.3390/ijms19041166

Qiu, Z., Ji, J., Xu, Y., Zhu, Y., Gao, C., Wang, G., et al. (2022). Common DNA methylation changes in biliary tract cancers identify subtypes with different immune characteristics and clinical outcomes. *BMC Med.* 20 (1), 64. doi:10.1186/s12916-021-02197-w

Schröder, M. S., Culhane, A. C., Quackenbush, J., and Haibe-Kains, B. (2011). survcomp: an R/Bioconductor package for performance assessment and comparison of survival models. *Bioinformatics* 27 (22), 3206–3208. doi:10.1093/bioinformatics/btr511

Sengupta, S., Weeraratne, S. D., Sun, H., Phallen, J., Rallapalli, S. K., Teider, N., et al. (2014). α5-GABAA receptors negatively regulate MYC-amplified medulloblastoma growth. *Acta Neuropathol.* 127 (4), 593–603. doi:10.1007/s00401-013-1205-7

Singh, S., Narayanan, S. P., Biswas, K., Gupta, A., Ahuja, N., Yadav, S., et al. (2017). Intragenic DNA methylation and BORIS-mediated cancer-specific splicing contribute to the Warburg effect. *Proc. Natl. Acad. Sci. U. S. A.* 114 (43), 11440–11445. doi:10.1073/pnas.1708447114

Sui, J., Li, Y-H., Zhang, Y-Q., Li, C-Y., Shen, X., Yao, W-Z., et al. (2016). Integrated analysis of long non-coding RNA-associated ceRNA network reveals potential lncRNA biomarkers in human lung adenocarcinoma. *Int. J. Oncol.* 49 (5), 2023–2036. doi:10.3892/ijo.2016.3716

Szejniuk, W. M., Robles, A. I., McCulloch, T., Falkmer U G I Røe, O. D., and Roe, O. D. (2019). Epigenetic predictive biomarkers for response or outcome to platinum-based chemotherapy in non-small cell lung cancer, current state-of-art. *Pharmacogenomics J.* 19 (1), 5–14. doi:10.1038/s41397-018-0029-1

Tozzi, R., Masci, F., and Pezzopane, M. (2020). A stress test to evaluate the usefulness of Akaike information criterion in short-term earthquake prediction. *Sci. Rep.* 10 (1), 21153. doi:10.1038/s41598-020-77834-0

Uhan, S., Zidar, N., and Tomažič A Hauptman, N. (2020). Hypermethylated promoters of genes UNC5D and KCNA1 as potential novel diagnostic biomarkers in colorectal cancer. *Epigenomics* 12 (19), 1677–1688. doi:10.2217/epi-2020-0118

Verma, S., Shankar, E., Chan, E. R., and Gupta, S. (2020). Metabolic reprogramming and predominance of solute carrier genes during acquired enzalutamide resistance in prostate cancer. *Cells* 9 (12), E2535. doi:10.3390/cells9122535

Wang, R., Löhr, C. V., Fischer, K., Dashwood, W. M., Greenwood, J. A., Ho, E., et al. (2013). Epigenetic inactivation of endothelin-2 and endothelin-3 in colon cancer. *Int. J. Cancer* 132 (5), 1004–1012. doi:10.1002/ijc.27762

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinformatics* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Xu, H., Sun, Y., Ma, Z., Xu, X., Qin, L., and Luo, B. (2018). LOC134466 methylation promotes oncogenesis of endometrial carcinoma through LOC134466/hsa-miR-196a-5p/TAC1 axis. *Aging (Albany NY)* 10 (11), 3353–3370. doi:10.18632/aging.101644

Yamashita, K., Hosoda, K., Nishizawa, N., Katoh, H., and Watanabe, M. (2018). Epigenetic biomarkers of promoter DNA methylation in the new era of cancer treatment. *Cancer Sci.* 109 (12), 3695–3706. doi:10.1111/cas.13812

Yang, S., Huang, Y., and Zhao, Q. (2022). Epigenetic alterations and inflammation as emerging use for the advancement of treatment in non-small cell lung cancer. *Front. Immunol.* 13, 878740. doi:10.3389/fimmu.2022.878740

Yee, N. S. (2016). TRPM8 ion channels as potential cancer biomarker and target in pancreatic cancer. *Adv. Protein Chem. Struct. Biol.* 104, 127–155. doi:10.1016/bs.apcsb.2016.01.001

Zhang, H., Kong, Q., Wang, J., Jiang, Y., and Hua, H. (2020). Complex roles of cAMP-PKA-CREB signaling in cancer. *Exp. Hematol. Oncol.* 9 (1), 32. doi:10.1186/s40164-020-00191-1

Zhang, M., Di Martino, J. S., Bowman, R. L., Campbell, N. R., Baksh, S. C., Simon-Vermot, T., et al. (2018). Adipocyte-Derived lipids mediate melanoma progression via FATP proteins. *Cancer Discov.* 8 (8), 1006–1025. doi:10.1158/2159-8290.Cd-17-1371

Zhao, J., Guo, C., Ma, Z., Liu, H., Yang, C., and Li, S. (2020). Identification of a novel gene expression signature associated with overall survival in patients with lung adenocarcinoma: A comprehensive analysis based on TCGA and geo databases. *Lung Cancer* 149, 90–96. doi:10.1016/j.lungcan.2020.09.014

Zhong, X., Yang, Y., Li, B., Liang, P., Huang, Y., Zheng, Q., et al. (2021). Downregulation of SLC27A6 by DNA hypermethylation promotes proliferation but suppresses metastasis of nasopharyngeal carcinoma through modulating lipid metabolism. *Front. Oncol.* 11, 780410. doi:10.3389/fonc.2021.780410

Zhu, L., Sun, H., Tian, G., Wang, J., Zhou, Q., Liu, P., et al. (2021). Development and validation of a risk prediction model and nomogram for colon adenocarcinoma based on methylation-driven genes. *Aging* 13 (12), 16600–16619. doi:10.18632/aging.203179

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 02 September 2022
DOI 10.3389/fgene.2022.919857

OPEN ACCESS

EDITED BY
Yingming Sun,
Fujian Medical University, China

REVIEWED BY
Zhentao Yu,
Chinese Academy of Medical Sciences and Peking Union Medical College, China
Snehal Shukla,
Pfizer, United States
Zhengqing Yan,
3D Medicines Co., China

*CORRESPONDENCE
Yue Yu,
yuyue@tmu.edu.cn
Jie Ge,
gejie1980@tmu.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 14 April 2022
ACCEPTED 08 August 2022
PUBLISHED 02 September 2022

CITATION
Hu J, Zhang P-J, Zhang D, Chen Z-H, Cao X-C, Yu Y and Ge J (2022), An autophagy-associated lncRNAs model for predicting the survival in non-small cell lung cancer patients.
Front. Genet. 13:919857.
doi: 10.3389/fgene.2022.919857

An autophagy-associated lncRNAs model for predicting the survival in non-small cell lung cancer patients

Jing Hu[1,2,3,4,5†], Pei-Jin Zhang[1,2,3,4†], Di Zhang[1,2,3,4], Zhao-Hui Chen[1,2,3,4], Xu-Chen Cao[1,2,3,4], Yue Yu[1,2,3,4*] and Jie Ge[1,2,3,4*]

[1]The First Department of Breast Cancer, Tianjin Medical University Cancer Institute and Hospital, National Clinical Research Center for Cancer, Tianjin, China, [2]Key Laboratory of Cancer Prevention and Therapy, Tianjin, China, [3]Tianjin's Clinical Research Center for Cancer, Tianjin, China, [4]Key Laboratory of Breast Cancer Prevention and Therapy, Tianjin Medical University, Ministry of Education, Tianjin, China, [5]Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, China

Long non-coding RNAs (lncRNAs) can influence the proliferation, autophagy, and apoptosis of non-small cell lung cancer (NSCLC). LncRNAs also emerge as valuable prognostic factors for NSCLC patients. Consequently, we set out to discover more autophagy-associated lncRNAs. We acquired autophagy-associated genes and information on lncRNAs from The Cancer Genome Atlas database (TCGA), and the Human Autophagy Database (HADb). Then, the prognostic prediction signature was constructed through using co-expression and Cox regression analysis. The signature was constructed including 7 autophagy-associated lncRNAs (ABALON, NKILA, LINC00941, AL161431.1, AL691432.2, AC020765.2, MMP2-AS1). After that, we used univariate and multivariate Cox regression analysis to calculate the risk score. The survival analysis and ROC curve analysis confirmed good performances of the signature. GSEA indicated that the high-risk group was principally enriched in the adherens junction pathway. In addition, biological experiments showed that ABALON promoted the proliferation, metastasis and autophagy levels of NSCLC cells. These findings demonstrate that the risk signature consisting of 7 autophagy-associated lncRNAs accurately predicts the prognosis of NSCLC patients and should be investigated for potential therapeutic targets in clinic.

KEYWORDS

autophagy, lncRNAs, NSCLC, TCGA, prognostic signature

Introduction

Lung cancer is the most common malignant disease. Among these cases, 84% are classified as NSCLC with the remaining 16% classified as small cell lung cancer (SCLC) (Torre et al., 2012). Because of the advent of new targeted drugs and advances in therapy, clinical treatment of lung cancer has made great progress. However, the attendant problems also increased, such as inherent resistance to both chemotherapy and radiation therapy (Hirsch et al., 2017); (Yang et al., 2019), and it is becoming increasingly challenging to assess the prognosis of NSCLC patients. Therefore, it is imperative to confirm biomarkers of prognosis in NSCLC.

As an intracellular catabolic degradation process, autophagy contributes to normal cell physiology by eliminating damaged proteins and other cell components (Li et al., 2020). There is a great deal of studies showing that autophagy also participates in various pathological processes, for example liver disorders and infectious diseases (Ueno and Komatsu, 2017); (Levine and Kroemer, 2019); (Choi et al., 2018). Moreover, a growing body of research suggests autophagy plays a dual role in cancer. Autophagy can suppress chronic tissue damage to inhibit tumorigenesis. It could promote longevity of normal cells *via* regulating the quality of proteins and organelles, promoting the stability of the genome, or a combination of these factors (Barnard et al., 2016); (Mizushima and Levine, 2020). However, autophagy can also maintain the function of mitochondria and reduce DNA damage to enhance the ability to resist stress and apoptosis in cancer cells (White et al., 2015); (Bravo-San Pedro et al., 2017). Many studies have also shown that modulators of autophagy can prevent NSCLC from developing (Bai et al., 2019). However, other research shows that upregulation of autophagy can promote tumorigenesis and immune escape of cancer cells (Ma et al., 2020). Ma et al. (2013) SKIL promoted tumorigenesis and immune escape of NSCLC cells through upregulation of TAZ/autophagy axis and inhibition on downstream STING pathway. Therefore, it is important to find autophagy-associated transcripts which are considered as valuable biomarkers for diagnosis and prognosis in NSCLC.

LncRNAs, more than 200 nucleotides, is a class of RNA transcripts and does not code proteins. LncRNAs participate in some fundamental cancer-related processes through transcriptional for or post-transcriptional regulation, such as proliferation, migration, survival, and metastasis (Fang and Fullwood, 2016); (Slack and Chinnaiyan, 2019). Furthermore, a great number of studies identified that lncRNAs regulated autophagy (Xu et al., 2019); (Wu et al., 2021); (Kopp and Mendell, 2018). For example, Wang et al. (2019a) proved the lncRNA LINRIS could block the degradation of IGF2BP2 and suppress the proliferation of CRC *via* the ubiquitination-autophagy pathway. The lncRNA NBAT1 inhibits autophagy by inhibiting ATG7 in NSCLC (Zheng et al., 2018). Hence, it is meaningful to confirm major lncRNAs connected with autophagy and prognosis of NSCLC.

In our study, we systematically analyzed lncRNAs data of NSCLC patients in TCGA. We also founded an accurate prognostic signature of 7 autophagy-associated lncRNAs and assessed their ability to precisely predict the prognosis. It was verified that the downregulation of ABALON affected proliferation, metastasis, and autophagy of NSCLC cells by experimental validation. We provide a novel prognostic signature consisting of 7 autophagy-associated lncRNAs, which may also be potential therapeutic targets.

Materials and methods

Patient data sets

The transcriptome profiles and corresponding clinical date of 1,145 NSCLC patients (1,037 cases NSCLC patients and 108 healthy controls) were extracted from TCGA (https://portal.gdc.cancer.gov/), and all genes ID were transformed. LncRNAs and protein-coding genes were annotated and classified *via* the Ensembl human genome browser. We acquired data of autophagy-associated genes from the HADb (https://www.autophagy.lu/). We extract potential autophagy-associated lncRNAs *via* pearson correlation analysis, $|R^2| > 0.3$ and p-value < 0.001 were defined as thresholds.

Construction and evaluation of an autophagy-associated long non-coding RNAs prognostic signature

We combined the expression level of autophagy-associated lncRNAs with the corresponding survival results in TCGA. Autophagy-associated lncRNAs were confirmed by univariate Cox regression analysis. The calculation formula of risk score for every patient was: risk score = $\sum_{k=1}^{n} coef(k)*lncRNA(k)$, where coef (k) and lncRNA (k) respectively represent the regression coefficient and expression level of corresponding autophagy-associated lncRNA. These lncRNAs data were divided into two groups based on the median risk score. Then we constructed the best prognostic risk model of autophagy-associated lncRNAs *via* multivariate Cox regression analysis. Subsequently, we evaluated the survival difference in these two groups *via* Kaplan-Meier (KM) survival analysis. We explored the correlation between clinical factors and risk score. *R* packages "survival" and "forestplot" were performed to visualize the forest plot respectively. Finally, we draw ROC curves to estimate the predictive value of different clinical pathological factors.

Establishment of the long non-coding RNA-mRNA co-expression network

The relationship between autophagy-associated lncRNAs and their corresponding mRNAs was explored through a co-

expression network and Sankey diagram. We extracted the mRNAs that were associated with autophagy-associated lncRNAs *via* Pearson correlation coefficients, the absolute threshold coefficient value >0.3. The network was visually analyzed by cytoscape software (v 3.7.1) and ggalluvial *R* package.

Estimation and construction of nomogram

A nomogram survival prediction model of NSCLC patients was constructed using the "survival" and "rms" packages by combining risk score and expression of autophagy-associated lncRNAs. We then constructed the ROC curves and calculated the AUC values of this nomogram.

Functional enrichment analysis

Perform GSEA to find expression changes in predefined genomes, rather than individual genes. We verified the enrichment of differentially expressed gene sets between the low- and high-risk groups by GSEA (v4.1.0). These two groups were enriched in different signaling pathway.

Cell culture

Beas-2B, A549, NCI-H292, NCI-H460, and NCI-H1299 cell lines were acquired from American Tissue Culture Collection (Beijing, China). Beas-2B cells were cultured in Leibovitz's L-15 medium with 10% fetal bovine serum (FBS, Corning Incorporated) at 37°C in a 5% CO_2 incubator. Other cell lines were cultured in RPMI 1640 with 10% FBS at 37°C in 5% CO_2 incubator.

RNA extraction and qRT-PCR

Total RNA was isolated from cells using the TRIzol reagent (Beyotime, Shanghai, China). The total RNA was reversed to cDNA by the PrimeScript RT reagent Kit (Takara, Japan). QRT-PCR was performed to detect the expression of ABALON using SYBR Green Mixture (Tli RNaseH Plus) (Takara, Japan) and gene specific primers. We collected data from a Roche LightCycler 480 PCR system. The results were normalized with GAPDH as an internal control. The primer sequences were as follows: GAPDH (forward: CGGAGTCAACGGATTTGGTCGTAT; reverse: AGCCTTCTCCATGGTGGTGAAGAC), ABALON (forward: CTCTCTCTTGCACGCCCCTTG; reverse: CCT GGGCTGGTGCTTAAATAGA).

Transient transfection and small interfering RNAs

The siRNA targeting ABALON and Control were purchased from RiboBio (Guangzhou, China). When the cell confluency reaches 35%, we use FuGENE HD Transfection Reagent (Promega) to transiently transfect siRNAs (100 nm) in the cells. The medium containing 10% FBS was refreshed after 12 h.

Proliferation and metastasis assays

EDU, MTT and colony formation assays were used to detect cell proliferation. Transwell and wound healing/scratch assays were used to estimate cell metastasis. Detailed procedures were listed in the Supplementary Material.

Western blot analysis

Cellular protein was extracted using RIPA lysis buffer (Solarbio, China). Proteins were measured by the BCA method (Thermo scientific, United States) and separated on the SDS-PAGE gels. The protein bands were then transferred into PVDF membranes. Then the bands were blocked with 5% skim milk at room temperature for 1 h and incubated with primary antibodies overnight at 4°C. The bands were incubated with a corresponding secondary antibody for 1 h at room temperature. The protein bands were then detected using the ECL reagent (Millipore, United States). The antibodies were listed in the Supplementary Material.

Statistical analysis

Data are presented as the means ± SD. Each experiment has a minimum of 3 replicates. Statistical analyses were executed using *R* Studio (version 4.0.3) and SPSS software (IBM Corp, United States). p-values < 0.05 were regarded as statistically significant.

Results

Identification of prognostic autophagy-associated long non-coding RNAs in non-small cell lung cancer patients

The analysis framework of this research is performed in Figure 1. We extracted 141432 lncRNAs data sets from TCGA and 232 autophagy-associated genes from HADb in NSCLC. Then, we identified 1,496 autophagy-associated lncRNAs *via* conducting

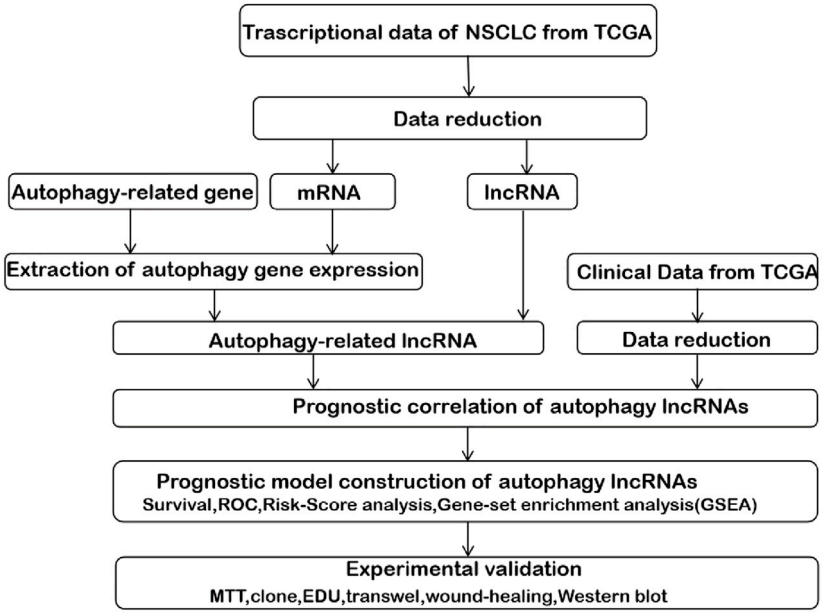


FIGURE 1
The comprehensive prognostic analysis framework of autophagy-associated lncRNAs in NSCLC based on the TCGA database.

TABLE 1 The prognostic effect of autophagy-related lncRNAs by univariate Cox analysis.

lncRNA	HR	HR.95L	HR.95H	*p*-value
AC020765.2	0.7915	0.6782	0.9238	0.003
AP000695.1	1.151	1.051	1.2605	0.0024
MMP2-AS1	0.8172	0.7124	0.9373	0.0039
AC068338.3	0.7071	0.5624	0.889	0.003
NKILA	1.0806	1.0339	1.1294	0.0006
AL691432.2	0.9098	0.8539	0.9693	0.0035
LINC00941	1.1098	1.0712	1.1499	0.0003
ABALON	1.4231	1.1691	1.7323	0.0004
AL161431.1	1.0057	1.002	1.0095	0.0025
AC135050.6	0.9733	0.9552	0.9917	0.0047
CRNDE	0.9616	0.9388	0.9849	0.0014
AC012615.1	0.8814	0.8098	0.9592	0.0035

pearson correlation analysis ($p < 0.001$). Within them, 12 autophagy-associated lncRNAs were closely correlated with the survival of NSCLC patients from TCGA *via* univariate Cox regression analysis ($p < 0.005$; Table 1). Multivariate Cox regression analysis further revealed 7 autophagy-associated lncRNAs were good candidates for constructing the diagnostic signature. Among these 7 autophagy-associated lncRNAs, AC020765.2, MMP2-AS1, and AL691432.2 were regarded as protective factors [hazard ration (HR) < 1], while the remaining 4 lncRNAs, NKILA, LINC00941, ABALON, and AL161431.1, were considered to be risk factors (HR > 1) (Figure 2A). The relationships between these autophagy-associated lncRNAs and mRNAs are displayed in Figures 2B,C.

Prognosis evaluation of the autophagy-associated long non-coding RNA signature in non-small cell lung cancer

We used the prognostic risk scoring method to construct a model consisting of 7 autophagy-associated lncRNAs. The formula of the risk score was as follows: risk score= (0.0674 × NKILA) + (−0.1483 × AC020765.2) + (−0.1896 × MMP2-AS1) + (−0.0692 × AL691432.2) + (0.0555 × LINC00941) + (0.3039 × ABALON) + (0.0049 × AL161431.1) ($p < 0.05$; Table 2). Subsequently, we figured out the risk score for NSCLC patients. These samples from 1,145 NSCLC were divided into high- and low-risk groups *via* this median risk score as the cutoff. KM survival curve analysis revealed the OS of high-risk NSCLC patients was obviously shorter than of low-risk NSCLC patients ($p < 0.001$). The 3-year OS rates of the high and low-risk groups were 51.6% and 65.9%. Similarly, the 5-year OS rates of these two groups were 36.2% and 46.9% (Figure 3A), respectively. We next evaluated the prognostic power of the autophagy-associated lncRNAs model by ROC analysis. The AUC values were respectively 0.658, 0.625, and 0.581 for 1-, 3-, and 5-year OS (Figure 3B). Thus, the signature demonstrated a precise prognostic value. The risk curve of the prognostic signature and scatterplot indicated that the mortality

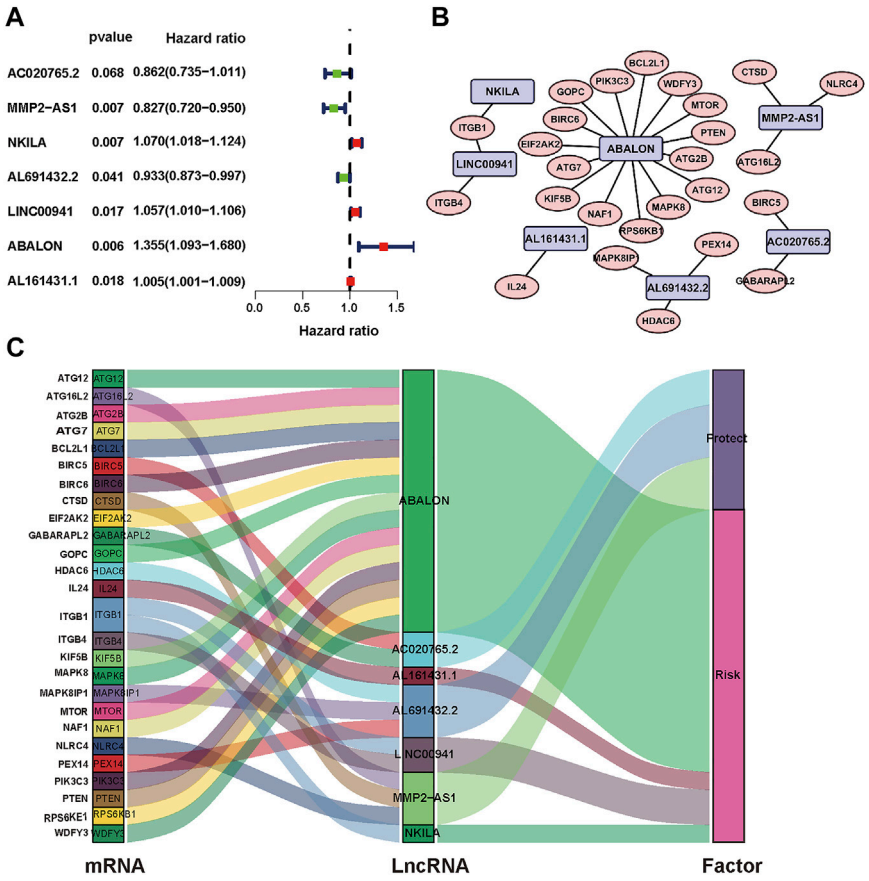


FIGURE 2
Identification of prognostic autophagy-associated lncRNAs in NSCLC patients. **(A)** Multivariate Cox analysis to establish the prognostic signature. **(B)** Establishment of the co-expression network in NSCLC. Blue represents lncRNAs and pink represents mRNAs. **(C)** The Sankey diagram shows the connection degree between 26 mRNAs and 7 autophagy-associated lncRNAs.

was correlated with the risk score. The heatmap showing the expression of 7 autophagy-associated lncRNAs in NSCLC samples exhibited that NKILA, LINC00941, ABALON, and AL161431.1 were high in the high-risk group, whereas AC020765.2, MMP2-AS1, and AL691432.2 were low in the low-risk group. The risk curve, the scatterplot, and the heat map were shown in Figure 3C.

To establish an accurate prognosis for NSCLC patients, we evaluated the survival rate about 7 autophagy-associated lncRNAs at 1, 2, and 3 years *via* calculating a nomogram. And the nomogram may help specialists make individualized clinical therapy for NSCLC patients (Figure 4A).

Clinical value of the prognostic autophagy-associated long non-coding RNAs signature

To assess its clinical value, we calculated risk scores about the autophagy-associated lncRNAs model and clinical data from TCGA, including age, gender, grade, and TNM stage. Then, we used univariate and multivariate Cox regression analyses (Figures 4B,C, respectively). The univariate analysis displayed that stage, T stage, N stage, M stage, and the prognostic risk score were closely associated with OS. Then the multivariate analysis revealed that age and risk score were closely associated to OS. We performed multiple ROC curve to assess the accuracy of prediction about the signature. The results displayed that the AUC value of the prognostic risk-related model was 0.660, which was higher than other clinic factors (Figure 4D). In general, the ROC curves indicated that the predictive accuracy of the prognostic lncRNAs model in NSCLC was acceptable.

Then, to determine the correlation between autophagy-associated lncRNAs and OS, we drew Kaplan Meier curves. And we identified the seven lncRNAs (AC020765.2, MMP2-AS1, NKILA, LINC00941, ABALON, AL691432.2, and AL161431.1). These results determined the seven

TABLE 2 Correlation between autophagy genes and lncRNAs in NSCLC.

Autophagy genes	lncRNA	Correlation	*p*-value
BIRC5	AC020765.2	0.35	3.32E-31
GABARAPL2	AC020765.2	0.30	4.15E-23
ATG16L2	MMP2-AS1	0.38	1.23E-37
CTSD	MMP2-AS1	0.34	6.07E-30
NLRC4	MMP2-AS1	0.37	2.17E-34
ITGB1	NKILA	0.30	4.58E-23
HDAC6	AL691432.2	0.31	8.37E-24
MAPK8IP1	AL691432.2	0.38	3.69E-36
PEX14	AL691432.2	0.39	4.15E-39
ITGB1	LINC00941	0.32	1.16E-26
ITGB4	LINC00941	0.31	1.03E-24
ATG12	ABALON	0.32	3.30E-26
ATG2B	ABALON	0.39	8.16E-40
ATG7	ABALON	0.35	6.50E-31
BCL2L1	ABALON	0.33	6.61E-28
BIRC6	ABALON	0.55	2.05E-82
EIF2AK2	ABALON	0.31	5.94E-24
GOPC	ABALON	0.31	2.04E-24
KIF5B	ABALON	0.31	6.05E-25
MAPK8	ABALON	0.32	6.24E-26
MTOR	ABALON	0.32	6.32E-27
NAF1	ABALON	0.33	4.50E-28
PIK3C3	ABALON	0.36	1.15E-32
PTEN	ABALON	0.44	2.77E-50
RPS6KB1	ABALON	0.33	2.34E-28
WDFY3	ABALON	0.53	1.88E-76
IL24	AL161431.1	0.34	6.86E-29

$p < 0.05$ was regarded as a significant difference.

autophagy-associated lncRNAs were closely related to the OS of NSCLC patients (Figures 5A–G).

GSEA enrichment

According to the prognostic model of autophagy-associated lncRNA, GSEA divided the gene set into high-risk and low-risk groups (FDR q-value < 0.05). The result displayed that the high-risk group was primarily enriched in stromal pathways, such as ECM receptor interaction and regulation of actin cytoskeleton. Furthermore, the O-glycan biosynthesis pathway was also important. It is worth noting that the adherens junction pathway was closely associated with the autophagy-associated lncRNAs. These results further confirmed that these autophagy-associated lncRNAs could regulate NSCLC and autophagy by some special pathways, which may inspire new approaches of therapy in NSCLC (Figure 5H).

Downregulation of apoptotic BCL2L1-antisense long non-coding RNA inhibited proliferation and metastasis in non-small cell lung cancer cells

Survival analysis displayed the ABALON was a risk factor with a poor prognosis. We explored the biological signaling pathways by GSEA, and the adherens junction pathway showed a crucial enrichment score (NES = 2.08, FDR q-val = 0.020). Consequently, we select ABALON for experimental validation. Based on the result of qRT-PCR, we chose two representative NSCLC cell lines (NCI-H292 and A549) to carry out experiments *in vitro* (Figure 6A).

First, the siRNA of ABALON and Control were transfected into NCI-H292 and A549 cells. The results of MTT, colony formation, and EDU assays showed low-expression of ABALON inhibited the proliferation of NCI-H292 and A549 cells (Figures 6C–E, 7B–D). The results of transwell and wound healing assays further displayed low expression of ABALON inhibited the metastasis of A549 cells (Figures 6F,G). GSEA analysis indicated that ABALON may influence the behavior of NSCLC *via* the adherens junction pathway (Figure 7E). Therefore, the results of western blotting displayed knockdown of ABALON led to decrease expression of ß-catenin and increase expression of E-cadherin (Figure 7F). In summary, these results show that depletion of ABALON can inhibit NSCLC progression.

Downregulation of apoptotic BCL2L1-antisense long non-coding RNA inhibits autophagy in non-small cell lung cancer cells

To assess changes of autophagy after depletion of ABALON, we used western blotting and tandem mRFP-GFP fluorescence microscopy as the most common approaches. Western blotting was performed to evaluate protein levels of autophagy markers, including LC3 and p62. Depletion of ABALON compared with siControl significantly reduced the LC3-II/LC3-I ratio and simultaneously increased the level of the autophagy substrate p62 ($p < 0.05$) (Figure 7F).

However, the changes in the abundance of these proteins as assessed by western blot analysis cannot fully reflect real alteration of the autophagic flux alterations, which are mainly reflected in autophagosome docking and fusion with lysosomes. Hence, we transferred NSCLC cells with the mRFP-GFP-LC3B plasmid to visualize phagosome-lysosome fusion *via* fluorescence microscopy. Under the acidic and/or proteolytic conditions of the lysosomal cavity, GFP signal becomes sensitive, while mRFP becomes more stable. Hence, when autophagosomes fuse with

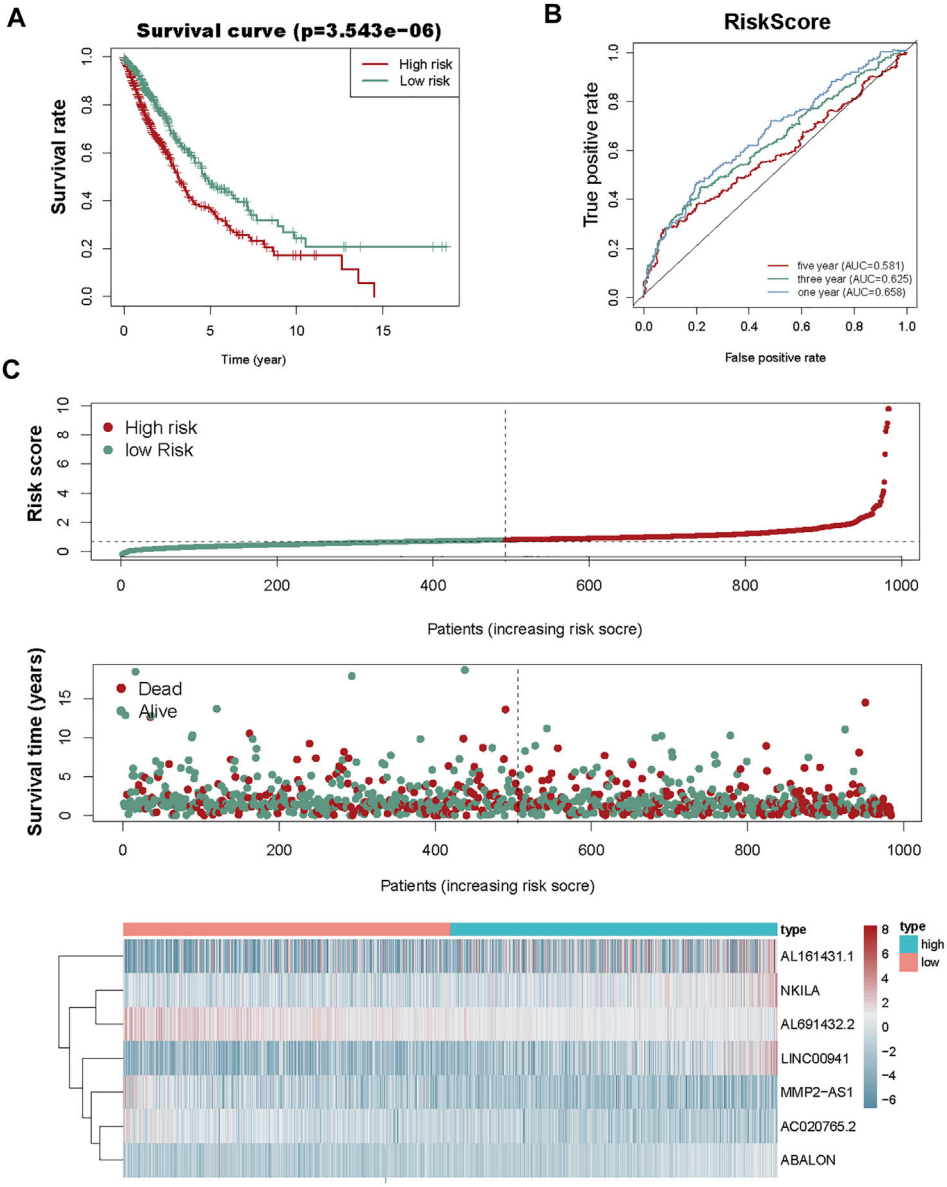


FIGURE 3
Risk score analysis of the prognostic model of these seven autophagy-associated lncRNAs. **(A)** KM survival analysis for high- and low-risk groups. Red, high risk. Green, low-risk. **(B)** ROC analysis to evaluate the predictive ability of the model. **(C)** Risk score distribution (top), survival status distribution (middle), and heat map of 7 autophagy-associated lncRNAs (bottom).

lysosom, the index of PH in cells would down and the GFP signal is quenched. Colocalization of GFP and mRFP fluorescence in cells means autophagosomes have not fused with lysosomes, indicating that autophagy is blocked. Conversely, if only mRFP fluorescence is visible, there is highly active autophagy. NCI-H292 and A549 cells transfected siABALON only exhibited yellow fluorescence, which confirmed the results of western blotting (Figures 7F,G). These results indicated that siABALON restrained autophagy in NSCLC cells *in vitro*.

Discussion

Autophagy is the process in which damaged proteins and organelles are engulfed and transported to lysosomes for degradation and circulation. The effect of autophagy seems to be extremely complicated and is difficult to summarize in cancer (Mizushima, 2007; Morel et al., 2017). It can play a tumor suppressor as well as an oncogenic effect. Through regulating functions of proteins and organelles, autophagy can keep the genome stable, reduce normal cell death and prevent

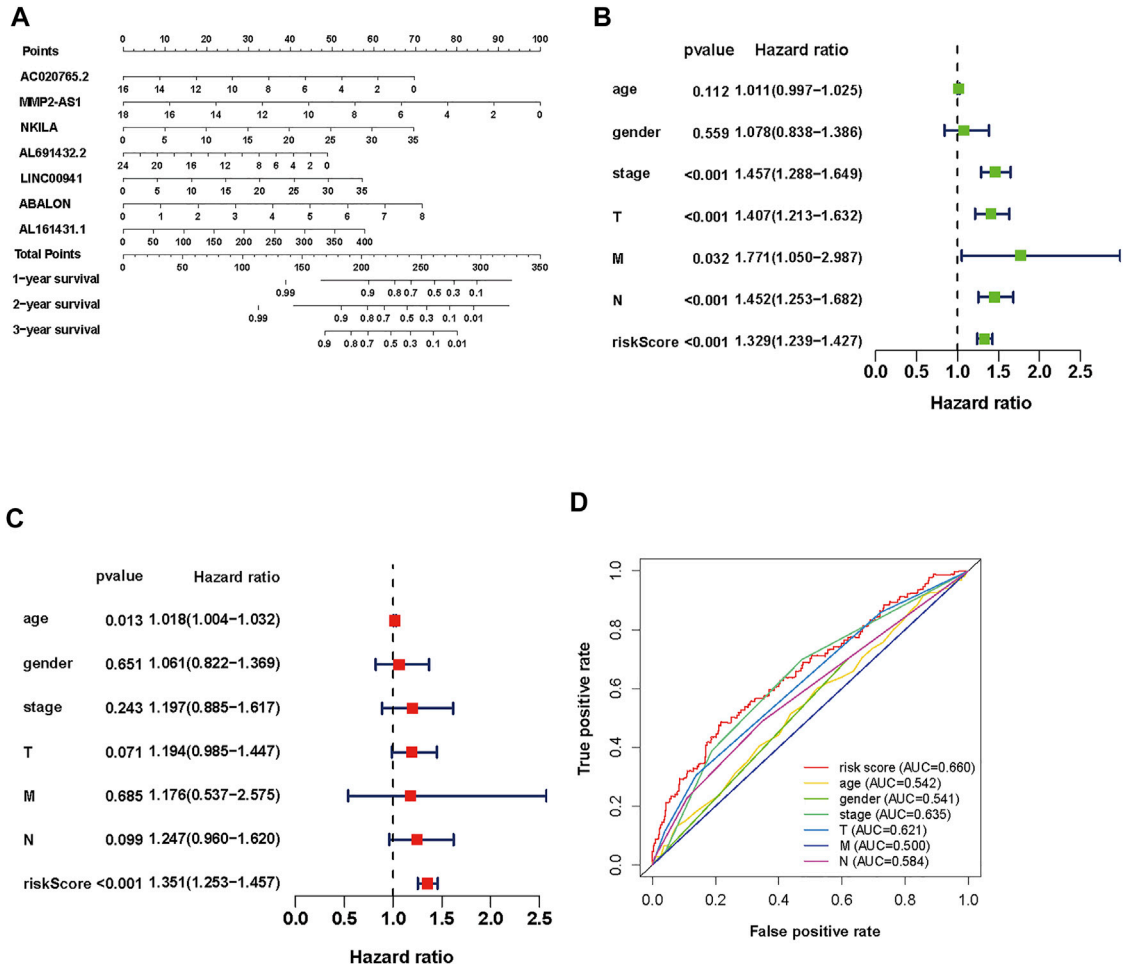


FIGURE 4
Evaluation of the prognostic risk model of the 7 autophagy-associated lncRNAs in NSCLC. **(A)** Nomogram of the 7 autophagy-associated lncRNAs. **(B,C)** The univariate **(B)** and multivariate **(C)** Cox analysis of risk score model and clinical features. **(D)** The ROC curve analysis displays the prognostic accuracy of clinical features such as age, gender, T stage, M stage, N stage, and risk score.

tumorigenesis. For example, autophagy can remove senescent organelles and defective proteins (Wang and Klionsky, 2011; Wu et al., 2012). However, autophagy also maintains cancer cell metabolism, promotes tumor genesis and development, and resistance to therapeutic drugs (White, 2012; Poillet-Perez and White, 2019). Hence, we consider that autophagy can improve the therapeutic effects of tumor-targeting drugs and reduce drug resistance.

In this study, we determined 27 autophagy-associated mRNAs in NSCLC, which were associated with the 7 autophagy-associated lncRNAs. Among them, only ATG12, MTOR, ATG7, and PTEN are known to be related to autophagy in NSCLC. Autophagy-related (ATG) proteins play crucial roles in cancer. The numerous ATG proteins and their core complexes including the ULK/Atg1 kinase core complex, ATG9A/Atg9 trafficking system, ATG12/Atg12-conjugation system and LC3/Atg8-conjugation system, exert multiple activities in the autophagy pathway and take part in all processes of autophagy (Li et al., 2020). Among the 4 autophagy-associated genes identified in this study, ATG7, which acts as an essential protein to promote autophagy, was informed to attenuate the proliferation, invasion, and metastasis of NSCLC (Zheng et al., 2018; Cao et al., 2020). He et al. (2020) identified ATG12 was a target of miR-372-3p. ATG12 can promote autophagy and radiosensitivity in lung adenocarcinoma by suppressing miR-372-3p. Furthermore, we found that PTEN and mTOR exhibited a high correlation coefficient with autophagy. PTEN is an important tumor suppressor and the main antagonist of PI3K, which can promote the degradation of AKT (Lim et al., 2015; Jamaspishvili et al., 2018). In addition, miR-181 was found to mediate cisplatin-resistance and attenuate autophagy through the PTEN/PI3K/AKT pathway in NSCLC (Liu et al., 2018).

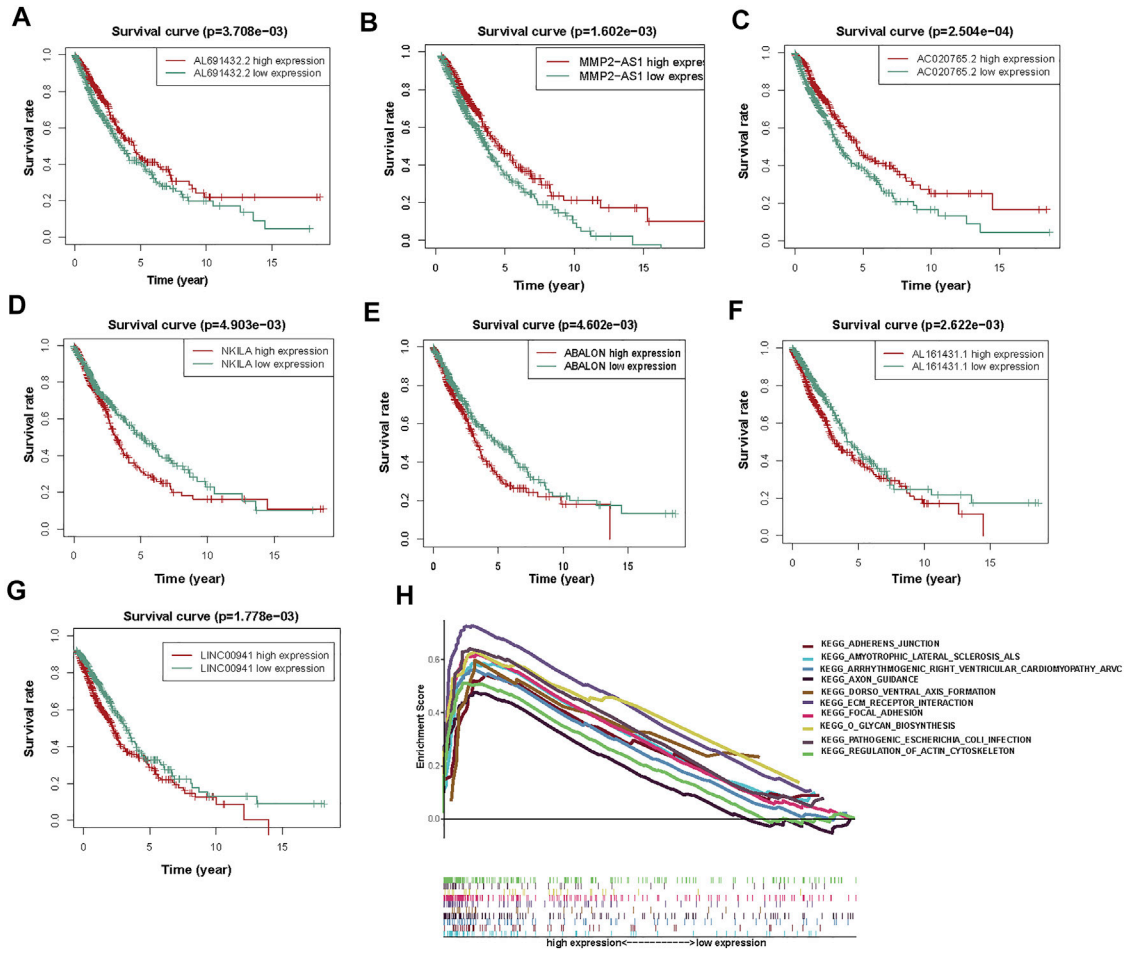


FIGURE 5
The KM survival curves and GSEA analysis of 7 prognostic autophagy-associated lncRNAs. **(A)** AL691432.2, **(B)** MMP2-AS1, and **(C)** AC020765.2 were favorable prognostic factors. **(D)** NKILA, **(E)** ABALON, **(F)** AL161431.1, and **(G)** LINC00941 were harmful prognostic factors. **(H)** Top ten primarily enriched categories of 7 autophagy-associated lncRNAs in the high-risk group by GSEA.

However, the roles of other autophagy-associated mRNAs in NSCLC remain unclear.

Additionally, we identified 7 autophagy-associated lncRNAs in NSCLC patients from TCGA. Notably, lncRNAs have become a key regulatory factor in various cellular processes. Studies show that lncRNAs can also regulate autophagy in cancer. The specific mechanisms through which lncRNAs regulate autophagy can be segmented into three classes: 1) LncRNAs familiarly modulate autophagy through regulating the expression of ATG proteins. 2) LncRNAs regulate tumor genesis through the AKT/mTOR signaling pathway. mTOR forms two different signal complexes, mTOR complex 1(mTORC1) and mTORC2, by binding to a variety of companion proteins. mTORC1 inhibits the initiation of autophagy by phosphorylating ATG13 and autophagy activating kinase (ULK) (Kim and Guan, 2015). 3) LncRNAs can also act as competing endogenous RNAs to modulate miRNAs. Many studies investigated the effect of autophagy and lncRNAs in other cancers, but less research has been done in NSCLC.

Further bioinformatics analysis confirmed that these 7 autophagy-associated lncRNAs were associated with the OS of NSCLC patients. Among them, NKILA, LINC00941, ABALON, and AL161431.1 were risk factors for prognosis of NSCLC, and other lncRNAs (AC020765.2, MMP2-AS1, and AL691432.2) had the opposite effect. Previous studies have shown that LINC00941, also called MSC upregulated factor (lncRNA-MUF), was negatively associated with OS and phosphorylation of the PI3K/AKT signaling pathway in lung adenocarcinoma patients (Wang et al., 2019b). Another study also found that depletion of LINC00941 inhibited EMT and activated Wnt/ß-catenin signaling in hepatocellular carcinoma (Yan et al., 2017). Additionally, Ren et al. informed that LINC00941 promoted the progression of

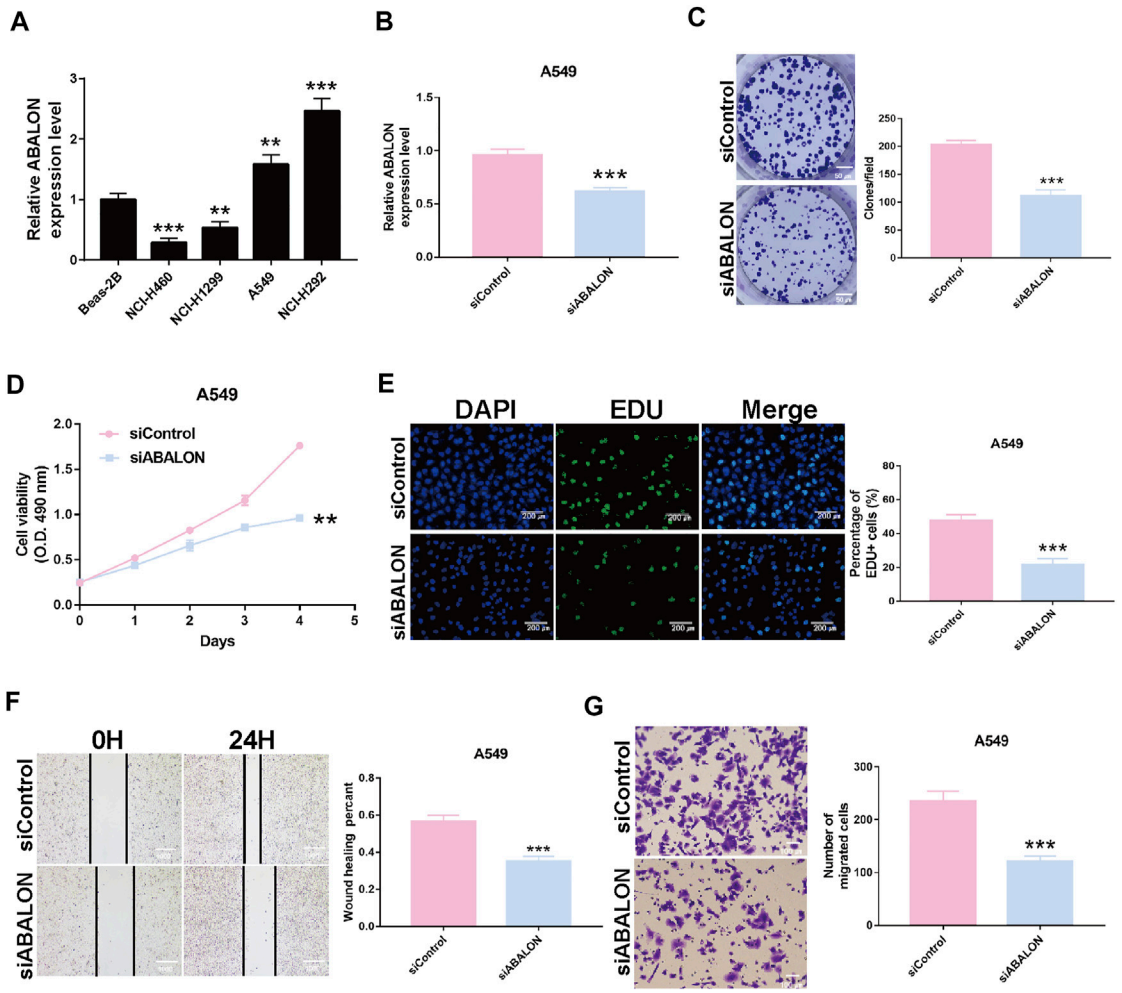


FIGURE 6
Downregulation of ABALON suppressed the proliferation and metastasis in A549 cells. **(A)** The expression of ABALON in corresponding cell lines identified by RT-qPCR. **(B)** RT-qPCR for the expression of ABALON in siControl and siABALON A549 cells. **(C–E)** Colony formation **(C)**, MTT **(D)**, and EDU **(E)** assays identified proliferation was suppressed. **(F,G)** Cell metastasis suppression was identified by wound healing **(F)**, and transwell **(G)** assays. $^{**}p < 0.01$, $^{***}p < 0.001$.

NSCLC through the miR-877-3p/VEGFA axis (Ren et al., 2021). AL161431.1 facilitated the proliferation and metastasis by regulating miR-1252-5p in endometrial carcinoma (Gu and Liu, 2020). Similarly, Qiang et al. found that AL161431.1 was also negatively associated with the OS of patients in lung squamous cell carcinoma (Ju et al., 2020). Matrix metalloproteinases (MMPs) are a family of zinc-dependent endopeptidases, and some studies indicated MMP2 promotes invasion and metastasis of NSCLC cells (Hsieh et al., 2019). Thus, we speculate that MMP2-AS1, encoding MMP2 antisense RNA1, may restrain the function of MMP2 in NSCLC. Consistent with our finding, LINC00941 tended to be a high-risk factor and AL691432.2 tended to be a low-risk factor in the construction of a prognostic model for NSCLC.

E-cadherin is an important cyctomembrane component and plays an important role in adherhens junction. E-cadherin and the Catenins (β-catenin, α-catenin, p120-catenin) are the substates of important kinases and phosphatases for regulating adherens junction (Coopman and Djiane, 2016). In this study, we demonstrated that apoptotic BCL2L1-antisense LncRNA (ABALON) acts as an oncogene in NSCLC *via* adherens junction pathway. We also confirmed that ABALON promoted autophagy by western blotting and fluorescence microscopy. However, the mechanisms of ABALON need further study. In summary, the results reveal the 7 identified autophagy-associated lncRNAs have a prognostic value in NSCLC.

However, our study has certain limitations. First, we applied traditional statistical analysis methods to establish

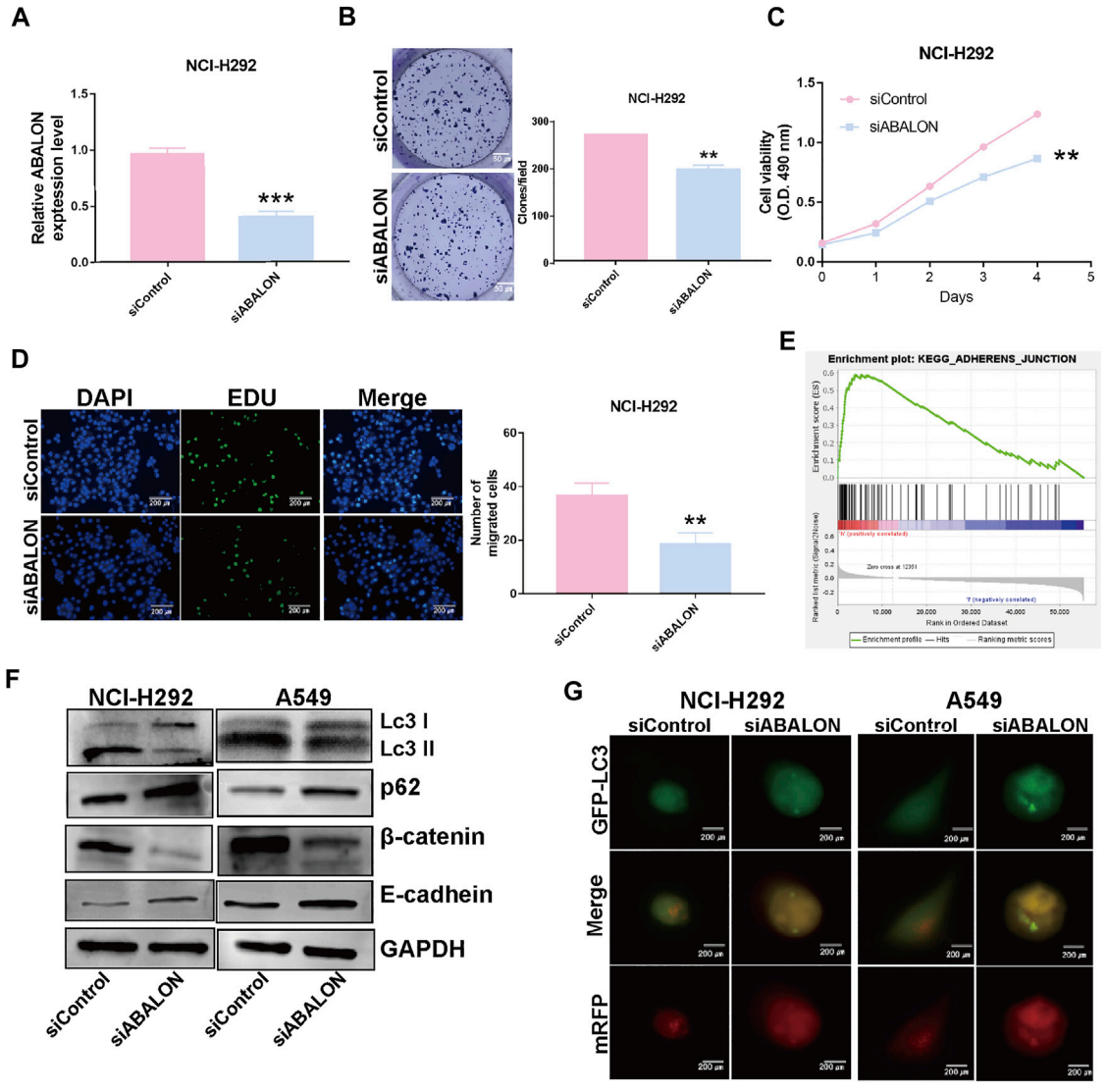


FIGURE 7
Downregulation of ABALON suppressed the proliferation and metastasis in NCI-H292 cells and promoted autophagy level. **(A)** RT-qPCR for the expression of ABALON in siControl and siABALON NCI-H292 cells. **(B–D)** Colony formation **(B)**, MTT **(C)**, EDU **(D)** assays revealed cell proliferation promotion *via* downregulation of ABALON in NCI-H292 cells. **(E)** A significant enrichment results between ABALON-high and ABALON-low groups in NSCLC. **(F)** Western blotting to illustrate the expression levels of LC3-II, P62, E-cadherin and β-catenin. **(G)** Tandem mRFP-GFP fluorescence microscopy to illustrate the levels of autophagy. $^{**}p < 0.001$.

and evaluate prognostic risk models for 7 autophagy-associated lncRNAs. Although these methods have been applied and validated in many studies, we need to refine our further studies with more advanced methods and techniques in the future. In the study, we confirmed ABALON promotes the proliferation, metastasis and autophagy in NSCLC cells *via* biological experiments. However, the mechanism by which ABALON regulates NSCLC cells remains unclear. To further verify our prediction results, in-depth studies about the molecular mechanisms are needed.

Conclusion

In conclusion, this study identified 26 autophagy-associated genes in NSCLC, and constructed a prognostic risk signature of 7 autophagy-associated lncRNAs. The signature was accurate to predict prognosis of NSCLC patients with high reliability. Moreover, experimental validation confirmed that ABALON promotes the proliferation, metastasis, and autophagy in NSCLC cells. This signature provides a basis for further studies on the clinical application of these autophagy-associated lncRNAs. In the future, with prospective validation,

the 7 autophagy-associated lncRNAs signature may improve predictive accuracy and guide individualized therapy for non-small cell lung cancer 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 authors.

Author contributions

YY and GJ designed the study; HJ and ZJ preformed the experiments; HJ analyze data. ZJ drafted the paper; YY and GJ reviewed and revised the paper. All authors read and approved the final manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (No. 81702624 and 81502518).

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.2022.919857/full#supplementary-material

References

Bai, Y., Liu, X., Qi, X., Liu, X., Peng, F., Li, H., et al. (2019). PDIA6 modulates apoptosis and autophagy of non-small cell lung cancer cells via the MAP4K1/JNK signaling pathway. *EBioMedicine* 42, 311–325. doi:10.1016/j.ebiom.2019.03.045

Barnard, R. A., Regan, D. P., Hansen, R. J., Maycotte, P., Thorburn, A., and Gustafson, D. L. (2016). Autophagy inhibition delays early but not late-stage metastatic disease. *J. Pharmacol. Exp. Ther.* 358, 282–293. doi:10.1124/jpet.116.233908

Bravo-San Pedro, J. M., Kroemer, G., and Galluzzi, L. (2017). Autophagy and mitophagy in cardiovascular disease. *Circ. Res.* 120, 1812–1824. doi:10.1161/CIRCRESAHA.117.311082

Cao, Q., You, X., Xu, L., Wang, L., and Chen, Y. (2020). PAQR3 suppresses the growth of non-small cell lung cancer cells via modulation of EGFR-mediated autophagy. *Autophagy* 16, 1236–1247. doi:10.1080/15548627.2019.1659654

Choi, Y., Bowman, J. W., and Jung, J. U. (2018). Autophagy during viral infection - a double-edged sword. *Nat. Rev. Microbiol.* 16, 341–354. doi:10.1038/s41579-018-0003-6

Coopman, P., and Djiane, A. (2016). Adherens Junction and E-Cadherin complex regulation by epithelial polarity. *Cell. Mol. Life Sci.* 73, 3535–3553. doi:10.1007/s00018-016-2260-8

Fang, Y., and Fullwood, M. J. (2016). Roles, functions, and mechanisms of Long non-coding RNAs in cancer. *Genomics Proteomics Bioinforma.* 14, 42–54. doi:10.1016/j.gpb.2015.09.006

Gu, Z. R., and Liu, W. (2020). The LncRNA AL161431.1 targets miR-1252-5p and facilitates cellular proliferation and migration via MAPK signaling in endometrial carcinoma. *Eur. Rev. Med. Pharmacol. Sci.* 24, 2294–2302. doi:10.26355/eurrev_202003_20495

He, H., Song, X., Yang, Z., Mao, Y., Zhang, K., Wang, Y., et al. (2020). Upregulation of KCNQ1OT1 promotes resistance to stereotactic body radiotherapy in lung adenocarcinoma by inducing ATG5/ATG12-mediated autophagy via miR-372-3p. *Cell Death Dis.* 11, 883. doi:10.1038/s41419-020-03083-8

Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J., Jr., Wu, Y. L., et al. (2017). Lung cancer: Current therapies and new targeted treatments. *Lancet* 389, 299–311. doi:10.1016/S0140-6736(16)30958-8

Hsieh, N. T., Huang, C. Y., Li, C. C., Wang, I. C., and Lee, M. F. (2019). MED28 and forkhead box M1 (FOXM1) mediate matrix metalloproteinase 2 (MMP2)-dependent cellular migration in human nonsmall cell lung cancer (NSCLC) cells. *J. Cell. Physiol.* 234, 11265–11275. doi:10.1002/jcp.27784

Jamaspishvili, T., Berman, D. M., Ross, A. E., Scher, H. I., De Marzo, A. M., Squire, J. A., et al. (2018). Clinical implications of PTEN loss in prostate cancer. *Nat. Rev. Urol.* 15, 222–234. doi:10.1038/nrurol.2018.9

Ju, Q., Zhao, Y. J., Ma, S., Li, X. M., Zhang, H., Zhang, S. Q., et al. (2020). Genome-wide analysis of prognostic-related lncRNAs, miRNAs and mRNAs forming a competing endogenous RNA network in lung squamous cell carcinoma. *J. Cancer Res. Clin. Oncol.* 146, 1711–1723. doi:10.1007/s00432-020-03224-8

Kim, Y. C., and Guan, K. L. (2015). mTOR: a pharmacologic target for autophagy regulation. *J. Clin. Invest.* 125, 25–32. doi:10.1172/JCI73939

Kopp, F., and Mendell, J. T. (2018). Functional classification and experimental dissection of Long noncoding RNAs. *Cell* 172, 393–407. doi:10.1016/j.cell.2018.01.011

Levine, B., and Kroemer, G. (2019). Biological functions of autophagy genes: A disease perspective. *Cell* 176, 11–42. doi:10.1016/j.cell.2018.09.048

Li, X., He, S., and Ma, B. (2020). Autophagy and autophagy-related proteins in cancer. *Mol. Cancer* 19, 12. doi:10.1186/s12943-020-1138-4

Lim, H. J., Crowe, P., and Yang, J. L. (2015). Current clinical regulation of PI3K/PTEN/Akt/mTOR signalling in treatment of human cancer. *J. Cancer Res. Clin. Oncol.* 141, 671–689. doi:10.1007/s00432-014-1803-3

Liu, J., Xing, Y., and Rong, L. (2018). miR-181 regulates cisplatin-resistant non-small cell lung cancer via downregulation of autophagy through the PTEN/PI3K/AKT pathway. *Oncol. Rep.* 39, 1631–1639. doi:10.3892/or.2018.6268

Ma, F., Ding, M. G., Lei, Y. Y., Luo, L. H., Jiang, S., Feng, Y. H., et al. (2020). SKIL facilitates tumorigenesis and immune escape of NSCLC via upregulating TAZ/autophagy axis. *Cell Death Dis.* 11, 1028. doi:10.1038/s41419-020-03200-7

Ma, Y., Galluzzi, L., Zitvogel, L., and Kroemer, G. (2013). Autophagy and cellular immune responses. *Immunity* 39, 211–227. doi:10.1016/j.immuni.2013.07.017

Mizushima, N. (2007). Autophagy: Process and function. *Genes Dev.* 21, 2861–2873. doi:10.1101/gad.1599207

Mizushima, N., and Levine, B. (2020). Autophagy in human diseases. *N. Engl. J. Med.* 383, 1564–1576. doi:10.1056/NEJMra2022774

Morel, E., Mehrpour, M., Botti, J., Dupont, N., Hamaï, A., Nascimbeni, A. C., et al. (2017). Autophagy: A druggable process. *Annu. Rev. Pharmacol. Toxicol.* 57, 375–398. doi:10.1146/annurev-pharmtox-010716-104936

Poillet-Perez, L., and White, E. (2019). Role of tumor and host autophagy in cancer metabolism. *Genes Dev.* 33, 610–619. doi:10.1101/gad.325514.119

Ren, M. H., Chen, S., Wang, L. G., Rui, W. X., and Li, P. (2021). LINC00941 promotes progression of non-small cell lung cancer by sponging miR-877-3p to regulate VEGFA expression. *Front. Oncol.* 11, 650037. doi:10.3389/fonc.2021.650037

Slack, F. J., and Chinnaiyan, A. M. (2019). The role of non-coding RNAs in oncology. *Cell* 179, 1033–1055. doi:10.1016/j.cell.2019.10.017

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2012). Global cancer statistics, 2012. *Ca. Cancer J. Clin.* 65, 87–108. doi:10.3322/caac.21262

Ueno, T., and Komatsu, M. (2017). Autophagy in the liver: Functions in health and disease. *Nat. Rev. Gastroenterol. Hepatol.* 14, 170–184. doi:10.1038/nrgastro.2016.185

Wang, K., and Klionsky, D. J. (2011). Mitochondria removal by autophagy. *Autophagy* 7, 297–300. doi:10.4161/auto.7.3.14502

Wang, L., Zhao, H., Xu, Y., Li, J., Deng, C., Deng, Y., et al. (2019). Systematic identification of lincRNA-based prognostic biomarkers by integrating lincRNA expression and copy number variation in lung adenocarcinoma. *Int. J. Cancer* 144, 1723–1734. doi:10.1002/ijc.31865

Wang, Y., Lu, J. H., Wu, Q. N., Jin, Y., Wang, D. S., Chen, Y. X., et al. (2019). LncRNA LINRIS stabilizes IGF2BP2 and promotes the aerobic glycolysis in colorectal cancer. *Mol. Cancer* 18, 174. doi:10.1186/s12943-019-1105-0

White, E. (2012). Deconvoluting the context-dependent role for autophagy in cancer. *Nat. Rev. Cancer* 12, 401–410. doi:10.1038/nrc3262

White, E., Mehnert, J. M., and Chan, C. S. (2015). Autophagy, metabolism, and cancer. *Clin. Cancer Res.* 21, 5037–5046. doi:10.1158/1078-0432.CCR-15-0490

Wu, Q., Ma, J., Wei, J., Meng, W., Wang, Y., and Shi, M. (2021). lncRNA SNHG11 promotes gastric cancer progression by activating the wnt/β-catenin pathway and oncogenic autophagy. *Mol. Ther.* 29, 1258–1278. doi:10.1016/j.ymthe.2020.10.011

Wu, W. K., Coffelt, S. B., Cho, C. H., Wang, X. J., Lee, C. W., Chan, F. K., et al. (2012). The autophagic paradox in cancer therapy. *Oncogene* 31, 939–953. doi:10.1038/onc.2011.295

Xu, S., Wang, P., Zhang, J., Wu, H., Sui, S., Zhang, J., et al. (2019). Ai-lncRNA EGOT enhancing autophagy sensitizes paclitaxel cytotoxicity via upregulation of ITPR1 expression by RNA-RNA and RNA-protein interactions in human cancer. *Mol. Cancer* 18, 89. doi:10.1186/s12943-019-1017-z

Yan, X., Zhang, D., Wu, W., Wu, S., Qian, J., Hao, Y., et al. (2017). Mesenchymal stem cells promote hepatocarcinogenesis via lncRNA-MUF interaction with ANXA2 and miR-34a. *Cancer Res.* 77, 6704–6716. doi:10.1158/0008-5472.CAN-17-1915

Yang, S., Zhang, Z., and Wang, Q. (2019). Emerging therapies for small cell lung cancer. *J. Hematol. Oncol.* 12, 47. doi:10.1186/s13045-019-0736-3

Zheng, T., Li, D., He, Z., Feng, S., and Zhao, S. (2018). Long noncoding RNA NBAT1 inhibits autophagy via suppression of ATG7 in non-small cell lung cancer. *Am. J. Cancer Res.* 8, 1801

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 06 September 2022
DOI 10.3389/fgene.2022.921902

OPEN ACCESS

EDITED BY
Yingming Sun,
Fujian Medical University, China

REVIEWED BY
Jin-Zhou Xu,
Huazhong University of Science and Technology, China
Zhi-qian Zhang,
Southern University of Science and Technology, China
Song Xu,
Tianjin Medical University General Hospital, China

*CORRESPONDENCE
Jiancheng Li,
jianchengli_jack@126.com

†These authors have contributed equally to this work

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

RECEIVED 16 April 2022
ACCEPTED 25 July 2022
PUBLISHED 06 September 2022

CITATION
Zheng J, Chen X, Huang B and Li J (2022), A novel immune-related radioresistant lncRNAs signature based model for risk stratification and prognosis prediction in esophageal squamous cell carcinoma.
Front. Genet. 13:921902.
doi: 10.3389/fgene.2022.921902

A novel immune-related radioresistant lncRNAs signature based model for risk stratification and prognosis prediction in esophageal squamous cell carcinoma

Jianqing Zheng[1,2,3†], Xiaohui Chen[1,3,4†], Bifen Huang[5] and Jiancheng Li[1,3,6*]

[1]Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou, Fujian, China, [2]Department of Radiation Oncology, The Second Affiliated Hospital of Fujian Medical University, Quanzhou, Fujian, China, [3]The Graduate School of Fujian Medical University, Fuzhou, Fujian, China, [4]Department of Thoracic Surgery, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou, Fujian, China, [5]Department of Obstetrics and Gynecology, Quanzhou Medical College People's Hospital Affiliated, Fuzhou, Fujian, China, [6]Department of Radiation Oncology, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou, Fujian, China

Background and purpose: Radioresistance remains a major reason of radiotherapeutic failure in esophageal squamous cell carcinoma (ESCC). Our study is to screen the immune-related long non-coding RNA (ir-lncRNAs) of radiation-resistant ESCC (rr-ESCC) *via* Gene Expression Omnibus (GEO) database and to construct a prognostic risk model.

Methods: Microarray data (GSE45670) related to radioresistance of ESCC was downloaded from GEO. Based on pathologic responses after chemoradiotherapy, patients were divided into a non-responder (17 samples) and responder group (11 samples), and the difference in expression profiles of ir-lncRNAs were compared therein. Ir-lncRNA pairs were constructed for the differentially expressed lncRNAs as prognostic variables, and the microarray dataset (GSE53625) was downloaded from GEO to verify the effect of ir-lncRNA pairs on the long-term survival of ESCC. After modelling, patients are divided into high- and low-risk groups according to prognostic risk scores, and the outcomes were compared within groups based on the COX proportional hazards model. The different expression of ir-lncRNAs were validated using ECA 109 and ECA 109R cell lines *via* RT-qPCR.

Results: 26 ir-lncRNA genes were screened in the GSE45670 dataset with differential expression, and 180 ir-lncRNA pairs were constructed. After matching with ir-lncRNA pairs constructed by GSE53625, six ir-lncRNA pairs had a significant impact on the prognosis of ESCC from univariate analysis model, of which three ir-lncRNA pairs were significantly associated with prognosis in multivariate COX analysis. These three lncRNA pairs were used as prognostic indicators to construct a prognostic risk model, and the predicted risk scores were calculated. With a median value of 2.371, the patients were

divided into two groups. The overall survival (OS) in the high-risk group was significantly worse than that in the low-risk group ($p < 0.001$). The 1-, 2-, and 3-year prediction performance of this risk-model was 0.666, 0.702, and 0.686, respectively. In the validation setting, three ir-lncRNAs were significantly up-regulated, while two ir-lncRNAs were obviouly down-regulated in the responder group.

Conclusion: Ir-lncRNAs may be involved in the biological regulation of radioresistance in patients with ESCC; and the prognostic risk-model, established by three ir-lncRNAs pairs has important clinical value in predicting the prognosis of patients with rr-ESCC.

KEYWORDS

radioresistance, esophageal squamous cell carcinoma, lncRNA, prognostic model, bioinformatics

Introduction

Esophageal cancer (EPC) is one of the most lethal tumors in China and worldwide (Sung et al., 2021). Esophageal squamous cell carcinoma (ESCC) is the major pathologic type in Chinese population, accounting for more than 90% cases, and more than 45%–60% of them were diagnosed at an advanced stage (Yang et al., 2018). In general, the prognosis of EPC is very poor, the cancer-specific mortality rate ranks fourth, and the 5-year overall survival rate is less than 30% (Chen et al., 2016; Zeng et al., 2018). Radiotherapy is one of the main treatments for EPC, especially in patients at an advanced stage. However, after radiotherapy with/without chemotherapy, the total objective response rate (ORR) is only 60%–76%, and radiotherapy resistance is an important reason for the local failure of radiotherapy for EPC (Chen et al., 2021). Although multiple factors may affect the radiosensitivity, the specific mechanisms of long non-coding RNA (lncRNAs) in radioresistance are still worth exploring. The outcomes of radiotherapy are heterogeneous, and no clinical or pathological method could predict tumor response of radiotherapy.

LncRNA referred to a type of RNAs whose length is greater than 200 bp and does not encode or translate proteins after transcription. LncRNAs are involved in various physiological and pathological processes, such as the cellular replication, transcription, translation and so on. In the process of gene expression and transcription, about 70% of human genes are regulated by lncRNAs (Batista and Chang, 2013; Shi et al., 2013). It's reported that lncRNAs are widely involved in the biological processes of cancer, including tumor radioresistance (Li and Chen, 2013). Many studies believed that the prognosis of tumors is closely related to the tumor microenvironment (TME), and lncRNAs may participate in the regulation of TME through molecular biological functions such as chromatin modification, transcription, and post-transcriptional processing (Sounni and Noel, 2013; Singh et al., 2016). In addition, some lncRNAs are also involved in the regulation of immune function in the TME, where may involve various types of cells and cytokines. Such lncRNAs are often referred to as immune-related lncRNAs (ir-lncRNAs) (Bremnes et al., 2016; Chen M. M. et al., 2020). In order to explore whether ir-lncRNAs are involved in the biological process of radio-resistance in patients with ESCC, our study analyzed the differences in the expression of ir-lncRNAs between ESCC-patients with radioresistance and complete tumor remission based on gene expression profiling data in the Gene Expression Omnibus (GEO) database. Ir-lncRNA pairs, which had high correlation with immune genes, were selected to establish a prognostic risk-model to explore the value of ir-lncRNAs in predicting the prognosis of ESCC patients with radioresistance, and further to find potential biomarkers of radioresistance.

Materials and methods

Dataset collection and preparation

Using "esophageal cancer, esophagus cancer, and radioresistance" as keywords, mRNA and lncRNA expression profiling data related to radioresistance of esophageal cancer was searched in the GEO database (https://www.ncbi.nlm.nih.gov/). Two public esophageal cancer microarray profiling datasets (GSE45670 and GSE53625) were downloaded from GEO and were finally selected for data mining. The GSE45670 microarray profiling datasets were provided by Wen et al., where gene expression analyses were performed on pretreatment cancer biopsies from 28 ESCCs who received neoadjuvant chemoradiotherapy (CRT) and surgery and 10 normal esophageal epithelia using Affymetrix U133 Plus 2.0 arrays (Wen et al., 2014). After preoperative chemoradiotherapy among 28 ESCCs, complete remission of tumor occurred in 11 patients, who were divided into responder group, while no obvious tumor regression occurred in the other 17 patients, who were considered to be radioresistant and divided into

non-responder group. The average age of the patients in the non-responder group was (55.65 ± 5.53) years old, of whom 16 patients were male (accounting for 94.11%); four patients were in T2N1M0 stage, and 13 patients were in T3N1M0 stage; while average age of the patients in the responder group was (57.45 ± 6.80) years old, of whom nine patients were male (accounted for 81.82%); four patients were in T2N1M0 stage and seven patients were in T3N1M0 stage.

The GSE53625 microarray profiling datasets were provided by Li et al.(Shi et al., 2016; Li et al., 2017; Liu et al., 2020), including 179 patients with esophageal cancer. The profiling datasets were derived from cancer tissue and adjacent normal tissue. The dataset contained detailed clinical information of 179 ESCC patients that can be used for prognostic validation analysis. Both datasets were annotated with the platform files provided by GEO to obtain Ensemble ID, and then were annotated with the "org.Hs.eg.db" package to obtain gene symbols, thus ensuring that the two datasets have similar annotation conditions.

Screening, differential expression analysis and matched pairs of ir-lncRNAs

The immune gene list was download from the ImmuPORT database (https://www.immport.org/). Subsequently, the subsets of immune genes or lncRNAs were extracted from the GSE45670 dataset and GSE53625, respectively. The correlation test was performed on the immune gene expression matrix and the lncRNA expression matrix, and the related lncRNAs were screened as ir-lncRNAs. The screening criteria were: correlation coefficient $\rho \geq 0.4$ or $\rho \leq -0.4$ and $p \leq 0.0001$. Differential expression analysis (DEA) was performed on eligible ir-lncRNAs *via* the "limma" and "SVA" packages, and the expression difference was defined as: the absolute value of the log2 value (fold change) of the difference fold was greater than (mean ± 2 times the standard deviation of expression); p value ≤ 0.05. The differentially expressed lncRNAs were paired for each other. Taking the LINC01121|FAM167A-AS1 gene as an example, if the expression of LINC01121 gene was greater than that of FAM167A-AS1 gene, the matched index was recorded as 1, otherwise as 0. The ir-lncRNA pairs were introduced as independent variables into the COX prediction model to establish a prognostic risk-model.

Construction of the predictive model

The clinical survival data of ESCC samples were extracted from the GSE53625 dataset, and the ir-lncRNA pairs at the common intersection between the GSE45670 dataset and the GSE53625 dataset were considered as components of the survival data.

Univariate and multivariate Cox regression survival analyses were performed using the "survival" package in R 4.1.2 software to screen ir-lncRNAs with significant impact on prognosis. The least absolute shrinkage and selection operator (LASSO) regression analysis was carried out to narrow down the prognostically significant ir-lncRNA pairs. An ir-lncRNAs pairs-based risk prediction model was established with ir-lncRNAs of statistically significant differences in both univariate and multivariate Cox regression survival analysis.

The model building formula is (Huang et al., 2021):

$$\text{Riskscore} = \sum_{i=1}^{n} (lncRNApairsi \times coefi).$$

Where n is the counts of ir-lncRNA pairs, and $lncRNApairsi$ and $coefi$ represent the related matched results (values of 1 or 0) and coefficients of modeled ir-lncRNA pairs, respectively.

After the riskscores of all the included samples in GSE53625 dataset were calculated, the samples were divided into low-risk group and high-risk group according to the median value of riskscores. Furthermore, the Kaplan-Meier survival curve was applied to analyze the difference of survival prognosis between the high- and low-risk groups with Survival package. The ROC curves of 1-, 2-, and 3-year of overall survival were plotted with ROC package. Univariate and multivariate Cox regression survival analyses were performed with ESCC baseline data and riskscores *via* the "survival" package in R 4.1.2 software, respectively, to explore the independent prognostic factors of ESCC.

Differential analysis of riskscore for different clinical characteristics

The clinical characteristics such as gender, age, tumor TNM stage, etc. of ESCC samples were extracted from GSE53625 dataset and combined with the corresponding risk scores. Wilcoxon rank sum test was performed *via* the "ggpubr" package and the boxplots was drawn to show the correlations between risk scores and clinical characters.

Differential expression analysis of immune-related genes in different radiosensitivity groups

Based on those ir-lncRNAs screened from the GSE53625 dataset with prediction model, the immune-related genes in the GSE45670 dataset were reversely extracted, and the differential expression analysis of immune-related genes were compared between the non-responder group and the responder group. A heat map and a volcano plot were drawn.

Gene ontology, Kyoto encyclopedia of genes and genomes functional enrichment analysis

For the purpose of exploring the molecular mechanism of ir-lncRNAs and ir-genes related to radio-resistance in ESCC, we implemented gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) functional enrichment analysis *via* the "clusterProfiler" R package. In these analyses, a p value < 0.05 was considered statistically significant. GO enrichment analysis was composed of cellular components (CC), molecular functions (MF), and biological processes (BP), which showed the biological functions of genes at different levels, respectively. KEGG pathway enrichment analysis was used to evaluate the enrichment degree of genes in different pathways. The function of selected ir-genes whose expression were significantly different between the non-responder group and the responder group could indirectly speculate on the biological roles and mechanisms of immune-related lncRNAs.

lncRNAs isolation, cDNA synthesis, and RT-qPCR

Cellular total RNAs were isolated via TRIzol reagent (Thermo, United States) from ECA-109 cell lines and ECA-109R cell lines, where ECA-109R were identified as non-responder esophageal cancer and ECA-109 was identified as responder esophageal cancer. To increase the specificity of the real-time quantitative PCR (RT-qPCR), first-strand cDNA was synthesized from 1 mg total RNA *via* RevertAid First Strand cDNA Synthesis Kit (Invitrogen, China). The relative expression level of the target lncRNAs was detected by three-step RT-qPCR *via* the Fast SYBR Green Master Mix (Applied Biosystems Inc., CA, United States). The cycling conditions were 5 min of pre-degeneration at 95°C and 30 s of denaturation (polymerase activation) at 95°C followed by 40 cycles of annealing of primers at 95°C for 5 s and extension of primers at 60°C for 30 s. GAPDH was applied as an internal reference control. The relative expression level was calculated by the relative quantification $2^{-\Delta\Delta CT}$ method. The experiment was repeated three times and the primer sequences were listed in Supplementary Table S1.

Statistical analysis

All statistical analyses were performed *via* R software (version 4.1.2). The R packages involved the "Biobase" package, the "GEOquery" package, the "survival" package, the "ggpubr" package, and so on. The one-way ANOVA, Student's t-test (2-tailed) methods or Wilcoxon rank sum test were applied to assess the statistical significance. Quantitative data were shown as the mean ± SE. A p value < 0.05 was considered statistically significant.

Results

Screening results of ir-lncRNAs in GSE45670 and GSE53625 datasets

Flow chart of data collection and analysis is shown in Figure 1. According to the screening criteria, a total of 681 ir-lncRNAs were screened in the GSE45670 dataset. Further DEA showed that 12 ir-lncRNAs were up-regulated and 14 ir-lncRNAs were down-regulated in the non-responder group ($p < 0.05$). The results are indicated in Table 1 and Figure 2.

The 26 differentially expressed ir-lncRNAs genes constituted 325 ir-lncRNA pairs, of which 180 ir-lncRNA pairs were extracted according to the matching rate [pairRaito in the range of (0.2–0.8)], and the data are shown in Supplementary Table S2. Subsequently, we used the same method to construct the ir-lncRNA pairs in the GSE53625 dataset. Finally, 74 ir-lncRNA pairs successfully matched between the GSE45670 and the GSE53625 dataset after intersection. The detailed information of the 74 ir-lncRNA pairs is shown in Supplementary Table S3.

Clinical characteristics of esophageal squamous cell carcinoma samples in the GSE53625 dataset

The detailed clinical characteristics of GSE53625 ESCC are shown in Table 2.

Univariate and multivariate COX regression analysis of differential ir-lncRNAs

The results of univariate COX analysis showed that there were six ir-lncRNA pairs of statistical significance associated with survival outcomes ($p < 0.05$). In order to verify the positive variables obtained by univariate COX regression, LASSO regression (Least absolute shrinkage and selection operator) was further used to screen significant ir-lncRNAs as independent variables for multivariate COX analysis. As a result, five ir-lncRNA pairs, that is, LINC01121|FAM167A-AS1, ADAMTS9-AS2|MGC12916, MIR124-2HG|FAM167A-AS1, LINC00942|ADAMTS9-AS2, and PURPL|FAM167A-AS1, were the most powerful explanatory set of ir-lncRNAs as potential independent variables in multivariate COX model. A further multivariate COX regression was

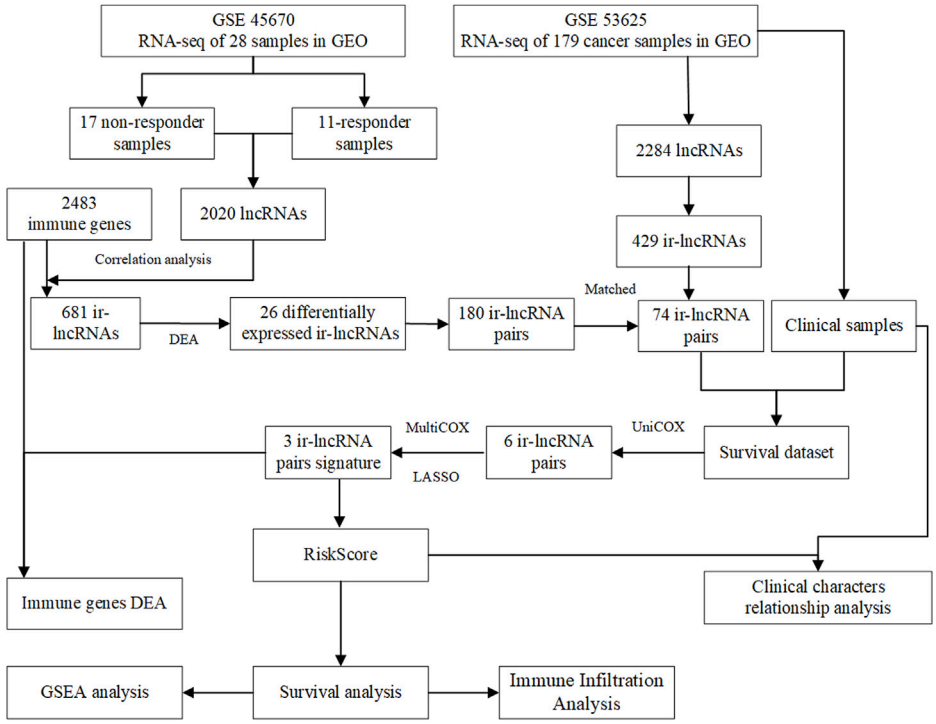


FIGURE 1
Flow chart of data collection and analysis.

performed, and the results showed that three ir-lncRNA pairs were statistically significant ($p < 0.05$), namely LINC01121|FAM167A-AS1, ADAMTS9-AS2|MGC12916, and MIR124-2HG|FAM167A-AS1. The above analysis revealed that the three ir-lncRNA pairs could serve as independent prognostic factors for ESCC. Since the hazard ratios of the three ir-lncRNAs were all greater then 1, they were considered to be risk factors for the prognosis of ESCC (as shown in Table 3 and Figure 3). Lasso regression road map and selection map were shown in Supplementary Figures S1, S2. ROC curves were used to assess the predictive ability and accuracy of the predictive models based on the three ir-lncRNA pairs. The results showed that the best cut-off point was 2.371 and the area under the curve (AUC) at 1-, 2-, and 3-year were 0.666, 0.702, and 0.686, respectively, as shown in Figures 4A,B.

Additionally, Supplementary Figure S3 provides differential distribution of riskscore with different clinical characters.

Construction of ir-lncRNAs-based risk model and survival analysis of high- and low-risk groups

The model formula obtained from multivariate COX regression analysis was:

$$\begin{aligned}\text{Riskscore} = {} & 0.687 \times \text{LINC01121|FAM167A} - \text{AS1} \\ & + 0.863 \times \text{ADAMTS9} - \text{AS2|MGC12916} \\ & + 0.419 \times \text{MIR124} - \text{2HG|FAM167A} - \text{AS1}\end{aligned}$$

The median riskscore was 1.989, and according to the median riskscore, 179 cases were divided into the high- (N = 76) and low-risk (N = 103) groups in the GSE53625 cohort. The Kaplan-Meier survival curves shows that patients in the high-risk group had a worse OS than patients in the low-risk group with more death ($p < 0.001$, Figure 5). The 1-, 3-, and 5-year survival rates in the low-risk group were higher than those in the high-risk group.

Differential expression analysis of ir-lncRNAs in high- and low-risk groups

Based on the aforementioned results, independent sample *t*-test were further applied to assess the expression of ir-lncRNAs in different riskscore groups. The distribution of risk scores is presented in Figure 6. Among the five significant lncRNAs, only FAM167A-AS1 had a decreased expression in the high-risk group, suggesting that FAM167A-AS1 is a protective factor.

TABLE 1 Differential expression results of immune-related lncRNAs in GSE45670.

Genes	log FC	Ave Expr	*t*	*p* value	adj *p* value	Change
LINC01121	−1.196	3.898	−3.679	0.001	0.318	Down
LINC00592	−1.306	7.669	−2.93	0.006	0.589	Down
CTD-3080P12.3	−1.076	3.852	−2.595	0.015	0.589	Down
WAKMAR2	−1.041	7.665	−2.569	0.016	0.589	Down
H19	−1.902	8.538	−2.536	0.017	0.589	Down
DLGAP4-AS1	−0.982	5.319	−2.497	0.018	0.589	Down
SCAT1	−1.459	5.612	−2.487	0.019	0.589	Down
LOC101928557	−1.139	4.483	−2.417	0.022	0.589	Down
PURPL	−1.619	3.829	−2.403	0.023	0.589	Down
IQCF5-AS1	−1.061	3.662	−2.302	0.029	0.641	Down
ELFN2	−1.126	6.26	−2.291	0.029	0.641	Down
MIR124-2HG	−0.991	4.119	−2.217	0.034	0.69	Down
LINC01102	−1.018	2.774	−2.088	0.046	0.825	Down
LOC101928389	−1.09	3.768	−2.048	0.05	0.825	Down
ADAMTS9-AS2	1.714	5.225	4.129	0	0.187	Up
SOX2-OT	1.731	5.417	2.935	0.006	0.589	Up
GRK3-AS1	1.601	3.371	2.864	0.008	0.589	Up
DELEC1	1.11	3.089	2.691	0.012	0.589	Up
FAM167A-AS1	1.391	3.184	2.536	0.017	0.589	Up
ZNF503-AS1	1.112	7.336	2.532	0.017	0.589	Up
RNF217-AS1	0.984	5.366	2.486	0.019	0.589	Up
MGC12916	1.079	4.97	2.446	0.021	0.589	Up
LOC101927798	1.036	5.289	2.409	0.022	0.589	Up
LINC00551	1.432	6.052	2.391	0.023	0.589	Up
LINC00942	1.545	6.705	2.109	0.044	0.824	Up
FSIP2-AS2	1.335	5.226	2.081	0.046	0.825	Up

Univariate and multivariate COX regression analysis of clinical characteristics

The clinical characteristics parameters in Table 2 combined with riskscore were selected to perform univariate and multivariate COX regression analysis. Univariate COX regression analysis demonstrated that age (HR = 1.681, 95% CI = 1.147–2.463, and $p = 0.008$), N-stage (Lymph node staging, N2 vs. N0, HR = 2.051, 95% CI = 1.137–3.702, and $p = 0.017$; N3 vs. N0, HR = 2.973, 95% CI = 1.426–6.200, and $p = 0.004$), TNM-stage (Stage III vs. Stage I, HR = 3.626, 95% CI = 1.138–11.548, and $p = 0.029$) and risk score (HR = 1.394, 95% CI = 1.225–1.587, and $p < 0.001$), which were all negative prognostic factors of OS in the GSE53625 cohort (Supplementary Table S4; Figure 7A). After adjusting for age, tumor grade, N-stage and TNM-stage, multivariate Cox analysis demonstrated that only risk score was a negative prognostic factor of OS (HR = 1.305, 95% CI = 1.139–1.495, and $p < 0.001$) (Supplementary Table S5; Figure 7B).

CIBERSORT immune infiltration analysis

To explore the difference of tumor immunity landscape between patients in the non-responder group and patients in the responder group, the CIBERSORT algorithm was utilized to evaluate immunity infiltration in the GSE 45670 dataset. The main results are shown in Figures 8A,B. According to the 22-classification method (Newman et al., 2019), the proportions of infiltrating activated mast cells were significantly higher in the responder group ($p < 0.05$). In addition, the proportions of infiltrating macrophages. M0 were also increased in the responder group, although the difference was not statistically significant ($p > 0.05$). The 4-classification method showed that the infiltration level of macrophages in the responder group was significantly higher than that in the non-responder group ($p < 0.05$) (Li B. et al., 2019).

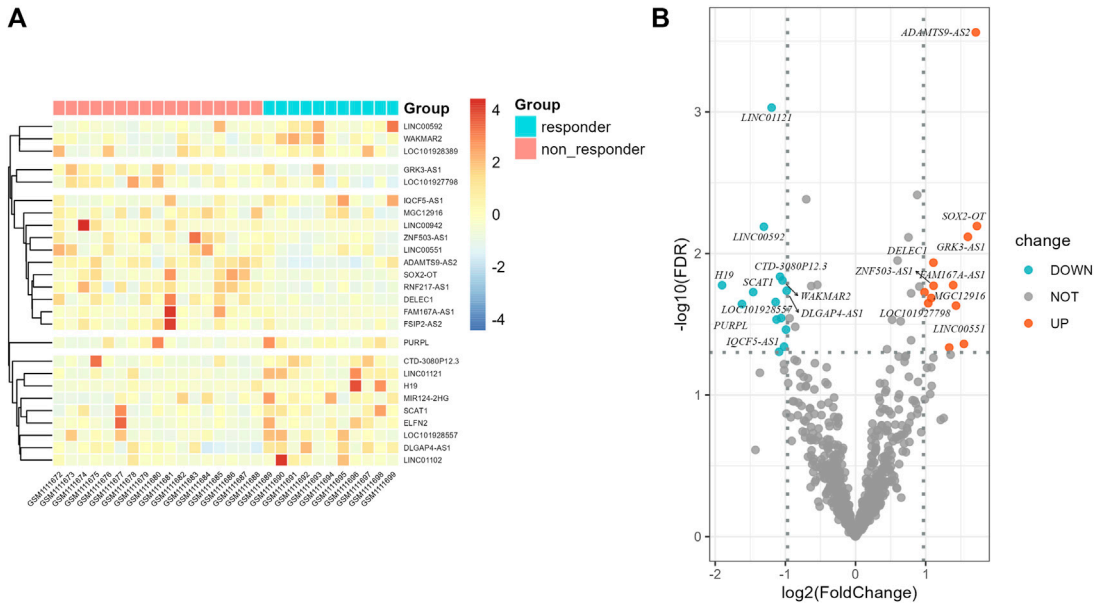


FIGURE 2
Differential expression results of immune-related lncRNAs in GSE45670 cohort. **(A)** Heatmap, **(B)** Volcano map.

TABLE 2 Clinical characteristics of esophageal squamous cell carcinoma samples in GSE53625.

Items	Overall (n = 179, %)	Items	Overall (n = 179, %)
Gender		Age	
Female	33 (18.44)	Mean (SD)	53.9 (9.03)
Male	146 (81.56)	Median (min, max)	59.7 (36.0,82.0)
Tobacco		Alcohol	
Yes	114 (63.69)	yes	106 (59.22)
T stage		N stage	
T1	12 (6.7)	N0	83 (46.37)
T2	27 (15.08)	N1	62 (34.64)
T3	110 (61.45)	N2	22 (12.29)
T4	30 (16.76)	N3	12 (6.7)
TNM stage		Tumor grade	
Stage I	10 (5.59)	Poorly	49 (27.37)
Stage II	77 (43.02)	Moderately	98 (54.75)
Stage III	92 (51.4)	Well	32 (17.88)

Differential expression analysis of immune genes

The correlation coefficient $\rho = \pm 0.4$ and $p < 0.05$ were used to screen the immune genes related to the above five ir-lncRNAs, and a total of 137 immune genes were screened, of which 11 genes had significant differences in expression levels between the non-responder group and the responder group ($p < 0.05$). Six genes (IL12RB2, IL32, MMP9, NGF, OASL, and TNFRSF12A) were down-regulated and 5 genes (BMP4, CHP2, OSGIN1, PAK5, and POMC) were up-regulated. The distribution of immune genes expression was presented in Figures 9 and 10. Further, we analyzed the expression differences of 5 ir-lncRNAs from GSE45670 cohort. The distribution of ir-lncRNAs expression was presented in Figure 11. Among the above five ir-lncRNAs, ADAMTS9-AS2,

TABLE 3 Univariate and multivariate COX regression analysis of differential ir-lncRNAs.

Ir-lncRNA pairs	Model	β	se	HR	HR.95L	HR.95H	p
LINC01121|FAM167A-AS1	Univariate	0.789	0.237	2.201	1.385	3.5	0.001
ADAMTS9-AS2|MGC12916	Univariate	0.664	0.288	1.943	1.105	3.418	0.021
MIR124-2HG|FAM167A-AS1	Univariate	0.567	0.196	1.764	1.202	2.588	0.004
LINC00942|ADAMTS9-AS2	Univariate	−0.84	0.424	0.432	0.188	0.991	0.048
PURPL|FAM167A-AS1	Univariate	0.598	0.249	1.819	1.117	2.961	0.016
ZNF503-AS1|H19	Univariate	2.662	1.041	14.327	1.863	110.187	0.011
LINC01121|FAM167A-AS1	Multivariate	0.687	0.252	1.989	1.213	3.259	0.006
ADAMTS9-AS2|MGC12916	Multivariate	0.863	0.293	2.371	1.335	4.211	0.003
MIR124-2HG|FAM167A-AS1	Multivariate	0.419	0.209	1.52	1.009	2.291	0.045

Note: HR.95L: lower limit of 95% confidence interval; HR.95H: upper limit of 95% confidence interval.

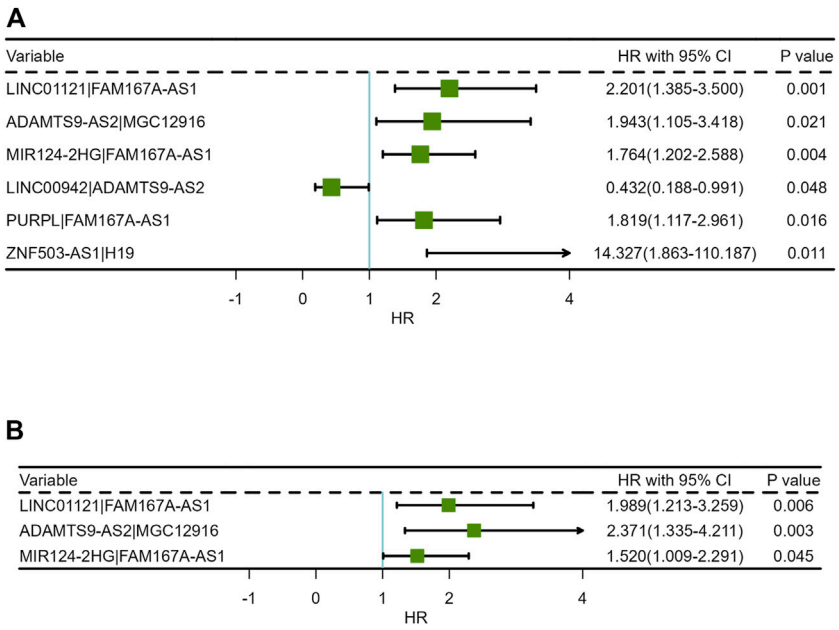


FIGURE 3
Forest plots of Cox regression analysis for ir-lncRNA pairs: **(A)** Univariate Cox regression analysis, **(B)** Multivariate Cox regression analysis.

and FAM167A-AS1 were up-regulated in the non-responder group, while LINC01121 and MIR124-2HG were down-regulated ($p < 0.05$).

Functional gene ontology and Kyoto encyclopedia of genes and genomes enrichment analysis of immune genes

The GO and KEGG enrichment analysis results are shown in Table 4 and Figure 12. In the GO molecular function enrichment analysis, the differentially expressed immune genes were mainly enriched in receptor ligand activity, signaling receptor activator activity, growth factor activity, and other pathways. In biological process enrichment analysis, differential immune genes were mainly enriched in extrinsic apoptotic signaling pathway, regulation of apoptotic signaling pathway, regulation of extrinsic apoptotic signaling pathway, etc. In the cellular component enrichment analysis, the differential immune genes were mainly enriched in endosome lumen, tertiary granule lumen, and Golgi lumen, but the corrected p value was not statistically significant, suggesting that these pathways were not significant. In the KEGG enrichment analysis, the differential immune genes were mainly enriched in

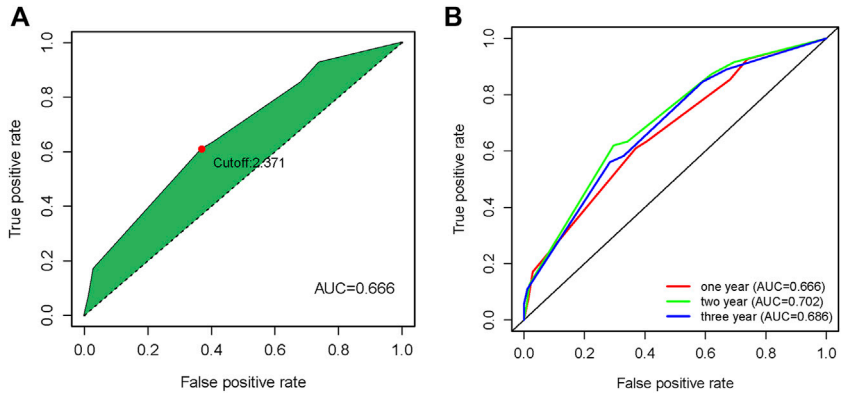


FIGURE 4
Time ROC curves on overall survival prediction in the GSE53625 cohort **(A)** Time ROC curves with cutoff value, **(B)** Time ROC curves at one year, two year and three year.

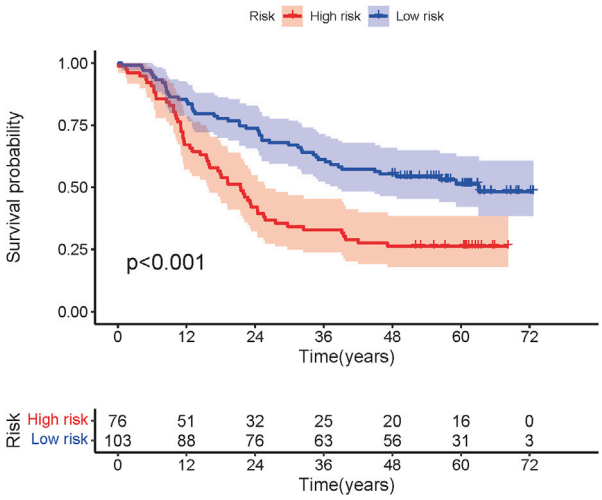


FIGURE 5
Kaplan-Meier curves for the overall survival of patients in the high- and low-risk group.

Cytokine-cytokine receptor interaction, Estrogen signaling pathway, Fluid shear stress and atherosclerosis. The immune genes mainly involved in enrichment analysis included NGF, IL32, BMP4, TNFRSF12A, and IL12RB2.

lncRNAs isolation, cDNA synthesis, and RT-qPCR

The expression levels of ADAMTS9-AS2, FAM167A-AS1, LINC01121, MIR124-2HG, MGC12916 and the internal reference RNA were detected by RT-qPCR using three of the above-mentioned cell lines. Compared with those lncRNAs in the non-responder group, expression levels of ADAMTS9- AS2, FAM167A-AS1, and MGC12916 lncRNA in the responder group were significantly up-regulated, while expression levels of LINC01121 and MIR124-2HG were down-regulated. The results are shown in Supplementary Figure S4 and Supplementary Table S6.

Discussion

Many recent studies have focused on establishing the signatures of coding genes with or without non-coding RNAs to assess prognosis in patients with malignancies (Zhu et al., 2016; Qu et al., 2018; Hong et al., 2020). This research strategy divides prognostic groupings based on the absolute expression level of certain genes of interest, of which the simplicity of operation is an important advantage, as the prognostic model was set up on quantifying the expression levels of transcripts (Hong et al., 2020). However, the shortcomings are also obvious, because the expression of most genes varies greatly, especially for non-coding RNAs. This means that some basic experiments (such as rt-qPCR) should be used for practical verification to ensure that the genes of interest had sufficient expression (Zhu et al., 2021). To overcome the shortcomings of the above model, new prognostic models based on the strategy of the relative expression of immune-related gene pairing were proposed, which became a mainstream bioinformatics research method (Sun et al., 2020; Wu et al., 2020; Wang Y. et al., 2021). In this study, we were inspired by the new strategy and attempted to construct a reasonable model with ir-lncRNA pairs assess the prognostic value of ir-lncRNAs. Therefore, we did not apply rt-qPCR to

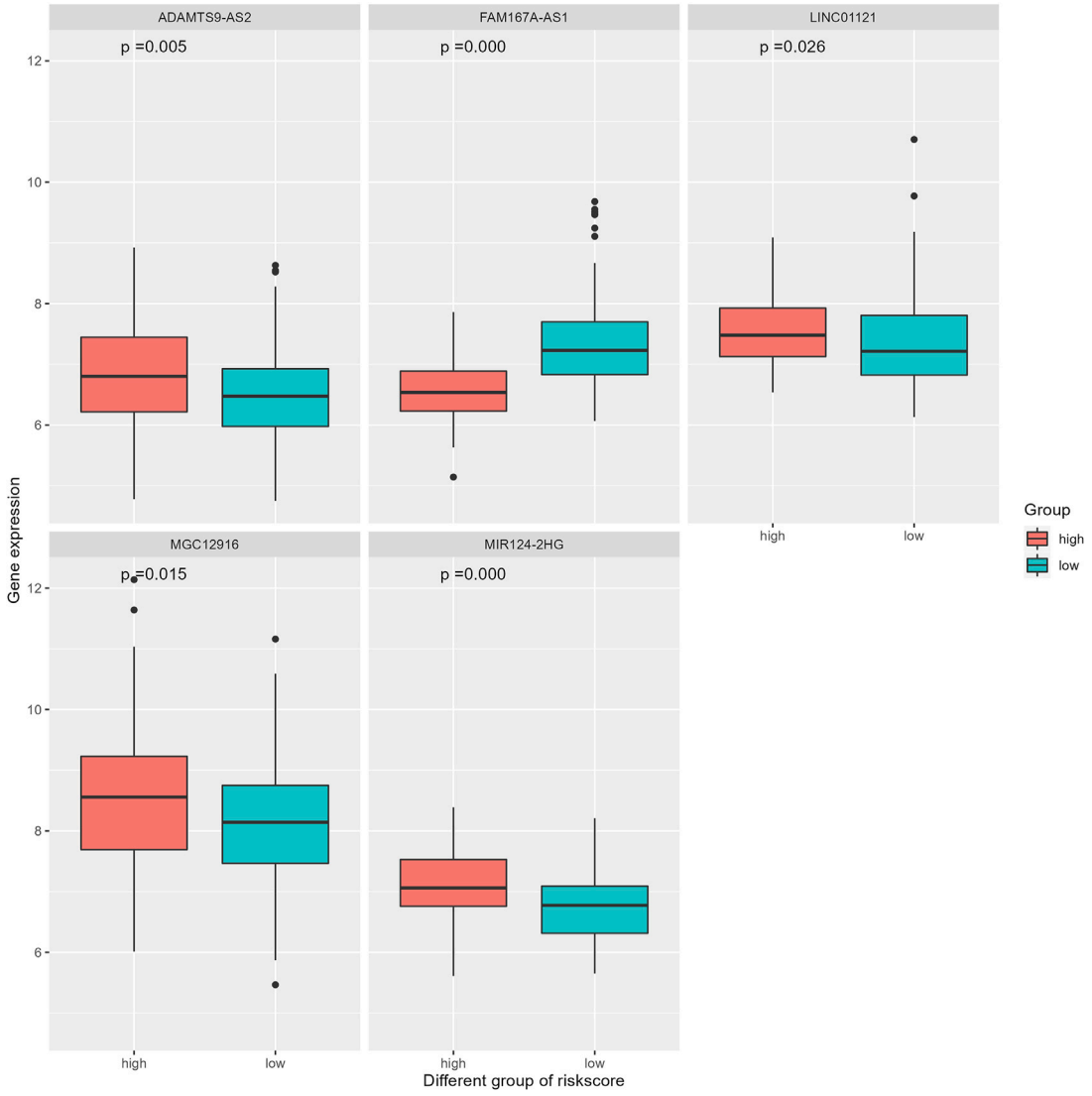


FIGURE 6
Different expression of five ir-lncRNAs between the low- and the high-risk group.

verify their exact expression levels in the signatures in our study. To the best of our knowledge, there are no similar studies investigating the radioresistance of esophageal cancer.

First, we explored the expression level of ir-lncRNAs between non-responder ESCC patients and responder ESCC patients to investigate the potential mechanism of radioresistance in ESCC after chemoradiotherapy. We first screened 26 differentially expressed ir-lncRNAs from the GEO45670 dataset. The differential expression of these lncRNAs indicates that ir-lncRNAs had played an important role in the radioresistance of ESCC. Second, previous studies have shown that lncRNAs achieve biological functions through a variety of target genes, which may involve immune-related genes (Wang Y. et al., 2021). Therefore, we further explored the differential expression of immune genes, which were not only differentially expressed in esophageal cancer tissues with different response characteristics of radiotherapy, but also correlated with the expression of immune lncRNAs, and thus could be considered as target genes of ir-lncRNAs. Third, differential co-expression analyses were performed to classify ir-lncRNAs and immune genes, and the prognostic performance of ir-lncRNAs were validated *via* single-pairing approach along with a 0 or 1 matrix by cyclical calculation. Fourth, we performed univariate analysis combined with Lasso penalized regression to identify the significant irlncRNA pairs. Next, multivariate regression was used to identify ir-lncRNAs with independent effects. Fifth, the best

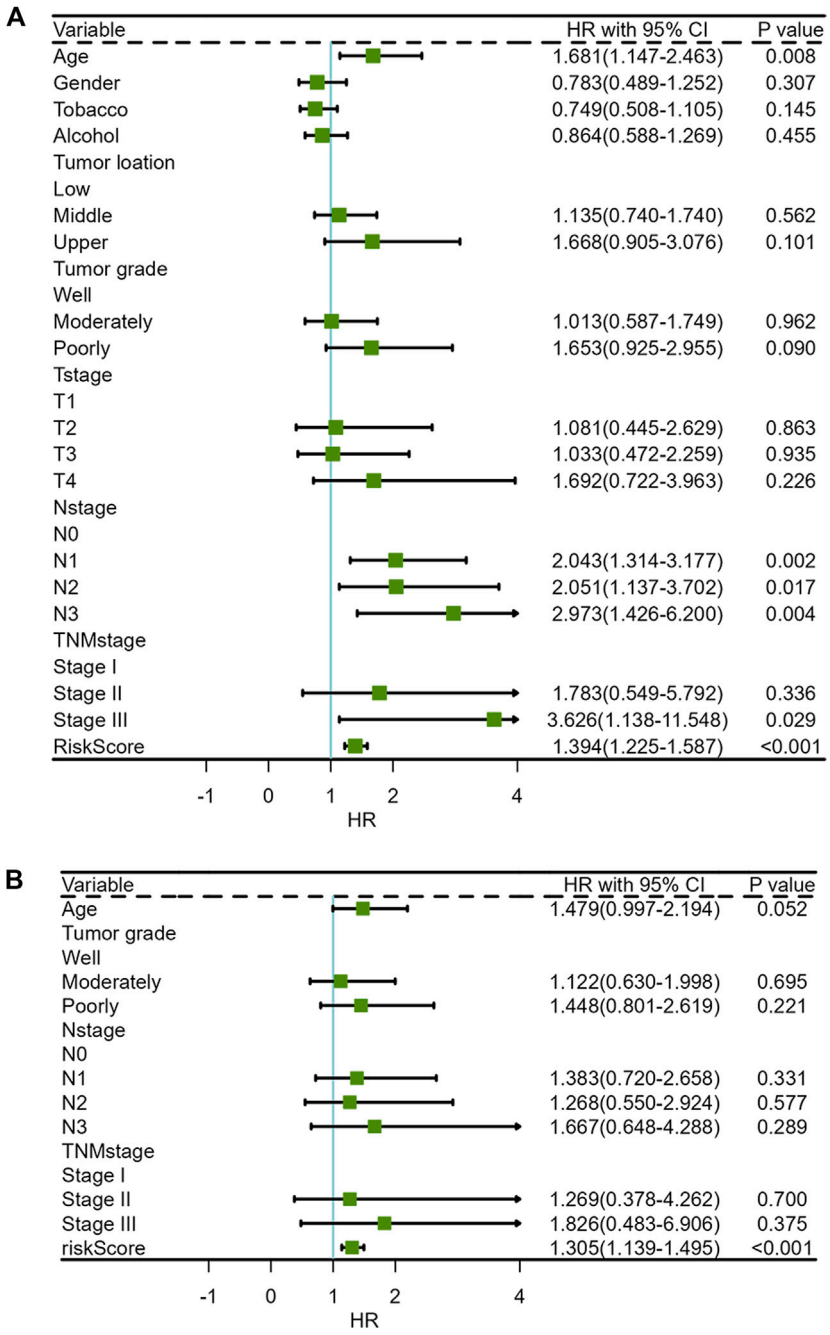


FIGURE 7
Forest plots of clinical characteristics and risk score for univariate and multivariate COX regression analysis **(A)** Univariate Cox regression analysis, **(B)** Multivariate Cox regression analysis.

model was set up by calculating AUC value with ROC, where the best cut-off point was applied to distinguish high- or low-risk groups of EPC patients. Sixth, we further evaluated this new model in a variety of clinical settings, including survival, clinical or pathological characters. At last, tumor-infiltrating immune cells between different radioresistance of esophageal cancer.

Radiotherapy is one of the main and effective treatments for advanced EPC, but the therapeutic outcomes of which are still unsatisfactory (Li et al., 2016; Yan et al., 2022), because the complete response rate of radiotherapy is less than 35%–40% (Mori et al., 2021). Radioresistance is the biggest problem and obstacle faced by radiotherapy in esophageal cancer, as poor

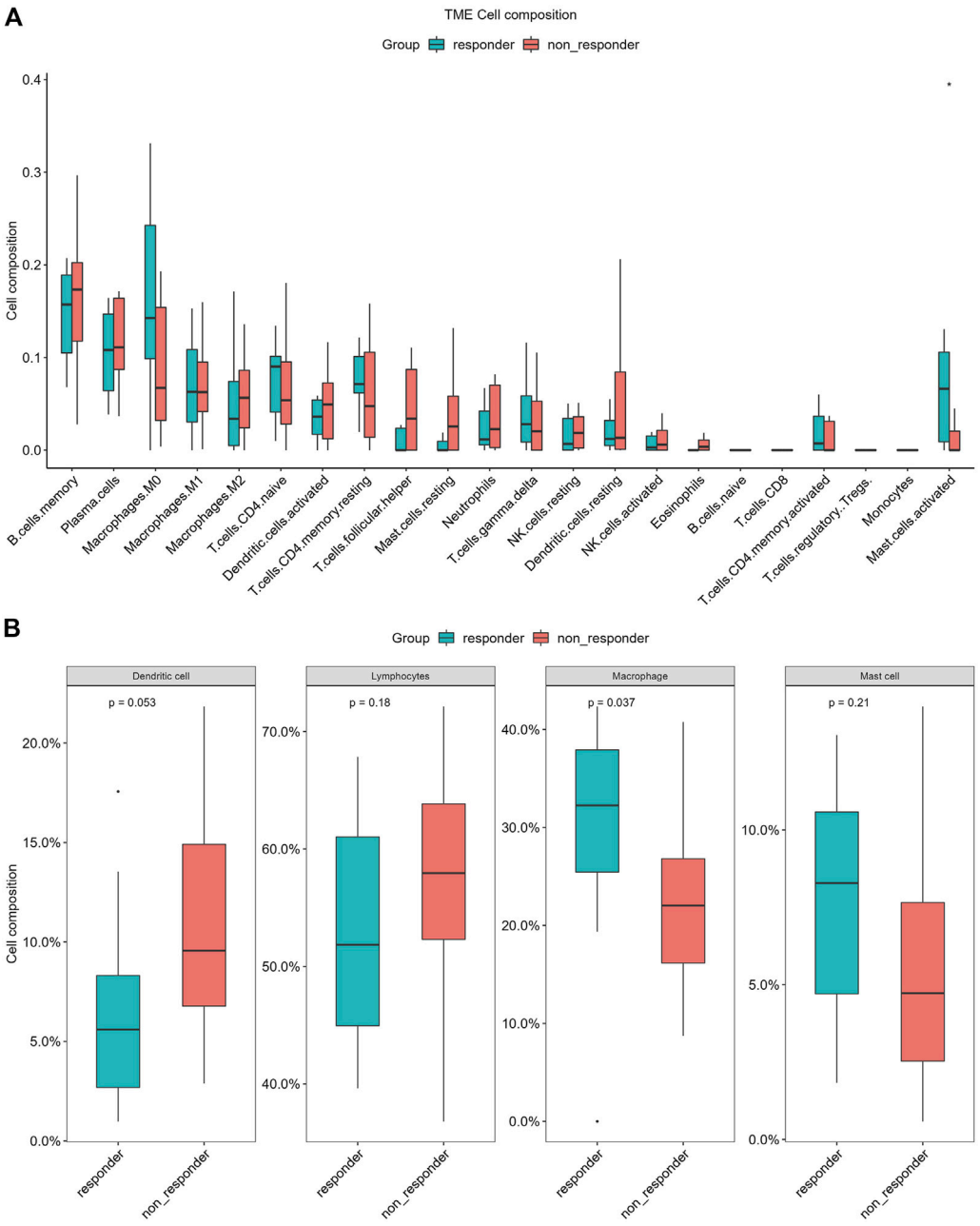


FIGURE 8
Differences of infiltrating immune cell types between the non-responder group and the responder group of CIBERSORT in GSE45670 cohort. **(A)** 22-classification method, **(B)** 4-classification method.

response to radiotherapy could result in local failure after radiotherapy (Chen H. et al., 2020). Therefore, it is particularly important to search for radioresistance-related molecular markers and new radiosensitizers to enhance radiosensitivity in ESCC cells and improve the survival of ESCC patients with radioresistance. Radioresistance is the focus and difficulty in cancer radiobiology research, which is also a clinically urgent issue (Wang et al., 2019; Yu et al., 2020; Sun et al., 2021; Liu et al., 2022). Genes are an intrinsic determinant of tumor radioresistance, and studies have shown that multiple genes can affect the radioresistance of esophageal cancer (Huang et al., 2019; Wang et al., 2019; Hua et al., 2020; Yu et al., 2020; Han et al., 2021; Sun et al., 2021; Liu et al., 2022). Practice has shown that it is common in clinical practice that even patients

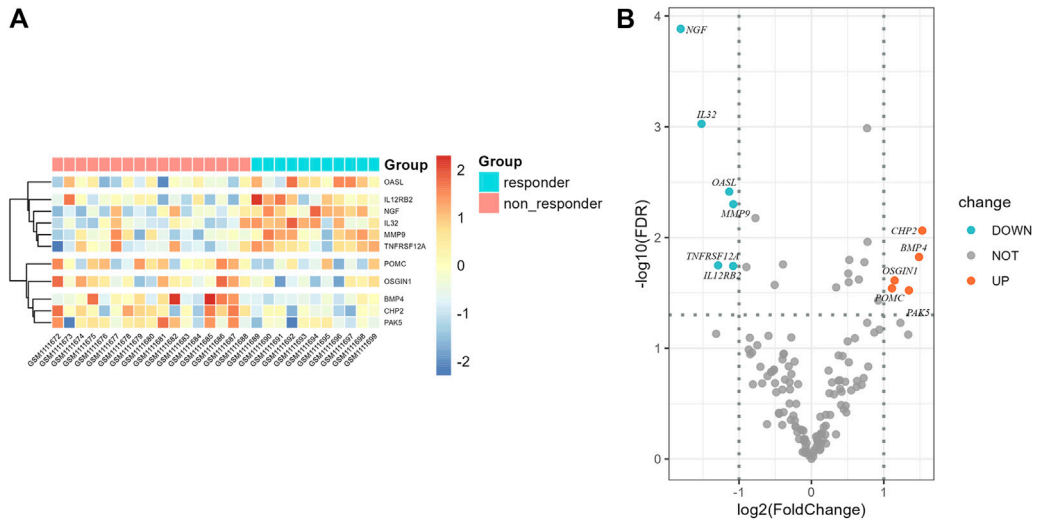


FIGURE 9
Differential expression results of immune genes in GSE45670 cohort. **(A)** Heatmap, **(B)** Volcano map.

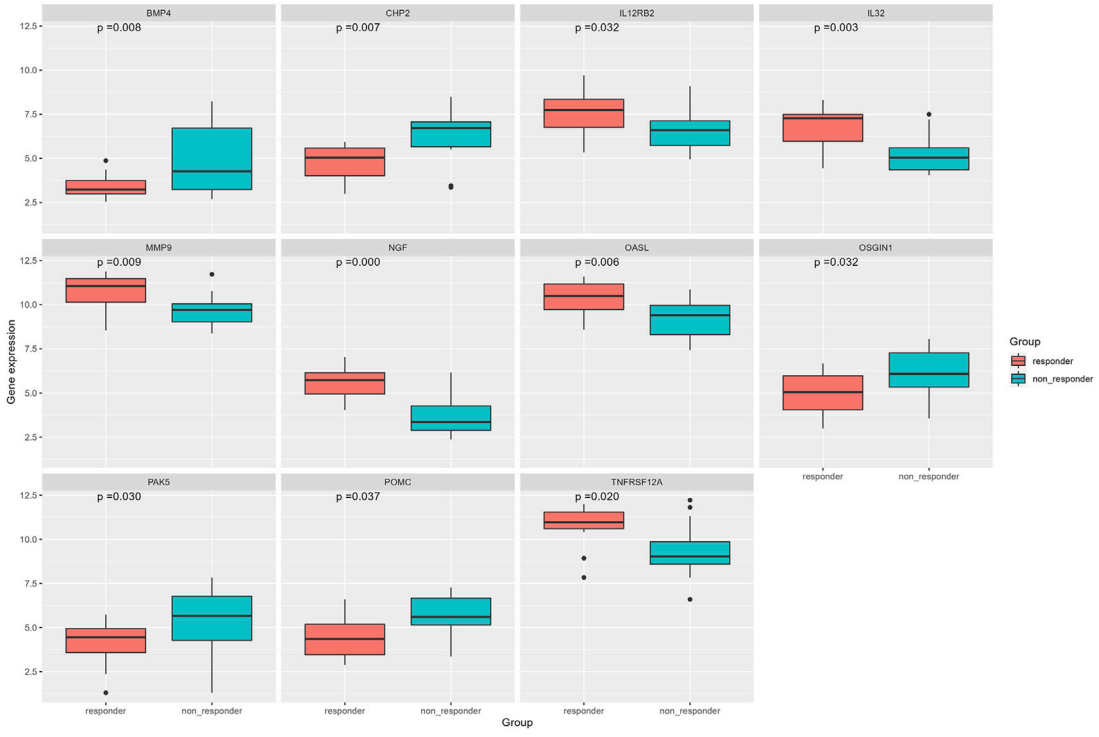


FIGURE 10
Differential expression analysis of immune genes in GSE45670 cohort.

with esophageal cancer of the same stage and the same pathological type have great differences in the effect of radiotherapy after receiving the same radiotherapy regimen. The fundamental reason is the genetic differences between different individuals (Huang et al., 2019; Wang et al., 2019; Hua et al., 2020; Yu et al., 2020; Han et al., 2021; Sun et al.,

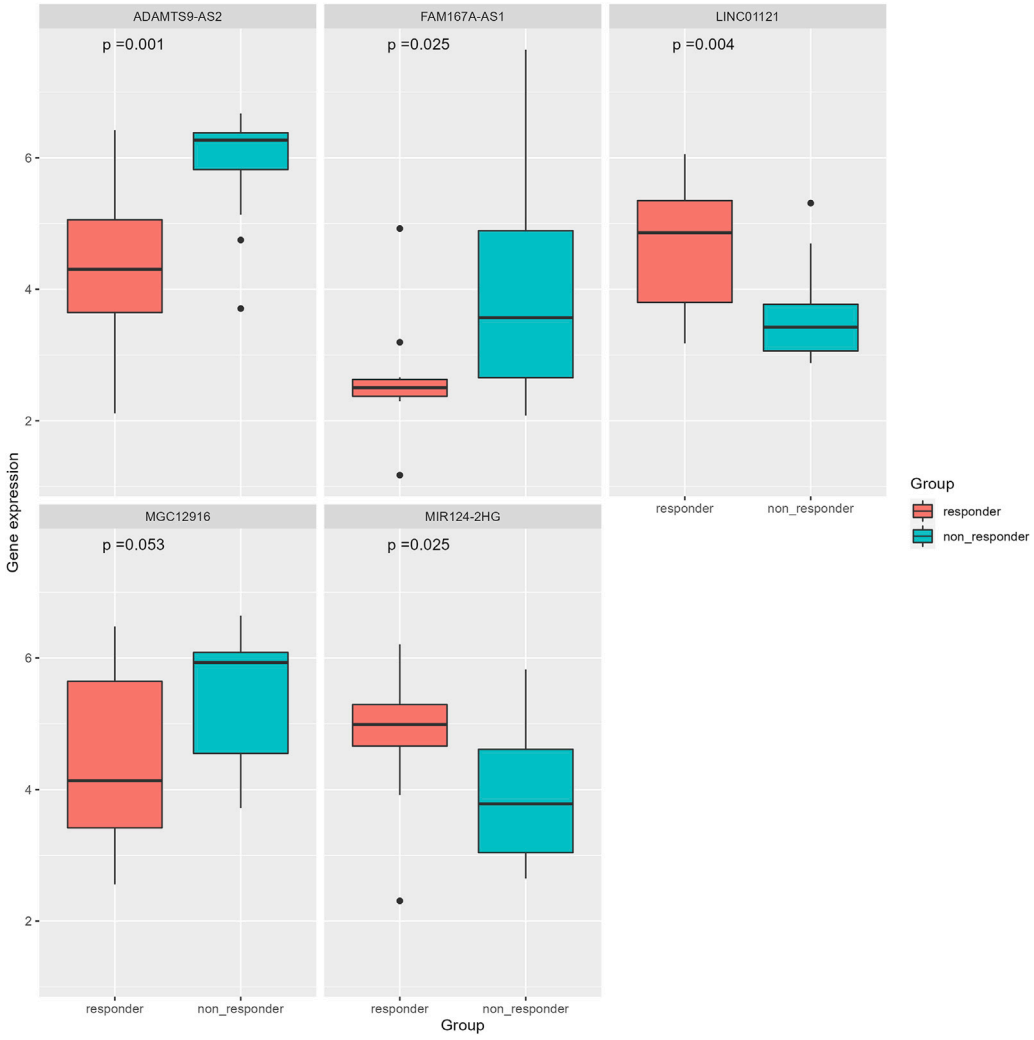


FIGURE 11
Differential expression analysis of five ir-lncRNAs in GSE45670 cohort.

2021; Liu et al., 2022). If the decisive genes of radioresistance can be screened out, it is of great significance for the study of radiosensitization, targeted therapy, and prediction of radiotherapy effect to guide individualized therapy. Although more molecular studies have been reported on the radioresistance or radiosensitivity of ESCC, the clinical significance of most molecular markers remains unclear and inconsistent due to the complexity and variability of detection methods.

Recent studies had found that a variety of lncRNAs can affect the radiosensitivity of EPC by regulating gene expression and key signal transduction pathways (Liu et al., 2021). Many lncRNAs are involved in the regulation of the tumor immune microenvironment, which are often referred to as immune-related lncRNAs (Huang et al., 2018). However, there is few reports on the relationship between ir-lncRNAs and radioresistance and prognosis of ESCC. Furthermore, the mechanism of ir-lncRNAs involved in the radioresistance is still unclear. In recent years, the use of bioinformatics methods for data mining at the molecular level provides new ideas for the study of molecular pathogenesis of various diseases including tumors (Zheng et al., 2020). Our present study used bioinformatics methods to re-analyze the radioresistance-related microarray data of esophageal cancer from GEO, and screened for differentially expressed ir-lncRNA genes. Through biological process annotation and signal pathway enrichment analysis, we have excavated some ir-lncRNA genes and immune-genes and potential signal pathways related to radioresistance to investigate the mechanism of radioresistance at the molecular level in ESCC. Zhu et al. (2021) had identified that a total of 111 immune-related

TABLE 4 Main results of GO and KEGG enrichment analysis for significantly expressed immune genes.

Type	Description	*p*	*p* adjust	Q value	gene ID	Count
GO_BP	extrinsic apoptotic signaling pathway	0.0000	0.004322	0.0021	NGF/BMP4/TNFRSF12A/PAK5	4
GO_BP	regulation of apoptotic signaling pathway	0.0000	0.014634	0.0071	MMP9/BMP4/TNFRSF12A/PAK5	4
GO_BP	regulation of extrinsic apoptotic signaling pathway	0.0001	0.020688	0.0100	BMP4/TNFRSF12A/PAK5	3
GO_MF	receptor ligand activity	0.0000	0.000115	0.0001	NGF/IL32/BMP4/OSGIN1/POMC	5
GO_MF	signaling receptor activator activity	0.0000	0.000115	0.0001	NGF/IL32/BMP4/OSGIN1/POMC	5
GO_MF	growth factor activity	0.0001	0.001442	0.0009	NGF/BMP4/OSGIN1	3
GO_CC	endosome lumen	0.0195	0.211736	0.1932	NGF	1
GO_CC	tertiary granule lumen	0.0305	0.211736	0.1932	MMP9	1
GO_CC	Golgi lumen	0.0570	0.211736	0.1932	NGF	1
KEGG	Cytokine-cytokine receptor interaction	0.0000	0.000317	0.0003	NGF/IL32/BMP4/TNFRSF12A/IL12RB2	5
KEGG	Estrogen signaling pathway	0.0095	0.150833	0.1419	MMP9/POMC	2
KEGG	Fluid shear stress and atherosclerosis	0.0096	0.150833	0.1419	MMP9/BMP4	2

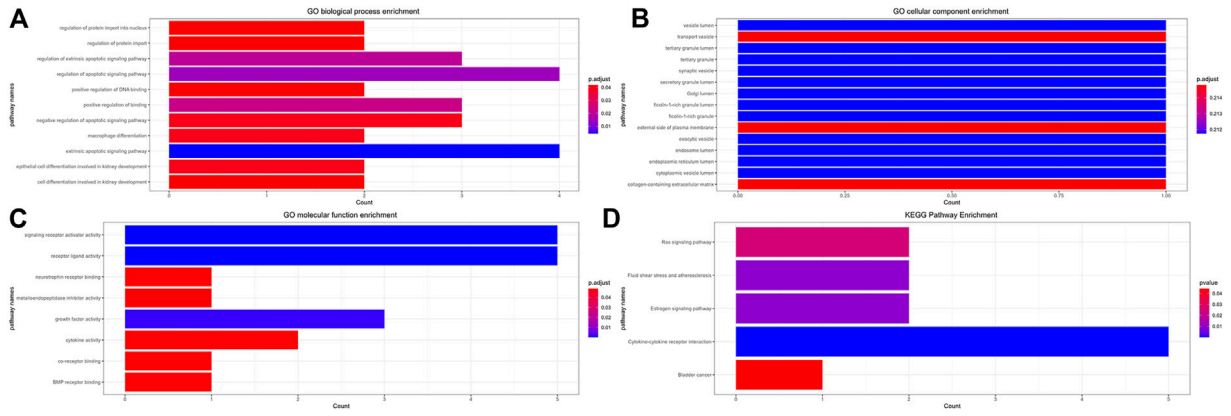


FIGURE 12
Functional GO and KEGG Enrichment Analysis of Immune Genes. **(A)** GO biological process enrichment, **(B)** GO cellular component enrichment, **(C)** GO molecular function enrichment, **(D)** KEGG pathway enrichment.

lncRNAs were different expressed in ESCC, in which,14 lncRNAs markedly related to prognosis of ESCC were identified *via* univariate analysis. Finally, a model based on the 8-lncRNA signature was identified in the multiple regression model, which demonstrated that the 8-lncRNA signature has certain power in predicting the prognosis of ESCC patients (Zhu et al., 2021). Since the expression abundance of ir-lncRNAs is extremely low, which is different from the expression of protein genes, we used the relative expression status between different ir-lncRNAs to explore the biological functions of lncRNAs. Based on 26 differentially expressed ir-lncRNAs, we constructed 325 ir-lncRNA pairs in the GSE45670 dataset, and 180 of which were extracted to further study. As seen in our study, we found that there were differences in gene expression between non-responder ESCC and responder ESCC in terms of LINC01121, FAM167A-AS1, ADAMTS9-AS2, MGC12916, MIR124-2HG. ADAMTS9-AS2 and FAM167A-AS1 were up-regulated in the non-responder group, while LINC01121 and MIR124-2HG were down-regulated ($p < 0.05$). Among the five lncRNAs we identified; some studies have demonstrated their prognostic effects on malignant tumors. Shen et al. (2020) had found that a worse 5-year overall survival was detected in ESCC-patients with low-expressed ADAMTS9-AS2. Similar results were seen in clear cell renal cell carcinoma and bladder cancer (Song et al., 2019; Zhang et al., 2020), which indicates that ADAMTS9-AS2 is a tumor suppressor gene. However, the opposite results were found in patients with tongue squamous cell carcinoma, where high-expression of LncRNA ADAMTS9-AS2 promotes proliferation, migration and epithelial-mesenchymal transition (EMT) with poor prognosis, and low-expression was detected in patient with lymph node metastasis (Li Y. et al., 2019). In our study, the expression of ADAMTS9-AS2 was increased in non-

responsive EPC patients, suggesting that ADAMTS9-AS2 was involved in the radioresistance of EPC, but the specific mechanism has not been reported. Some scholars have found that ADAMTS9-AS2 have been shown to play essential roles in temozolomide (TMZ) resistance in glioblastoma (GBM) (Yan et al., 2019). It should be mentioned that, although multiple lncRNAs, such as HOTAIR (Lv et al., 2013), CCAT2 (Zhang et al., 2015) and MALAT1 (Deng et al., 2016), have shown potential prognostic value in ESCC, the role of immune-related lncRNA signatures in prognosis has not been elucidated in the literature.

A few studies had addressed the role of immune-related lncRNAs in survival and prognosis of ESCC. A previous study had established an immune gene-based prognostic model for ESCC and esophageal adenocarcinoma (EAC) (Fei et al., 2021). Prognosis-related immune-gene-based model based on BMP1, EGFR, S100A12, HLA-B, TNFSF18, IL1B, and MAPT had proved to be useful for prognosis in ESCC (Fei et al., 2021). To verify the relationship between ir-lncRNA pairs and survival from ESCC, we adopted the GES53625 dataset as validation cohort. We constructed a novel prognostic prediction model consisting of three ir-lncRNA pairs, including LINC01121|FAM167A-AS1, ADAMTS9-AS2|MGC12916, and MIR124-2HG|FAM167A-AS1, which involved five ir-lncRNAs. We further confirmed by Wilcoxon rank sum test that there were significant differences in the expression levels of the above ir-lncRNAs in different risk score groups in the GES53625 cohort. In addition, the prognostic value of the novel model was confirmed by the multivariate Cox analysis. The overall survival in the high-risk group was significantly worse than that in the low-risk group ($p < 0.001$). The 1-year, 2-year, and 3-year prediction performance of this risk-model was 0.666, 0.702, and 0.686, respectively. Univariate and multivariate Cox analysis confirmed that prognostic riskscores based on three ir-lncRNA pairs were important prognostic factors for ESCC. Univariate and multivariate COX regression analysis indicated that the three ir-lncRNA pairs were associated with the prognosis of ESCC, and in the high-risk group ADAMTS9-AS2, LINC01121, MGC12916, and MIR124-2HG was up-regulated, FAM167A-AS1 was up-regulated. All the three ir-lncRNA pairs were markers of poor prognosis with worse overall survival for ESCC. A previous research showed that ADAMTS9-AS2 is a prognostic biomarker correlated with immune infiltrates and predicted a poorer overall survival when it was low expressed in lung adenocarcinoma (Lin et al., 2021). Li W. et al. (2020) had identified the expression levels of eight lncRNAs to establish a signature for predicting the survival of patients with ESCC. Li Z. Y. et al. (2020) found that the expression of lncRNA Rpph1 in patients with EPC was significantly higher than that in healthy participants ($p < 0.05$), and was positively correlated with cancer tissues ($r = 0.681$, $p < 0.05$). *In vitro* experiments confirmed that silencing lncRNARpph1 could up-regulate radio-induced pro-apoptotic-related proteins such as Bax, down-regulate anti-apoptotic-related proteins such as Bcl-2, thereby increasing radiotherapy-induced apoptosis of EPC cells. In addition, silencing lncRNA Rpph1 can also improve the radiosensitivity of EPC cells by reducing radiation-induced G2/M phase arrest and epithelial-mesenchymal transition (EMT) (Li Z. Y. et al., 2020). Wang et al. found that lncRNA CCAT2 was highly expressed in EPC cells, which can negatively regulate the expression of miR-145 and inhibit the phosphorylation of Akt, ERK, and p70 s6K1 to increase radioresistance. *In vitro* experiments confirmed that knockout of lncRNA CCAT2 can significantly increase radiation-induced apoptosis, thereby increasing radiosensitivity in EPC (Wang et al., 2020). In addition to mediating radioresistance, some lncRNAs also have radiosensitizing effects. Lin et al. reported that compared with the radioresistant ESCC cell line TE-1-R, the expression of lncRNA GAS5, and RECK was higher in the radiosensitive cell line TE-1, while the expression of miR-21 was lower in cell line TE-1 (Lin et al., 2020). Further research found that up-regulation of lncRNA GAS5 can increase the expression of RECK by inhibiting miR-21, reduce the viability and colony formation ability of EPC cells under radiation exposure, and increase radiation-induced apoptosis of cancer cells (Lin et al., 2020). All these evidences had suggested that lncRNAs play an important role in radioresistance or radiosensitivity of ESCC. In general, high-abundance lncRNAs have significant biological functions, especially lncRNAs with significant differences in expression (Yan et al., 2021). Different from the above studies, we focused on the ir-lncRNAs related to the immune environment or immune regulation of radioresistance. As expected, we found that ir-lncRNAs are involved in the bidirectional regulation of radiosensitivity, that is, some lncRNAs could promote radioresistance, while others may increase radiosensitivity.

In recent years, the study of the tumor immune microenvironment has taken a leading role in field of cancer research (Wang et al., 2022). The differential expression of immune genes is an important molecular mechanism leading to changes in the tumor immune microenvironment (Anderson, 2020; Peña-Romero and Orenes-Piñero, 2022). Several previous studies reported the prognostic value of a single immune-related gene in EPC or lung cancer, such as FGFR1, TNFRSF10B, and IL1B (Schabath et al., 2013; Takase et al., 2016; Zhang Y. et al., 2021). However, the study focused on the prognostic role of the ir-lncRNAs with microenvironment and immune cells in ESCC is lacking, especially for ESCC with radioresistance. Identification of prognostic value of the tumor microenvironment in esophageal cancer is necessary (Tan et al., 2021). In our study, we identified 11 immune genes with differential expression in the GSE 45670 cohort, and our findings further confirmed the involvement of immune alterations in the biological process of radiation resistance in ESCC. To further explore the changes and biological

pathways in the immune microenvironment between patients in the non-responder group and patients in the responder group, CIBERSORT was carried out in this study as well, which was an analytical tool developed by Newman et al. (2015) to provide an estimation of the abundances of member cell types in a mixed cell population *via* gene expression data, and provided the possibility of identifying immune biomarkers for diagnosis and prognosis. The GSE 45670 dataset was utilized to evaluate immunity infiltration *via* the CIBERSORT algorithm. Growing evidence pointed to the underlying mechanisms by which the local immune microenvironment and immune cells drive tumorigenesis in many cancers (Greten and Grivennikov, 2019; Wang Q. et al., 2021; Kumar et al., 2022). Previous studies have shown that DNA damage repair-related genes, apoptosis-related genes, cellular hypoxia-related genes, cell cycle-related genes, and autophagy genes play important roles in radiosensitivity by changing the microenvironment (Chen et al., 2017; Tang et al., 2018; Zhang H. et al., 2021). Tumor-infiltrating immune cells (TIICs) in esophageal cancer tissue may be an important determinant of prognosis and therapy response (Lu et al., 2020). In our study, we found that there were some differences in infiltrating immune cells between the non-responder group and the responder group, which mainly manifested in differences in infiltration of activated mast cell and macrophage. Mast cells are an important member of innate immune cells. Circulating mast cells contribute to the growth and metastasis of many tumors, while mast cell infiltration in tumor tissue is closely related to tumor survival in some cancer (Welsh et al., 2005). In a study of lung cancer, it has been shown that more human mast cells infiltrated in cancer tissue improved the survival of cancer, suggesting that mast cells can participate in the antitumor immune process (Welsh et al., 2005). In ESCC cases with tumor complete remission, the number of mast cells infiltrated was higher, indicating that mast cells were involved in the therapeutic effect of radiotherapy.

There are a large number of tumor-associated macrophages (TAMs) in the tumor microenvironment, which have a high degree of interaction with tumor cells, tumor stem cells, epidermal cells, fibroblasts, T/B cells, and NK cells (Welsh et al., 2005). Although macrophages theoretically have the ability to destroy tumors, there is growing experimental evidence that TAMs promote tumor progression (Mantovani and Allavena, 2015). Resilience and diversity are two characteristics of macrophages, which means that macrophages have dual effects in tumor development (Cassetta and Pollard, 2020). According to the activation type of macrophages and their different roles in the tumor microenvironment, TAMs are generally classified into two functionally opposite subtypes, classically activated M1 macrophages and alternately activated M2 macrophages, both of which represent one of the main tumor-infiltrating immune cell types (Cassetta and Pollard, 2020). Basic research has found that a large number of TAMs proliferate after radiation, and at the same time release a large number of inflammatory signals (IL-1) and immunosuppressive signals (TGF-b). Unfortunately, the massive accumulation of macrophages can lead to tumor recurrence, which is very similar to TAMs-guided tissue damage repair after chemotherapy. M1 macrophages have antitumor effects, while M2 macrophages mainly play a role in promoting tumor growth, invasion and metastasis. Our study found that macrophages infiltrated more in esophageal cancer tissues in responder group, among which M0 macrophages were the main infiltrating cells. The infiltrating number of M2 macrophages was more often in the non-responder group, although there was no statistical difference compared with that in responder group. It is worth noting that both M1 and M2 macrophages have high degree of plasticity, which makes it possible to design appropriate methods to re-induce and re-educate them, thereby becoming an effective weapon against tumors. Different types of macrophages can be converted into each other upon tumor microenvironment changes or therapeutic interventions. In view of the important role of macrophages in tumor radioresistance, it needs to do in-depth research on the realization of their functions and their regulatory mechanisms, in order to find new anti-tumor targets.

Some limitations should be mentioned in our study. First, the prognostic model based on ir-lncRNA pairs was established through bioinformatics analyses from data available in the GEO databases. Hence, some further prospective trials or experimental data should be performed to validate the findings of this study. Second, our study found that most of clinical characters were not correlated with the riskscore, and tumor grade had a weak correlation with the riskscore. We speculate that preoperative adjuvant therapy altered gene expression status, resulting in clinical factors not correlated with gene expression. In addition, insufficient sample size may also be an important reason. Third, our study only initially explored the potential relationship between ir-lncRNAs risk signatures and immune cell infiltration, so further studies are needed to reveal the underlying mechanisms. Fourth, the effect of ir-lncRNAs on the immune microenvironment was indirectly speculated through immune genes, and there is currently a lack of direct evidence. More experiments are needed to confirm the impact of ir-lncRNAs on the radioresistance-related microenvironment of ESCC. Finally, as a preliminary exploratory study, it only has a qualitative role in the identification of ir-lncRNAs and the prognosis risk of ESCC patients. The relationship between ir-lncRNAs and the prognosis of ESCC patients has not been accurately quantified, which still needs to be further verified by multicenter studies with large samples.

In conclusion, ir-lncRNAs may be involved in the biological regulation of radioresistance in patients with ESCC. Changes in macrophage infiltration and immune gene expression are potential mechanisms of radiotherapy resistance, which are worthy of further study. This study had successfully established a prognostic risk model based on three ir-lncRNAs pairs, which is an important attempt to identify and predict the prognosis of ESCC. More importantly, it is a useful supplementary method to predict the prognosis of patients with esophageal squamous cell carcinoma based on TNM staging.

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

Conception and design: JZ and XC; Provision of study materials or patients: JZ and BH; Collection and assembly of data: JZ and BH; Data analysis and interpretation: JZ and JL; Manuscript writing and editing: JZ, XC, BH, and JL. All authors contributed to the article and approved the submitted version.

Funding

This study was supported in part by the National Clinical Key Specialty Construction Program (Grant No. 2021), and the Fujian Provincial Clinical Research Center for Cancer Radiotherapy and Immunotherapy (Grant No. 2020Y2012), Training and Nurturing Project for Young and Middle-aged Leading Talents in Healthcare in Fujian Province (Fujian Healthcare Personnel Document 2022-954 to XC), Joint Funds for the Innovation of Science and Technology, Fujian Province (Grant No: 2020Y9036 to XC), Fujian provincial health technology project (Youth Scientific Research Project, 2019-1-50 to JZ) and the Nursery Fund Project of the Second Affiliated Hospital of Fujian Medical University (Grant No: 2021MP05 to JZ).

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.2022.921902/full#supplementary-material

SUPPLEMENTARY FIGURE S1
Selection of the optimal candidate genes in the LASSO model.

SUPPLEMENTARY FIGURE S2
LASSO coefficients of prognosis-associated ir-lncRNA pairs, each curve represents a gene.

SUPPLEMENTARY FIGURE S3
Differential distribution of risk score with different clinical characters.

SUPPLEMENTARY FIGURE S4
The expression levels of ir-lncRNAs detected by RT-qPCR, $*p < 0.05$, $**p < 0.01$, $***p < 0.001$.

References

Anderson, G. (2020). Tumour microenvironment: roleis of the aryl hydrocarbon receptor, O-GlcNAcylation, acetyl-CoA and melatonergic pathway in regulating dynamic metabolic interactions across cell types-tumour microenvironment and metabolism. *Int. J. Mol. Sci.* 22 (1), 141. doi:10.3390/ijms22010141

Batista, P. J., and Chang, H. Y. (2013). Long noncoding RNAs: cellular address codes in development and disease. *Cell* 152 (6), 1298–1307. doi:10.1016/j.cell.2013.02.012

Bremnes, R. M., Busund, L. T., Kilvær, T. L., Andersen, S., Richardsen, E., Paulsen, E. E., et al. (2016). The role of tumor-infiltrating lymphocytes in development, progression, and prognosis of non-small cell lung cancer. *J. Thorac. Oncol.* 11 (6), 789–800. doi:10.1016/j.jtho.2016.01.015

Cassetta, L., and Pollard, J. W. (2020). Tumor-associated macrophages. *Curr. Biol.* 30 (6), R246–r248. doi:10.1016/j.cub.2020.01.031

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 (2), 115–132. doi:10.3322/caac.21338

Chen, G. Z., Zhu, H. C., Dai, W. S., Zeng, X. N., Luo, J. H., and Sun, X. C. (2017). The mechanisms of radioresistance in esophageal squamous cell carcinoma and current strategies in radiosensitivity. *J. Thorac. Dis.* 9 (3), 849–859. doi:10.21037/jtd.2017.03.23

Chen, H., Yao, X., Di, X., Zhang, Y., Zhu, H., Liu, S., et al. (2020a). MiR-450a-5p inhibits autophagy and enhances radiosensitivity by targeting dual-specificity phosphatase 10 in esophageal squamous cell carcinoma. *Cancer Lett.* 483, 114–126. doi:10.1016/j.canlet.2020.01.037

Chen, M. M., Zeng, G. P., Li, J., Fu, J. H., Long, Y. Y., Pan, J. Y., et al. (2020b). High infiltration of CD20(+) B lymphocytes in extranodal natural killer/T-cell lymphoma

is associated with better prognosis. *Br. J. Haematol.* 191 (5), e116–e120. doi:10.1111/bjh.17069

Chen, Y., Tang, J., Li, L., and Lu, T. (2021). Effect of Linc-POU3F3 on radiotherapy resistance and cancer stem cell markers of esophageal cancer cells. *Zhong Nan Da Xue Xue Bao Yi Xue Ban.* 46 (6), 583–590. doi:10.11817/j.issn.1672-7347.2021.190758

Deng, H. Y., Wang, Y. C., Ni, P. Z., Lin, Y. D., and Chen, L. Q. (2016). Long noncoding RNAs are novel potential prognostic biomarkers for esophageal squamous cell carcinoma: an overview. *J. Thorac. Dis.* 8 (8), E653–E659. doi:10.21037/jtd.2016.07.01

Fei, Z., Xie, R., Chen, Z., Xie, J., Gu, Y., Zhou, Y., et al. (2021). Establishment of a novel risk score system of immune genes associated with prognosis in esophageal carcinoma. *Front. Oncol.* 11, 625271. doi:10.3389/fonc.2021.625271

Greten, F. R., and Grivennikov, S. I. (2019). Inflammation and cancer: Triggers, mechanisms, and consequences. *Immunity* 51 (1), 27–41. doi:10.1016/j.immuni.2019.06.025

Han, G., Ling, R., Sun, C., Wang, X., Zhou, Y., Yu, L., et al. (2021). HMGB1 knockdown increases the radiosensitivity of esophageal squamous cell carcinoma by regulating the expression of molecules involved in DNA repair. *Oncol. Lett.* 22 (1), 503. doi:10.3892/ol.2021.12764

Hong, W., Liang, L., Gu, Y., Qi, Z., Qiu, H., Yang, X., et al. (2020). Immune-related lncRNA to construct novel signature and predict the immune landscape of human hepatocellular carcinoma. *Mol. Ther. Nucleic Acids* 22, 937–947. doi:10.1016/j.omtn.2020.10.002

Hua, Y., Wang, W., Zheng, X., Yang, L., Wu, H., Hu, Z., et al. (2020). NVP-BSK805, an inhibitor of JAK2 kinase, significantly enhances the radiosensitivity of esophageal squamous cell carcinoma *in vitro* and *in vivo*. *Drug Des. devel. Ther.* 14, 745–755. doi:10.2147/DDDT.S203048

Huang, D., Chen, J., Yang, L., Ouyang, Q., Li, J., Lao, L., et al. (2018). NKILA lncRNA promotes tumor immune evasion by sensitizing T cells to activation-induced cell death. *Nat. Immunol.* 19 (10), 1112–1125. doi:10.1038/s41590-018-0207-y

Huang, H. Z., Yin, Y. F., Wan, W. J., Xia, D., Wang, R., and Shen, X. M. (2019). Up-regulation of microRNA-136 induces apoptosis and radiosensitivity of esophageal squamous cell carcinoma cells by inhibiting the expression of MUC1. *Exp. Mol. Pathol.* 110, 104278. doi:10.1016/j.yexmp.2019.104278

Huang, H. Y., Wang, Y., Wang, W. D., Wei, X. L., Gale, R. P., Li, J. Y., et al. (2021). A prognostic survival model based on metabolism-related gene expression in plasma cell myeloma. *Leukemia* 35 (11), 3212–3222. doi:10.1038/s41375-021-01206-4

Kumar, V., Ramnarayanan, K., Sundar, R., Padmanabhan, N., Srivastava, S., Koiwa, M., et al. (2022). Single-cell atlas of lineage States, tumor microenvironment, and subtype-specific expression programs in gastric cancer. *Cancer Discov.* 12 (3), 670–691. doi:10.1158/2159-8290.CD-21-0683

Li, C. H., and Chen, Y. (2013). Targeting long non-coding RNAs in cancers: Progress and prospects. *Int. J. Biochem. Cell Biol.* 45 (8), 1895–1910. doi:10.1016/j.biocel.2013.05.030

Li, D. J., Li, H. W., He, B., Wang, G. M., Cai, H. F., Duan, S. M., et al. (2016). Patterns of failure after involved field radiotherapy for locally advanced esophageal squamous cell carcinoma. *J. BUON* 21 (5), 1268–1273.

Li, Y., Lu, Z., Che, Y., Wang, J., Sun, S., Huang, J., et al. (2017). Immune signature profiling identified predictive and prognostic factors for esophageal squamous cell carcinoma. *Oncoimmunology* 6 (11), e1356147. doi:10.1080/2162402x.2017.1356147

Li, B., Cui, Y., Nambiar, D. K., Sunwoo, J. B., and Li, R. (2019a). The immune subtypes and landscape of squamous cell carcinoma. *Clin. Cancer Res.* 25 (12), 3528–3537. doi:10.1158/1078-0432.Ccr-18-4085

Li, Y., Wan, Q., Wang, W., Mai, L., Sha, L., Mashrah, M., et al. (2019b). LncRNA ADAMTS9-AS2 promotes tongue squamous cell carcinoma proliferation, migration and EMT via the miR-600/EZH2 axis. *Biomed. Pharmacother.* 112, 108719. doi:10.1016/j.biopha.2019.108719

Li, W., Liu, J., and Zhao, H. (2020a). Identification of a nomogram based on long non-coding RNA to improve prognosis prediction of esophageal squamous cell carcinoma. *Aging (Albany NY)* 12 (2), 1512–1526. doi:10.18632/aging.102697

Li, Z. Y., Li, H. F., Zhang, Y. Y., Zhang, X. L., Wang, B., and Liu, J. T. (2020b). Value of long non-coding RNA Rpph1 in esophageal cancer and its effect on cancer cell sensitivity to radiotherapy. *World J. Gastroenterol.* 26 (15), 1775–1791. doi:10.3748/wjg.v26.i15.1775

Lin, J., Liu, Z., Liao, S., Li, E., Wu, X., and Zeng, W. (2020). Elevation of long non-coding RNA GAS5 and knockdown of microRNA-21 up-regulate RECK expression to enhance esophageal squamous cell carcinoma cell radio-sensitivity after radiotherapy. *Genomics* 112 (3), 2173–2185. doi:10.1016/j.ygeno.2019.12.013

Lin, Z., Huang, W., Yi, Y., Li, D., Xie, Z., Li, Z., et al. (2021). LncRNA ADAMTS9-AS2 is a prognostic biomarker and correlated with immune infiltrates in lung adenocarcinoma. *Int. J. Gen. Med.* 14, 8541–8555. doi:10.2147/ijgm.S340683

Liu, J., Wang, Y., Chu, Y., Xu, R., Zhang, D., and Wang, X. (2020). Identification of a TLR-induced four-lncRNA signature as a novel prognostic biomarker in esophageal carcinoma. *Front. Cell Dev. Biol.* 8, 649. doi:10.3389/fcell.2020.00649

Liu, J., Zhou, R., Deng, M., Xue, N., Li, T., Guo, Y., et al. (2021). Long non-coding RNA DIO3OS binds to microRNA-130b to restore radiosensitivity in esophageal squamous cell carcinoma by upregulating PAX9. *Cancer Gene Ther.* 29, 870. doi:10.1038/s41417-021-00344-2

Liu, Z., Lu, X., Wen, L., You, C., Jin, X., and Liu, J. (2022). Hsa_circ_0014879 regulates the radiosensitivity of esophageal squamous cell carcinoma through miR-519-3p/CDC25A axis. *Anticancer. Drugs* 33 (1), e349–e361. doi:10.1097/CAD.0000000000001213

Lu, G., Chen, L., Wu, S., Feng, Y., and Lin, T. (2020). Comprehensive analysis of tumor-infiltrating immune cells and relevant therapeutic strategy in esophageal cancer. *Dis. Markers* 2020, 8974793. doi:10.1155/2020/8974793

Lv, X. B., Lian, G. Y., Wang, H. R., Song, E., Yao, H., and Wang, M. H. (2013). Long noncoding RNA HOTAIR is a prognostic marker for esophageal squamous cell carcinoma progression and survival. *PLoS One* 8 (5), e63516. doi:10.1371/journal.pone.0063516

Mantovani, A., and Allavena, P. (2015). The interaction of anticancer therapies with tumor-associated macrophages. *J. Exp. Med.* 212 (4), 435–445. doi:10.1084/jem.20150295

Mori, K., Sugawara, K., Aikou, S., Yamashita, H., Yamashita, K., Ogura, M., et al. (2021). Esophageal cancer patients' survival after complete response to definitive chemoradiotherapy: a retrospective analysis. *Esophagus.* 18 (3), 629–637. doi:10.1007/s10388-021-00817-1

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37 (7), 773–782. doi:10.1038/s41587-019-0114-2

Peña-Romero, A. C., and Orenes-Piñero, E. (2022). Dual effect of immune cells within tumour microenvironment: Pro- and anti-tumour effects and their triggers. *Cancers* 14 (7), 1681. doi:10.3390/cancers14071681

Qu, L., Wang, Z. L., Chen, Q., Li, Y. M., He, H. W., Hsieh, J. J., et al. (2018). Prognostic value of a long non-coding RNA signature in localized clear cell renal cell carcinoma. *Eur. Urol.* 74 (6), 756–763. doi:10.1016/j.eururo.2018.07.032

Schabath, M. B., Giuliano, A. R., Thompson, Z. J., Amankwah, E. K., Gray, J. E., Fenstermacher, D. A., et al. (2013). TNFRSF10B polymorphisms and haplotypes associated with increased risk of death in non-small cell lung cancer. *Carcinogenesis* 34 (11), 2525–2530. doi:10.1093/carcin/bgt244

Shen, F. F., Zhang, F., Yang, H. J., Li, J. K., Su, J. F., Yu, P. T., et al. (2020). ADAMTS9-AS2 and CADM2 expression and association with the prognosis in esophageal squamous cell carcinoma. *Biomark. Med.* 14 (15), 1415–1426. doi:10.2217/bmm-2020-0432

Shi, X., Sun, M., Liu, H., Yao, Y., and Song, Y. (2013). Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* 339 (2), 159–166. doi:10.1016/j.canlet.2013.06.013

Shi, X., Chen, Z., Hu, X., Luo, M., Sun, Z., Li, J., et al. (2016). AJUBA promotes the migration and invasion of esophageal squamous cell carcinoma cells through upregulation of MMP10 and MMP13 expression. *Oncotarget* 7 (24), 36407–36418. doi:10.18632/oncotarget.9239

Singh, S. R., Rameshwar, P., and Siegel, P. (2016). Targeting tumor microenvironment in cancer therapy. *Cancer Lett.* 380 (1), 203–204. doi:10.1016/j.canlet.2016.04.009

Song, E. L., Xing, L., Wang, L., Song, W. T., Li, D. B., Wang, Y., et al. (2019). LncRNA ADAMTS9-AS2 inhibits cell proliferation and decreases chemoresistance in clear cell renal cell carcinoma via the miR-27a-3p/FOXO1 axis. *Aging (Albany NY)* 11 (15), 5705–5725. doi:10.18632/aging.102154

Sounni, N. E., and Noel, A. (2013). Targeting the tumor microenvironment for cancer therapy. *Clin. Chem.* 59 (1), 85–93. doi:10.1373/clinchem.2012.185363

Sun, J., Zhang, Z., Bao, S., Yan, C., Hou, P., Wu, N., et al. (2020). Identification of tumor immune infiltration-associated lncRNAs for improving prognosis and immunotherapy response of patients with non-small cell lung cancer. *J. Immunother. Cancer* 8 (1), e000110. doi:10.1136/jitc-2019-000110

Sun, X., Sun, Y., Li, J., Zhao, X., Shi, X., Gong, T., et al. (2021). SOCS6 promotes radiosensitivity and decreases cancer cell stemness in esophageal squamous cell carcinoma by regulating c-Kit ubiquitylation. *Cancer Cell Int.* 21 (1), 165. doi:10.1186/s12935-021-01859-2

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 (3), 209–249. doi:10.3322/caac.21660

Takase, N., Koma, Y., Urakawa, N., Nishio, M., Arai, N., Akiyama, H., et al. (2016). NCAM- and FGF-2-mediated FGFR1 signaling in the tumor microenvironment of esophageal cancer regulates the survival and migration of tumor-associated macrophages and cancer cells. *Cancer Lett.* 380 (1), 47–58. doi:10.1016/j.canlet.2016.06.009

Tan, L., Cheng, D., Wen, J., Huang, K., and Zhang, Q. (2021). Identification of prognostic hypoxia-related genes signature on the tumor microenvironment in esophageal cancer. *Math. Biosci. Eng.* 18 (6), 7743–7758. doi:10.3934/mbe.2021384

Tang, L., Wei, F., Wu, Y., He, Y., Shi, L., Xiong, F., et al. (2018). Role of metabolism in cancer cell radioresistance and radiosensitization methods. *J. Exp. Clin. Cancer Res.* 37 (1), 87. doi:10.1186/s13046-018-0758-7

Wang, X., An, D., Liu, X., Wang, X., and Li, B. (2019). MicroRNA-27a downregulates the expression of Hsp90 and enhances the radiosensitivity in esophageal squamous cell carcinoma. *Onco. Targets. Ther.* 12, 5967–5977. doi:10.2147/OTT.S197456

Wang, M., Wang, L., He, X., Zhang, J., Zhu, Z., Zhang, M., et al. (2020). lncRNA CCAT2 promotes radiotherapy resistance for human esophageal carcinoma cells via the miR-145/p70S6K1 and p53 pathway. *Int. J. Oncol.* 56 (1), 327–336. doi:10.3892/ijo.2019.4929

Wang, Q., Li, X., Wang, Y., Qiu, J., Wu, J., He, Y., et al. (2021a). Development and validation of a three-gene prognostic signature based on tumor microenvironment for gastric cancer. *Front. Genet.* 12, 801240. doi:10.3389/fgene.2021.801240

Wang, Y., Liu, J., Ren, F., Chu, Y., and Cui, B. (2021b). Identification and validation of a four-long non-coding RNA signature associated with immune infiltration and prognosis in colon cancer. *Front. Genet.* 12, 671128. doi:10.3389/fgene.2021.671128

Wang, T., Zhang, J., Li, N., Li, M., Ma, S., Tan, S., et al. (2022). Spatial distribution and functional analysis define the action pathway of Tim-3/Tim-3 ligands in tumor development. *Mol. Ther.* 30 (3), 1135–1148. doi:10.1016/j.ymthe.2021.11.015

Welsh, T. J., Green, R. H., Richardson, D., Waller, D. A., O'Byrne, K. J., and Bradding, P. (2005). Macrophage and mast-cell invasion of tumor cell islets confers a marked survival advantage in non-small-cell lung cancer. *J. Clin. Oncol.* 23 (35), 8959–8967. doi:10.1200/jco.2005.01.4910

Wen, J., Yang, H., Liu, M. Z., Luo, K. J., Liu, H., Hu, Y., et al. (2014). Gene expression analysis of pretreatment biopsies predicts the pathological response of esophageal squamous cell carcinomas to neo-chemoradiotherapy. *Ann. Oncol.* 25 (9), 1769–1774. doi:10.1093/annonc/mdu201

Wu, M., Shang, X., Sun, Y., Wu, J., and Liu, G. (2020). Integrated analysis of lymphocyte infiltration-associated lncRNA for ovarian cancer via TCGA, GTEx and GEO datasets. *PeerJ* 8, e8961. doi:10.7717/peerj.8961

Yan, Y., Xu, Z., Chen, X., Wang, X., Zeng, S., Zhao, Z., et al. (2019). Novel function of lncRNA ADAMTS9-AS2 in promoting temozolomide resistance in glioblastoma via upregulating the FUS/MDM2 ubiquitination Axis. *Front. Cell Dev. Biol.* 7, 217. doi:10.3389/fcell.2019.00217

Yan, S., Xu, J., Liu, B., Ma, L., Feng, H., Tan, H., et al. (2021). Long non-coding RNA BCAR4 aggravated proliferation and migration in esophageal squamous cell carcinoma by negatively regulating p53/p21 signaling pathway. *Bioengineered* 12 (1), 682–696. doi:10.1080/21655979.2021.1887645

Yan, K., Wei, W., Shen, W., Du, X., Zhu, S., Zhao, H., et al. (2022). Combining the systemic inflammation response index and prognostic nutritional index to predict the prognosis of locally advanced elderly esophageal squamous cell carcinoma patients undergoing definitive radiotherapy. *J. Gastrointest. Oncol.* 13 (1), 13–25. doi:10.21037/jgo-21-784

Yang, Z., Zeng, H., Xia, R., Liu, Q., Sun, K., Zheng, R., et al. (2018). Annual cost of illness of stomach and esophageal cancer patients in urban and rural areas in China: A multi-center study. *Chin. J. Cancer Res.* 30 (4), 439–448. doi:10.21147/j.issn.1000-9604.2018.04.07

Yu, Y., Guan, H., Jiang, L., Li, X., Xing, L., and Sun, X. (2020). Nimotuzumab, an EGFRtargeted antibody, promotes radiosensitivity of recurrent esophageal squamous cell carcinoma. *Int. J. Oncol.* 56 (4), 945–956. doi:10.3892/ijo.2020.4981

Zeng, H., Chen, W., Zheng, R., Zhang, S., Ji, J. S., Zou, X., et al. (2018). Changing cancer survival in China during 2003-15: a pooled analysis of 17 population-based cancer registries. *Lancet. Glob. Health* 6 (5), e555–e567. doi:10.1016/S2214-109X(18)30127-X

Zhang, X., Xu, Y., He, C., Guo, X., Zhang, J., He, C., et al. (2015). Elevated expression of CCAT2 is associated with poor prognosis in esophageal squamous cell carcinoma. *J. Surg. Oncol.* 111 (7), 834–839. doi:10.1002/jso.23888

Zhang, Z., Jia, J.-P., Zhang, Y.-J., Liu, G., Zhou, F., and Zhang, B.-C. (2020). Long noncoding RNA ADAMTS9-AS2 inhibits the proliferation, migration, and invasion in bladder tumor cells. *Onco. Targets. Ther.* 13, 7089–7100. doi:10.2147/OTT.S245826

Zhang, H., Si, J., Yue, J., and Ma, S. (2021a). The mechanisms and reversal strategies of tumor radioresistance in esophageal squamous cell carcinoma. *J. Cancer Res. Clin. Oncol.* 147 (5), 1275–1286. doi:10.1007/s00432-020-03493-3

Zhang, Y., Zhu, M., Mo, J., and Xian, L. (2021b). Tumor microenvironment characterization in esophageal cancer identifies prognostic relevant immune cell subtypes and gene signatures. *Aging (Albany NY)* 13 (24), 26118–26136. doi:10.18632/aging.203800

Zheng, Y., Chen, Z., Han, Y., Han, L., Zou, X., Zhou, B., et al. (2020). Immune suppressive landscape in the human esophageal squamous cell carcinoma microenvironment. *Nat. Commun.* 11 (1), 6268. doi:10.1038/s41467-020-20019-0

Zhu, X., Tian, X., Yu, C., Shen, C., Yan, T., Hong, J., et al. (2016). A long non-coding RNA signature to improve prognosis prediction of gastric cancer. *Mol. Cancer* 15 (1), 60. doi:10.1186/s12943-016-0544-0

Zhu, T., Ma, Z., Wang, H., Wei, D., Wang, B., Zhang, C., et al. (2021). Immune-related long non-coding RNA signature and clinical nomogram to evaluate survival of patients suffering esophageal squamous cell carcinoma. *Front. Cell Dev. Biol.* 9, 641960. doi:10.3389/fcell.2021.641960

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 September 2022
DOI 10.3389/fgene.2022.982888

OPEN ACCESS

EDITED BY
Yan Gong,
Wuhan University, China

REVIEWED BY
Ling Han,
Naval Medical University, China
Guo Zu,
Dalian Municipal Central Hospital, China

*CORRESPONDENCE
Tianci Yao,
ytc0128@163.com
Hongqi Huo,
huohongqi77@126.com
Jing Huang,
jolietjing@163.com

†These authors have contributed equally to this work

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

RECEIVED 30 June 2022
ACCEPTED 16 August 2022
PUBLISHED 09 September 2022

CITATION
Yu W, Huo H, You Z, Lu R, Yao T and Huang J (2022), Identification of cuproptosis-associated lncRNAs signature and establishment of a novel nomogram for prognosis of stomach adenocarcinoma. *Front. Genet.* 13:982888. doi: 10.3389/fgene.2022.982888

Identification of cuproptosis-associated lncRNAs signature and establishment of a novel nomogram for prognosis of stomach adenocarcinoma

Wei Yu[1†], Hongqi Huo[2*†], Zhixin You[2], Rong Lu[3], Tianci Yao[4*] and Jing Huang[1*]

[1]Department of Pharmacy, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital, Fuzhou, China, [2]Nuclear Medicine Department, HanDan Central Hospital, Handan, China, [3]Department of Laboratory Medicine, The First Affiliated Hospital of Xiamen University, Xiamen Key Laboratory of Genetic Testing, School of Medicine, Xiamen University, Xiamen, China, [4]Department of Pharmacy, The First Affiliated Hospital of Xiamen University, Xiamen, China

Purpose: Stomach adenocarcinoma (STAD) is one of the common cancers globally. Cuproptosis is a newly identified cell death pattern. The role of cuproptosis-associated lncRNAs in STAD is unknown.

Methods: STAD patient data from TCGA were used to identify prognostic lncRNAs by Cox regression and LASSO. A nomogram was constructed to predict patient survival. The biological profiles were evaluated through GO and KEGG.

Results: We identified 298 cuproptosis-related lncRNAs and 13 survival-related lncRNAs. Patients could be categorized into either high risk group or low risk group with 9-lncRNA risk model with significantly different survival time ($p < 0.001$). ROC curve and nomogram confirmed the 9-lncRNA risk mode had good prediction capability. Patients in the lower risk score had high gene mutation burden. We also found that patients in the two groups might respond differently to immune checkpoint inhibitors and some anti-tumor compounds.

Conclusion: The nomogram with 9-lncRNA may help guide treatment of STAD. Future clinical studies are necessary to verify the nomogram.

KEYWORDS

risk score, prognosis, cuproptosis, immune, tumor

1 Introduction

Stomach adenocarcinoma (STAD) is frequently found in the digestive tract (Bray et al., 2018). It is mostly reported in eastern Asia and South America. There are several risk factors for STAD, including *Helicobacter pylori* infection, adenomatous gastric polyps, diet low in fruits and vegetables and diet high in cured or smoked foods (Wroblewski et al., 2010; Rawla

and Barsouk, 2019; Akshatha et al., 2021). Although STAD is treatable surgically in its early stages, advanced STAD has a poor prognosis. Innovative therapeutics and prognostic models are both needed to improve prognosis of advanced STAD (Ajani et al., 2017; Ivey et al., 2022).

Metal micronutrients, especially iron (Fe), zinc (Zn), and copper (Cu), are essential for life. For example, Zn is involved in regulation of gene expression, and approximately 2,800 proteins may bind Zn *in vivo* (Andreini et al., 2006). Cu is catalyst or structural cofactor in many cellular activities, including mitochondrial respiration, immune function, and free radical scavenging (Festa and Thiele, 2011; Cobine et al., 2021). Despite its role for normal life, high serum copper level has been linked to increased risk of cancer (Brady et al., 2014; Tsang et al., 2020) and atherosclerotic diseases (Reunanen et al., 1992; Ford, 2000; Chen et al., 2015).

Mostly recently Tsvetkov et al. (2022) demonstrated a novel mechanism of Cu-induced cell death that is related to mitochondria dysfunction. This novel form of regulated cell death was termed "cuproptosis". Cuproptosis may happen when mitochondrial enzymes aggregate and leads to mitochondrial stress. Cuproptosis is different from apoptosis, ferroptosis, or necroptosis. This discovery suggests mitochondrial Cu homeostasis may be exploited for cancer therapy.

Here we explored whether cuproptosis-related lncRNAs may be involved in STAD patient prognosis. The results might help understand the roles of cuproptosis in the development and progression of STAD.

2 Materials and methods

2.1 TCGA data

We downloaded RNA sequencing (RNA-seq) and expression files and mutation files from the Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/repository). The data included tumor tissues of 343 STAD patients and 30 matched normal tissues. Data were downloaded and handled according to TCGA guidelines.

2.2 Identification of cuproptosis-related lncRNAs

According to the study by Tsvetkov et al. (2022), 19 cuproptosis-associated genes were evaluated (Supplementary Table S1). Correlation between cuproptosis-related genes and differentially expressed lncRNAs was evaluated. Pearson's correlation coefficients (R) of gene expression patterns were used as a measure of gene coexpression. The PCC threshold to retrieve cuproptosis-related lncRNAs was 0.4 ($|R| > 0.4$), with a p value < 0.001.

2.3 Cuproptosis-related lncRNAs signature for STAD prognosis

The downloaded clinical and demographic data of STAD patients were analyzed with univariate Cox regression analysis to identify lncRNAs associated with patient overall survival (OS) and those associated with cuproptosis were further identified as candidate lncRNAs for the construction of prognostic signature. Lasso regression was performed to screen lncRNAs that were truly correlated with a patient's survival on the basis of 10-fold cross-validation. Based on the nine optimal lncRNAs identified, the risk scores of patients were calculated according to the following formula:

$$\text{risk score} = \sum_{i}^{n} Xi^{*}Yi$$

Where X was regression coefficient and Y was expression level of cuproptosis-related lncRNAs.

A total of 343 STAD patients were allocated to either the training cohort or the test cohort randomly in a 1:1 ratio for constructing and validating the cuproptosis-related lncRNAs signature. Patients in each cohort were classified into either low--risk group or high-risk group according to the cut-off value, which was the median risk score (Meng et al., 2019; Hong et al., 2020). The Chi-square test and the receiver operating characteristics (ROC) curves were used to help determine if observed OS was in line with expected OS, and the 1-year, 3-years, and 5-years OS rates were compared between the low-risk group and the high-risk group by Kaplan–Meier analysis. We further constructed a nomogram with cuproptosis-related lncRNA risk score and established clinical risk factors to calculated patient survival time. Then concordance index (C-index) and calibration curves were used to evaluate the prediction power of the nomogram. Finally, stratified analysis was used to assess whether the signature retained its predictive ability in subgroups of patients (stages I–II and stages III–IV). The "survival", "rms", "survminer" and "timeROC" R packages were used.

2.4 Principal component analysis, gene ontology and gene set enrichment analysis

We used principal component analysis (PCA) to characterize cuproptosis-related lncRNAs expression patterns. PCA is a common unsupervised method for the analysis of gene expression data. 3D scatter plots were used to visualize the relationship between the three variables of samples. The analysis of differentially expressed genes (DEGs) was performed with the glm method of the "edgeR" R package. We set the threshold value of log fold change (log2FC) at $|\log2FC| \geq 1$, with a false discovery rate (FDR) < 0.05, to identify important DEGs. Gene Ontology (GO) was used to interpret DEGs

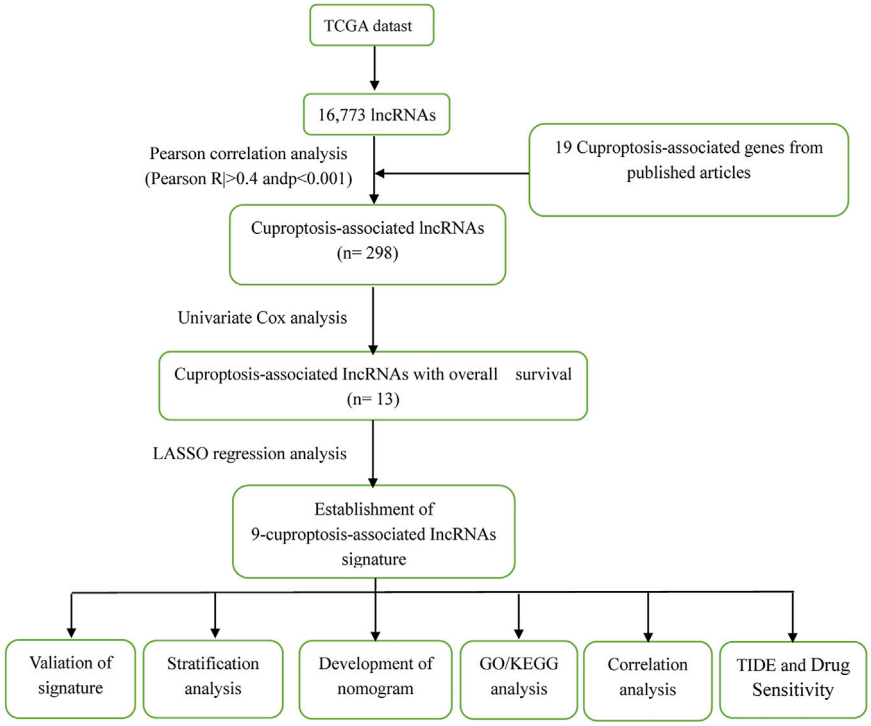


FIGURE 1
The process of the study.

for the relevant cellular components, biological processes, and molecular functions. Differential Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways between the high-risk group and the low-risk group were screened using Gene Set Enrichment Analysis (GSEA), with a FDR <0.25.

2.5 Immune function

Single-sample GSEA (ssGSEA), an extension of GSEA, was used to calculate separate enrichment scores for immunological pathways by the normalized enrichment score (NES) (Subramanian et al., 2005). Each ssGSEA enrichment score represents the degree to which the genes are coordinately upregulated or downregulated within a sample.

2.6 Tumor mutation burden

We downloaded the somatic mutation file and calculated each patient's tumor mutation burden (TMB) score. The influence of TMB on patient OS was evaluated by Kaplan–Meier analysis and compared between the high- and low-risk groups by *t*-test. Maftools R package was used.

2.7 Tumor immune dysfunction and exclusion score and drug sensitivity prediction

To predict treatment response of immune checkpoint blockades (ICBs), tumor immune dysfunction and exclusion (TIDE) algorithm was used to identify signatures of T cell dysfunction and signatures that exclude T cell infiltration into tumors (Jiang et al., 2018). To predict treatment response of the most important groups of drugs again STAD, the half-maximal inhibitory concentrations (IC_{50}) were calculated using pRRophetic as described in Genomics of Drug Sensitivity in Cancer (GDSC) (Geeleher et al., 2014).

3 Results

3.1 lncRNAs data

Figure 1 illustrates the results of the search and the process of screening. A total of 16,773 lncRNAs that may be associated with 19 cuproptosis-associated genes were found. Among these lncRNAs, 298 lncRNAs met the pre-defined criteria ($(|R| > 0.4)$. All 298 lncRNAs upregulated the expression of cuproptosis genes in the Sankey diagram (Figure 2A. Univariate Cox regression analysis found that 13 lncRNAs were prognostic factors of patient survival (Figure 2B).

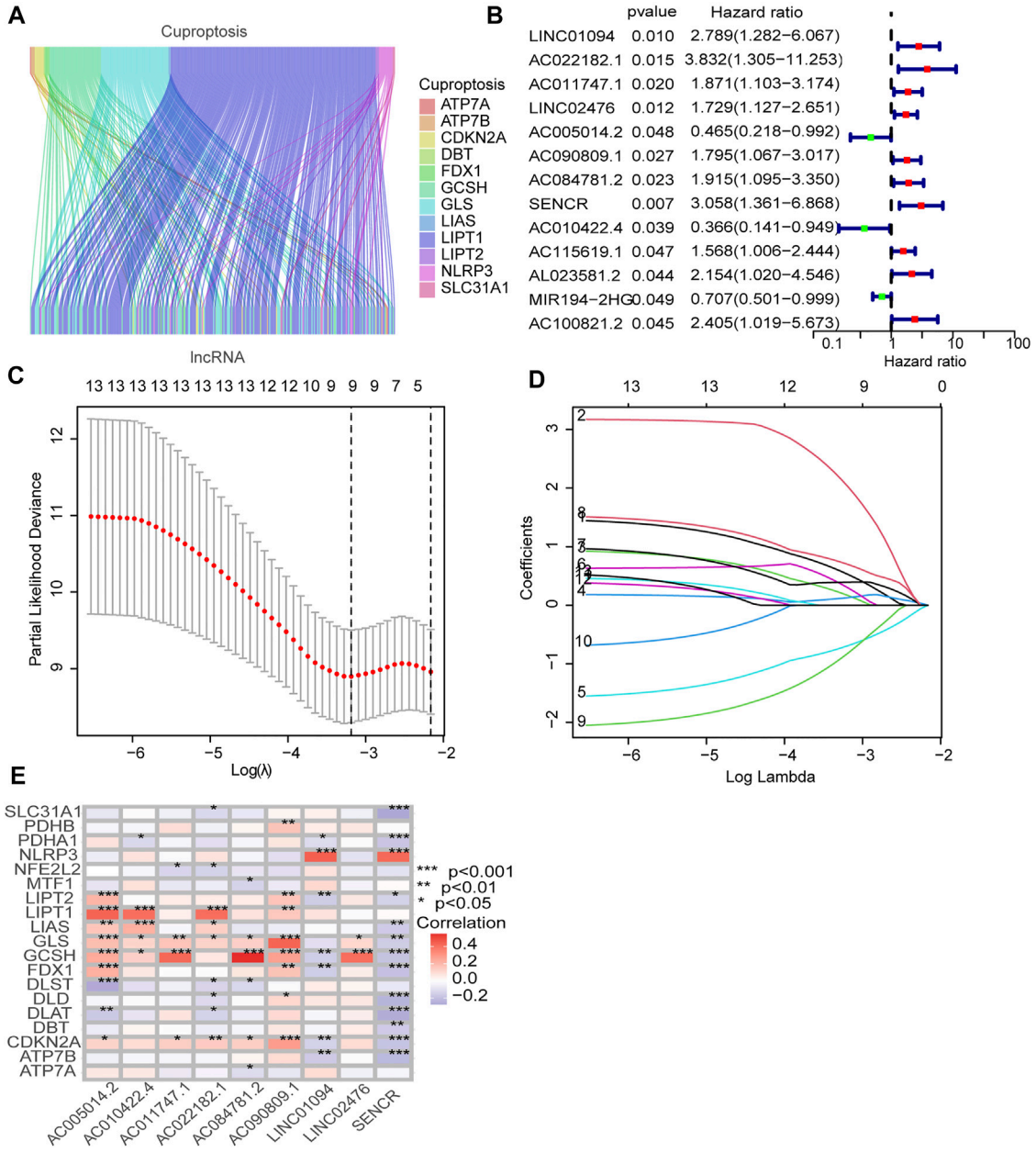


FIGURE 2
Identification of prognostic cuproptosis-related lncRNAs in STAD. **(A)** The Sankey diagram demonstrates correlation between cuproptosis-related lncRNAs and cuproptosis-related genes. **(B)** The prognostic lncRNAs identified by uni-Cox regression analysis. **(C)** LASSO model, with a 10-fold cross-validation. **(D)** The coefficient profile of nine lncRNAs screened by the LASSO model. **(E)** Correlations between lncRNAs in the risk model and cuproptosis-related genes.

3.2 Risk model

To construct a risk model with cuproptosis-related lncRNAs in STAD, we randomly allocated 343 STAD cases into the training set and the test set at 1:1 ratio. The chi-square test showed that the two groups were comparable in terms of both clinicopathologic and demographic parameters (Table 1).

To avoid overfitting, nine lncRNAs were further identified by LASSO regression method (Figures 2C,D. A formula was established with the expression levels of nine lncRNAs:

TABLE 1 Clinicopathologic and demographic characteristics of STAD patients in the training and test cohorts.

Variable	Total	Training cohort	Test cohort	*p* value
≤65	72 (42.6%)	38 (44.71%)	34 (40.48%)	0.6889
>65	97 (57.4%)	47 (55.29%)	50 (59.52%)	
Female	69 (40.83%)	30 (35.29%)	39 (46.43%)	0.1882
Male	100 (59.17%)	55 (64.71%)	45 (53.57%)	
G1	3 (1.78%)	2 (2.35%)	1 (1.19%)	0.5907
G2	71 (42.01%)	33 (38.82%)	38 (45.24%)	
G3	92 (54.44%)	49 (57.65%)	43 (51.19%)	
Unknown	3 (1.78%)	1 (1.18%)	2 (2.38%)	
Stage I	24 (14.2%)	10 (11.76%)	14 (16.67%)	0.8193
Stage II	46 (27.22%)	23 (27.06%)	23 (27.38%)	
Stage III	68 (40.24%)	35 (41.18%)	33 (39.29%)	
Stage IV	20 (11.83%)	11 (12.94%)	9 (10.71%)	
Unknown	11 (6.51%)	6 (7.06%)	5 (5.95%)	
T1	13 (7.69%)	8 (9.41%)	5 (5.95%)	0.1533
T2	29 (17.16%)	9 (10.59%)	20 (23.81%)	
T3	74 (43.79%)	39 (45.88%)	35 (41.67%)	
T4	47 (27.81%)	25 (29.41%)	22 (26.19%)	
Unknown	6 (3.55%)	4 (4.71%)	2 (2.38%)	
N0	50 (29.59%)	20 (23.53%)	30 (35.71%)	0.0402
N1	45 (26.63%)	27 (31.76%)	18 (21.43%)	
N2	31 (18.34%)	12 (14.12%)	19 (22.62%)	
N3	34 (20.12%)	22 (25.88%)	12 (14.29%)	
Unknown	9 (5.33%)	4 (4.71%)	5 (5.95%)	
M0	148 (87.57%)	75 (88.24%)	73 (86.9%)	1
M1	14 (8.28%)	7 (8.24%)	7 (8.33%)	
Unknown	7 (4.14%)	3 (3.53%)	4 (4.76%)	

Risk score = LINC01094 × (0.5250) + AC022182.1 × (2.02146) + AC011747.1 × (0.1655) + LINC02476 × (0.1295) + AC005014.2 × (−0.6903) + AC090809.1 × (0.2959) + AC084781.2 × (0.3942) + SENCR × (0.6958) + AC010422.4 × (−0.8166) (Meng et al., 2019).

As expected, the high-risk group had worse survival in each sample set (Figure 3).

3.3 Assessment of the risk model

The areas under the l-, 3- and 5-years ROC curves (AUC) were 0.719, 0.773, and 0.755 respectively (Figure 4A). The AUC of risk score was 0.719 and the C-index in the risk model was 0.726, indicting a perfect predictive ability (Figures 4B,C). In the uni-Cox regression, the hazard ratios (HR) of the risk score was 1.0726 ($p < 0.001$), and in the multi-Cox regression, HR of the risk score was 1.092 ($p < 0.001$) (Figures 4D,E).

3.4 Nomogram

A nomogram model was drawn to predict OS of patients (Figure 5A). The calibration plots showed the predicted l-, 3- and 5-years OS was consistent with the actual OS (Figure 5B). Thus the nomogram was well calibrated, with good prediction of patient survival. The high value of C index (0.726) indicated that the nomogram has excellent discriminative ability.

The results of decision curve analyses to compare the performance of the nomogram are shown in Figure 5C. The nomogram has greater net benefit than other clinical parameters in all patients.

3.4 PCA and biological pathways analyses

The 3D scatter diagram showed the low-risk group and the high-risk group had distinct aggregation features of PCA (Figures 6A–C). GO analysis indicated related biological processes included B cell activation signaling pathway, antigen receptor–mediated signaling pathway, and immune response–regulating signaling pathway; related cellular components included immunological synapse, endocytic vesicle membrane, endocytic vesicle, T cell receptor complex, and immunoglobulin complex, and related molecular functions included immune receptor activity,

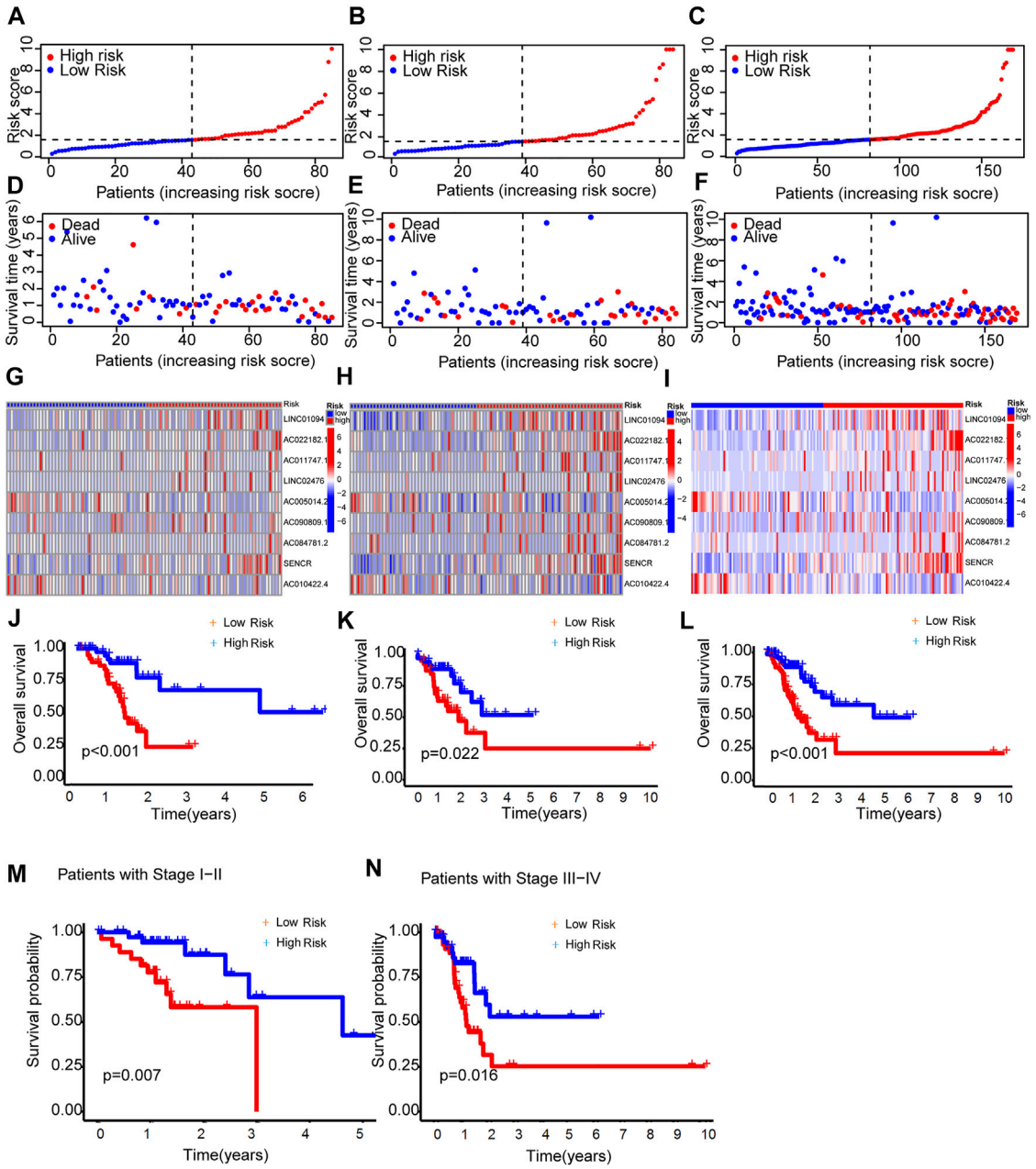


FIGURE 3
Prognosis capability of the model in the three patient sets. **(A–C)** Distribution of patient with different scores. **(D–F)** Distribution of patient survival time. **(G–I)** The heatmap of nine lncRNAs expression. **(J–L)** Comparison of OS curves of patients between the two groups of each set. **(M,N)** OS curves of stratified by clinicopathologic characteristics in the entire set.

heparin binding, glycosaminoglycan binding, sulfur compound binding, immunoglobulin receptor binding, and antigen binding (Figures 6D,E). GSEA identified genes involved in PI3K–Akt signaling pathway, cell adhesion, cytokine–cytokine receptor interaction and chemokine signaling pathway were differentially expressed between the low--risk group and high-risk group (Figures 6F,G).

3.5 Correlation analysis between risk scores and gene mutations

Somatic mutations between the two groups were compared. The ten most mutated genes were *TP53*, *TTN*, *PCLO*, *ZFHX4*, *CSMD3*, *SYNE1*, *ARID1A*, *LRP18*, *MUC16*, and *ACVR2A*. The high-risk group had more frequent *TP53* mutation (Figures

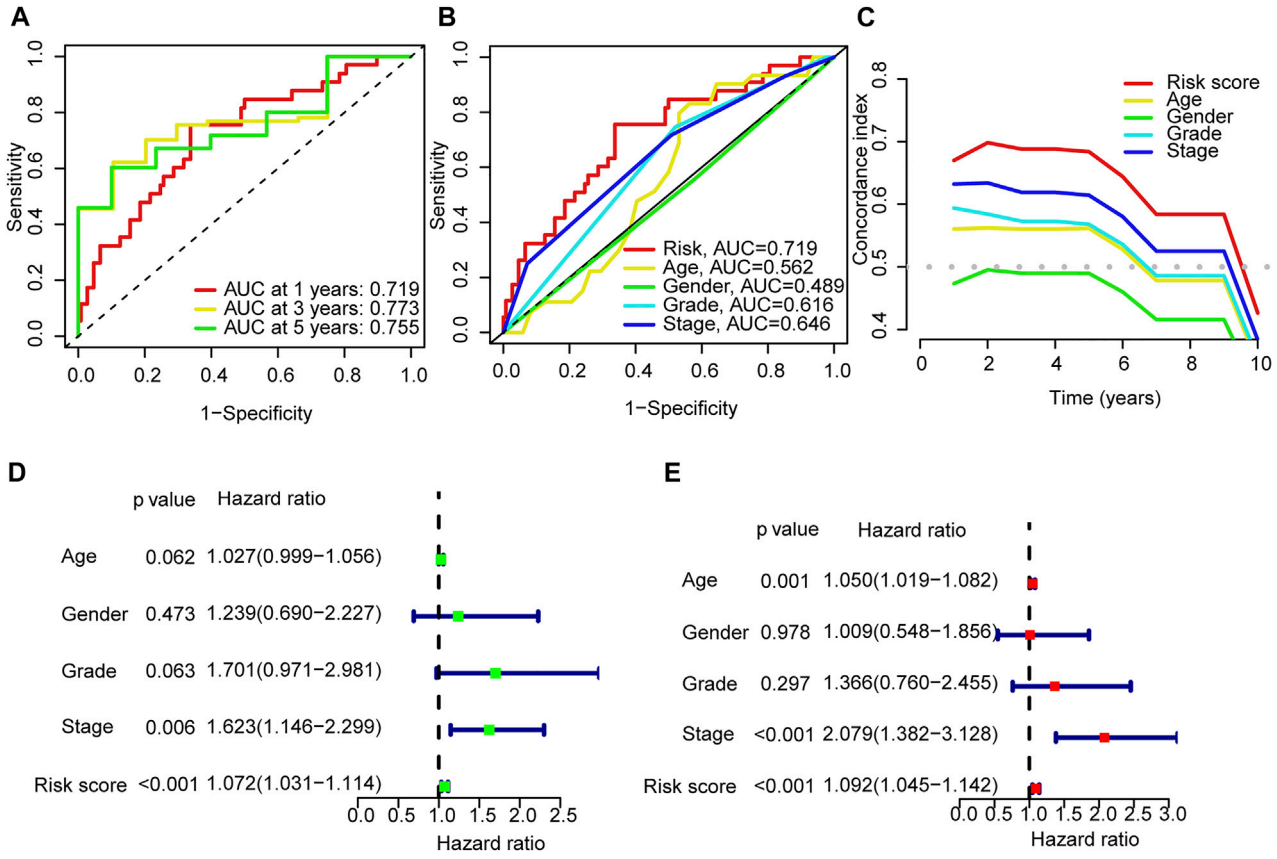


FIGURE 4
Validation of the model. **(A)** Time-dependent ROC curve analyses for survival of all patients based on the risk score model. **(B)** Comparision of the ROC curves of risk score, patient age, patient gender, tumor grade and tumor stage. **(C)** The C-index curves of risk score, patient age, patient gender, tumor grade and tumor stage. **(D,E)** Uni-Cox and multi-Cox analyses of overall survival for risk score, patient age, patient gender, tumor grade and tumor stage.

7A,B) but overall lower TMB (Figure 7C). Patients with higher scores and lower TMB had the worst prognosis among the four groups (Figures 7D,E).

3.6 TIDE, immune functions and prediction of clinical treatment response

The TIDE scores were significantly higher in the high-risk group compared to the low-risk group. This indicated that TIDE could be used to evaluate sensitivity to ICB therapy for STAD patients (Figure 8A). Indeed, several immune-related pathways had different activities between the two groups. Patients in the high-risk group had higher activities in terms of T cell co−inhibition and check−point (Figure 8B). Drug sensitivity comparison showed most drugs have similar IC_{50} between the two groups, and there were eight drugs that had lower IC_{50} in the high-risk group: PD−173,074, AZD8055, BEZ235, CGP-60474, Dasatinib, Pazopanib, TGX221, and HG-6-64-1 (Figure 8C).

4 Discussion

STAD is a common malignancy worldwide. Although the mortality of STAD has declined due to earlier detection and treatment advancement including targeted therapy, the OS of STAD patients remains low due to delayed diagnoses that makes tumor unresectable. The copper level has been reported to be increased in cancer patients, which could promote tumor angiogenesis, progression and metastasis. Recently Tsvetkov et al. reported cuproptosis, a novel form of regulated cell death (Tsvetkov et al., 2022). Investigation of cuproptosis-related genes in cancer could help understand mechanisms of tumor development. The identification of cuproptosis may also promote innovations in the development new anti-cancer agents.

Biomarkers, including genetic and epigenetic ones, are playing a crucial role in cancer treatment and prognosis (Mishra and Verma, 2010). For example, the TCGA project classify STAD into four major subtypes with different genomic profiles to guide targeted therapy (2014). Non-

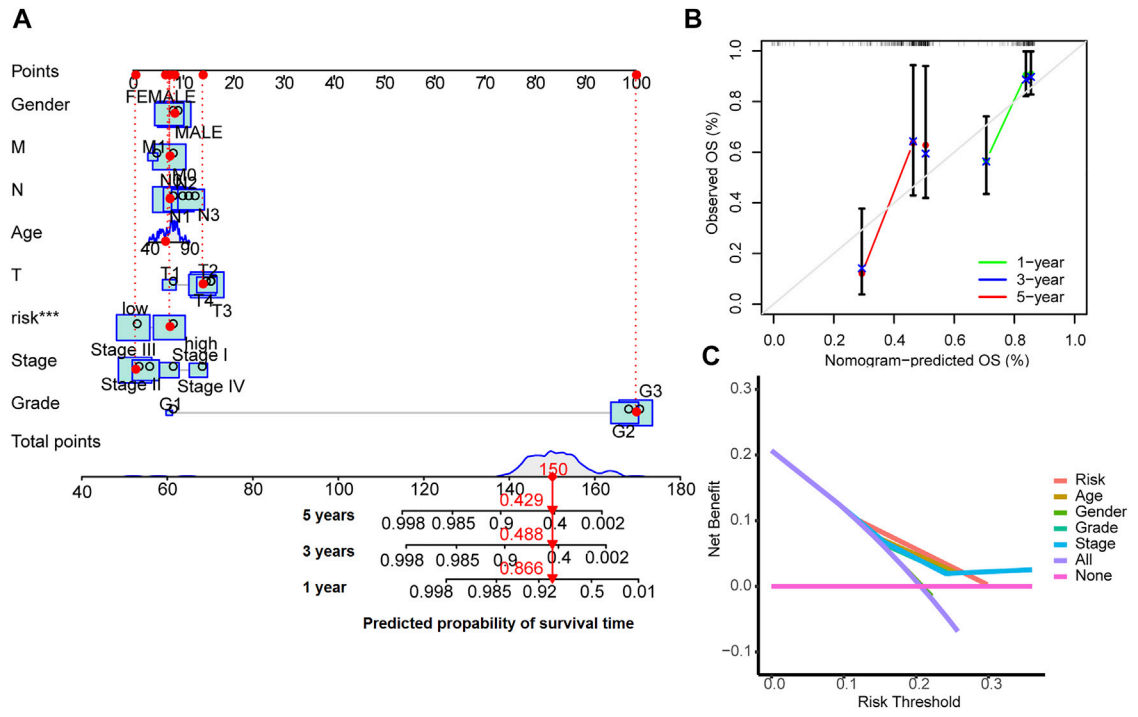


FIGURE 5
Nomogram for survival prediction **(A)**, the calibration curves **(B)** and the decision curves **(C)**.

coding RNA transcripts, such as lncRNAs, can also be used as biomarkers because lncRNAs might regulate cancer development (Djebali et al., 2012; Lee, 2012; Huarte, 2015; Marchese et al., 2017; Mattick, 2018). With the abundant novel lncRNAs identified recently, the annotation of these lncRNAs is urgently needed. We found that nine cuproptosis-related lncRNAs were related to survival of STAD patients. On one hand, LINC01094, AC022182.1, AC011747.1, LINC02476, AC090809.1, AC084781.2, and SENCR were risk factors for STAD patients. On the other hand, AC010422.4 and AC005014.2 were protective factors for STAD patients. The underlying mechanisms for the relationship between STAD prognosis and expression levels of LINC01094, AC022182.1, AC011747.1, LINC02476, AC090809.1, AC084781.2, SENCR, AC010422.4 and AC005014.2 are unknown presently.

Several studies have reported that LINC01094 was associated with diverse tumors. Jiang et al. (2020) found that LINC01094 expression was upregulated in clear cell renal cell carcinoma (ccRCC) in the TCGA database and ccRCC cell lines. LINC01094 knockdown inhibited ccRCC cell growth and metastasis via binding miR-224-5p. Increased expression of LINC01094 was also found in glioma, and was associated with glioma grade. LINC01094 bound to miR-330-3p in glioma (Zhu et al., 2020). In ovarian cancer, LINC01094 expression was elevated and was related to FIGO stage and lymph node metastasis. LINC01094 expression was also a risk factor for ovarian cancer patient survival. In ovarian cancer cells, LINC01094 bound to miR-577 and increased cell proliferation, migration, and the expressions of β-catenin, c-Myc and cyclin D1 (Xu et al., 2020). In colorectal cancer, LINC01094 was also highly expressed and correlated with lymph node metastasis and TNM stage. LINC01094 promoted proliferation, invasion, and migration of colorectal cancer cells by sponging miR-1266-5p (Zhang et al., 2022). Thus LINC01094 is an oncogene in an array of tumors.

SENCR (Smooth Muscle And Endothelial Cell Enriched Migration/Differentiation-Associated LncRNA) is a super enhancer lncRNA originally reported to be overexpressed in smooth muscle cells and endothelial cells. SENCR promoted proliferation, differentiation, and migration of endothelial cells (Bell et al., 2014; Boulberdaa et al., 2016; Sun et al., 2018). Studies have found that SENCR is closely related to the progress of several human cancers. Non-small cell lung cancer (NSCLC) had higher expression of SENCR. Knockdown of SENCR inhibited the growth and metastasis of NSCLC through miR-1-3p. SENCR increased CDK4 and CDK6 expression by binding to miR-1-3p (Cheng et al., 2021). Knockdown of SENCR in cisplatin-resistant

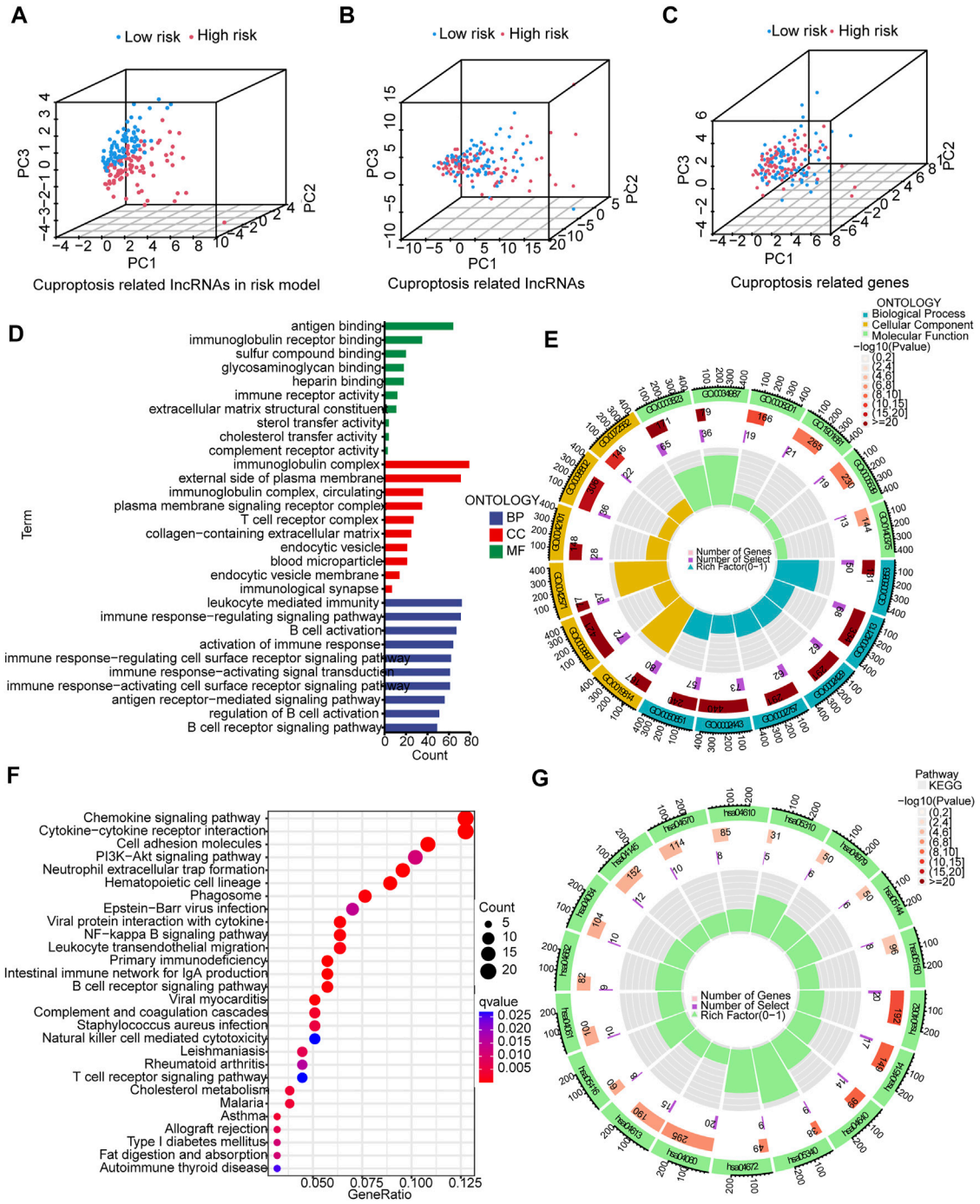


FIGURE 6
PCA, GO, and KEGG analyses. **(A–C)** 3D scatter plots of sample distribution. **(D,E)** GO analysis of biological processes, cellular components and molecular functions. **(F,G)** KEGG analysis of PI3K–Akt signaling pathway, cell adhesion, cytokine–cytokine receptor interaction and chemokine signaling pathway.

A549 cell reduced cell proliferation, accompanied by decreased levels of proteins PCNA, MDMX, and P-gp and increased apoptosis. Overexpressing SENCR could increase FLI1 expression (Shen et al., 2022).

When we compared somatic mutations between the two groups, we found mutations were more frequent in the high-risk group. *TP53* mutations are very common in cancers, ranging from 38% to 50% in a variety of solid tumors and in about 5% of

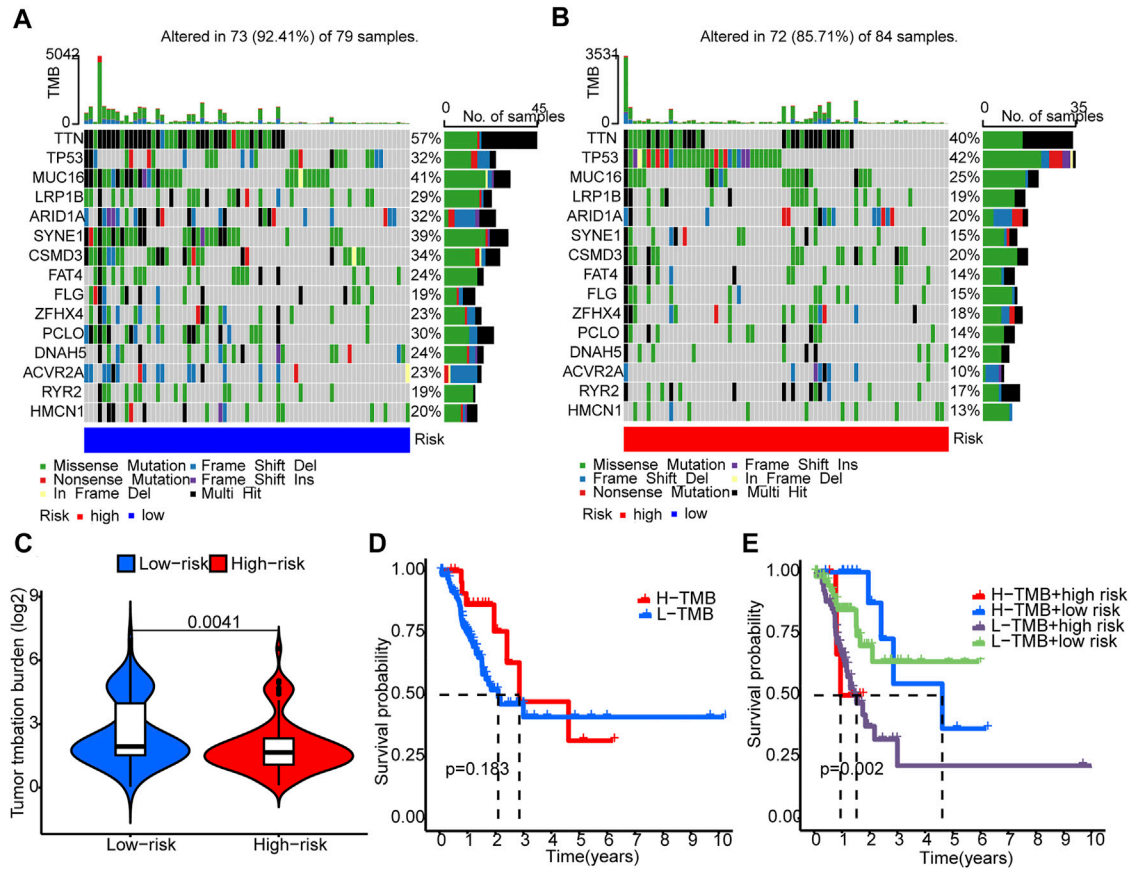


FIGURE 7
Tumor mutation burden (TMB). **(A,B)** The waterfall plots illustrates the frequencies of mutations of genes with different colors representing different types of mutations. **(C)** There were significantly higher TMB in the low-risk group compared to the high-risk group. **(D)** K-M survival curves show similar patient survival between the high- and the low-TMB groups. **(E)** K-M survival curves show different patient survival among the four groups.

primary leukemia. Germline mutations of *TP53* are the underlying cause of Li-Fraumeni syndrome with early-onset cancers. *TP53* (Correa, 2016) mutations may caused by chemical damage induced by particular mutagens, including environmental agents. We propose that higher level of Cu in cancer patients may induce *TP53* mutations, which may related to cuproptosis.

GSEA identified genes of PI3K–Akt signaling pathway might be differentially expressed between the low-risk group and the high-risk group. The PI3K/AKT signaling pathway regulates cell survival and proliferation. Aberrant activation of the pathway is often associated with tumor progression and resistance to cancer therapies (LoRusso, 2016). Thus the relationship between PI3K–Akt signaling pathway and cuproptosis deserves further studies.

We predicted treatment response of the drugs again STAD using pRRophetic (Geeleher et al., 2014) and found that cuproptosis may be related to drug sensitivity. Indeed, Tsvetkov et al. (2022) reported the hydrophilic antioxidant glutathione (GSH) blocked the toxicity of elesclomol (ES)-Cu by chelating intracellular Cu. They also found that NCIH2030 lung cancer cells that rely on galactose-mediated mitochondrial respiration were much more sensitive to ES-Cu-induced growth inhibition than cells that rely on glucose-induced glycolysis. The depletion of GSH by buthionine sulfoximine also increased susceptibility to cuproptosis in A549 lung cancer cells. Thus it is reasonable to expect that drugs involved in galactose regulation pathways may have different effects on cancer cells with different expression of cuproptosis-related genes.

To conclude, we constructed a nomogram exploiting cuproptosis-associated lncRNA expression to predict survival of patients with STAD. Cu is a crucial metal with redox properties. Depending on it's concentration in cells, Cu may be either beneficial or toxic to the cell. Further studies of the roles of Cu in cancer development will lead to more innovative therapies (Ge et al., 2022). The usefulness of

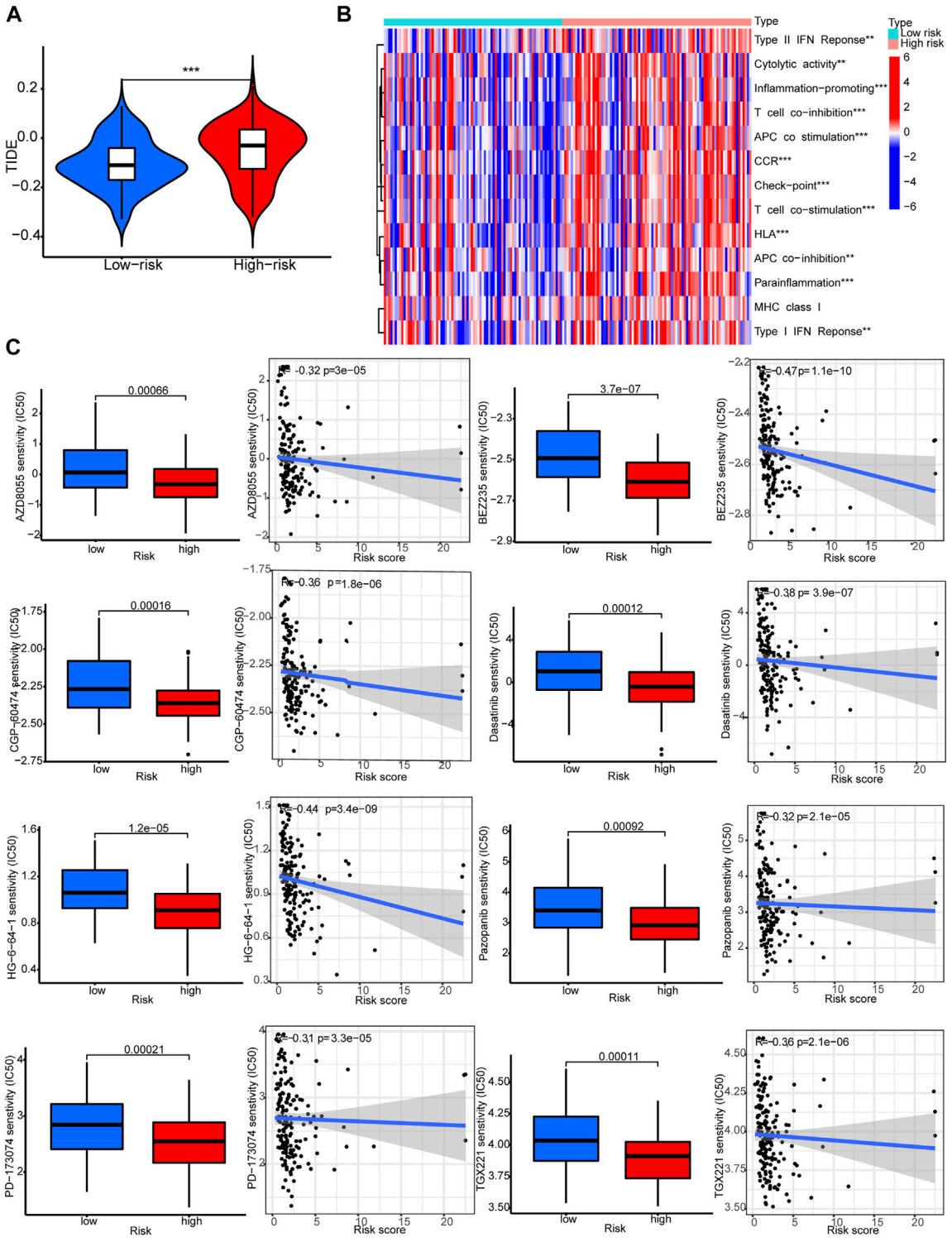


FIGURE 8
Immune functions and prediction of clinical treatment response. **(A)**TIDE scores. **(B)** Immune function heat maps. **(C)** IC_{50} of eight drugs.

this nomogram in predicting patient survival and in treatment decision-making need to be explored in the future studies.

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

WY, JH, HH, and TY contributed to conception and design of the study. WY and JH organized the database. TY and HH performed the statistical analysis. WY wrote the first draft of the manuscript. JH, HH, ZY, RL, and TY wrote sections of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.982888/full#supplementary-material

References

Ajani, J. A., Lee, J., Sano, T., Janjigian, Y. Y., Fan, D., and Song, S. (2017). Gastric adenocarcinoma. *Nat. Rev. Dis. Prim.* 3, 17036. doi:10.1038/nrdp.2017.36

Akshatha, C. R., Bhat, S., Sindhu, R., Shashank, D., Rose Sommano, S., Tapingkae, W., et al. (2021). Current therapeutic options for gastric adenocarcinoma. *Saudi J. Biol. Sci.* 28 (9), 5371–5378. doi:10.1016/j.sjbs.2021.05.060

Andreini, C., Banci, L., Bertini, I., and Rosato, A. (2006). Counting the zinc-proteins encoded in the human genome. *J. Proteome Res.* 5 (1), 196–201. doi:10.1021/pr050361j

Bell, R. D., Long, X., Lin, M., Bergmann, J. H., Nanda, V., Cowan, S. L., et al. (2014). Identification and initial functional characterization of a human vascular cell-enriched long noncoding RNA. *Arterioscler. Thromb. Vasc. Biol.* 34 (6), 1249–1259. doi:10.1161/atvbaha.114.303240

Boulberdaa, M., Scott, E., Ballantyne, M., Garcia, R., Descamps, B., Angelini, G. D., et al. (2016). A role for the long noncoding RNA SENCR in commitment and function of endothelial cells. *Mol. Ther.* 24 (5), 978–990. doi:10.1038/mt.2016.41

Brady, D. C., Crowe, M. S., Turski, M. L., Hobbs, G. A., Yao, X., Chaikuad, A., et al. (2014). Copper is required for oncogenic BRAF signalling and tumorigenesis. *Nature* 509 (7501), 492–496. doi:10.1038/nature13180

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

Chen, A., Li, G., and Liu, Y. (2015). Association between copper levels and myocardial infarction: A meta-analysis. *Inhal. Toxicol.* 27 (5), 237–246. doi:10.3109/08958378.2015.1030480

Cheng, R., Zhang, G., Bai, Y., Zhang, F., and Zhang, G. (2021). LncRNA SENCR promotes cell proliferation and progression in non-small-cell lung cancer cells via sponging miR-1-3p. *Cell Cycle* 20 (14), 1402–1414. doi:10.1080/15384101.2021.1924958

Cobine, P. A., Moore, S. A., and Leary, S. C. (2021). Getting out what you put in: Copper in mitochondria and its impacts on human disease. *Biochim. Biophys. Acta. Mol. Cell Res.* 1868 (1), 118867. doi:10.1016/j.bbamcr.2020.118867

Correa, H. (2016). Li-fraumeni syndrome. *J. Pediatr. Genet.* 5 (2), 84–88. doi:10.1055/s-0036-1579759

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., et al. (2012). Landscape of transcription in human cells. *Nature* 489 (7414), 101–108. doi:10.1038/nature11233

Festa, R. A., and Thiele, D. J. (2011). Copper: An essential metal in biology. *Curr. Biol.* 21 (21), R877–R883. doi:10.1016/j.cub.2011.09.040

Ford, E. S. (2000). Serum copper concentration and coronary heart disease among US adults. *Am. J. Epidemiol.* 151 (12), 1182–1188. doi:10.1093/oxfordjournals.aje.a010168

Ge, E. J., Bush, A. I., Casini, A., Cobine, P. A., Cross, J. R., DeNicola, G. M., et al. (2022). Connecting copper and cancer: From transition metal signalling to metalloplasia. *Nat. Rev. Cancer* 22 (2), 102–113. doi:10.1038/s41568-021-00417-2

Geeleher, 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

Hong, W., Liang, L., Gu, Y., Qi, Z., Qiu, H., Yang, X., et al. (2020). Immune-related lncRNA to construct novel signature and predict the immune landscape of human hepatocellular carcinoma. *Mol. Ther. Nucleic Acids* 22, 937–947. doi:10.1016/j.omtn.2020.10.002

Huarte, M. (2015). The emerging role of lncRNAs in cancer. *Nat. Med.* 21 (11), 1253–1261. doi:10.1038/nm.3981

Ivey, A., Pratt, H., and Boone, B. A. (2022). Molecular pathogenesis and emerging targets of gastric adenocarcinoma. *J. Surg. Oncol.* 125 (7), 1079–1095. doi:10.1002/jso.26874

Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1

Jiang, Y., Zhang, H., Li, W., Yan, Y., Yao, X., and Gu, W. (2020). FOXM1-Activated LINC01094 promotes clear cell renal cell carcinoma development via MicroRNA 224-5p/CHSY1. *Mol. Cell. Biol.* 40 (3), e00357-19. doi:10.1128/mcb.00357-19

Lee, J. T. (2012). Epigenetic regulation by long noncoding RNAs. *Science* 338 (6113), 1435–1439. doi:10.1126/science.1231776

LoRusso, P. M. (2016). Inhibition of the PI3K/AKT/mTOR pathway in solid tumors. *J. Clin. Oncol.* 34 (31), 3803–3815. doi:10.1200/jco.2014.59.0018

Marchese, F. P., Raimondi, I., and Huarte, M. (2017). The multidimensional mechanisms of long noncoding RNA function. *Genome Biol.* 18 (1), 206. doi:10.1186/s13059-017-1348-2

Mattick, J. S. (2018). The state of long non-coding RNA biology. *Noncoding. RNA* 4 (3), E17. doi:10.3390/ncrna4030017

Meng, T., Huang, R., Zeng, Z., Huang, Z., Yin, H., Jiao, C., et al. (2019). Identification of prognostic and metastatic alternative splicing signatures in kidney renal clear cell carcinoma. *Front. Bioeng. Biotechnol.* 7, 270. doi:10.3389/fbioe.2019.00270

Mishra, A., and Verma, M. (2010). Cancer biomarkers: Are we ready for the prime time? *Cancers (Basel)* 2 (1), 190–208. doi:10.3390/cancers2010190

Rawla, P., and Barsouk, A. (2019). Epidemiology of gastric cancer: Global trends, risk factors and prevention. *Prz. Gastroenterol.* 14 (1), 26–38. doi:10.5114/pg.2018.80001

Reunanen, A., Knekt, P., and Aaran, R. K. (1992). Serum ceruloplasmin level and the risk of myocardial infarction and stroke. *Am. J. Epidemiol.* 136 (9), 1082–1090. doi:10.1093/oxfordjournals.aje.a116573

Shen, Q., Zhou, H., Zhang, M., Wu, R., Wang, L., Wang, Y., et al. (2022). Super enhancer-LncRNA SENCR promoted cisplatin resistance and growth of NSCLC through upregulating FLI1. *J. Clin. Lab. Anal.* 36 (6), e24460. doi:10.1002/jcla.24460

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, H., Wang, S., and Song, M. (2018). Long non-coding RNA SENCR alleviates the inhibitory effects of rapamycin on human umbilical vein endothelial cells. *Mol. Med. Rep.* 18 (2), 1405–1414. doi:10.3892/mmr.2018.9094

The Cancer Genome Atlas Research Network (2014). Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 513 (7517), 202–209. doi:10.1038/nature13480

Tsang, T., Posimo, J. M., Gudiel, A. A., Cicchini, M., Feldser, D. M., and Brady, D. C. (2020). Copper is an essential regulator of the autophagic kinases ULK1/2 to drive lung adenocarcinoma. *Nat. Cell Biol.* 22 (4), 412–424. doi:10.1038/s41556-020-0481-4

Tsvetkov, P., Coy, S., Petrova, B., Dreishpoon, M., Verma, A., Abdusamad, M., et al. (2022). Copper induces cell death by targeting lipoylated TCA cycle proteins. *Science* 375 (6586), 1254–1261. doi:10.1126/science.abf0529

Wroblewski, L. E., Peek, R. M., Jr., and Wilson, K. T. (2010). *Helicobacter pylori* and gastric cancer: Factors that modulate disease risk. *Clin. Microbiol. Rev.* 23 (4), 713–739. doi:10.1128/cmr.00011-10

Xu, J., Zhang, P., Sun, H., and Liu, Y. (2020). LINC01094/miR-577 axis regulates the progression of ovarian cancer. *J. Ovarian Res.* 13 (1), 122. doi:10.1186/s13048-020-00721-9

Zhang, G., Gao, Y., Yu, Z., and Su, H. (2022). Upregulated long intergenic non-protein coding RNA 1094 (LINC01094) is linked to poor prognosis and alteration of cell function in colorectal cancer. *Bioengineered* 13 (4), 8526–8537. doi:10.1080/21655979.2022.2051839

Zhu, B., Liu, W., Liu, H., Xu, Q., and Xu, W. (2020). LINC01094 down-regulates miR-330-3p and enhances the expression of MSI1 to promote the progression of glioma. *Cancer Manag. Res.* 12, 6511–6521. doi:10.2147/cmar.S254630

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 09 September 2022
DOI 10.3389/fgene.2022.937069

OPEN ACCESS

EDITED BY
Xuehui Yang,

REVIEWED BY
Xiangrong Cui,
Affiliated of Shanxi Medical University, China
Shixin Zhang,
Third Military Medical University, China

*CORRESPONDENCE
Yi-Qing Qu,
quyiqing@sdu.edu.cn
Xiu-Li Ji,
xiuliji@163.com

†These authors have contributed equally to this work

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

RECEIVED 05 May 2022
ACCEPTED 15 August 2022
PUBLISHED 09 September 2022

CITATION
Hu D-M, Zhang W-D, Shi Z-E, Zhang M-Y, Li R, Wang Q-X, Ji X-L and Qu Y-Q (2022), FOXP family DNA methylation correlates with immune infiltration and prognostic value in NSCLC.
Front. Genet. 13:937069.
doi: 10.3389/fgene.2022.937069

FOXP family DNA methylation correlates with immune infiltration and prognostic value in NSCLC

Dong-Mei Hu[1,2], Wen-Di Zhang[1,2], Zhuang-E Shi[1,2], Meng-Yu Zhang[1,2], Rui Li[1,2], Qing-Xiang Wang[1,2], Xiu-Li Ji[3*†] and Yi-Qing Qu[1*†]

[1]Shandong Key Laboratory of Infectious Respiratory Diseases, Department of Pulmonary and Critical Care Medicine, Qilu Hospital of Shandong University, Jinan, China, [2]Laboratory of Basic Medical Sciences, Qilu Hospital of Shandong University, Jinan, China, [3]Department of Pulmonary Disease, Jinan Traditional Chinese Medicine Hospital, Jinan, China

Background: Forkhead box P (FOXP) family was introduced as a double-edged sword in tumorigenesis and influenced immunotherapy response by modulating host immunity. This study aimed to summarize the involvement of the FOXP family in non-small cell lung cancer (NSCLC).

Methods: The UALCAN, Gene Expression Profiling Interactive Analysis (GEPIA), and Reverse transcription-quantitative polymerase chain reaction (RT–qPCR) were used to analyse the expression levels of the FOXP family in NSCLC. The prognostic impact was evaluated using Kaplan-Meier Plotter. MethSurv, UALCAN, and cBioPortal were applied to analyse the DNA methylation and mutation status of the FOXP family respectively. COEXPEDIA, STRING, and GeneMANIA were used to explore the interaction mechanism. Finally, TISIDB was used to investigate all of the immune-related characteristics regulated by the FOXP family.

Results: The expression levels of FOXP1/3/4 were dysregulated in NSCLC tissues than that in normal tissues. Groups with low expression levels of FOXP1/4 and high expression levels of FOXP2/3 were associated with poor prognosis in NSCLC. The transcriptional levels of FOXP2/3/4 were correlated with DNA methylation in NSCLC. FOXP1/3/4 DNA methylation were correlated with prognosis. Pathway enrichment analysis indicated the FOXP family was

Abbreviations: FOXP, Forkhead box P; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; NSCLC, non-small cell lung cancer; GEPIA, Gene Expression Profiling Interactive Analysis; GEO, Gene Expression Omnibus; TCGA, The Cancer Genome Atlas; DEG, differentially expressed genes; GTEx, Genotype-Tissue Expression; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; RT-qPCR, Reverse transcription-quantitative polymerase chain reaction; EMT, epithelial-mesenchymal transition; GO, Gene Ontology; CC, cellular component; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes; MHC, Major Histocompatibility Complex; OS, The Overall Survival; ICIs, immune checkpoint inhibitors; PD-1, programmed cell death receptor 1; PD-L1, programmed cell death 1 ligand 1; CTLA-4, cytotoxic T-lymphocyte-associated protein 4; PPI, protein-protein interaction; LLS, log-likelihood scores; TILs, tumour-infiltrating lymphocytes; TKI, tyrosine kinase inhibitor; HR, hazard ratio.

mainly related to immune-related pathways. After DNA methylation, the correlations between FOXP family and immune factors were opposite to that before alteration in NSCLC.

Conclusion: This study elucidated FOXP family could serve as vital diagnostic and prognostic biomarkers in NSCLC. Our study highlighted novel potential functions of FOXP family DNA methylation in regulation of immune-related signatures in NSCLC.

KEYWORDS

non-small cell lung cancer, FOXP family, prognostic value, immune infiltration, DNA methylation

Introduction

Lung cancer accounts for a large proportion of malignant tumours in the world, of which non-small cell lung cancer (NSCLC) accounts for approximately 85% (Hirsch et al., 2017). According to diverse histological subtypes, NSCLC can be divided into lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) (Ruiz-Cordero and Devine, 2020). At present, surgery, cisplatin-based therapy, stereotactic body radiation therapy, definitive concurrent chemotherapy, and radiation therapy have significantly reduced the risk of death in NSCLC. However, these treatments are only suitable for a very small proportion of NSCLC patients. Meanwhile, according to data from recent years, the long-term survival rate of NSCLC patients is still very poor (Evison and AstraZeneca, 2020). Most recently, immune checkpoint inhibitors (ICIs), including inhibitors of the programmed cell death receptor 1 (PD-1) axis, have apparently altered the NSCLC management landscape (Camidge et al., 2019). However, effective biomarkers for guiding NSCLC patients to use ICI drugs are still lacking (Zhang et al., 2021). Therefore, investigating the molecular mechanisms that drive NSCLC initiation and progression, searching for more sensitive biomarkers, and identifying biomarkers for ICI efficacy are the current research hotspots. The forkhead box P (FOXP) family consists of four members, including FOXP1, FOXP2, FOXP3, and FOXP4 (Kim et al., 2019). The FOXP family is responsible for the occurrence of many tumours. For example, FOXP1 is related to the occurrence of drug resistance in patients with ovarian cancer during treatment (Hu et al., 2020). FOXP1 also has a function in the occurrence of cancer cachexia that causes weakness (Neyroud et al., 2021). FOXP2 promotes tumour progression in triple-negative breast cancer through the mechanisms of targeting specific molecules (Wu et al., 2018). FOXP3 is involved in the regulation of autophagy-related proteins in gastric cancer (Li et al., 2020a). Overexpression of FOXP4 is closely implicated in the malignant prognosis of breast cancer by promoting the biological process of EMT (Ma and Zhang, 2019). Therefore, we know that the FOXP family plays a role in tumour suppressor genes and oncogenes in tumours (Kim et al., 2019). However, the roles of the FOXP family in the effect and mechanism of immune infiltration have not yet been determined. In this article, we comprehensively analysed FOXP family mRNA expression/DNA methylation signatures, mutations, functional pathways of coexpression networks, survival value, epigenetic alterations, and relationships with immune-related factors. Furthermore, we performed real-time quantitative PCR (RT–qPCR) to detect the expression levels of the FOXP family.

Materials and methods

UALCAN

UALCAN (http://ualcan.path.uab.edu/), an online website was used to compare the difference in the mRNA expression levels of the FOXP family between NSCLC tissues and normal tissues obtained from The Cancer Genome Atlas (TCGA) (Chandrashekar et al., 2017). Then, we explored the changes in FOXP family expression levels in different pathological stages with this tools. In addition, we used UALCAN to analyse the effect of DNA methylation on the translational levels of the FOXP family.

Gene expression profiling interactive analysis

GEPIA (http://gepia.cancer-pku.cn) was used to analyse the mRNA levels of the FOXP family in NSCLC tissues compared to normal tissues using the open public data from TCGA (Tang et al., 2017). Under the condition of selecting the corresponding cancer species, the website can automatically output the corresponding scatter diagrams, bar charts, and box plots according to the input gene name.

Kaplan–Meier plotter

Kaplan–Meier plotter (http://kmplot.com/analysis/) provided data and algorithms for analysing the prognostic

significance of patients with expression imbalances of the FOXP family (Peng et al., 2017). All the patients were divided into two groups according to the median expression levels of FOXP family genes to measure the difference in survival time between the above two groups. Kaplan–Meier curves were plotted to explore the overall survival (OS) analysis by the log-rank test. p values < 0.05 were defined as statistically significant.

MethSurv database

The MethSurv database (https://biit.cs.ut.ee/methsurv/) was used to perform survival analysis of DNA methylation of the FOXP family in NSCLC by selecting a specific gene name and cancer type using the TCGA dataset. The "Region-based analysis" module was used by choosing "LUAD TCGA March 2017" and "LUSC TCGA March 2017".

R/Bioconductor package

We visualized the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis results of coexpressed genes of the FOXP family using R/Bioconductor packages ("BiocManager," "DOSE," "cluster Profiler," "org.Hs.eg.db," "enrichplot" and "ggplot2") which were downloaded from Bioconductor (http://www.bioconductor.org/packages/release/bioc/html/). The enrichment analysis results with a p value < 0.05 were demonstrated to have great significance.

TISIDB

TISIDB (http://cis.hku.hk/TISIDB/index.php) was applied to infer the relative abundance of immune-related characteristics of 28 tumour-infiltrating lymphocyte (TIL) types, immunomodulators, chemokines, and receptors regulated by the FOXP family in NSCLC tissues. On the foundation of the mRNA expression of the FOXP family profiles, gene set variation analysis (GSVA) examined which types of immune-related characteristics were regulated by the current genes with epigenetic alterations (copy number alteration and DNA methylation). In addition, TISIDB provided data on the degree of infiltration of immune-related characteristics in NSCLC tissues to infer the regulatory effect of the FOXP family. Finally, TISIDB was applied to explore the expression of the FOXP family in different immune subtypes (Ru et al., 2019).

Cancer single-cell state atlas

Cancer single-cell state atlas (CancerSEA) (http://biocc.hrbmu.edu.cn/CancerSEA/) provided datasets that was applied to assess the functional roles of the FOXP family in NSCLC. The CancerSEA supported the evaluation of 14 functional states at the single-cell level using public datasets including epithelial-mesenchymal transition (EMT), DNA damage, and so on.

COEXPEDIA

COEXPEDIA (https://www.coexpedia.org) is an online database. The corresponding predicted target genes were obtained according to the coexpression trend of consistency and the common pathways involved in the regulation of the occurrence and development of disease. COEXPEDIA offered a network reflecting clear interactions between the members of the FOXP family and the corresponding coexpressed genes.

cBioPortal

cBioPortal (https://www.cbioportal.org/) was used to ascertain the consequence of alteration frequency and mutation type of the FOXP family in NSCLC (Gao et al., 2013). cBioPortal precisely presented the details of all forms of mRNA dysregulation, gene amplification, and deep deletion with the FOXP family in NSCLC patients by the OncoPrint module.

STRING

STRING (https://string-db.org) was used to construct a protein–protein interaction (PPI) network for the retrieval of interacting genes (Szklarczyk et al., 2017). In this article, STRING was used to examine the interactions among the FOXP family and determine the hub regulatory genes. The genes not only required a minimum interaction score ≥ 0.4, but were also imported into Cytoscape (version 3.7.2) with the cytoHubba app to screen the modules of the top 10 hub genes.

GeneMANIA

GeneMANIA (http://www.genemania.org) administers data on protein and genetic interactions, pathways, and coexpression to predict gene clusters with similar functions (Warde-Farley et al., 2010). This site relies on credible evidence sources of literature to forecast functionally identical genes of the FOXP family to clarify the interaction mechanism of the FOXP family.

Cell lines and culture conditions and reverse transcription-quantitative polymerase chain reaction

A human lung epithelial cell line (BEAS-2B Cell Article: No. CL-0496), LUAD cell lines (A549 Cell Article: No. CL-0016, NCI-H1299 Cell Article: No. CL-0165, and PC9 Cell Article: No. CL-0298), and LUSC cell line (NCI-H226 Cell Article: SNL-388) were purchased from Procell Life Science & Technology Co. Ltd. (Wuhan, China) on 10 December 2021. All cell lines were identified by short tandem repeat (STR) analysis. The human lung epithelial cell line BEAS-2B and the LUAD cell line PC9 were cultured in Dulbecco's modified Eagle's medium (DMEM, Gibco). NSCLC cell lines (A549, NCI-H1299, and NCI-H226) were cultured in RPMI 1640. The two types of culture media both contain 10% heat-inactivated foetal bovine serum (FBS). The gas concentration in the incubator was set to 5% CO_2, and the temperature was set to 37°C. The method of evaluating the gene expression was RT–qPCR. TRIzol reagent (Invitrogen) was applied to extract total RNA. After the concentration of extracted RNA reached the appropriate standard, we used miRNA reverse transcription and complementary DNA (cDNA) reverse transcription kits to carry out reverse transcription. Then, RT–qPCR was performed on a Bio-Rad after the corresponding steps were executed according to the manufacturer's instructions for TB Green Premix Ex Taq II (Takara). Finally, we used the $2^{-\Delta\Delta Ct}$ method to calculate relative mRNA expression. The reference gene was glyceraldehyde 3-phosphate dehydrogenase (GAPDH), and the sequences of the primers for the GAPDH and FOXP family are listed in Supplementary Table S1.

Statistical analysis

The statistical data were analysed using GraphPad Prism 9.3.1 by Student's test and ordinary one-way ANOVA to evaluate the differential expression. The statistical data are presented as the mean ± SEM. Kaplan–Meier Plotter was used to explore the overall survival (OS) analysis by the log-rank test. The prognostic values of single CpGs in DNA methylation analysis were assessed *via* the likelihood-ratio test. p values < 0.05 obtained from all the above analyses were defined as statistically significant.

Results

The mRNA expression levels of the forkhead box P family in NSCLC

A flowchart was created to illustrate our study (Figure 1A). UALCAN was used to compare the difference in the mRNA expression levels of the FOXP family between normal samples and NSCLC samples. The summary of the transcriptional levels of the FOXP family is shown in the form of heatmaps (Figures 2A,B). Moreover, the GEPIA database was applied to verify the expression of the FOXP family between NSCLC tissues and normal tissues (Figures 2C–J). Compared to normal tissues, there were lower expression levels of FOXP1 in LUAD and LUSC, a lower expression level of FOXP2 in LUAD, higher expression levels of FOXP3 in LUAD and LUSC, a higher expression level of FOXP4 in LUAD, and a lower expression level of FOXP4 in LUSC. In addition, the expression level of FOXP2 was not significantly different in LUSC. We examined the mRNA expression levels of the FOXP family in cell lines (BEAS-2B, A549, NCI-H1299, PC9, and NCI-H226) (Figures 2K–N). The outcomes of RT-qPCR showed that the mRNA expression levels of FOXP1, FOXP3, FOXP4 did have statistical differences between LUSC cell line (NCI-H226) and normal human lung epithelial cell line (BEAS-2B). However, when tested individually to verify the differential expression levels of FOXP family between LUAD cell lines and normal control, we found that only two members (FOXP1 and FOXP3) were statistically significant between LUAD cell lines (A549, PC9, and NCI-H1299) and normal human lung epithelial cell line (BEAS-2B), which were consistent with analysis of GEPIA database. In order to find the source of this difference, we conducted meta-analysis to explore the expression difference of FOXP family from different database using LUNG CANCER EXPLORER (https://lce.biohpc.swmed.edu/lungcancer/index.php#page-top) database. The results showed that the different expression trends of FOXP2 and FOXP4 objectively existed in LUAD among different data sets. After meta-analysis, it was more likely that the expression of FOXP2 was no statistically significant, and the expression of FOXP4 was upregulated in LUAD patients compared with normal controls (Supplementary Figure S1). Therefore, we concluded that FOXP1 was downregulated, and FOXP3 was upregulated between LUAD patients compared with normal controls, while the expression levels of FOXP2 and FOXP4 in LUAD compared with normal controls need to be verified by more clinical samples. Besides, FOXP1 and FOXP4 were downregulated, FOXP3 was upregulated, and FOXP2 was not statistically significant between LUSC patients with normal controls.

Relationship between the transcriptional levels of the forkhead box P family and clinicopathological stages in non-small cell lung cancer

Next, the inconsistency of the transcriptional expression levels of the FOXP family members among the clinicopathological parameters of NSCLC patients was analysed by UALCAN (Supplementary Figures S2–S5). The clinicopathological parameters included histological subtypes,

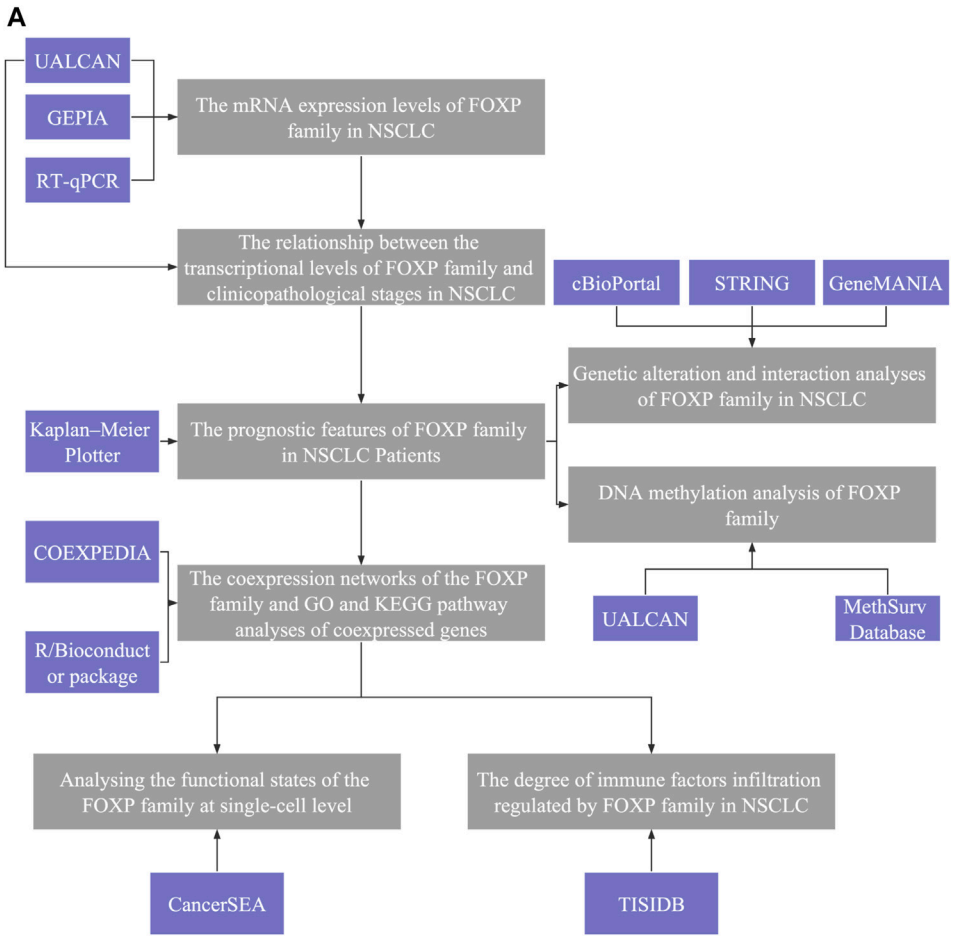


FIGURE 1
Analysis explanation with a detailed flow diagram of this study. **(A)** The study comprised eight parts: I The mRNA expression levels of FOXP family in NSCLC; II The relationship between the transcriptional levels of FOXP family and clinicopathological stages in NSCLC; III The prognostic features of FOXP family in NSCLC patients; IV The coexpression networks of the FOXP family and GO and KEGG pathway analyses of coexpressed genes; V Analysing the functional states of the FOXP family at single-cell level; VI The degree of immune factors infiltration regulated by FOXP family in NSCLC; VII Genetic alteration and interaction analyses of FOXP family in NSCLC; VIII DNA methylation analysis of FOXP family.

individual cancer stages, patient age, patient smoking habits, nodal metastasis status, and TP53 mutation status. As shown in the histograms in Figures 3A,B, the transcriptional levels of FOXP1/3/4 were basically markedly correlated with the above six clinicopathological stages in LUAD. However, there was no discernible difference in the relationship between the transcriptional level of FOXP2 and the six clinicopathological stages in LUAD (Figure 3A). The transcriptional levels of FOXP1/2/3 were markedly correlated with the above six clinicopathological stages in LUSC, while the difference in FOXP4 was unremarkable (Figure 3B). In brief, the above results preliminarily suggested that the FOXP family was involved in characteristics that included age factors, inducements, progression, metastasis, and mutation types in NSCLC patients.

Prognostic features of the forkhead box P family in non-small cell lung cancer patients

In this step, Kaplan–Meier Plotter was used to explore the prognostic value of the FOXP family in NSCLC. Survival curves were generated to present the association between the overall survival (OS) rate of NSCLC patients and the corresponding gene expression levels of the FOXP family. All results are shown in Figures 4A–H; Supplementary Figure S6. Upon stratification

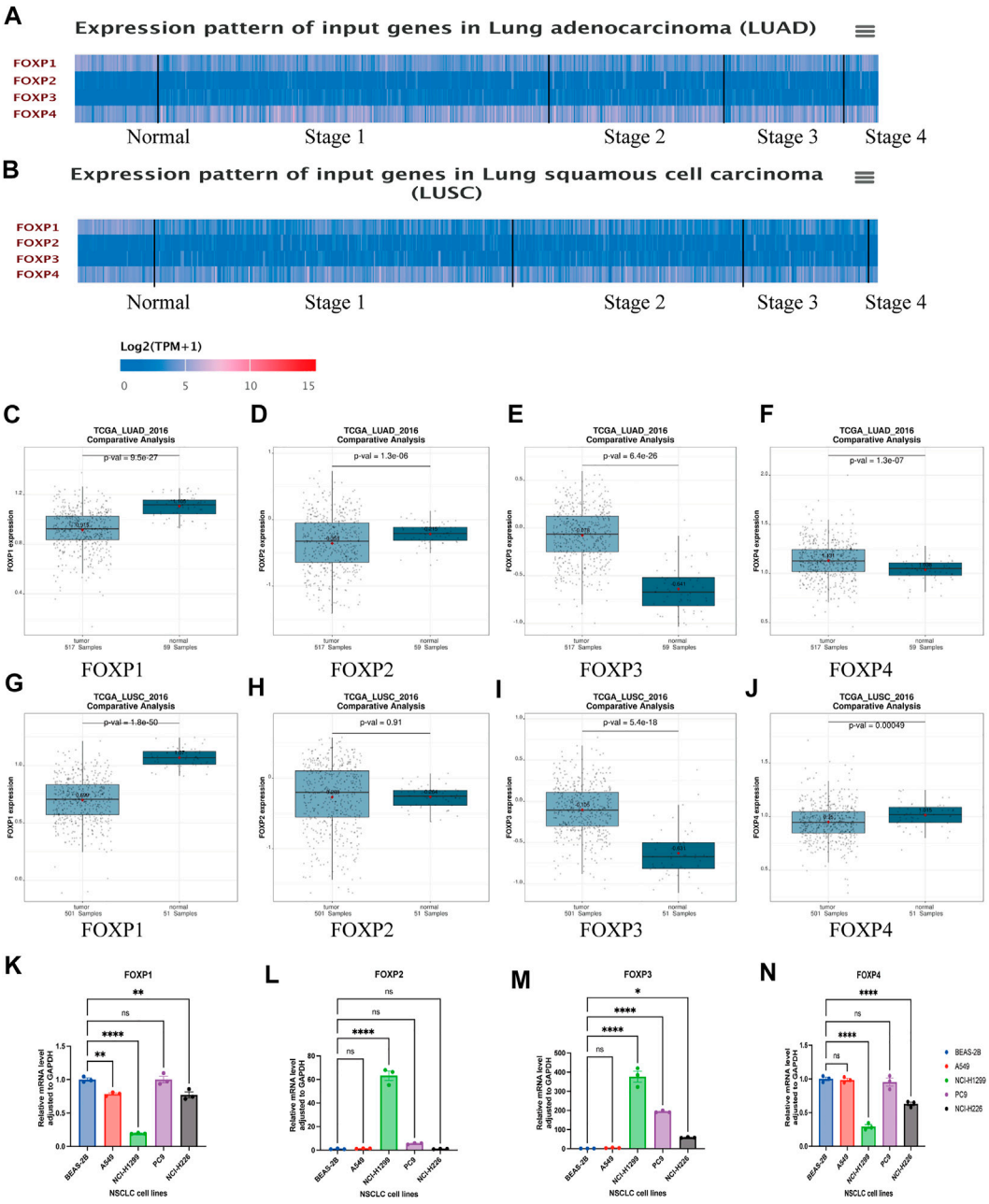


FIGURE 2
The differential expression of FOXP family in NSCLC. **(A)** The heatmap represented the transcriptional levels of FOXP family in patients with LUAD compared with normal samples using UALCAN. **(B)** The heatmap represented the transcriptional levels of FOXP family in patients with LUSC compared with normal samples using UALCAN. **(C–J)** The compare the mRNA expression of FOXP1/2/3/4 between LUAD/LUSC and normal tissue samples by using GEPIA dataset; The box plot showed the relative expression levels of family in normal tissue and NSCLC tissue. $p < 0.05$ was defined as statistically significant. **(K–N)** The mRNA levels of FOXP family between LUAD cell lines (A549, NCI-H1299, and PC9)/LUSC cell line (NCI-H226) and normal human lung epithelial cell line (BEAS-2B) by RT-qPCR. (Legend: ***$p \leq 0.001$; **$p \leq 0.01$; *$p \leq 0.05$; ns. $p > 0.05$; LUAD, Lung adenocarcinoma; LUSC, Squamous cell carcinoma of lung; FOXP, Forkhead box P; RT-qPCR, Reverse transcription-quantitative polymerase chain reaction).

according to the median expression level, higher FOXP1 expression was correlated with better prognosis of LUAD (Figure 4A, n = 336, hazard ratio (HR) = 0.66, 95% CI 0.52–0.84, log-rank p = 0.00075). Higher FOXP1 expression was correlated with better prognosis of NSCLC (Figure 4B, n = 572, HR = 0.69, 95% CI 0.58–0.81, log-rank p = 9e-06). Lower

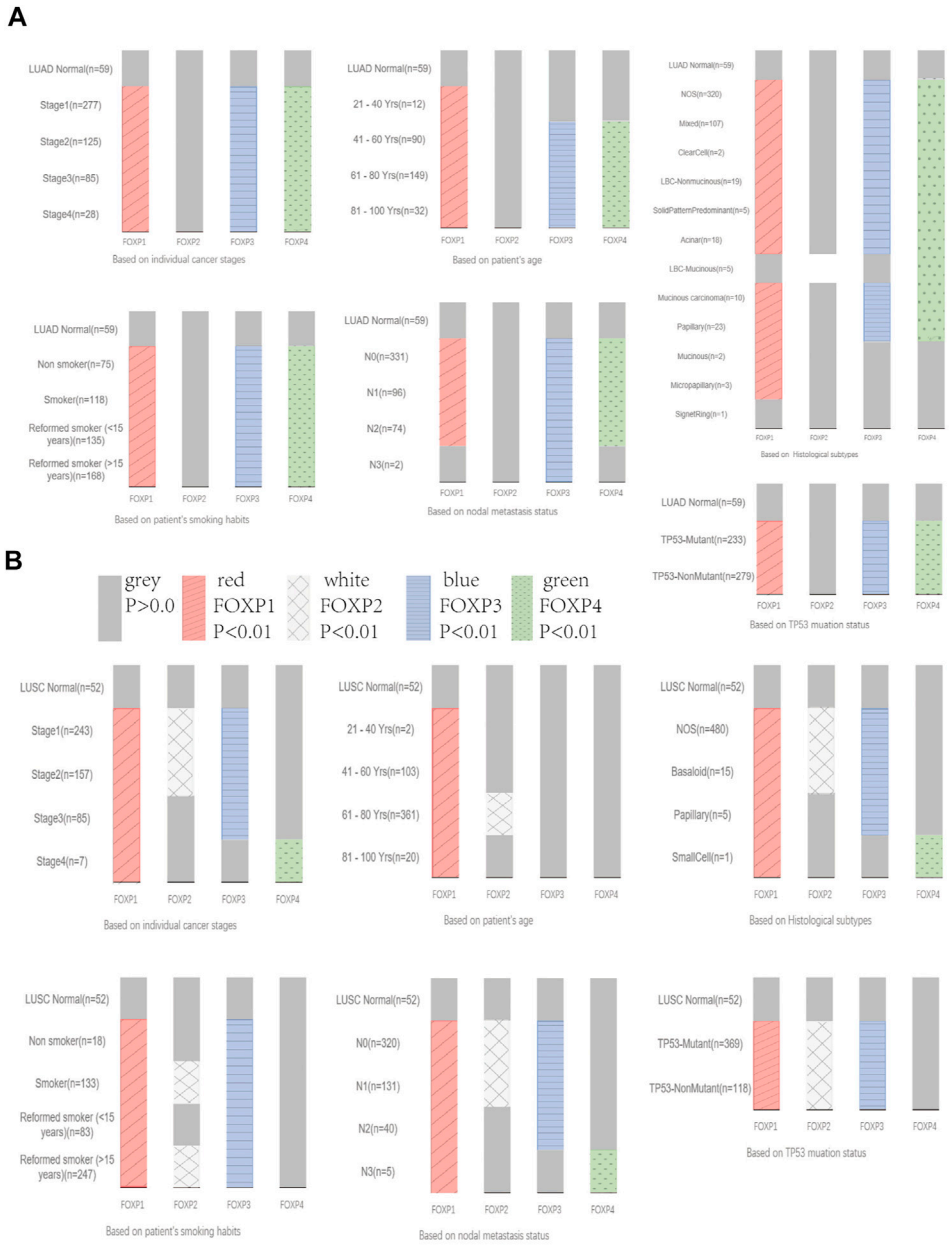


FIGURE 3
The relationship between the expression levels of FOXP family and clinicopathological stages. **(A)** The Bar graphs showing the expression of FOXP family differences between the clinicopathological stages of LUAD and normal tissues. **(B)** The Bar graphs showing the expression of FOXP family differences between the clinicopathological stages of LUSC and normal tissues.

FOXP2 expression was correlated with better prognosis of LUAD (Figure 4C, n = 348, HR = 1.31, 95% CI 1.03–1.67, log-rank p = 0.027). Lower FOXP2 expression was correlated with better prognosis of NSCLC (Figure 4D, n = 596, HR = 1.38, 95% CI 1.17–1.63, log-rank p = 0.00012). Lower FOXP3 expression was correlated with better prognosis of LUAD (Figure 4E, n = 372, HR = 1.37, 95% CI 1.09–1.73, log-rank p = 0.0072). Lower FOXP3 expression was correlated with better prognosis of NSCLC (Figure 4F, n = 984, HR = 1.25, 95% CI 1.1–1.41, log-rank p = 0.00065). Higher FOXP4 expression was correlated with better prognosis of LUAD (Figure 4G, n = 336, HR = 0.71, 95% CI 0.56–0.9, log-rank p = 0.0053). Higher FOXP4 expression was correlated with better prognosis of NSCLC (Figure 4H, n = 569, HR = 0.77, 95% CI 0.65–0.91, log-rank p = 0.0017). Groups with FOXP1/2/3/4 expression were not associated with

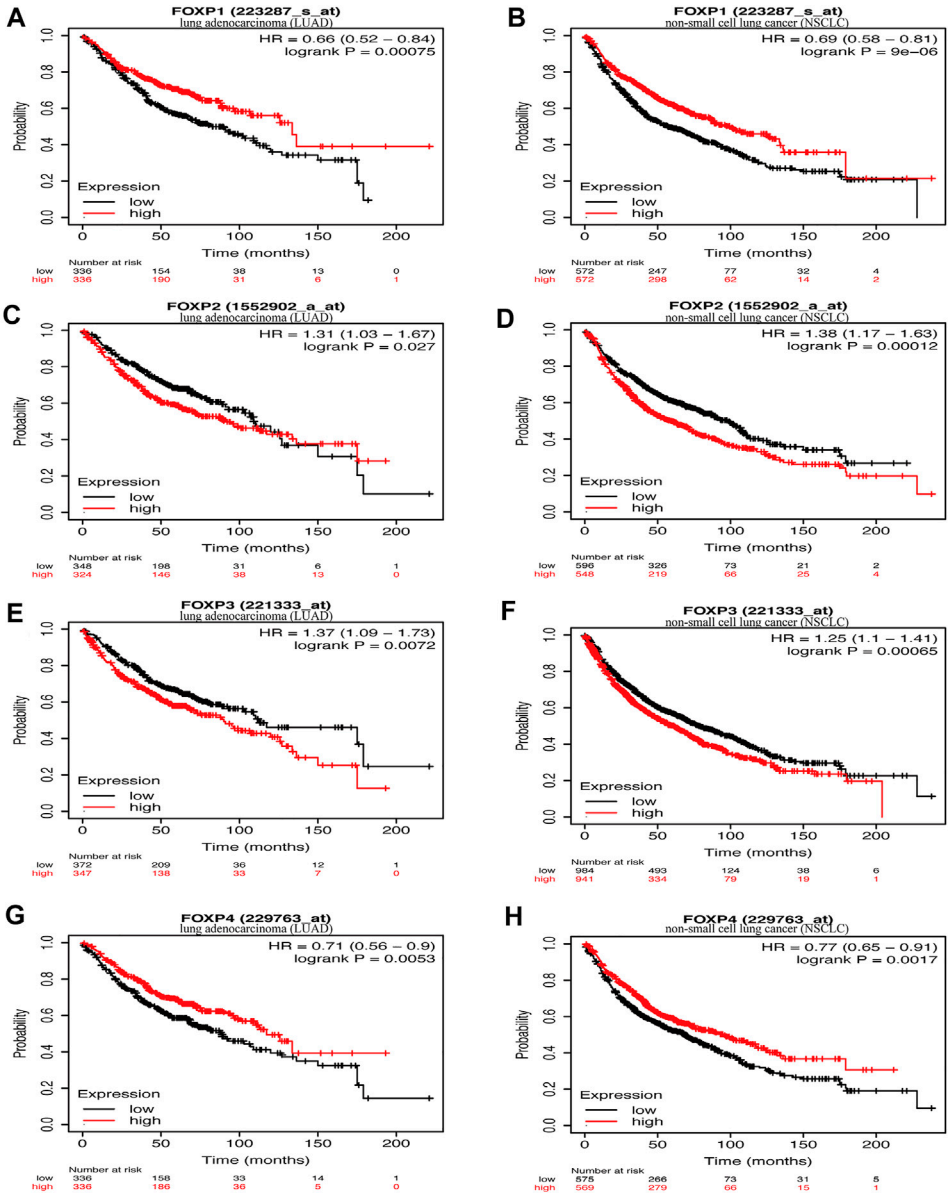


FIGURE 4
The relationship between the expression of FOXP family and survival. **(A)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP1 in LUAD. **(B)** he survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP1 in NSCLC. **(C)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP2 in LUAD. **(D)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP2 in NSCLC. **(E)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP3 in LUAD. **(F)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP3 in NSCLC. **(G)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP4 in LUAD. **(H)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP3 in NSCLC.

prognosis in LUSC patients (Supplementary Figure S6). Overall, groups with low FOXP1/4 and high FOXP2/3 expression were associated with poor prognosis (p value < 0.005). Both the high mRNA expression of FOXP1/4 and the low mRNA expression of FOXP2/3 were related to improved prognosis (p value < 0.05) in NSCLC patients.

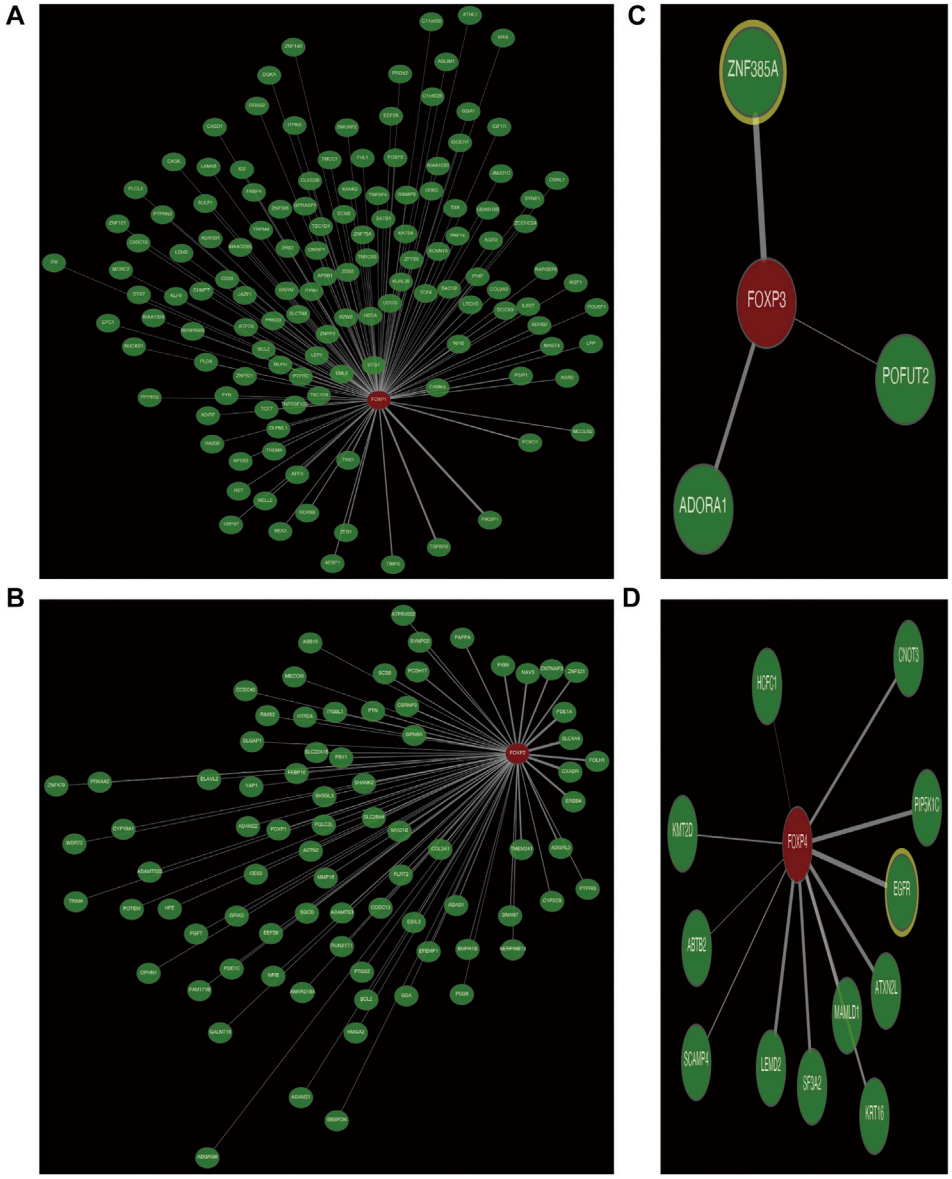


FIGURE 5
The coexpression network of FOXP family. **(A)** The coexpression network presented the coexpressed genes of FOXP1. **(B)** The coexpression network presented the coexpressed genes of FOXP2. **(C)** The coexpression network presented the coexpressed genes of FOXP3. **(D)** The coexpression network presented the coexpressed genes of FOXP4.

Coexpression networks of the forkhead box P family and gene ontology and kyoto encyclopedia of genes and genomes pathway analyses of coexpressed genes

Genes coexpressed with the FOXP family were investigated by the COEXPEDIA website. The coexpression networks of the FOXP family are displayed in Figures 5A–D. The log-likelihood score (LLS score) was used to evaluate the correlations between the FOXP family and its linked genes. The larger the LLS score, the more relevant the coexpression trend of the FOXP family member and its linked genes. The LLS scores of all coexpressed genes are summarized in Supplementary Table S2. GO and KEGG enrichment analyses for coexpressed genes related to the FOXP family were implemented to analyse biological functions and pathways associated with the FOXP family. The biological process (BP), molecular function (MF), and cellular

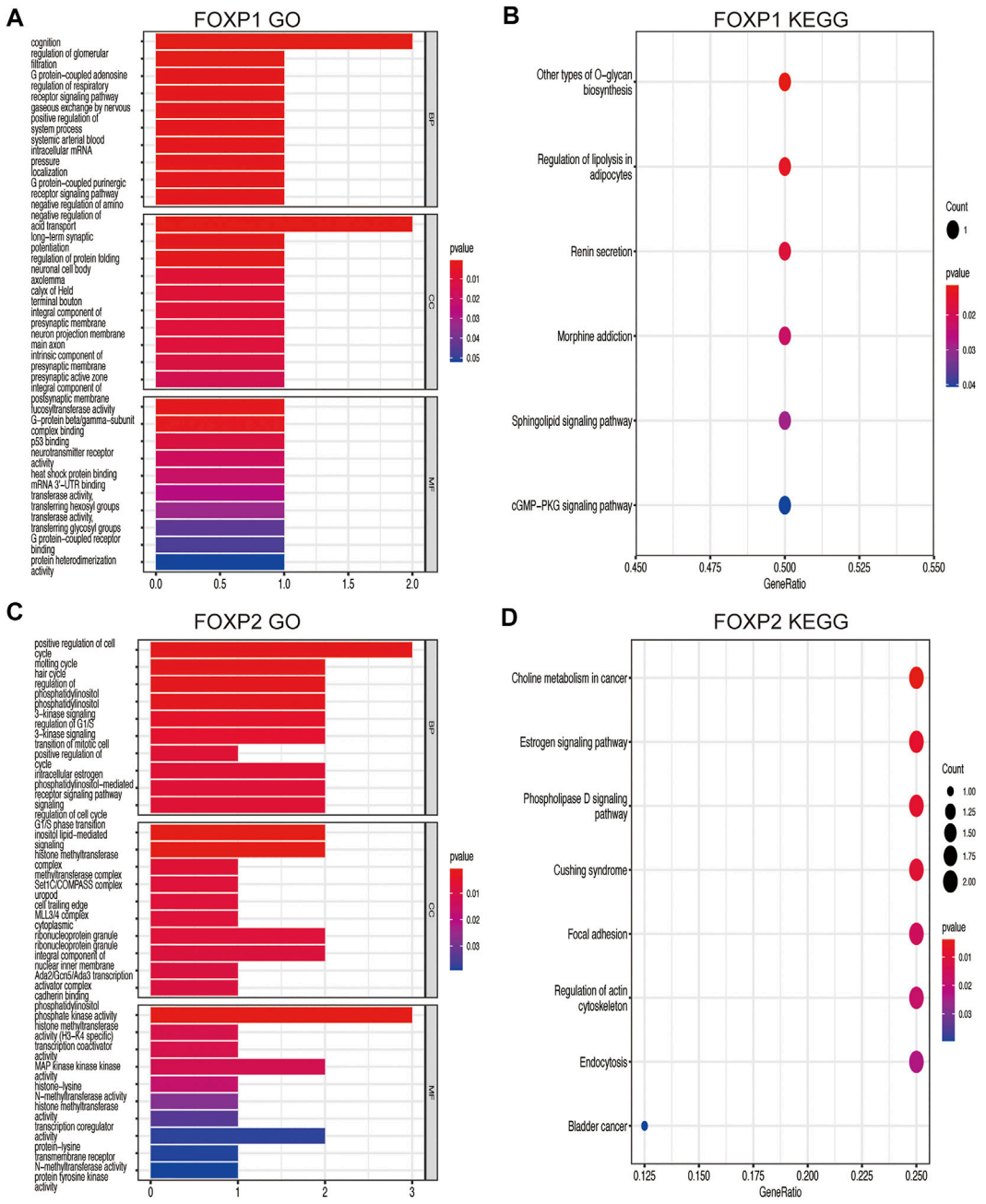


FIGURE 6
GO functional and KEGG pathway enrichment analyses were performed on the coexpressed genes. **(A)** The GO functional enrichment analysis result on the coexpressed genes of FOXP1 using three annotation systems (BP; CC; MF). **(B)** The KEGG pathway enrichment analysis result on the coexpressed genes of FOXP1. **(C)** The GO functional enrichment analysis result on the coexpressed genes of FOXP2. **(D)** The KEGG pathway enrichment analysis result on the coexpressed genes of FOXP2. (BP, Biological process; MF, Molecular function; CC, Cellular component).

component (CC) of GO enrichment analysis are displayed in Figures 6A,C, 7A,C. In addition, the 20 most relevant KEGG pathways for coexpressed genes are presented in Figures 6B,D, 7B,D. Notably, GO enrichment results showed that the coexpressed genes of the FOXP family mainly acted on the immune process in MF, such as differentiation of immune cells

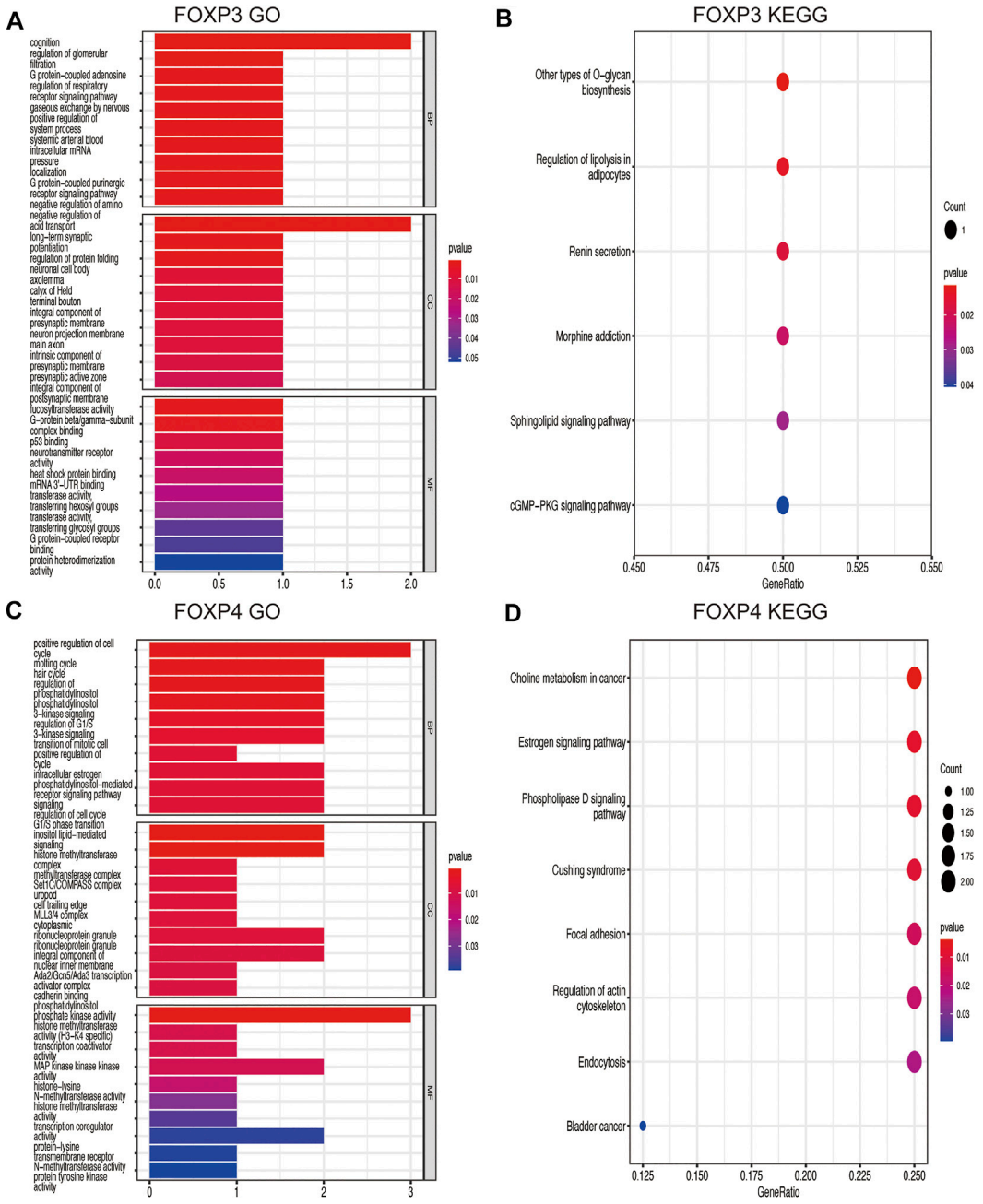


FIGURE 7
GO functional enrichment analysis and KEGG pathway enrichment analysis are performed on the coexpressed genes. **(A)** The result of GO functional enrichment analysis on the coexpressed genes of FOXP3. **(B)** The result of KEGG pathway enrichment analysis on the coexpressed genes of FOXP3. **(C)** The result of GO functional enrichment analysis on the coexpressed genes of FOXP4. **(D)** The result of KEGG pathway enrichment analysis on the coexpressed genes of FOXP4.

(lymphoid, monocyte, and T cell), fucosyltransferase activity, phosphatidylinositol phosphate kinase activity, transcription coactivator activity, and transcription costimulatory factor regulation. KEGG pathway analysis results showed that coexpressed genes clusters of the FOXP family acted on typical cancer- and immune-related signalling pathways including the T cell receptor, sphingolipid, cGMP-PKG,and phospholipase D signalling pathway. These results strongly implied that the FOXP family was involved in the process of immune regulation in NSCLC.

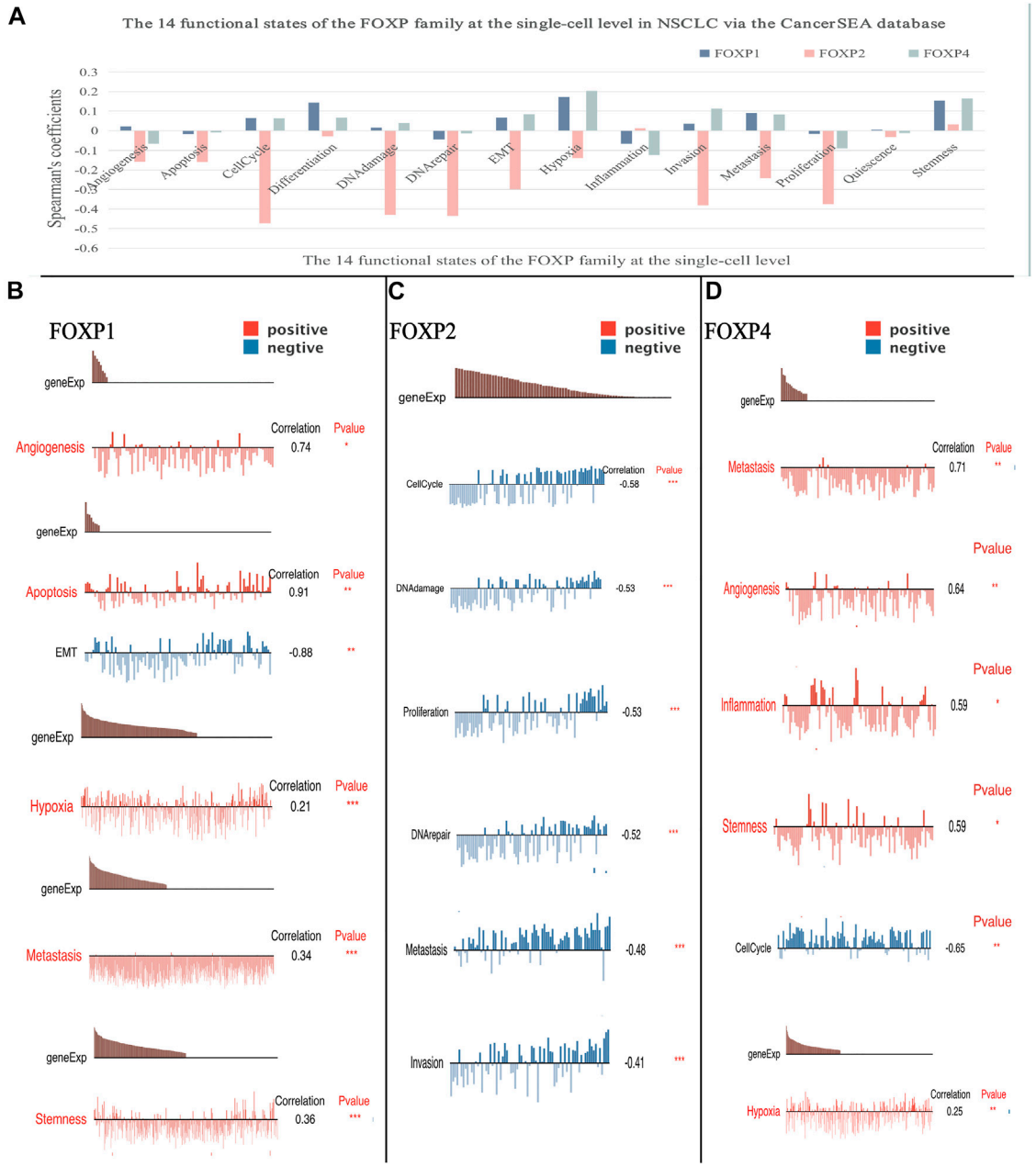


FIGURE 8
The correlation between the FOXP family and 14 functional states at single-cell level. **(A)** The result of the correlation between expression of the FOXP1/2/4 with functional states (including angiogenesis, apoptosis, invasion, EMT, differentiation, proliferation, DNA damage, metastasis, hypoxia, inflammation, cell cycle, DNA repair, stemness, and quiescence). **(B)** The sample Kim (Exp0068) showed the result of the correlation between the FOXP1 with functional states. **(C)** The sample Kim (Exp0066) showed the result of the correlation between the FOXP2 with functional states. **(D)** The sample Kim (Exp0068) showed the result of the correlation between the FOXP4 with functional states. (EMT, epithelial-mesenchymal transition).

Analysing the functional states of the forkhead box P family at the single-cell level

Enrichment analysis results showed that coexpressed gene clusters of the FOXP family acted on several typical cancer pathways. To better understand the relevance and underlying mechanisms of the FOXP family in NSCLC, we investigated the 14 functional states of the FOXP family at the single-cell level *via* the CancerSEA database (Figure 8). The results indicated that FOXP1 was mainly positively correlated with differentiation and hypoxia, FOXP2 was mainly negatively correlated with cell cycle,

DNA damage, DNA repair, invasion, metastasis, proliferation, and FOXP4 was mainly positively correlated with hypoxia, invasion, stemness (Figure 8A). Besides, the single-cell analysis result related to FOXP3 were not stated here for the CancerSEA database lacked the FOXP3 data at the single-cell level. we need to supplement this part in the future. In terms of functional relevance in different T cell groups, Kim (Exp0068) showed that FOXP1 had positive correlations with angiogenesis, apoptosis, metastasis, and stemness (Spearman's coefficients, 0.74, 0.91, 0.34, and 0.36 respectively; p value < 0.05) and a negative correlation with EMT (−0.88, p value < 0.01) in NSCLC. Kim (Exp0066) showed that FOXP2 had negative correlations with cell cycle, DNA damage, proliferation, DNA repair, metastasis, and invasion(Spearman's coefficients, −0.53, −0.53, −0.52, −0.48 and −0.41 respectively; p value < 0.05) in NSCLC. Kim (Exp0068) reported that high FOXP4 expression was positively correlated with metastasis, angiogenesis, inflammation, stemness and hypoxia, (Spearman's coefficients, 0.71, 0.64, 0.59, 0.59, and 0.25 respectively; p value < 0.05) and negatively associated with cell cycle (Spearman's coefficients, −0.65, p value < 0.01) in NSCLC. These discoveries indicate that the FOXP family may crucially affect the tumour progression of NSCLC.

The degree of immune factor infiltration regulated by the forkhead box P family in non-small cell lung cancer

To augment the understanding of the relationship between the FOXP family and immune infiltration, the connection between the FOXP family and various immune signatures, which included the immune-related characteristics of 28 TIL types, immunomodulators (immunoinhibitor, immunostimulator, and MHC molecules), chemokines and receptors, was investigated. All the heatmaps showing the correlation results are presented in Figures 9, 10; Supplementary Figure S7. It was obvious from the heatmaps that the FOXP family was related to immune signatures. To further analyse the relevant mechanisms of the FOXP family in regulating immunity, we selected two modules, the one with the most relevant expression and the other with the most relevant infiltration after copy number alteration and DNA methylation as representatives (the rho of the Spearman correlations test was the highest). When different immune molecules showed upregulation and downregulation trends under the same conditions, two modules were chosen to represent the upregulation and downregulation molecular clusters. The representative immune signatures regulated by FOXP1 were Act CD4 and Tem CD8 in LUAD, NK cells, and neutrophils in LUSC. The infiltration abundances of Act CD4 and Tem CD8 in LUAD tumour tissue were negatively correlated with the expression of FOXP1, and FOXP1 was low in LUAD tissue. That is, the abundances of Act CD4 and Tem CD8 infiltration increased in LUAD tissue. The correlation scores of the two were −0.258 and −0.042, respectively, and there were positive correlations due to copy number alteration and DNA methylation of FOXP1. The representative lymphocytes regulated by FOXP1 in LUSC were NK cells and neutrophils. As we confirmed above, FOXP1 was expressed at low levels in LUSC tissues and was positively correlated with the abundance of NK cells and neutrophil infiltration. The relative rho scores were 0.526 and 0.312, respectively, so NK cells and Neutrophil infiltration were abundant. The degree of decrease in LUSC and was negatively correlated due to the variations in copy number alteration and DNA methylation. By analogy, the regulation of copy number alteration and DNA methylation is shown in the Figures 11A–F. The downregulation of FOXP1 affected the results, including ActCD4, TemCD8, TGFBR1, TIGIT, TNFRSF25, ICOS, TAP1, TAP2, CCL14, CXCL10, CCL5, CX3CR1, CCR5, NK, neutrophils, KDR, ADORA2A, ENTPD1, TMEM1730, HLA-DOA, TAPBP, CCL12, CCL28, CCL26, CXCR4, and CXCR1. The immune infiltration coefficient of FOXP2 in NSCLC tissue was less than those of FOXP1/3/4. The upregulation of FOXP3 mainly affected TemCD8, ActCD4, TIGIT, CTLA4, ICOS, IL2RA, HLA-B, HLA-DPB1, HLA-DOB, CCL19, CCL11, CCR8, ImmB, ActB, TIGIT, IDO1, ICOS, CD27, HLA-DPB1, CCL5, CCL19, CCR8, and CCR7. The downregulation of FOXP4 was mainly associated with ActCD4, Th1, PDCD1LG2, HAVCR-2, TNFSF4, CD40, B2 M, HLA-B, CCL26, CCL14, CCR1, CD56bright, Eosinophil, PVRL2, ADORA2A, ICOSLG, CXCR4, TAPBP, HLA-DOA, CCL26, CCL28, CCR10, and CCR6. We also constructed rate scores to compare the influence of copy number alteration and DNA methylation on the FOXP family (Figures 11G–L). The results presented that both copy number alteration and DNA methylation on the FOXP family play effects on the infiltration correlation results of immune factors in NSCLC, and it was obvious that the changes of immune infiltration correlation after DNA methylation on the FOXP family were significant than those after FOXP family copy number alteration. The multiple influences were different due to different pathological types of NSCLC. Therefore, we could infer that the corresponding conclusion that copy number alteration and DNA methylation regulated the infiltration of corresponding immune factors by the FOXP family. In addition, except for FOXP2 in LUAD and FOXP4 in LUSC, the remaining FOXP family members had significantly different effects on immunophenotyping C1-C6 in NSCLC (Figure 12). Therefore, it was confirmed that the FOXP family participated widely in modulating various immune molecules to affect immune infiltration in NSCLC progression.

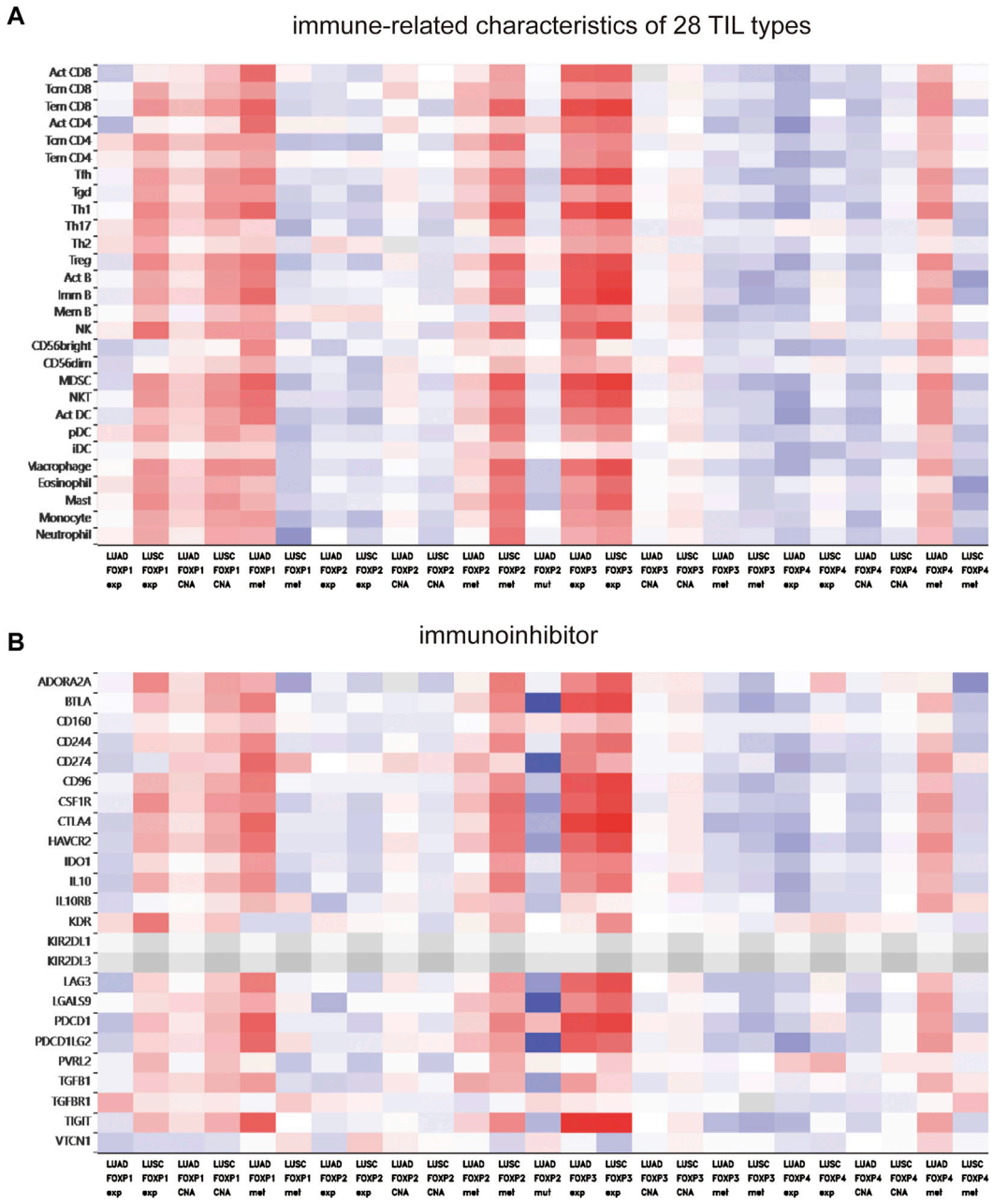


FIGURE 9
The relationship between the degree of immune factors infiltration in NSCLC and the expression of FOXP family. **(A)** The hotmap presented the correlations between the FOXP family and immune-related characteristics of 28 TIL types. **(B)** The hotmap presents the correlations between the FOXP family and immunoinhibitor.

Genetic alteration and interaction analyses of the forkhead box P family in non-small cell lung cancer

Upon analysis of the FOXP family in the OncoPrint module on cBioPortal, the results revealed that gene alterations in FOXP1/2/3/4 occurred in 3%, 3%, 2.2%, and 2.7% of the NSCLC samples, respectively (Figure 13A). The genetic alterations of structural variants, mutations, amplifications, deep deletions, and copy number alterations of the FOXP family all occurred in NSCLC (Figure 13B). The details of all mutations in NSCLC are summarized in Supplement Figure 6. FOXP1 had 15 missense mutations, 3 splice mutations, and one fusion mutation. FOXP2 had

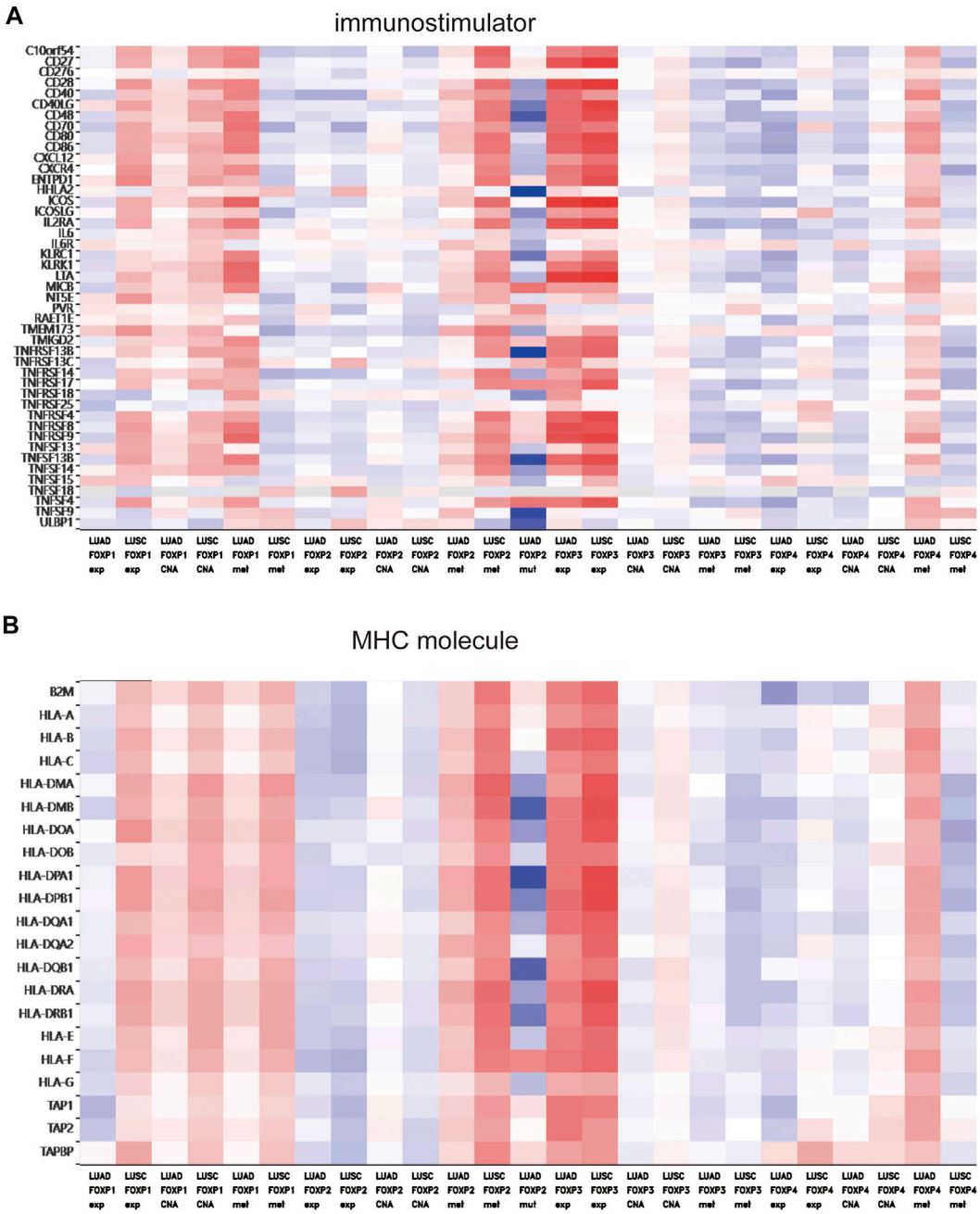


FIGURE 10
The relationship between the degree of immune factors infiltration in NSCLC and the expression of the FOXP family. **(A)** The heatmap presented the correlations between the FOXP family and immunostimulator. **(B)** The heatmap presented the correlations between the FOXP family and MHC molecule.

one missense mutation and one Fusion mutation. FOXP3 had one missense mutation. FOXP4 had no mutation. Only FOXP1 had domain mutations (Figure 13C), while the remaining FOXP2/3/4 had no domain mutations (Figure 13D). The abovementioned multiple alterations of the FOXP family might partially explain the mechanism of occurrence and progression in NSCLC. In addition, we conducted a PPI network analysis of the FOXP family by STRING to investigate the feasible interactions in their internal and related genes. Multiple nodes (34) and edges (212) are shown in the PPI network (Figure 13E). The STRING results mainly displayed the functions connected with

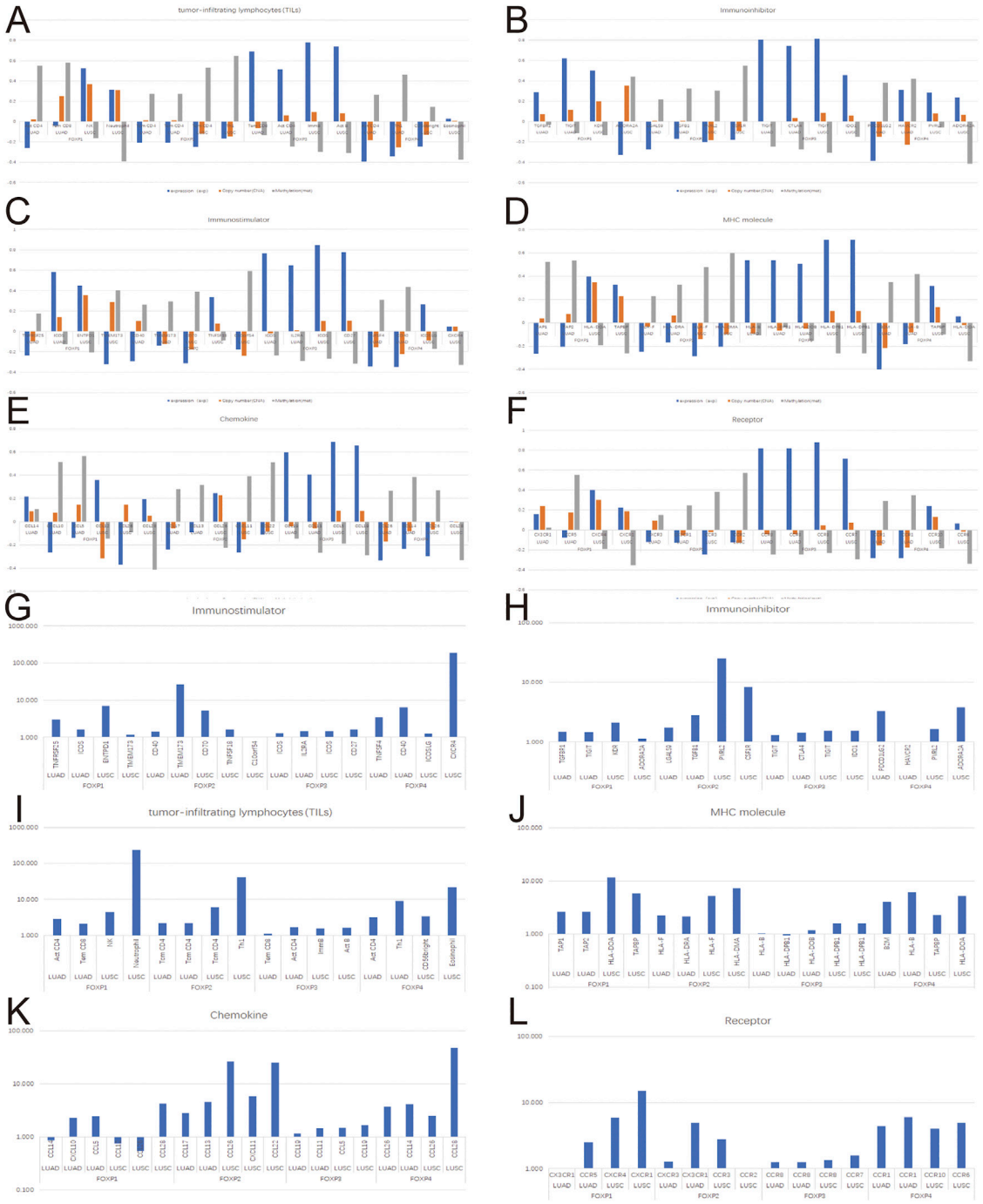


FIGURE 11
The effects of FOXP family expression, copy number alteration, and DNA methylation on immune factors. **(A–F)** These histograms present the correlation scores of the top two most relevant immune infiltration module and correlation scores modified by copy number alteration and DNA methylation. **(G–L)** These histograms present the fold relationship between copy number alteration and DNA methylation correlation scores of the FOXP family in NSCLC. (CNA, Copy number alteration; MET, DNA Methylation).

immunity, including regulation of T cell homeostatic proliferation, the activity of T-helper 17 cells, the signalling pathway mediated by interleukin-2, and the adjustment of regulatory T cell differentiation. We further investigated the results of STRING in Cytoscape and then curtained and locked out 10 hub genes (IL2, IFNG, FOXP3, CTLA4, STAT3, IRF4,

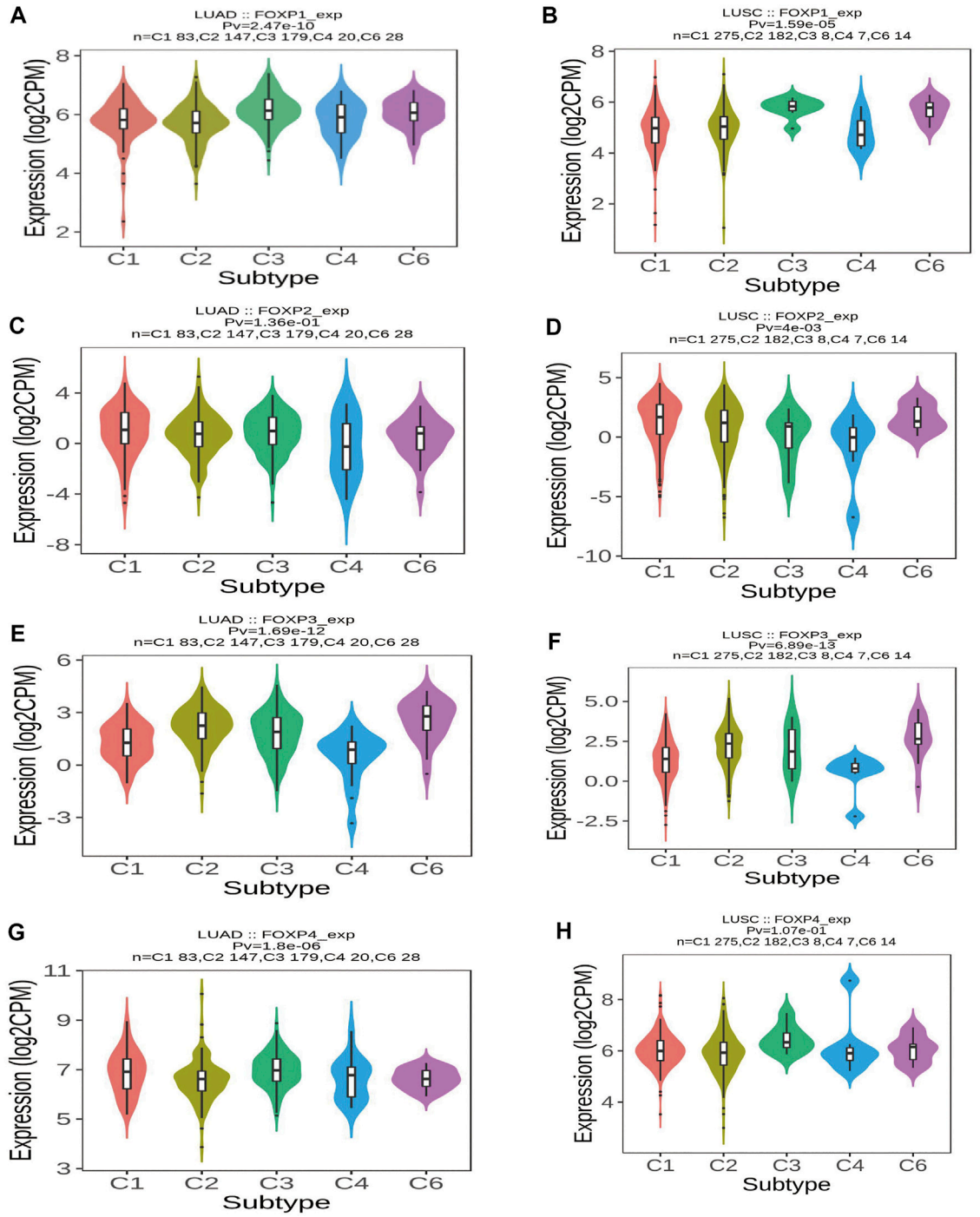


FIGURE 12
The relationship between immune types and the FOXP family. **(A)** The violins plot showed statistical relationships between FOXP1 and immune types C1–C6 in LUAD. **(B)** The violins plot showed statistical relationships between FOXP1 and immune types C1–C6 in LUSC. **(C)** The violins plot showed statistical relationships between FOXP2 and immune types C1–C6 in LUAD. **(D)** The violins plot showed statistical relationships between FOXP2 and immune types C1–C6 in LUSC. **(E)** The violins plot showed statistical relationships between FOXP3 and immune types C1–C6 in LUAD. **(F)** The violins plot showed statistical relationships between FOXP3 and immune types C1–C6 in LUSC. **(G)** The violins plot showed statistical relationships between FOXP4 and immune types C1–C6 in LUAD. **(H)** The violins plot showed statistical relationships between FOXP4 and immune types C1–C6 in LUSC. [C1 (wound healing); C2 (IFN-gamma dominant); C3 (inflammatory); C4 (lymphocyte depleted); C5 (immunologically quiet); C6 (TGF-b dominant)].

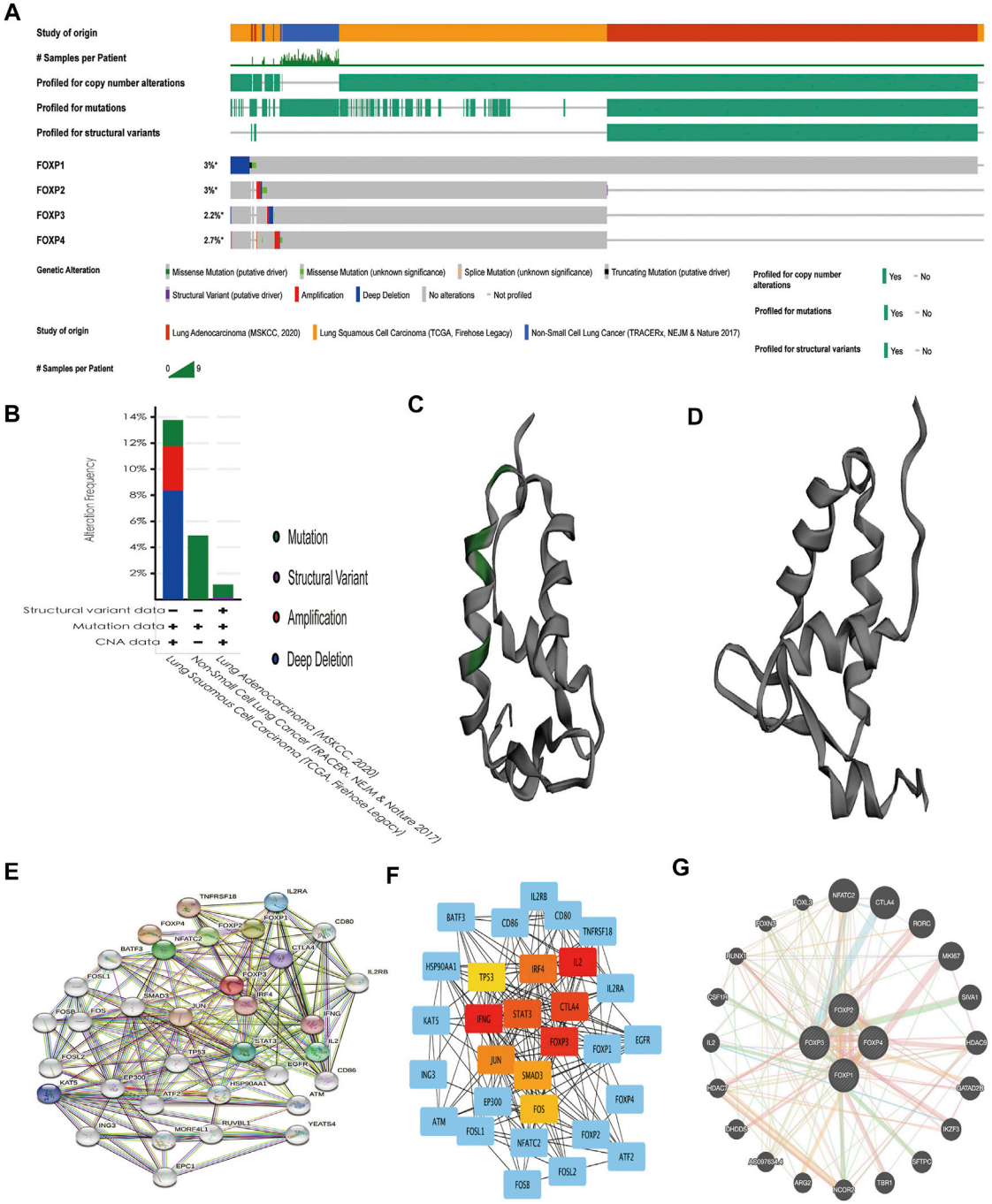


FIGURE 13
Genetic alteration, and interaction analyses of the FOXP family in NSCLC patients. **(A)** The heatmap showed the respective frequencies of Gene alterations occurring in the sequenced cases by the FOXP family in the data obtained from the OncoPrint schematic of cBioPortal. **(B)** The mutation types of NSCLC in three datasets. **(C)** The chromosomal structure of FOXP1 with domain mutations. **(D)** The chromosomal structure of FOXP2/3/4 without domain mutation. **(E)** The protein-protein interaction PPI network analysis of the FOXP family using STRING. **(F)** The top 10 hub genes were exported by Cytoscape (version 3.7.2) with the cytoHubba app. **(G)** The result of GeneMANIA reveals functionally similar genes of the FOXP family.

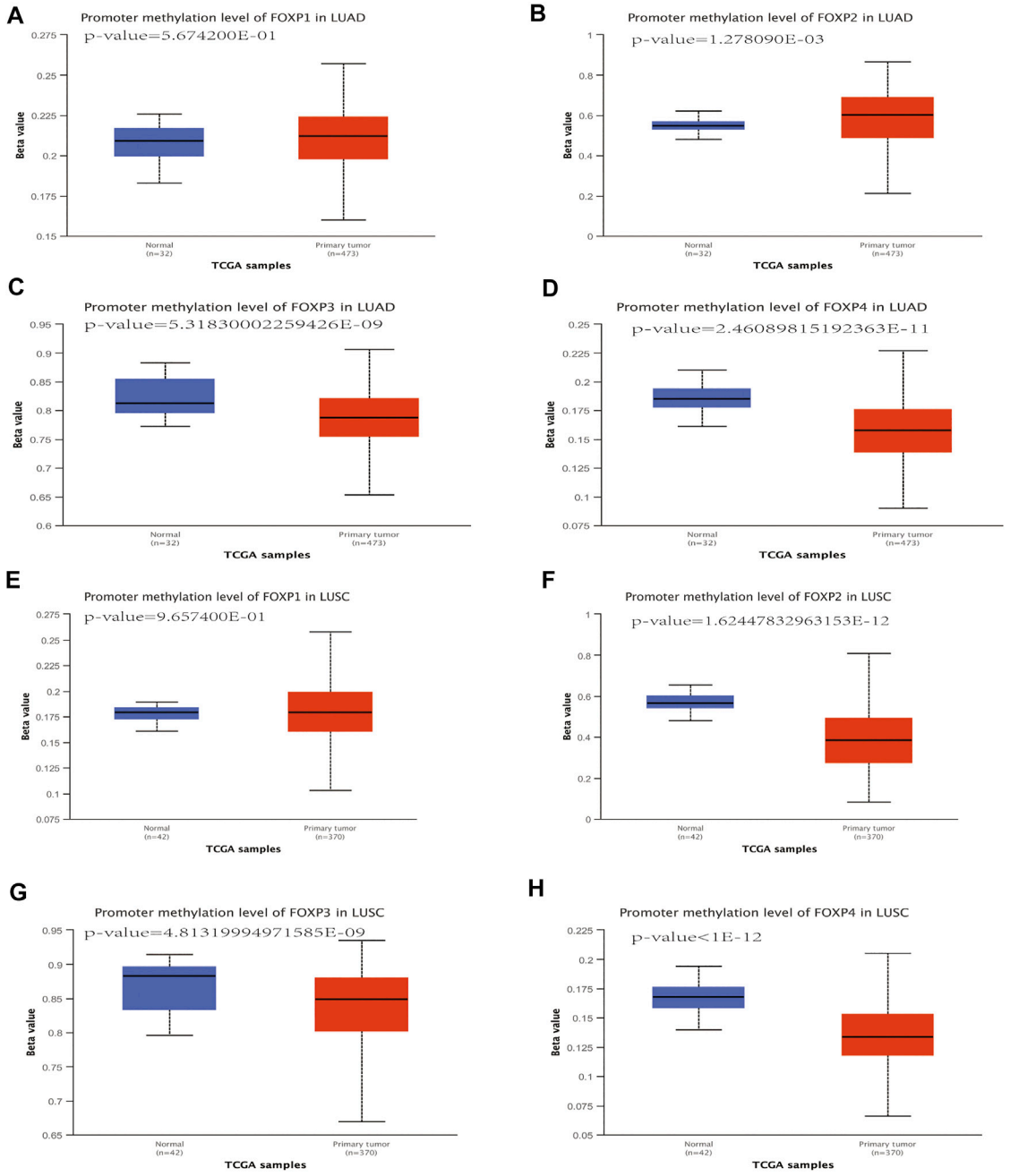


FIGURE 14
The DNA methylation analysis of FOXP family in NSCLC. **(A–H)** the promoter methylation levels of the FOXP family in LUAD/LUSC compared with normal samples.

JUN, SMAD3, FOS, TP53), as shown in Figure 13F. The FOXP family was input to the GeneMANIA website to link genes with similar functions. Functionally similar genes surrounded the outside of the FOXP family in the presentation (Figure 13G). The GeneMANIA results affirmed that the functions of the FOXP family and their related clusters were chiefly related to the differentiation of lymphocytes and T cells and the regulation of leukocytes. The above results support that the FOXP family participates in the immune process of NSCLC under the condition of interaction.

TABLE 1 The prognostic value of single CpG of FOXP family in NSCLC by MethSurv ($p < 0.05$).

Gene	Tissue	CpG	RefGene group	Relation to CpG island	HR	*p*-value
FOXP1	LUAD	cg00201568	Body	Open_Sea	1.097	0.0026
		cg00707452	TSS1500	S_Shore	0.839	0.0043
		cg01173432	5′UTR	Open_Sea	1.308	0.0066
		cg01186551	5′UTR	Open_Sea	0.923	0.0068
		cg01189917	TSS1500	S_Shore	0.965	0.0076
		cg01232145	Body	Open_Sea	0.746	0.015
		cg01331540	Body	Open_Sea	1.095	0.015
		cg01534217	5′UTR	Island	1.111	0.019
		cg02002523	Body	Open_Sea	0.847	0.02
		cg02220284	Body	Open_Sea	1.318	0.02
		cg02336104	5′UTR	Open_Sea	1.133	0.022
		cg02520804	5′UTR	N_Shore	1.101	0.027
		cg02862354	Body	Island	0.886	0.036
	LUSC	cg22798400	Body	Open_Sea	0.638	0.006
		cg02520804	5′UTR	N_Shore	0.682	0.02
		cg00052246	Body	Open_Sea	1.46	0.034
		cg01173432	5′UTR	Open_Sea	1.564	0.038
		cg00201568	Body	Open_Sea	1.409	0.046
		cg01189917	TSS1500	S_Shore	1.474	0.049
		cg25481160	Body	N_Shelf	0.673	0.05
FOXP2	LUAD	—				
	LUSC	—				
FOXP3	LUAD	—				
	LUSC	cg04920616	TSS200	Open_Sea	0.684	0.032
FOXP4	LUAD	cg12911122	5′UTR	S_Shore	1.759	0.00058
		cg26432961	5′UTR	S_Shore	1.891	0.0035
		cg08696640	5′UTR	S_Shelf	1.558	0.0068
		cg05734456	5′UTR	Island	1.489	0.014
		cg04617914	TSS1500	N_Shore	1.45	0.024
		cg17620505	5′UTR	N_Shelf	1.574	0.028
		cg01508045	5′UTR	Island	1.558	0.029
	LUSC	cg03442064	5′UTR	Island	1.568	0.0057
		cg00806680	Body	N_Shore	0.688	0.022
		cg08727957	TSS1500	Island	0.696	0.026
		cg05140895	TSS200	Island	0.644	0.038

Notes: HR, hazard ratio.

DNA methylation analysis of the forkhead box P family

In the process of using TISIDB to study the effect of the FOXP family on immune infiltration, we found that the FOXP family significantly changed the correlation degrees between immune signatures after undergoing epigenetic alterations of copy number alteration and DNA methylation. These cBioPortal findings suggested that copy number alteration of the FOXP family played a role in the progression of NSCLC. Therefore, we used DNA methylation as a representative epigenetic alteration to evaluate its effect on the expression levels of the FOXP family and patients prognosis. We first applied the UALCAN database to determine the relationship between DNA methylation and the expression of the FOXP family in NSCLC. The DNA methylation level of FOXP1 was no statistically significant in NSCLC than those in normal samples (Figures 14A,E). The DNA methylation level of FOXP2 was higher in LUAD but lower in LUSC (Figures 14B,F). The DNA methylation levels of FOXP3/4 in NSCLC were lower than those in normal samples (Figures 14C,D,G,H). According to these data, the expression levels of FOXP2/3/4 were significantly associated with DNA methylation in

NSCLC. In addition, the results showed that 20 CpGs of FOXP1, 1 CpGs of FOXP3, and 12 CpGs of FOXP4 presented important statistical significance related to prognosis. Moreover, with the occurrence of different CPG sites, FOXP family members' DNA methylation statuses were related to different prognoses. Specific details of the results including the types of CpG, RefGene groups, relationship to CpG islands, HRs, and *p* values, are listed in Table 1.

Discussion

The treatment of NCSLC with immunotherapy including ICIs, has improved the clinical benefits for patients and greatly innovated the traditional chemotherapy regimen (Cogdill et al., 2017). Nonetheless, many patients are still rejected for immunotherapy due to not meeting the inclusion criteria. As a result, research on advanced and effective modulators at immune-related critical points is in full swing. To further understand the molecular regulatory mechanism of the immune system in the management of NSCLC. Our article elaborates on the specific regulatory details of specific immune molecules from the perspective of the FOXP family.

At present, the FOXP family is observed to play negative or positive roles in particular cancers. For example, FOXP1 drives the occurrence of malignant behaviour by dominating the expression level of PKLR in gallbladder cancer (Wang et al., 2019). FOXP2 participates in the process of invasion and metastasis of breast cancer *via* the TGFβ/SMAD pathway (Chen et al., 2018). Aberrant expression of FOXP3 in colorectal cancer is related to immune overdrive in a high-risk subpopulation (Cui et al., 2021). FOXP4 directly acts on LEF-1 and gives impetus to the occurrence of laryngeal squamous cell carcinoma (Shi et al., 2021). At the same time, many recently published works in the literature show that the FOXP family participates in the process of immune system reconstruction of tumour tissue by activating or inhibiting the specific function of immune molecules (Fleskens and van Boxtel, 2014). The FOXP family, as a major contributor, can regulate tumour-associated inflammation and immune responses in tumour progression. For example, FOXP1 inhibits the behaviour of immune activation and the expression of MHC class II in diffuse large B-cell lymphomas (Brown et al., 2016). FOXP3 is defined as a manager to administer the immunosuppressive response of T cells (Klimenko, 2011). FOXP3 directly restrains CD44 breast cancer by participating in the corresponding regulatory role (Zhang et al., 2015). Furthermore, increasing evidence unambiguously confirms that epigenetic alteration plays a role in the process of cancer, and various epigenetic alterations can be used as maker molecules to evaluate the risk of tumour prognosis (Richardson et al., 2018; Zhang and Zhang, 2020). For instance, the regulation of immune cells is closely related to the copy number alteration of TRPV1 in renal cell carcinoma (Zheng et al., 2020). Abnormal DNA methylation impacts gene expression and survival time in breast cancer patients (Gyorffy et al., 2016). Therefore, the underlying mechanism of FOXP family expression/copy number alteration/DNA methylation in the regulation of immune-related signatures was initially elucidated in this paper.

Previous studies have shown that FOXP3 is overexpressed to facilitate the invasion and metastasis of NSCLC (Li et al., 2021). Our results showed there were different expression levels of the FOXP family according to different pathological types in NSCLC compared with normal tissue. UALCAN presented that the expression levels of the FOXP family had significant effects on the clinical parameters, including patient age, smoking habits, histological subtypes, individual cancer stages, nodal metastasis status, and TP53 mutation status. The Kaplan–Meier Plotter showed that the overexpression levels of FOXP1/4 were involved in the better prognosis, and the overexpression levels of FOXP2/3 were associated with poor prognosis of NSCLC. It may be an option to analyse the mRNA expression levels of the FOXP family members in NSCLC patients to provide powerful markers to define prognosis.

GO and KEGG pathway analyses of coexpressed genes of the FOXP family indicated that the FOXP family possessed roles in activating the Wnt, PI3K/AKT/mTOR, and FOCAD-FAK pathways to regulate tumourigenesis and the progression of relevant immune responses in NSCLC. Combined with previous research, the above typical pathways were associated with NSCLC progression (Heavey et al., 2014; Stewart, 2014; Liu et al., 2020). Our study further clarified the role of the FOXP family in the development of NSCLC. Likewise, CancerSEA showed the functional states of the FOXP family have a necessary relationship with the activity of the cell cycle, differentiation, apoptosis, angiogenesis, invasion, EMT, proliferation, hypoxia, inflammation, and stemness at the single-cell level. The results validated previous evidence that the FOXP family was involved in the progression of a variety of cancers. For example, the FOXP family regulates β cell proliferation in concert with NFATC2 (Simonett et al., 2021). Furthermore, FOXP2 targets GRP78 in breast cancer to promote tumour proliferation and metastasis (Wu et al., 2018). In addition, FOXP1 inhibits guidance proteins to promote angiogenesis in cell activity (Grundmann et al., 2013). Taken together, the functional states of the FOXP family accurately revealed that the FOXP family might crucially affect the progression of NSCLC. The results of PPI interaction and GeneMANIA analyses further demonstrated the occurrence of cooperation and interaction between FOXP members. These results implied that FOXP

members could function through alliance mechanisms in NSCLC. Our study showed that the FOXP family was prominently dysregulated in NSCLC, and we then carried out the an analysis of genetic alterations. Unsurprisingly, there was the evidence of fusion, mutation, and amplification of the FOXP family in NSCLC. These genetic alterations were undoubtedly involved in the molecular malignant behaviour of NSCLC.

The DNA methylation process of specific genes mediates different biological results of cancer. For example, DNA methylation is related to the occurrence of drug resistance in patients with glioblastoma during treatment (Lu et al., 2020). In addition, the absence of DNA methylation can cause immune evasion in various cancers (Jung et al., 2019). To explore the particular mechanism of the FOXP family in NSCLC, we investigated the connection between the promoter methylation levels and the expression levels of the FOXP family using UALCAN databases. The outcomes showed that the expression levels of FOXP2/3/4 were correlated with their promoter methylation levels in NSCLC. In addition, we analysed the relationship between DNA methylation modification behaviour at different sites of the FOXP family and patient survival time. Significant prognostic values (p value < 0.05) were observed for FOXP1/3/4. In short, analysis of FOXP family DNA methylation provides a new approach to the prognosis of NSCLC.

Data from recent years have shown that the combined use of ICIs has improved the survival time of NSCLC patients by blocking the checkpoint inhibition process. Our study presented the correlations between FOXP family expression/copy number alteration/DNA methylation and immune signatures. The results showed that the FOXP family without epigenetic alterations mainly controlled the degrees of infiltration of immune-related factors (Tem CD8, TXNDC5, TAP1, TAP2, CCL5, NK, KDR, ENTPD1, and HLA-DOA) in NSCLC. Previous studies have confirmed that Tem CD8 inhibits tumour growth in mouse models and plays a vital role in cancer immune surveillance and treatment (Wang et al., 2020). TXNDC5 promotes pulmonary fibrosis by augmenting TGFβ signalling through TGFBR1 stabilization (Lee et al., 2020). TAP1 and TAP2 are typical tumour predictors (Gostout et al., 2003; Henle et al., 2017). CCL5, as a receptor antagonist, plays a positive role in the process of tumour progression by attracting macrophages (Van Damme et al., 2004). The activation of NK cells is related to immune dysfunction and a harmful tumour microenvironment (Li et al., 2020b). There is currently a small-molecule tyrosine kinase inhibitor (Moulder, #137) for KDR that is effective for lung cancer (Dai et al., 2019; Song et al., 2020). FOXP3 regulates the expression and infiltration of ENTPD1 to promote the occurrence of tumours (Sun et al., 2010). HLA-DOA has confirmed that the degree of infiltration in the tissue is directly proportional to the degree of inflammation (Okada et al., 2016). After copy number alteration and DNA methylation, our results revealed that the correlations between the FOXP family and immune parameters were opposite to those before alteration in NSCLC. In addition, the influence of DNA methylation was stronger than that of copy number alteration. In addition, due to the different pathological types of NSCLC, the multiples of the influence intensity were also different. Altogether, our results partly showed that FOXP family expression/copy number alteration/DNA methylation regulated the infiltration of corresponding immunity in NSCLC. This paper provided more detailed molecular mechanisms for the development of new immune checkpoints from the perspective of FOXP family.

Our research has many details that need to be further verified. The data required for the content of bioinformatics analysis in this paper are from public databases. Further basic and clinical trials are still required to explore the detailed molecular mechanism of the FOXP family in NSCLC.

Conclusion

This paper systematically analysed molecular mechanism of FOXP family member regulation, including the expression levels, correlation with clinicopathological stages, DNA methylation levels, epigenetics alterations, prognostic values, relationship with immune regulation and functional analysis based on coexpression in NSCLC. Activation of FOXP family-related pathways could significantly change the patient's response to tumour immunity. Our article showed that the FOXP family members, as diagnostic and prognostic biomarkers, provide new information for the development of ICI drugs for patients with NSCLC.

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

All authors contributed to the study conception and design. D-MH and W-DZ conceived and designed the study; M-YZ and Z-ES collected the literature; D-MH drafted the manuscript; Q-XW reviewed statistical methods; RL collected the data; X-LJ and Y-QQ revised the manuscript. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the national natural science foundation of China (Grant No. 82173173), the Special Fund for Clinical Research of Jinan City (201912011).

Acknowledgments

We would like to thank the public databases for providing shared resources and algorithms, including UALCAN, GEPIA, Kaplan-Meier Plotter, MethSurv, cBioPortal, TISIDB, COEXPEDIA, STRING and GeneMANIA.

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.2022.937069/full#supplementary-material

SUPPLEMENTARY FIGURE S1
The meta-analysis results of the expression difference of FOXP family from different data sets using LUNG CANCER EXPLORER database. The meta-analysis result of the expression level of FOXP1 from different LUAD/LUSC data sets. **(B,F)** The meta-analysis result of the expression level of FOXP2 from different LUAD/LUSC data sets. **(C,G)** The meta-analysis result of the expression level of FOXP3 from different LUAD and LUSC data sets. **(D,H)** The meta-analysis result of the expression level of FOXP4 from different LUAD and LUSC data sets.

SUPPLEMENTARY FIGURE S2
The relationship between the expression levels of FOXP family and clinicopathological stages. **(A-L)** AuthorAnonymous, The Box plots show the differential expression of FOXP1 between different staged tissues of LUAD/ LUSC and normal tissues.

SUPPLEMENTARY FIGURE S3
The relationship between the expression levels of FOXP family and clinicopathological stages. AuthorAnonymous, **(A-L)** The Box plots show the differential expression of FOXP2 between different staged tissues of LUAD/ LUSC and normal tissues.

SUPPLEMENTARY FIGURE S4
The relationship between the expression levels of FOXP family and clinicopathological stages. AuthorAnonymous,**(A-L)** The Box plots show the differential expression of FOXP3 between different staged tissues of LUAD/ LUSC and normal tissues.

SUPPLEMENTARY FIGURE S5
The relationship between the expression levels of FOXP family and clinicopathological stages. AuthorAnonymous,**(A-L)** The Box plots show the differential expression of FOXP4 between different staged tissues of LUAD/ LUSC and normal tissues.

SUPPLEMENTARY FIGURE S6
The relationship between the expression of FOXP family and survival in LUSC.**(A)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP1 in LUSC. **(B)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP2 in LUSC. **(C)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP3 in LUSC. **(D)** The survival curves reflected the relationship between the patients' overall survival (OS) rate and the corresponding gene expression levels of FOXP4 in LUSC.

SUPPLEMENTARY FIGURE S7
The relationship between the degree of immune factor infiltration in NSCLC and the expression of FOXP family.**(A)** The hotmap presented the correlations between FOXP family and chemokine. **(B)** The hotmap presented the correlations between FOXP family and receptor.

References

Brown, P. J., Wong, K. K., Felce, S. L., Lyne, L., Spearman, H., Soilleux, E. J., et al. (2016). FOXP1 suppresses immune response signatures and MHC class II expression in activated B-cell-like diffuse large B-cell lymphomas. *Leukemia* 30 (3), 605–616. doi:10.1038/leu.2015.299

Camidge, D. R., Doebele, R. C., and Kerr, K. M. (2019). Comparing and contrasting predictive biomarkers for immunotherapy and targeted therapy of NSCLC. *Nat. Rev. Clin. Oncol.* 16 (6), 341–355. doi:10.1038/s41571-019-0173-9

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B., et al. (2017). Ualcan: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Chen, M. T., Sun, H. F., Li, L. D., Zhao, Y., Yang, L. P., Gao, S. P., et al. (2018). Downregulation of FOXP2 promotes breast cancer migration and invasion through TGFβ/SMAD signaling pathway. *Oncol. Lett.* 15 (6), 8582–8588. doi:10.3892/ol.2018.8402

Cogdill, A. P., Andrews, M. C., and Wargo, J. A. (2017). Hallmarks of response to immune checkpoint blockade. *Br. J. Cancer* 117 (1), 1–7. doi:10.1038/bjc.2017.136

Cui, K., Yao, S., Zhang, H., Zhou, M., Liu, B., Cao, Y., et al. (2021). Identification of an immune overdrive high-risk subpopulation with aberrant expression of FOXP3 and CTLA4 in colorectal cancer. *Oncogene* 40 (11), 2130–2145. doi:10.1038/s41388-021-01677-w

Dai, X., Chen, X., Hakizimana, O., and Mei, Y. (2019). Genetic interactions between ANLN and KDR are prognostic for breast cancer survival. *Oncol. Rep.* 42 (6), 2255–2266. doi:10.3892/or.2019.7332

Evison, M., and AstraZeneca, U. K. L. (2020). The current treatment landscape in the UK for stage III NSCLC. *Br. J. Cancer* 123 (1), 3–9. doi:10.1038/s41416-020-01069-z

Fleskens, V., and van Boxtel, R. (2014). Forkhead box P family members at the crossroad between tolerance and immunity: a balancing act. *Int. Rev. Immunol.* 33 (2), 94–109. doi:10.3109/08830185.2013.816698

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 6 (269), pl1. doi:10.1126/scisignal.2004088

Gostout, B. S., Poland, G. A., Calhoun, E. S., Sohni, Y. R., Giuntoli Ii, R. L., McGovern, R. M., et al. (2003). TAP1, TAP2, and HLA-DR2 alleles are predictors of cervical cancer risk☆. *Gynecol. Oncol.* 88 (3), 326–332. doi:10.1016/s0090-8258(02)00074-4

Grundmann, S., Lindmayer, C., Hans, F. P., Hoefer, I., Helbing, T., Pasterkamp, G., et al. (2013). FoxP1 stimulates angiogenesis by repressing the inhibitory guidance protein semaphorin 5B in endothelial cells. *PLoS One* 8 (9), e70873. doi:10.1371/journal.pone.0070873

Gyorffy, B., Bottai, G., Fleischer, T., Munkacsy, G., Budczies, J., Paladini, L., et al. (2016). Aberrant DNA methylation impacts gene expression and prognosis in breast cancer subtypes. *Int. J. Cancer* 138 (1), 87–97. doi:10.1002/ijc.29684

Heavey, S., O'Byrne, K. J., and Gately, K. (2014). Strategies for co-targeting the PI3K/AKT/mTOR pathway in NSCLC. *Cancer Treat. Rev.* 40 (3), 445–456. doi:10.1016/j.ctrv.2013.08.006

Henle, A. M., Nassar, A., Puglisi-Knutson, D., Youssef, B., and Knutson, K. L. (2017). Downregulation of TAP1 and TAP2 in early stage breast cancer. *PLoS One* 12 (11), e0187323. doi:10.1371/journal.pone.0187323

Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J., Wu, Y.-L., et al. (2017). Lung cancer: current therapies and new targeted treatments. *Lancet* 389 (10066), 299–311. doi:10.1016/s0140-6736(16)30958-8

Hu, Z., Cai, M., Zhang, Y., Tao, L., and Guo, R. (2020). miR-29c-3p inhibits autophagy and cisplatin resistance in ovarian cancer by regulating FOXP1/ATG14 pathway. *Cell Cycle* 19 (2), 193–206. doi:10.1080/15384101.2019.1704537

Jung, H., Kim, H. S., Kim, J. Y., Sun, J. M., Ahn, J. S., Ahn, M. J., et al. (2019). DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. *Nat. Commun.* 10 (1), 4278. doi:10.1038/s41467-019-12159-9

Kim, J. H., Hwang, J., Jung, J. H., Lee, H. J., Lee, D. Y., and Kim, S. H. (2019). Molecular networks of FOXP family: dual biologic functions, interplay with other molecules and clinical implications in cancer progression. *Mol. Cancer* 18 (1), 180. doi:10.1186/s12943-019-1110-3

Klimenko, O. V. (2011). Regulation of immune responses, apoptosis, and tumorigenesis by separate FOXP-3-dependent genes: connection with clinical manifestations. *J. Microbiol. Immunol. Infect.* 44 (6), 412–417. doi:10.1016/j.jmii.2011.04.007

Lee, T. H., Yeh, C. F., Lee, Y. T., Shih, Y. C., Chen, Y. T., Hung, C. T., et al. (2020). Fibroblast-enriched endoplasmic reticulum protein TXNDC5 promotes pulmonary fibrosis by augmenting TGFβ signaling through TGFBR1 stabilization. *Nat. Commun.* 11 (1), 4254. doi:10.1038/s41467-020-18047-x

Li, C., Wang, H., Fang, H., He, C., Pei, Y., and Gai, X. (2021). FOXP3 facilitates the invasion and metastasis of non-small cell lung cancer cells through regulating VEGF, EMT and the Notch1/Hes1 pathway. *Exp. Ther. Med.* 22 (3), 958. doi:10.3892/etm.2021.10390

Li, J., Li, H., Zhang, C., Zhang, C., and Wang, H. (2020b). Integrative analysis of genomic alteration, immune cells infiltration and prognosis of lung squamous cell carcinoma (LUSC) to identify smoking-related biomarkers. *Int. Immunopharmacol.* 89, 107053. doi:10.1016/j.intimp.2020.107053

Li, J. P., Zhang, H. M., Liu, M. J., Xiang, Y., Li, H., Huang, F., et al. (2020a). miR-133a-3p/FOXP3 axis regulates cell proliferation and autophagy in gastric cancer. *J. Cell. Biochem.* 121 (5-6), 3392–3405. doi:10.1002/jcb.29613

Liu, P., Wu, D., Duan, J., Xiao, H., Zhou, Y., Zhao, L., et al. (2020). NRF2 regulates the sensitivity of human NSCLC cells to cystine deprivation-induced ferroptosis via FOCAD-FAK signaling pathway. *Redox Biol.* 37, 101702. doi:10.1016/j.redox.2020.101702

Lu, C., Wei, Y., Wang, X., Zhang, Z., Yin, J., Li, W., et al. (2020). DNA-methylation-mediated activating of lncRNA SNHG12 promotes temozolomide resistance in glioblastoma. *Mol. Cancer* 19 (1), 28. doi:10.1186/s12943-020-1137-5

Ma, T., and Zhang, J. (2019). Upregulation of FOXP4 in breast cancer promotes migration and invasion through facilitating EMT. *Cancer Manag. Res.* 11, 2783–2793. doi:10.2147/CMAR.S191641

Neyroud, D., Nosacka, R. L., Callaway, C. S., Trevino, J. G., Hu, H., Judge, S. M., et al. (2021). FoxP1 is a transcriptional repressor associated with cancer cachexia that induces skeletal muscle wasting and weakness. *J. Cachexia Sarcopenia Muscle* 12 (2), 421–442. doi:10.1002/jcsm.12666

Okada, Y., Suzuki, A., Ikari, K., Terao, C., Kochi, Y., Ohmura, K., et al. (2016). Contribution of a non-classical HLA gene, HLA-DOA, to the risk of rheumatoid arthritis. *Am. J. Hum. Genet.* 99 (2), 366–374. doi:10.1016/j.ajhg.2016.06.019

Peng, W. X., Huang, J. G., Yang, L., Gong, A. H., and Mo, Y. Y. (2017). Linc-RoR promotes MAPK/ERK signaling and confers estrogen-independent growth of breast cancer. *Mol. Cancer* 16 (1), 161. doi:10.1186/s12943-017-0727-3

Richardson, T. E., Sathe, A. A., Kanchwala, M., Jia, G., Habib, A. A., Xiao, G., et al. (2018). Genetic and epigenetic features of rapidly progressing IDH-mutant astrocytomas. *J. Neuropathol. Exp. Neurol.* 77 (7), 542–548. doi:10.1093/jnen/nly026

Ru, B., Wong, C. N., Tong, Y., Zhong, J. Y., Zhong, S. S. W., Wu, W. C., et al. (2019). Tisidb: an integrated repository portal for tumor-immune system interactions. *Bioinformatics* 35 (20), 4200–4202. doi:10.1093/bioinformatics/btz210

Ruiz-Cordero, R., and Devine, W. P., (2020). Targeted therapy and checkpoint immunotherapy in lung cancer. *Surg. Pathol. Clin.* 13(1) 17–33. doi:10.1016/j.path.2019.11.002

Shi, J., Wang, J., Cheng, H., Liu, S., Hao, X., Lan, L., et al. (2021). FOXP4 promotes laryngeal squamous cell carcinoma progression through directly targeting LEF1. *Mol. Med. Rep.* 24 (6), 831. doi:10.3892/mmr.2021.12471

Simonett, S. P., Shin, S., Herring, J. A., Bacher, R., Smith, L. A., Dong, C., et al. (2021). Identification of direct transcriptional targets of NFATC2 that promote beta cell proliferation. *J. Clin. Invest.* 131 (21), e144833. doi:10.1172/JCI144833

Song, Z. Z., Zhao, L. F., Zuo, J., Fan, Z. S., Wang, L., and Wang, Y. D. (2020). Clinical outcomes and safety of apatinib mesylate in the treatment of advanced non-squamous non-small cell lung cancer in patients who progressed after standard therapy and analysis of the KDR gene polymorphism. *Onco. Targets. Ther.* 13, 603–613. doi:10.2147/OTT.S222985

Stewart, D. J. (2014). Wnt signaling pathway in non-small cell lung cancer. *J. Natl. Cancer Inst.* 106 (1), djt356. doi:10.1093/jnci/djt356

Sun, X., Wu, Y., Gao, W., Enjyoji, K., Csizmadia, E., Muller, C. E., et al. (2010). CD39/ENTPD1 expression by CD4+Foxp3+ regulatory T cells promotes hepatic metastatic tumor growth in mice. *Gastroenterology* 139 (3), 1030–1040. doi:10.1053/j.gastro.2010.05.007

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45, D362–D368. doi:10.1093/nar/gkw937

Tang, Z., Li, C., Kang, B., Gao, G., Li, C., and Zhang, Z. (2017). GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 45, W98–W102. doi:10.1093/nar/gkx247

Van Damme, J., Struyf, S., and Opdenakker, G. (2004). Chemokine-protease interactions in cancer. *Semin. Cancer Biol.* 14 (3), 201–208. doi:10.1016/j.semcancer.2003.10.007

Wang, S., Zhang, Y., Cai, Q., Ma, M., Jin, L. Y., Weng, M., et al. (2019). Circular RNA FOXP1 promotes tumor progression and Warburg effect in gallbladder cancer by regulating PKLR expression. *Mol. Cancer* 18 (1), 145. doi:10.1186/s12943-019-1078-z

Wang, T., Shen, Y., Luyten, S., Yang, Y., and Jiang, X. (2020). Tissue-resident memory CD8(+) T cells in cancer immunology and immunotherapy. *Pharmacol. Res.* 159, 104876. doi:10.1016/j.phrs.2020.104876

Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., et al. (2010). The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* 38, W214–W220. doi:10.1093/nar/gkq537

Wu, J., Liu, P., Tang, H., Shuang, Z., Qiu, Q., Zhang, L., et al. (2018). FOXP2 promotes tumor proliferation and metastasis by targeting GRP78 in triple-negative breast cancer. *Curr. Cancer Drug Targets* 18 (4), 382–389. doi:10.2174/1568009618666180131115356

Zhang, C., Xu, Y., Hao, Q., Wang, S., Li, H., Li, J., et al. (2015). FOXP3 suppresses breast cancer metastasis through downregulation of CD44. *Int. J. Cancer* 137 (6), 1279–1290. doi:10.1002/ijc.29482

Zhang, J., Zhou, N., Lin, A., Luo, P., Chen, X., Deng, H., et al. (2021). ZFHX3 mutation as a protective biomarker for immune checkpoint blockade in non-small cell lung cancer. *Cancer Immunol. Immunother.* 70 (1), 137–151. doi:10.1007/s00262-020-02668-8

Zhang, P., and Zhang, M. (2020). Epigenetic alterations and advancement of treatment in peripheral T-cell lymphoma. *Clin. Epigenetics* 12 (1), 169. doi:10.1186/s13148-020-00962-x

Zheng, L., Dou, X., Song, H., Gao, R., and Tang, X. (2020). TRPV1 acts as a tumor suppressor and is associated with immune cell infiltration in clear cell renal cell carcinoma: evidence from integrated analysis. *J. Cancer* 11 (19), 5678–5688. doi:10.7150/jca.45918

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 14 October 2022
DOI 10.3389/fgene.2022.914354

OPEN ACCESS

EDITED BY
Yingming Sun,
Fujian Medical University, China

REVIEWED BY
Manisha Sachan,
Motilal Nehru National Institute of Technology Allahabad, India
Shihai Liu,
The Affiliated Hospital of Qingdao University, China

*CORRESPONDENCE
Salem Baldi,
salimlaboratory@gmail.com
Yaping Gao,
gaoyp@tsinghua-sz.org.cn
Hui Tian,
tianhui@axbio.cn

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

RECEIVED 06 April 2022
ACCEPTED 21 July 2022
PUBLISHED 14 October 2022

CITATION
Baldi S, He Y, Ivanov I, Khamgan H, Safi M, Alradhi M, Shopit A, Al-Danakh A, Al-Nusaif M, Gao Y and Tian H (2022), Aberrantly hypermethylated ARID1B is a novel biomarker and potential therapeutic target of colon adenocarcinoma.
Front. Genet. 13:914354.
doi: 10.3389/fgene.2022.914354

Aberrantly hypermethylated ARID1B is a novel biomarker and potential therapeutic target of colon adenocarcinoma

Salem Baldi[1]*, Yun He[1], Igor Ivanov[1], Hassan Khamgan[2], Mohammed Safi[3], Mohammed Alradhi[4], Abdullah Shopit[5], Abdullah Al-Danakh[6], Murad Al-Nusaif[7], Yaping Gao[8]* and Hui Tian[1]*

[1]Research Center of Molecular Diagnostics and Sequencing, Axbio Biotechnology (Shenzhen) Co.,Ltd, Shenzhen, Guangdong, China, [2]Department of Molecular Diagnostics and Therapeutics, Genetic Engineering and Biotechnology Research Institute, University of Sadat City, Sadat, Egypt, [3]Department of respiratory, Shandong Second Provincial General Hospital, Shandong University, Jinan, China, [4]Department of Urology, The Affiliated Hospital of Qingdao Binhai University, Qingdao, China, [5]Academic Integrated Medicine and Collage of Pharmacy, School of Pharmacy, Department of Pharmacology, Dalian Medical University, Dalian, China, [6]Department of Urology, First Affiliated Hospital of Dalian Medical University, Dalian, China, [7]Center for Clinical Research on Neurological Diseases, First Affiliated Hospital, Dalian Medical University, Dalian, China, [8]Research Center of Molecular Diagnostics and Sequencing, Research Institute of Tsinghua University in Shenzhen, Shenzhen, Guangdong, China

Background and Objective: Understanding the tumor microenvironment (TME) and immune cell infiltration (ICI) may help guide immunotherapy efforts for colon cancer (COAD). However, whether ARID1B is truly regulated by hypermethylation or linked to immune infiltration remains unknown. The current work focused on the ARID1B gene expression and methylation in COAD, as well as its relation with ICI.

Methods and Results: Multiple tools based on TCGA were used to analyze the differences in the expression of the ARID1B gene, DNA methylation, and its association with various clinicopathological features, somatic mutations, copy number variation, and the prognosis of patients with COAD. According to the analysis results, patients with high mRNA, low methylation levels showed better overall survival than patients with low mRNA, high methylation levels. The correlation analysis of immune cell infiltration and immune checkpoint gene expression showed that the infiltration rates of the main ICI subtypes, cancer-associated fibroblast, and myeloid cells were significantly enriched and correlated with ARID1B in COAD. An association between ARID1B expression and immune infiltration in COAD was found by correlating ICI indicators with ARID1B expression in the immune cell composition of the COAD microenvironment. Notably, M2 chemokines were related to ARID1B expression, while M1 chemokines were not.

Conclusion: This study provided evidence that ARID1B may have a role in the pathogenesis of COAD. The specific underlying mechanisms that could be responsible for ARID1B's downregulation in COAD will need to be investigated in the future.

KEYWORDS

ARID1B, DNA Methylation, RNA expression, Immune Cell Infiltration, Colon adenocarcinoma, TCGA

Introduction

Colon cancer is a prevalent and worldwide malignant tumour due to the late diagnosis (Wang et al., 2020; Siegel et al., 2022). Colon adenocarcinoma (COAD) is also one of the most cancers that are affected by intertumoral heterogeneity due to immunosuppression factors, and, immune cell subset dysfunction such as M2 macrophage polarization, $CD8^+$ T cells, B cells, and natural killer cells (Zhang et al., 2020a). The main epigenetic modifications of gene expression include DNA methylation, histone modification, chromatin remodelling, and RNA regulation that can influence gene transcription processes related to cell activity, which ultimately produce tumours (Han and Yoon, 2012; Cheng et al., 2021; Zebley et al., 2021). DNA methylation leads to changes in chromatin structure, DNA conformation, DNA stability, and the way that DNA interacts with proteins, thereby controlling gene expression and cause cancer (Li et al., 2014) As mentioned in the literature review, colon cancer is diagnosed in the middle or late stage and an effective early diagnosis of COAD is limited, therefore, it represents a major cause of death worldwide (Li et al., 2014; Zhou et al., 2019). The important analysis of DNA aberrant methylation comes from its frequent and early occurrence during the initial stage of the tumour that can be used as a potential biomarker for cancer detection, monitoring, and timed treatment (Watts et al., 2008; Salta et al., 2018). The tumor microenvironment (TME) is mainly composed of tumour cells, tumour-infiltrating immune cells, and matrix components (Trikha et al., 2016; Jiang et al., 2020). Based on recent studies, tumour immune cell infiltration (ICI) is related to the sensitivity of immunotherapy and the prognosis of cancers including colon cancer (Oda et al., 2018; Zhang et al., 2020b; Sato et al., 2020). Defects in the chromatin remodelling factor, ARID1B, cause extensive dysregulation across different cancer types (Kadoch et al., 2013; Aso et al., 2015; Tessier-Cloutier et al., 2020). However, the prognosis significance of ARID1B and its methylation in colon cancer need to be clarified. Multiple tools based on TCGA were used, and correlation analysis was performed to assess ARID1B expression, CpG methylation, and its association with various clinicopathological features, somatic mutation, copy number variation, and the prognosis of patients with COAD. The correlation analysis was also carried out between ARID1B and the tumour immune infiltration level in COAD.

Materials and methods

ARID1B gene expression and its DNA methylation analysis

The expression level of ARID1B mRNA and protein in normal and tumor COAD tissues was first estimated using the TNMplot online database (https://www.tnmplot.com/) and The Human Protein Atlas (https://www.proteinatlas.org/), respectively (Á and Győrffy, 2021; Uhlen et al., 2017). UALCAN provides promoter DNA methylation data from the TCGA for most of the genes http://ualcan.path.uab.edu/. UALCAN was also used to determine the association between ARID1B methylation and clinicopathological variables in COAD patients, including age, gender, tumor stage, and lymph metastasis (Chandrashekar et al., 2017). MethSurv (https://biit.cs.ut.ee/methsurv/), the third online way, is a web tool to perform multivariable survival analysis using DNA methylation data (Modhukur et al., 2018). Subsequently, the significantly identified probes were tested by univariate and multivariate analysis-based on clinical variables. Additionally, we also explored the expression of DNA methyl transferases (DNMT1and DNMT3A) between ARID1B high and ARID1B low based on TCGA database.

Cell culture and qPCR analysis

Human colonic epithelial cells (HCoEpiC) and colon adenocarcinoma cell lines (HCT116 and LoVo) were cultured, supplied with 5% CO_2, and incubated at 37°C. Extraction of total RNA, synthesis of cDNA, and qPCR conditions using SYPR green analysis reagent were previously published (Baldi et al., 2021). ARID1B forward (CAATGCCACAGGAAAGAGGTTT) and reverse (CTGTCTGTTGAGGTCCATACTGA) primers were utilized.

Survival analysis and clinical value of ARID1B methylation

CanEvolve is a public online tool used to analyze and visualize TCGA clinical and phenotype data. Three gene expression TCGA COAD datasets, including overall survival (OS), disease-free survival (DFS), and disease-specific survival (DSS) were selected from the CanEvolve online tool (http://www.canevolve.org/) (Samur et al., 2013). According to the median cut-off, patients in each dataset were divided into high and low expression groups to reveal the correlation between ARID1B expression and overall survival (OS). Whereas the clinical role of ARID1B methylation in COAD and the relationship between ARID1B methylation and the above-mentioned clinical elements were estimated using UALCAN. Besides, MethSurv [MethSurv-A web tool to perform multivariable survival analysis using DNA methylation data (ut.ee)] also contributed to the connection between position distribution around CpG islands and the prognosis of COAD patients.

Epigenetic analysis of gene promotor regions of ARID1B

SMART tools were used to explore the relationship between ARID1B methylated sites, somatic mutation and copy number variation (Li et al., 2019).

TIMER analysis

The reliability and validity of TIMER2 website in the study of immune cell infiltration are supported by previous researches enabled us to estimate the correlation between ARID1B expression and immune subset infiltration levels in the COAD microenvironment (http://timer.cistrome.org/) (Li et al., 2017). TIMER2 was also conducted to determine the power of correlation between ARID1B and immune cell markers, chemokines, and immunomodulator levels in the TCGA COAD microenvironment.

ARID1B related pathway analysis

GSCAlite using TCGA data was conducted to identify ARID1B pathways in COAD tumor http://bioinfo.life.hust.edu.cn/web/GSCALite/(Liu et al., 2018). On the other hand, UALCAN identified potentially associated and co-expressed genes of ARID1B in the TCGA COAD. The heat map of the top 50 upregulated and downregulated genes was created. To identify ARID1B-associated functions in COAD, we performed Gene Set Enrichment Analysis (GSEA) for each set using OmicBeans tool, (http://www.omicsbean.cn/).

Results

TCGA analysis of the clinical prognostic value of ARID1B

The first set of questions aimed to measure ARID1B expression levels from the TCGA database. Differential analysis of total of 82 RNA-Seq data using TNMplot database showed that ARID1B had a lower mRNA level in tumor COAD tissues (41 patient) compared to normal tissues (41 patient) (Figure 1A). Furthermore, the HPA tool was used for protein expression analysis and ARID1B protein was decreased in colon tumor cells than endothelial cells and glandular cells (Figure 1B).

ARID1B is associated with poor prognosis in colon adenocarcinoma

The next question we asked whether ARID1B low expression has a clinical value for COAD patients. The significance of ARID1B gene expression in determining the overall survival (OS), disease-free survival (DFS), and disease-specific survival (DSS) of COAD patients were assessed, and a low level of ARID1B expression was associated with poor prognosis in all three survival analysis datasets (Figures 1C–E). Colon cancer cell lines showed a low mRNA ARID1B expression compared to colon normal cell line (Figure 1F).

ARID1B is hypermethylated in colon cancer

To explore the reasons for the low expression of ARID1B in colon cancer, the methylation level of the ARID1B gene was analyzed with UALCAN web tool. The expression of the ARID1B gene in colon cancer was hypermethylated in cancer samples (313) than normal samples (Mu et al., 2018) (Figure 2A). Differentially methylated regions were identified and a heat map was created (Figure 2B). Interestingly, hypermethylation of ARID1B was correlated with the downregulation of ARID1B expression. Additionally, the expression of ARID1B was positively correlated with the expression of three methyl transferases (DNMT1, DNMT3A, and DNMT3L) in colon cancer, although the difference was not significant (Figures 2C–E).

ARID1B methylation is related to the clinicopathological variables and prognosis of COAD

We analyzed the differences in DNA methylation levels in the TCGA COAD cohort and determined the association between ARID1B methylation and clinicopathological variables, including age, gender, tumor stage, lymph metastasis, distant metastasis, and clinical stage. Compared with the normal group, ARID1B was hypermethylated and was associated with all indicated variables as well as clinical with each stage (Figure 3). Significantly, the hypermethylation of ARID1B predicted a shorter overall survival, thus hypomethylation of the ARID1B gene is conducive to survival (Figure 4). This is an interesting consistent result. Univariate and multivariate Cox regression was performed and the data presented in Tables 1, 2 showed that the prognostic significance of hypermethylation was an independent factor of clinicopathological characteristics including age, gender, race, and stage.

Functional enrichment and PPI analysis

Based on the functional enrichment and correlation analysis with the representative molecules of the pathway, the expression

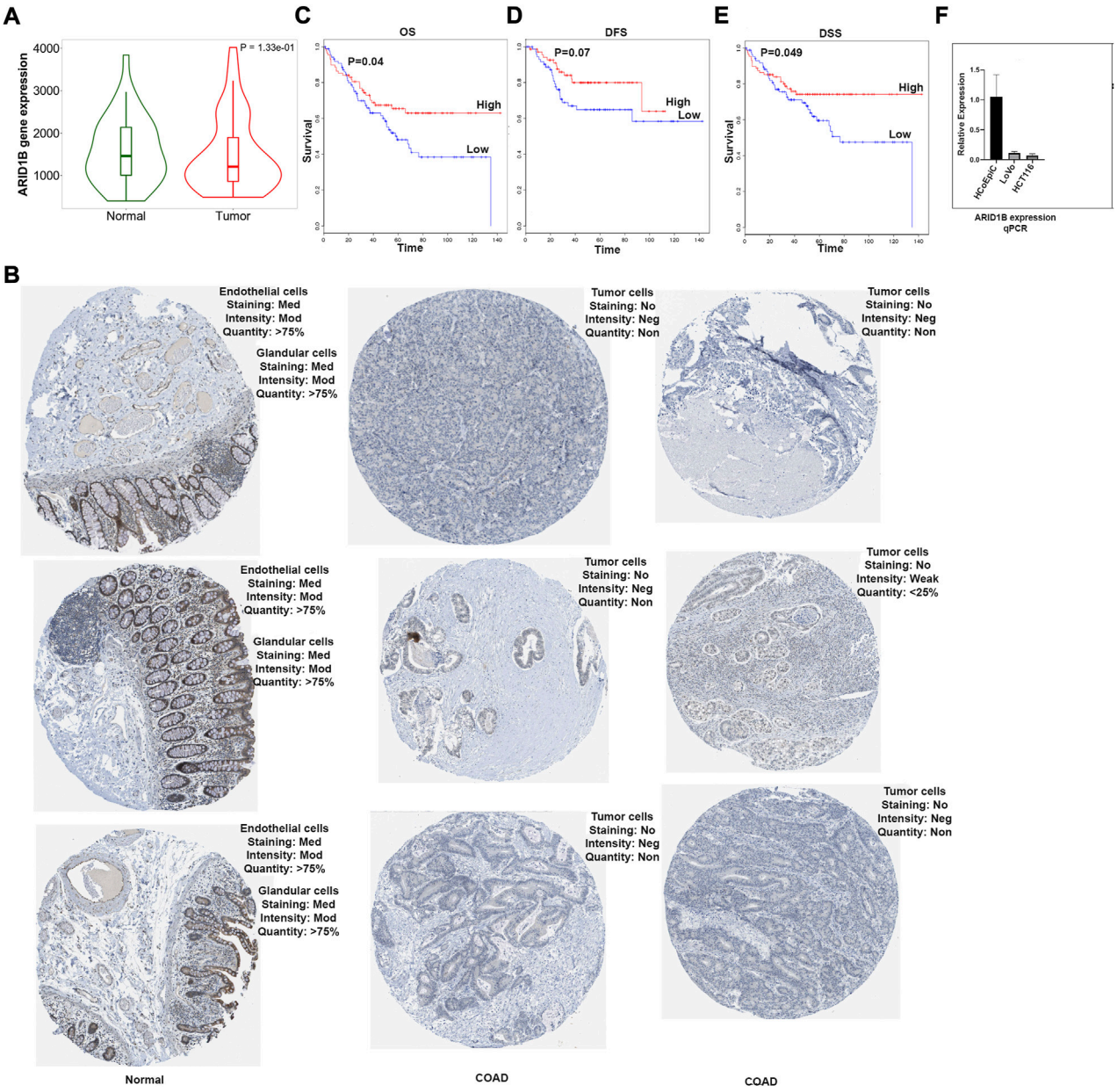


FIGURE 1
ARID1B mRNA analysis results. **(A)** ARID1B expression from TCGA, **(B)** ARID1B protein expression from HPA, **(C–E)** Kaplan Meir blot of Overall Survival, Disease Free Survival, and Disease Specific Survival. **(F)** qPCR results, ARID1B is downregulated in cancer cell line compared to normal cell line. HPA, Human Protein Atlas; OS, overall survival; DFS, disease-free survival, DSS; disease-specific survival; Med, Medium; Mod, Moderate; Neg, Negative; Non, None.

of the ARID1B gene was related to the PI3k/AKT, RTK, NA damage response, Hormone AR, RAS/MAPK, and TSC/mTOR pathways while it was inhibited in apoptosis, cell cycle, and EMT pathways (Figure 5A). Furthermore, the ARID1B gene exhibits high expression when the driving genes (APC and FBXW7) are mutated (Figure 5B). Contrary, there was a list of mutated gene decreased ARID1B expression in COAD (Figure 5C).

ARID1B methylated sites are associated with somatic variations

The association was evaluated between somatic mutation, CNV, and DNA methylation. Results showed a number of CpG sites associated with somatic mutation and copy number variations (Figure 6A,B).

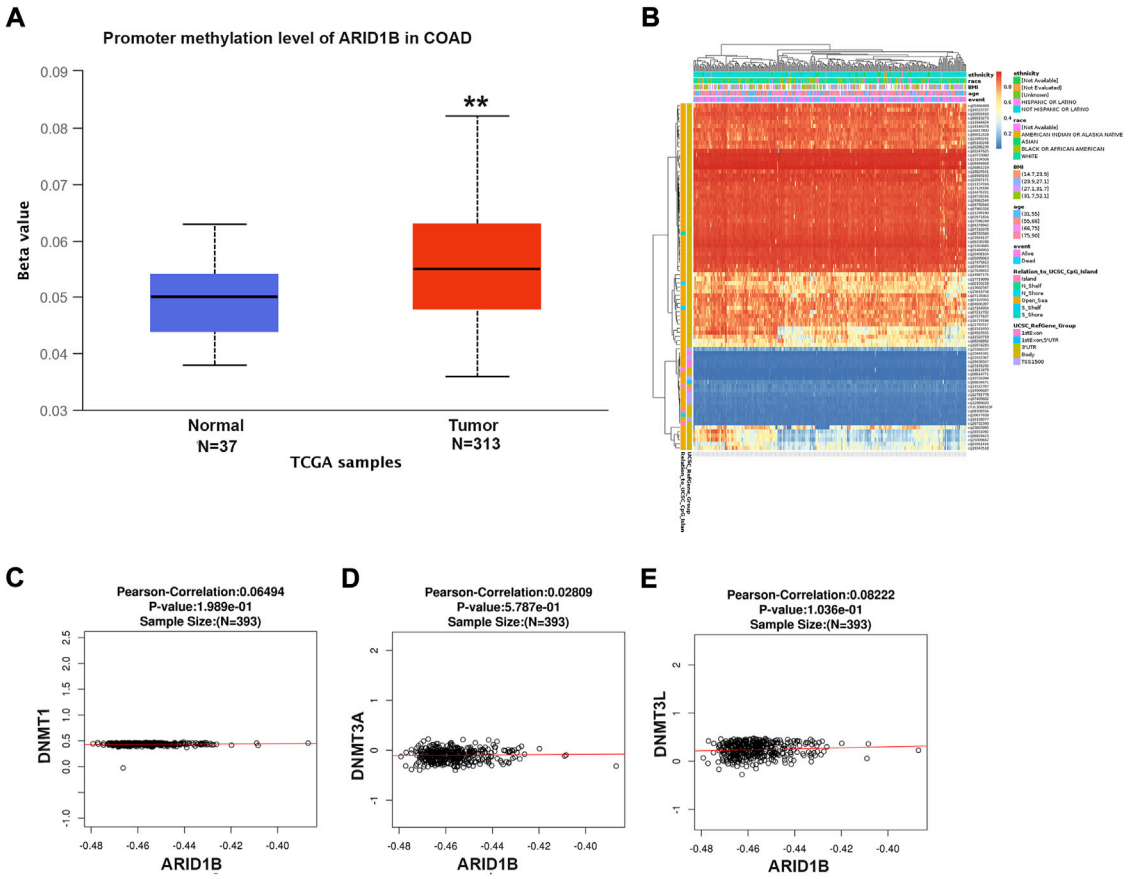


FIGURE 2
ARID1B methylation level analysis. **(A)** ARID1B methylation level from UALCAN **(B)** Different CpG probs of ARID1B from Methsurv **(C–E)** The positive correlation of ARID1B and methyltransferases. N, Normal; T, Tumor; S, Stage; Cau, Caucasian; Afr-Amr, African, American; Asi, Asian; M, Male, F, Female; y, year; N-W, Normal Wight; E-W, Extreme Wight; Ob, Obese; E-B, Extreme Obese; Ad, Adenocarcinoma; Muc, Mucinous Adenocarcinoma.

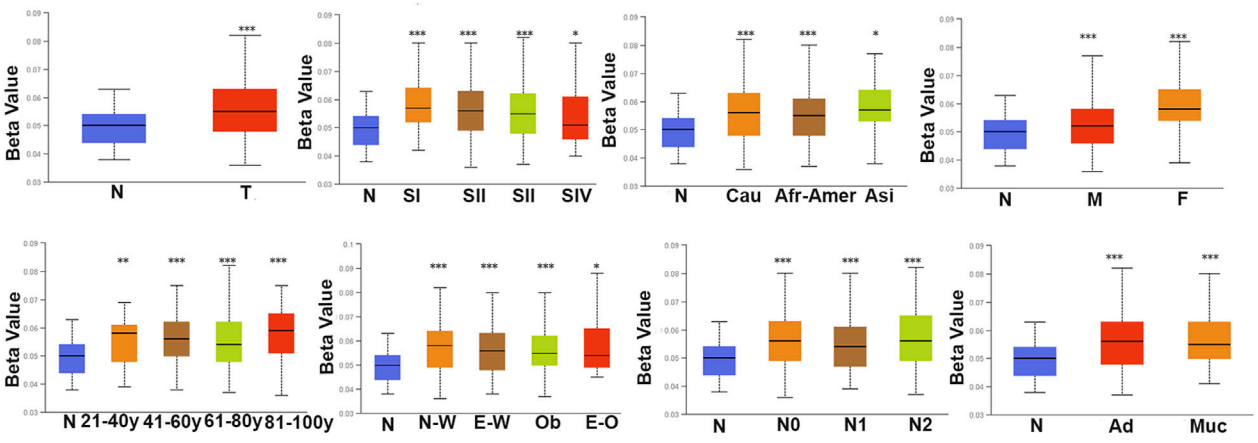


FIGURE 3
The methylation level of ARID1B across different clinicopathological parameters in COD patients. There is a significant difference between the defined subgroups containing gender, age, race, histological subtype, tumor grade, and nodal metastasis status of patients.

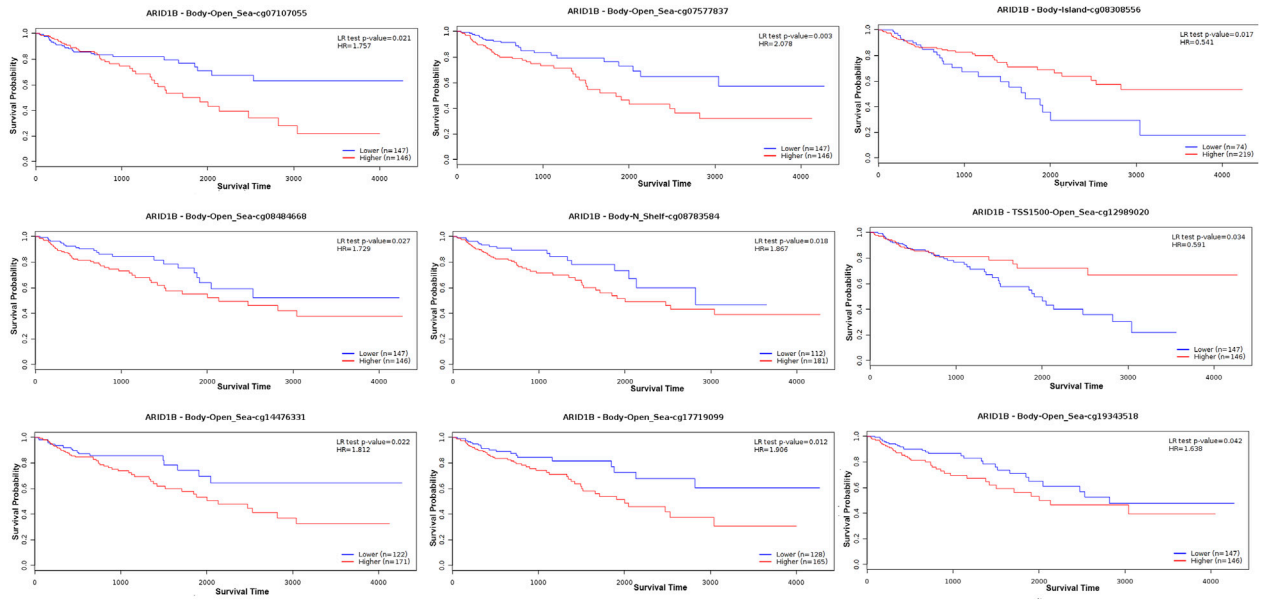


FIGURE 4
ARID1B hypermethylation impacts the prognosis of COAD patients.

TABLE 1 Univariate analysis.

Overall survival

Probe	Hazards. Ratio	CI.Lower	CI.Upper	*p*.value	Z
cg07107055	2.423	0.2873	20.432	0.4159	0.8136
cg07577837	18.6585	0.96	362.6319	0.0532	1.933
cg08484668	17109.36	4.00E-04	7.13E+11	0.2762	1.0888
cg08783584	11971363	4.4493	3.22E+13	0.031	2.1576
cg14476331	253056	0.9222	6.94E+10	0.0515	1.9473
cg17719099	3.4216	0.6514	17.9708	0.1461	1.4536
cg19343518	2.7602	0.7027	10.8419	0.1458	1.4545
cg21793517	181.0568	8.2055	3995.068	0.001	3.2933
cg23603995	1.2343	0.4007	3.8022	0.7138	0.3667

T lymphocyte infiltrates are associated with ARID1B expression

Surprisingly, as presented in Figure 7, there was a negative correlation of immune cells through TIMER analysis, including T cell, $CD8^+$, T cell, $CD4^+$, and NK cells and a positive correlation with B cells, macrophages, and neutrophils as well as tumor-associated fibroblast, dendritic cells, and cell regulatory cell in COAD. These data indicate that T cells $CD8^+$ tend to be depleted in the COAD microenvironment and that ARID1B may play a role in the COAD microenvironment. To verify these results, we performed a Pearson correlation between markers and ARID1B expression. Strong evidence of correlation analysis was found between a variety of the biomarkers and ARID1B expression in the COAD microenvironment (Table 3). These findings emphasized that ARID1B may regulate immune cell infiltration in COAD. Dendritic cells (DCS) are key participants in the antigen-specific immune response. Two different DC subsets (XCR1, CADM1, cDC1, and CD1A) and (CD172A+ cDC2) have been identified, which interact with $CD8^+$ and $CD4^+$ T cells, respectively. The current study detected evidence for some DCs including XCR1, CADM1, and CD1A in COAD.

TABLE 2 Multivariate analysis.

Variable	Coef	H	Z	95%CI lower	95%CI upper	*p*.value
Age	0.027525	1.027907	2.228524	1.003323	1.053095	0.025846
GenderMALE	0.178965	1.195979	0.586209	0.657443	2.175651	0.557735
RaceBLACK OR AFRICAN AMERICAN	0.252757	1.28757	0.281233	0.221186	7.495211	0.778531
RaceWHITE	−0.15422	0.857081	−0.18606	0.16884	4.350798	0.852397
StageStage II	0.75889	2.135903	1.185325	0.60899	7.491226	0.235889
StageStage III	0.971959	2.643117	1.517938	0.753493	9.271579	0.12903
StageStage IV	2.550279	12.81068	3.915276	3.573783	45.92153	9.03E-05
cg07107055	−2.70048	0.067173	−1.92926	0.004322	1.043918	0.053699
cg07577837	0.921898	2.514057	0.448105	0.044584	141.766	0.654077
cg08484668	0.227981	1.256061	0.023353	6.16E-09	2.56E+08	0.981368
cg08783584	6.649516	772.4103	0.694165	5.42E-06	1.1E+11	0.487579
cg14476331	9.892837	19788.12	1.269733	0.004618	8.48E+10	0.20418
cg17719099	0.271772	1.312288	0.244605	0.148692	11.58163	0.806762
cg19343518	1.370621	3.937793	1.528212	0.678936	22.83901	0.12646
cg21793517	3.309323	27.36659	1.517419	0.380929	1966.063	0.129161
cg23603995	0.059191	1.060978	0.087332	0.281051	4.00523	0.930408

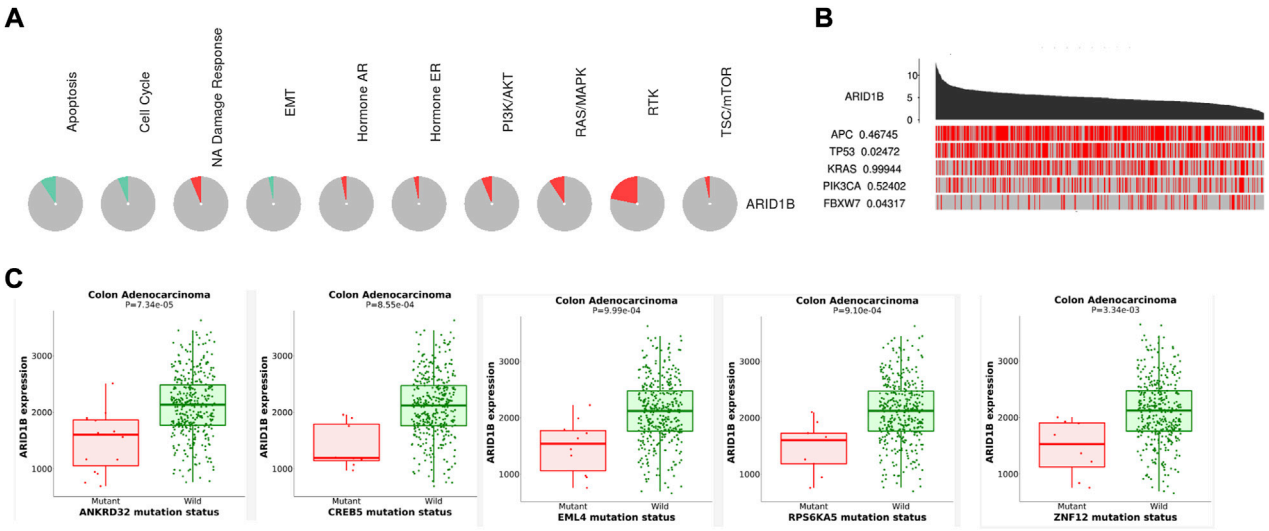


FIGURE 5
Enrichment pathways of ARID1B. **(A)** ARID1B is inhibited and activated in pathways indicated with green and red colors, respectively **(B)** A permutation test *p*-value of ARID1B between driver mutated and non-mutated samples **(C)** Analysis correlation between ARID1B expression and some mutated genes that cause a low level of ARID1B was explored.

ARID1B is associated with M2 macrophages, chemokines

Since the positive correlation was found between ARID1B and TAMs biomarkers, particularly M2-macrophages, and verified these findings, we immediately studied the correlation between ARID1B and the classical chemokines of both M1-macrophages (IL-12, IL-23, TNF, IFNG) and M2 macrophages (IL-10, TGF-b, IL-4, IL-13). The analysis whose reported positive results also reported a significant

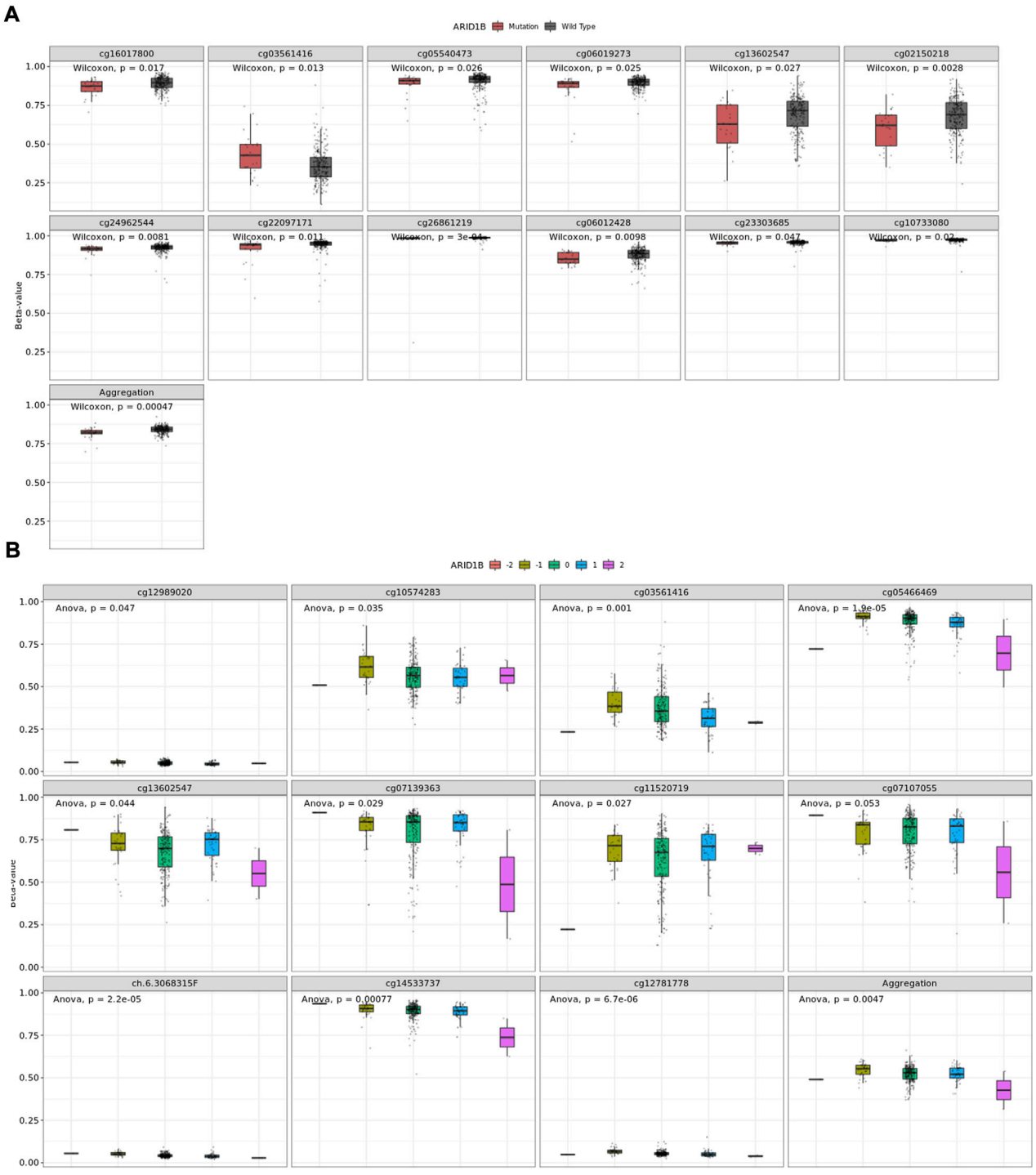


FIGURE 6
A number of CpGs are associate with ARID1B alterations **(A)** CpGs associated with somatic mutation, **(B)** CpGs associate with CNA.

positive correlation with M2 chemokines (TGF-B1, TGF-B3). Whereas none of the M1 macrophage chemokines were associated with ARID1B expression in the COAD microenvironment (Table 3). Therefore, these results support that ARID1B is closely related to M2-macrophages in the microenvironment of COAD.

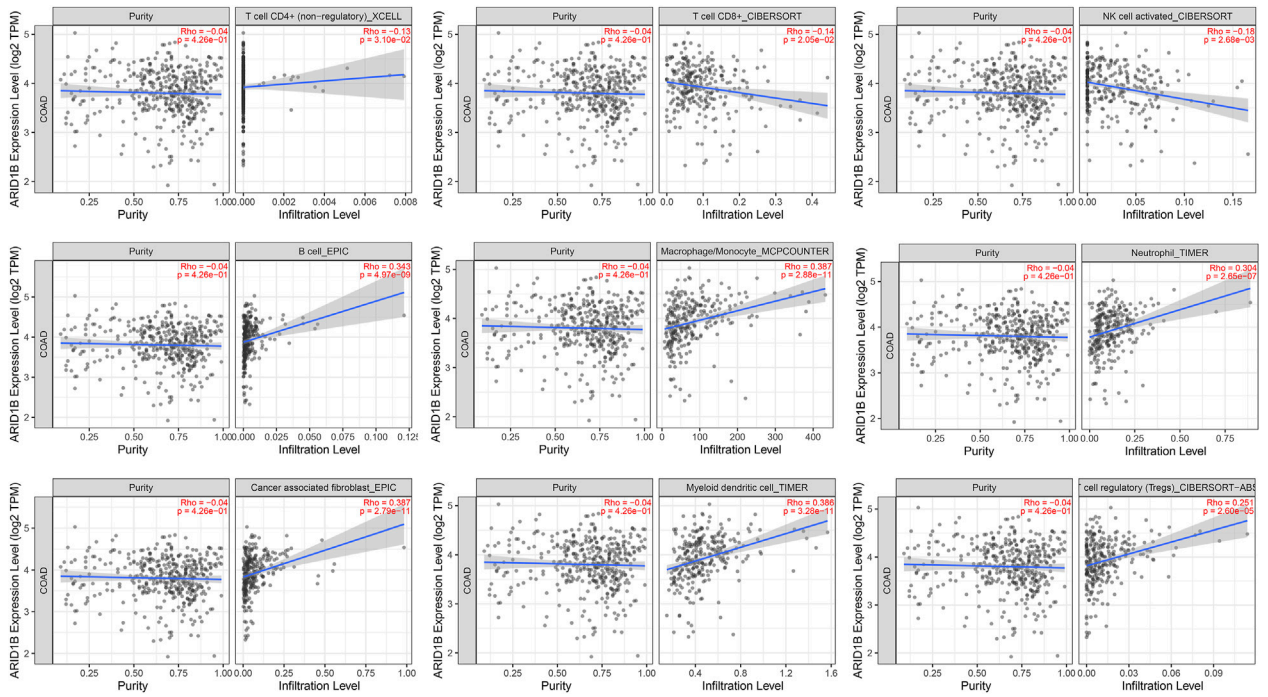


FIGURE 7
The correlation between ARID1B and immune cell infiltration. The significant negative-positive correlation between ARID1B and immune cell subtypes.

Checkpoint inhibitors are significantly correlated with ARID1B in the colon adenocarcinoma tumor microenvironment

A previous study reported that tumors may achieve immune escape by inhibiting the activity of CTLs. This is achieved by immunoinhibitory molecules that bind T cell surface ligands and inactivate its function. Therefore, we evaluated the immunomodulators level related to ARID1B expression in COAD and found a positive correlation of different immunosuppressors, including HAVCR2 (TIM3), LAG3, PDCD1 (PD-1), TIGIT, CD27, CTLA-4, and TNFRSF9 (markers of T cell failure) (Table 3). These results indicate that T cells may tend to be inactivated in the COAD microenvironment.

ARID1B is related to the signal pathway that regulates immune cells

To better understand the transcriptomic phenotype associated with infiltrated COAD tumors, ARID1B gene expression data were obtained *via* UALCAN online tool and applied gene set enrichment analysis of positive and negative correlated genes. Interestingly, RNA-seq data validated that ARID1B upregulated several immune pathways in colon cancer tissue. The enrichment results show that in ARID1B positively correlated genes, the main enrichment pathways are: (Wang et al., 2020) NF-κB signaling pathway; (Siegel et al., 2022) TNF signaling pathway; (Zhang et al., 2020a) TGF-beta signaling pathway; (Cheng et al., 2021) VEGF signaling pathway; (Han and Yoon, 2012) Wnt signaling pathway; (Zebley et al., 2021) Rap-1 signaling pathway; (Li et al., 2014) PI3K-Akt signaling pathway; (Zhou et al., 2019) mTOR signaling pathway; (Watts et al., 2008) JAK-STAT signaling pathway innate immune response; (Salta et al., 2018) ECM receptor interaction. The main enrichment pathways in ARID1B negatively associated genes were: (Wang et al., 2020) metabolic pathway; (Siegel et al., 2022) RNA polymerase, Huntington disease, Parkinson disease, Alzheimer disease; (Zhang et al., 2020a) Thermogenesis; (Cheng et al., 2021) non-alcoholic fatty liver disease; (Han and Yoon, 2012) spliceosome; (Zebley et al., 2021) oxidative phosphorylation; (Li et al., 2014) retrograde endocannabinoid signaling; (Zhou et al., 2019) oxidative phosphorylation (Figures 8A–D).

Discussion

Among the different types of molecule abnormalities, mRNA changes and DNA methylation, which are reported widely in

TABLE 3 Correlation between ARID1B and immune cell biomarkers in COAD.

Infiltrates type	Biomarker rho *P* value adj.p
T cell general	CD2 0.113171833 0.022569163 0.052159843
	CD3E 0.168758263 0.000639177 0.002480389
T cell CD8$^+$	CD8A 0.079836038 0.108217539 0.191405171
B cell	CD19 0.185617928 0.000169031 0.000796579
Treg cell	CD79A 0.233264327 2.02E-06 1.40E-05
	FOXP3 0.333580951 5.22E-12 3.29E-11
	STAT5B 0.641711169 1.72E-48 8.58E-47
	TGFB1 0.226418058 4.06E-06 1.37E-05
	CCR8 0.390194392 3.25E-16 2.98E-15
Natural killer cell	KIR3DL2 0.108363891 0.029023119 0.071883887
Neutrophils	CCR7 0.299913922 6.96E-10 1.05E-08
	ITGAM 0.292339406 1.93E-09 2.41E-08
Tumor asoociated macrohage (TAM)	CCL2 0.164064315 0.000906246 0.004096026
	CD68 0.25002831 3.33E-07 2.93E-06
	IL10 0.087827792 0.07712079 0.159422822
M1 macrophage	IRF5 0.175941418 0.000368018 0.001840091
M2 macrophage	CD163 0.299282134 7.59E-10 1.10E-08
Monocyte	MS4A1 0.22754202 3.63E-06 2.62E-05
	VSIG4 0.134237336 0.006753968 0.022053772
	VEGFA 0.349636033 4.06E-13 6.50E-12
	CD86 0.197668751 6.06E-05 0.000362082
	CSF1R 0.35744943 1.11E-13 3.70E-12
Dendiritic cell	HLA-DPA1 0.144785374 0.003458199 0.010703949
	HLA-DPB1 0.130777417 0.008332108 0.022218954
	TGAX 0.313424074 1.05E-10 1.89E-09
	NRP1 0.438944503 1.50E-20 7.09E-19
	CD1C 0.217792469 9.51E-06 5.49E-05
	ITGAX 0.313424074 1.05E-10 3.13E-09
T cell exhaustion	GZMB 0.033038754 0.506791971 0.599753812
	HAVCR2 0.160201284 0.001199674 0.002864894
	CTLA4 0.19495873 7.68E-05 0.000220185
	PDCD1 0.130017099 0.00872011 0.017353452
	CD27 0.183807214 0.000196114 0.00130743
Th 1	STAT1 0.279103955 1.06E-08 4.72E-08
	STAT4 0.167092289 0.000724253 0.001788279
	TNF 0.085459081 0.085470431 0.134599104
	IFNG 0.010326018 0.835675989 0.886658875
	TBX21 0.228770674 3.20E-06 1.10E-05
Th 2	GATA3 0.295890265 1.20E-09 5.96E-09
	STAT5A 0.357832565 1.04E-13 7.79E-13
	STAT6 0.42689993 2.06E-19 2.46E-18
Th 17	STAT3 0.520078905 1.63E-29 3.63E-28
Tfh	BCL6 0.465387725 3.26E-23 4.74E-22
Fibroblast	COL1A1 0.366780269 2.25E-14 6.01E-13
	FABP5 -0.383429108 1.14E-15 3.43E-14
	NNMT 0.133761138 0.006953953 0.017754774
	PDGFB 0.424214929 3.64E-19 2.91E-17

(Continued on following page)

TABLE 3 (*Continued*) Correlation between ARID1B and immune cell biomarkers in COAD.

Infiltrates type	Biomarker rho *P* value adj.p
M2 motif	BACH1 0.516162622 5.04E-29 3.25E-28
	MAF 0.423837711 3.94E-19 1.41E-18
	MAFG 0.521972992 9.41E-30 6.72E-29
	NFE2L2 0.517550043 3.39E-29 2.34E-28
	STAT2 0.461586024 8.14E-23 3.46E-22
Myloid cell	AREG 0.01689488 0.734310642 0.803393173
	CCL4 0.007207503 0.884883559 0.909905973
	CXCL2 -0.15809741 0.001394077 0.003620979
	CXCR4 0.244853698 5.89E-07 2.62E-06
	EGR1 0.304975 3.47E-10 2.24E-09
	IL1B-0.028009173 0.573612239 0.670891508
	NFKB1 0.51420058 8.83E-29 2.35E-27
	NFKB2 0.334349845 4.64E-12 3.79E-11
	NFKBIA 0.16438705 0.000885009 0.002391917
	NLRP3 0.344254286 9.72E-13 8.45E-12

almost all cancers. The main purpose of this study was to analyse the promotor methylation of ARID1B through preliminary difference analysis, correlation analysis, and prognosis analysis. Despite very little was found in the literature on the analysis of ARID1B expression and its DNA methylation in cancer, none of these studies was related to colon cancer (Aso et al., 2015; Cui et al., 2019; Tessier-Cloutier et al., 2020). Previous studies observed an increased expression of ARID1B in triple-negative breast cancer and breast invasive ductal carcinoma (Shao et al., 2015; Cui et al., 2019). In the current study, we evaluated DNA methylation and gene expression profiles of ARID1B in colon cancer samples from TCGA. The expression of ARID1B in TCGA COAD and normal tissues were first analysed using the TNMplot platform. To confirm the observed results experimentally, RNA from the colon cell line was extracted and detected by qPCR analysis. This method was repeated three times to ensure reliable results. Consistent with the significant downregulation of ARID1B expression in colon cell lines and TCGA tissues, ARID1B was hypermethylated in COAD tissues. Comparison of these findings with those of previous studies confirmed silenced expression of genes, especially ARID1B by aberrant methylation of DNA in cancer (Chuang and Jones, 2007; Khursheed et al., 2013; Liu et al., 2020; Tan et al., 2020). AKR1C2 mRNA, a gene previously identified as a Nrf2-target, was not found in HCT116 cells (Ebert et al., 2011). Importantly, survival analysis showed that the overall survival rate of patients in the low mRNA expression group was significantly shorter than that in the low-level group. ARID1B methylation level was related to clinicopathological characteristics and showed a significant negative correlation with its mRNA level in COAD. Besides, the prognostic analysis of a number of methylated sites was significantly associated with the patients' outcome in COAD patients. These consistent results of aberrantly silenced ARID1B provide strong evidence that low ARID1B expression is an important indicator of the poor prognosis of COAD and the epigenetic changes might be a potentially increased risk of colon cancer-related death. Knowing that DNMTs enzymes are responsible for establishing, maintaining, and mediating DNA methylation, we performed correlation analysis, and ARID1B was positively correlated with DNMTS in COAD (Chuang and Jones, 2007). The third question in this research aimed to determine the pathways of ARID1B expression on the basis of TCGA database. GSEA revealed that ARID1B is involved in multiple cancer-related pathways, among others PI3K-AKT and mTOR pathways, suggesting that ARID1B regulates these pathways and functions as a tumour suppressor in COAD.

Tumour infiltrates immune cells are an important immune cell type in the tumor microenvironment and have been reported to affect patients' overall survival (Kryczek et al., 2016; Yang et al., 2018). For this and to further determine the association between ARID1B and COAD, we used TIMER database to analyse the correlation between the expression of ARIDIB and immune cell infiltration. Immune cells and fibroblast infiltration demonstrated a link with ARID1B in the COAD microenvironment. Consistent with the obtained results, ARID1B showed a positive correlation with immune cell markers. It was reported that $CD8^+$ T cells could be inhibited by immunoinhibitory molecules *via* binding to the receptor PD-1 on the surface of $CD8^+$ T cells (Yang et al., 2018). TIMER analysis revealed a positive correlation between ARID1B and T-exhausted cell immunoinhibitory

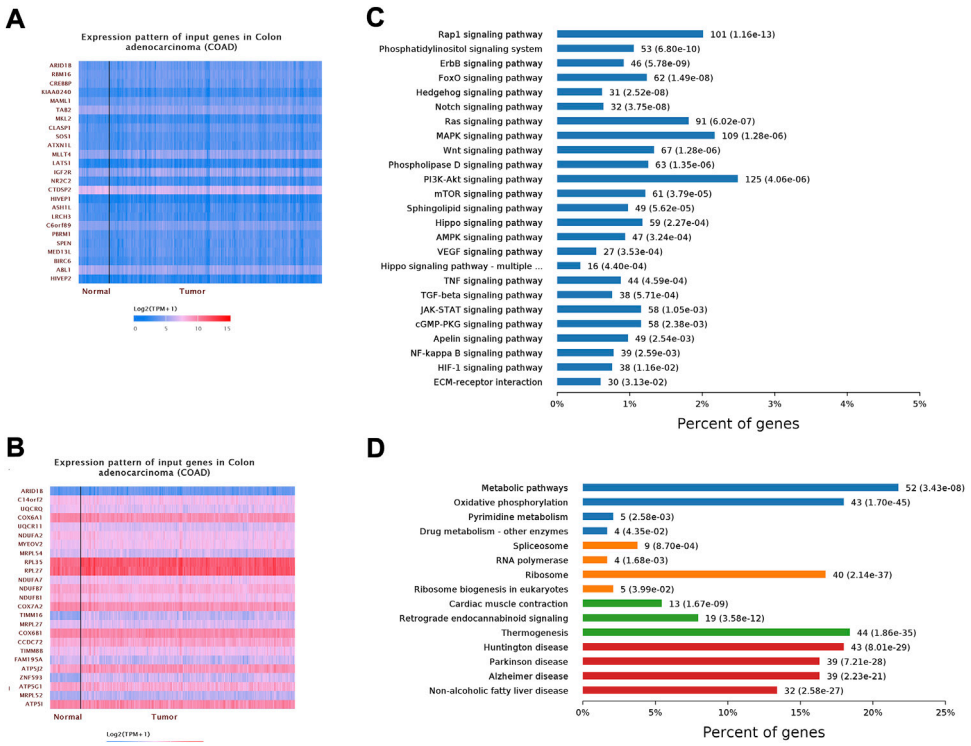


FIGURE 8
Shows gene set enrichment analysis of cooccurred genes. **(A)** Positively correlated genes **(B)** Negatively regulated genes **(C)** KEGG pathway **(D)** GO analysis.

molecules. The positive correlation of ARID1B and macrophages and depleted $CD8^+$ markers indicates the contribution of macrophages to the formation of the immunosuppressive microenvironment in COAD. Recently, The M2-polarization of TAMs was recognized as an immunosuppressive phenotype (Mu et al., 2018). Moreover, scientists found that monocytes recruited to tumour regions are reprogrammed to become tumour-associated macrophages. And the activity of motifs MAF, STAT1, STAT2 leads to M2 polarization, which provides potential targets for inhibiting or reversing the information of the immunosuppressive microenvironment (Zheng et al., 2021). In addition to TILs, ARID1B was positively correlated with TIM, including LAMP3, cDCs, LYVE1, resident tissue macrophages (RTMs), and FOLR2, TAM. In accordance with this result, recent work reported multiple tumour types of TIM cell subpopulations across the pan-cancer analysis (Cheng et al., 2021).

Finally, we explored the signal pathways that might regulate the immune cell infiltration and polarization of M2 macrophages in COAD. TGF-β signalling pathway is a cytokine signalling pathway involved in the development and progression of COAD. Hu et al. found that the transcription factors STAT3, HIVEP, NFAT, and other regulated genes are upregulated in depleted $CD8^+$ T cells. This result provides clues for identifying new candidate transcription factors for T cell dysfunction. Furthermore, STAT3 not only participates in the polarization of macrophages to M2 but also participates in the depletion of T cells, suggesting that STAT3 inhibition can be used as a new therapeutic strategy for the treatment of cancer (Hu et al., 2020). A pathway enrichment analysis of ARID1B-correlated genes also indicated the involvement of the STAT-JAK signalling pathway, which suggest that the TGF-β1/STAT-JAK pathways are associated with aggressive pathological features and poor clinical outcomes in COAD. While this report describes the function of ARID1B and its methylation in COAD, it also provides preliminary evidence about the role of ARID1B in the microenvironment of COAD immune cell infiltration, and M2 polarization and suggests that ARID1B may influence COAD immunotherapy. Our study has some limitation. The involvement of ARID1B in COAD was identified *via* bioinformatics. Even though the association between ARID1B and DNMTS expressions in COAD was not statistically significant, it remained unknown whether ARID1B was regulated by DNTMs enzymes. However, future experimental studies on the current topic are therefore recommended. In conclusion, the article comprehensively analysed ARID1B abnormalities (mRNA changes, DNA aberrant methylation) and its association with the immune cell infiltration of COAD.

The analysis revealed that ARID1B hypermethylation could serve as an early diagnostic biomarker for COAD treatment, and the difference in immune cell infiltration was found to be related to the ARID1B expression of the COAD tumour. In conclusion, the article comprehensively analysed ARID1B abnormalities (mRNA changes, DNA aberrant methylation) and its association with the immune cell infiltration of COAD. The analysis revealed that ARID1B hypermethylation could serve as an early diagnostic biomarker for COAD treatment, and the difference in immune cell infiltration was found to be related to the ARID1B expression of the COAD tumour.

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

SB; conceived the ideas, designed the study, and wrote the original draft. YH and II validated the results, MS, MA-N, and MA; performed statistical analysis, AS and AA-D.; collected literature review, YG funded the study, HT; reviewed and edited the paper. All authors discussed the results and contributed to the final manuscript.

Funding

The work is paid by 2021Szvup096, Shenzhen Science and Technology Innovation Strategic Research and Technology Transfer Promotion Center. This work is also supported by Grant no, LHTD20170002, Science and Technology Innovation Bureau of Nanshan district, Shenzhen.

Conflict of interest

Authors SB, YH, II, YG, and HT are employed by Research Center of Molecular Diagnostics and Sequencing, Axbio Biotechnology (Shenzhen) Co., Ltd. Shenzhen.

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

Publisher's note

References

Á, B., and Győrffy, B. (2021). TNMplot.com: A web tool for the comparison of gene expression in normal, tumor and metastatic tissues. *Int. J. Mol. Sci.* 22 (5), 2622. doi:10.3390/ijms22052622

Aso, T., Uozaki, H., Morita, S., Kumagai, A., and Watanabe, M. (2015). Loss of ARID1A, ARID1B, and ARID2 expression during progression of gastric cancer. *Anticancer Res.* 35 (12), 6819–6827.

Baldi, S., Khamgan, H., Qian, Y., Wu, H., Zhang, Z., Zhang, M., et al. (2021). Downregulated ARID1A by miR-185 is associated with poor prognosis and adverse outcomes in colon adenocarcinoma. *Front. Oncol.* 11, 679334. doi:10.3389/fonc.2021.679334

Chandrashekar, D. S., Bashel, B., Balasubramanya, S. A. H., Creighton, C. J., Ponce-Rodriguez, I., Chakravarthi, B., et al. (2017). Ualcan: A portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia (New York, NY)* 19 (8), 649–658. doi:10.1016/j.neo.2017.05.002

Cheng, S., Li, Z., Gao, R., Xing, B., Gao, Y., Yang, Y., et al. (2021). A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells. *Cell.* 184 (3), 792–809.e23. e23. doi:10.1016/j.cell.2021.01.010

Chuang, J. C., and Jones, P. A. (2007). Epigenetics and microRNAs. *Pediatr. Res.* 61, 24R–29R. doi:10.1203/pdr.0b013e3180457684

Cui, Y., Bai, X., Niu, M., Qin, Y., Zhang, X., and Pang, D. (2019). Upregulated expression of AT-rich interactive domain-containing protein 1B predicts poor prognosis in patients with triple-negative breast cancer. *Oncol. Lett.* 17 (3), 3289–3295. doi:10.3892/ol.2019.9961

Ebert, B., Kisiela, M., Wsól, V., and Maser, E. (2011). Proteasome inhibitors MG-132 and bortezomib induce AKR1C1, AKR1C3, AKR1B1, and AKR1B10 in human colon cancer cell lines SW-480 and HT-29. *Chem. Biol. Interact.* 191 (1-3), 239–249. doi:10.1016/j.cbi.2010.12.026

Han, J. W., and Yoon, Y. S. (2012). Epigenetic landscape of pluripotent stem cells. *Antioxid. Redox Signal.* 17 (2), 205–223. doi:10.1089/ars.2011.4375

Hu, J., Chen, Z., Bao, L., Zhou, L., Hou, Y., Liu, L., et al. (2020). Single-cell transcriptome analysis reveals intratumoral heterogeneity in ccRCC, which results in different clinical outcomes. *Mol. Ther.* 28 (7), 1658–1672. doi:10.1016/j.ymthe.2020.04.023

Jiang, X., Wang, J., Deng, X., Xiong, F., Zhang, S., Gong, Z., et al. (2020). The role of microenvironment in tumor angiogenesis. *J. Exp. Clin. Cancer Res.* 39 (1), 204. doi:10.1186/s13046-020-01709-5

Kadoch, C., Hargreaves, D. C., Hodges, C., Elias, L., Ho, L., Ranish, J., et al. (2013). Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* 45 (6), 592–601. doi:10.1038/ng.2628

Khursheed, M., Kolla, J. N., Kotapalli, V., Gupta, N., Gowrishankar, S., Uppin, S. G., et al. (2013). ARID1B, a member of the human SWI/SNF chromatin remodeling complex, exhibits tumour-suppressor activities in pancreatic cancer cell lines. *Br. J. Cancer* 108 (10), 2056–2062. doi:10.1038/bjc.2013.200

Kryczek, I., Wang, L., Wu, K., Li, W., Zhao, E., Cui, T., et al. (2016). Inflammatory regulatory T cells in the microenvironments of ulcerative colitis and colon carcinoma. *Oncoimmunology* 5 (8), e1105430. doi:10.1080/2162402X.2015.1105430

Li, M., Zhang, J., Liu, Q., Wang, J., and Wu, F. X. (2014). Prediction of disease-related genes based on weighted tissue-specific networks by using DNA methylation. *BMC Med. Genomics* 7, S4. doi:10.1186/1755-8794-7-S2-S4

Li, T., Fan, J., Wang, B., Traugh, N., Chen, Q., Liu, J. S., et al. (2017). Timer: A web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res.* 77 (21), e108–e110. doi:10.1158/0008-5472.CAN-17-0307

Li, Y., Ge, D., and Lu, C. (2019). The SMART app: An interactive web application for comprehensive DNA methylation analysis and visualization. *Epigenetics Chromatin* 12 (1), 71. doi:10.1186/s13072-019-0316-3

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. *Bioinforma. Oxf. Engl.* 34 (21), 3771–3772. doi:10.1093/bioinformatics/bty411

Liu, J., Liu, Z., Zhang, X., Yan, Y., Shao, S., Yao, D., et al. (2020). Aberrant methylation and microRNA-target regulation are associated with downregulated NEURL1B: A diagnostic and prognostic target in colon cancer. *Cancer Cell. Int.* 20, 342. doi:10.1186/s12935-020-01379-5

Modhukur, V., Iljasenko, T., Metsalu, T., Lokk, K., Laisk-Podar, T., and Vilo, J. (2018). MethSurv: A web tool to perform multivariable survival analysis using DNA methylation data. *Epigenomics* 10 (3), 277–288. doi:10.2217/epi-2017-0118

Mu, X., Shi, W., Xu, Y., Xu, C., Zhao, T., Geng, B., et al. (2018). Tumor-derived lactate induces M2 macrophage polarization via the activation of the ERK/STAT3 signaling pathway in breast cancer. *Cell. cycleGeorget. Tex)* 17 (4), 428–438. doi:10.1080/15384101.2018.1444305

Oda, K., Hamanishi, J., Matsuo, K., and Hasegawa, K. (2018). Genomics to immunotherapy of ovarian clear cell carcinoma: Unique opportunities for management. *Gynecol. Oncol.* 151 (2), 381–389. doi:10.1016/j.ygyno.2018.09.001

Salta, S., Nunes, S. P., Fontes-Sousa, M., Lopes, P., Freitas, M., Caldas, M., et al. (2018). A DNA methylation-based test for breast cancer detection in circulating cell-free DNA. *J. Clin. Med.* 7 (11), E420. doi:10.3390/jcm7110420

Samur, M. K., Yan, Z., Wang, X., Cao, Q., Munshi, N. C., Li, C., et al. (2013). canEvolve: a web portal for integrative oncogenomics. *PLoS One* 8 (2), e56228. doi:10.1371/journal.pone.0056228

Sato, J., Kitano, S., Motoi, N., Ino, Y., Yamamoto, N., Watanabe, S., et al. (2020). CD20(+) tumor-infiltrating immune cells and CD204(+) M2 macrophages are associated with prognosis in thymic carcinoma. *Cancer Sci.* 111 (6), 1921–1932. doi:10.1111/cas.14409

Shao, F., Guo, T., Chua, P. J., Tang, L., Thike, A. A., Tan, P. H., et al. (2015). Clinicopathological significance of ARID1B in breast invasive ductal carcinoma. *Histopathology* 67 (5), 709–718. doi:10.1111/his.12701

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *Ca. Cancer J. Clin.* 72 (1), 7–33. doi:10.3322/caac.21708

Tan, Y., Zhang, S., Xiao, Q., Wang, J., Zhao, K., Liu, W., et al. (2020). Prognostic significance of ARL9 and its methylation in low-grade glioma. *Genomics* 112 (6), 4808–4816. doi:10.1016/j.ygeno.2020.08.035

Tessier-Cloutier, B., Schaeffer, D. F., Bacani, J., Marginean, C. E., Kalloger, S., Köbel, M., et al. (2020). Loss of switch/sucrose non-fermenting complex protein expression in undifferentiated gastrointestinal and pancreatic carcinomas. *Histopathology* 77 (1), 46–54. doi:10.1111/his.14096

Trikha, P., Plews, R. L., Stiff, A., Gautam, S., Hsu, V., Abood, D., et al. (2016). Targeting myeloid-derived suppressor cells using a novel adenosine monophosphate-activated protein kinase (AMPK) activator. *Oncoimmunology* 5 (9), e1214787. doi:10.1080/2162402X.2016.1214787

Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., et al. (2017). A pathology atlas of the human cancer transcriptome. *Science* 357 (6352), eaan2507. doi:10.1126/science.aan2507

Wang, H., Liu, J., Li, J., Zang, D., Wang, X., Chen, Y., et al. (2020). Identification of gene modules and hub genes in colon adenocarcinoma associated with pathological stage based on WGCNA analysis. *Cancer Genet.* 242, 1–7. doi:10.1016/j.cancergen.2020.01.052

Watts, G. S., Futscher, B. W., Holtan, N., Degeest, K., Domann, F. E., and Rose, S. L. (2008). DNA methylation changes in ovarian cancer are cumulative with disease progression and identify tumor stage. *BMC Med. Genomics* 1, 47. doi:10.1186/1755-8794-1-47

Yang, H., Shi, J., Lin, D., Li, X., Zhao, C., Wang, Q., et al. (2018). Prognostic value of PD-L1 expression in combination with CD8(+) TILs density in patients with surgically resected non-small cell lung cancer. *Cancer Med.* 7 (1), 32–45. doi:10.1002/cam4.1243

Zebley, C. C., Abdelsamed, H. A., Ghoneim, H. E., Alli, S., Brown, C., Haydar, D., et al. (2021). Proinflammatory cytokines promote TET2-mediated DNA demethylation during CD8 T cell effector differentiation. *Cell. Rep.* 37 (2), 109796. doi:10.1016/j.celrep.2021.109796

Zhang, L., Li, Z., Skrzypczynska, K. M., Fang, Q., Zhang, W., O'Brien, S. A., et al. (2020). Single-cell analyses inform mechanisms of myeloid-targeted therapies in colon cancer. *Cell.* 181 (2), 442–459. e29. doi:10.1016/j.cell.2020.03.048

Zhang, X., Shi, M., Chen, T., and Zhang, B. (2020). Characterization of the immune cell infiltration landscape in head and neck squamous cell carcinoma to aid immunotherapy. *Mol. Ther. Nucleic Acids* 22, 298–309. doi:10.1016/j.omtn.2020.08.030

Zheng, Y., Huang, C., Lu, L., Yu, K., Zhao, J., Chen, M., et al. (2021). STOML2 potentiates metastasis of hepatocellular carcinoma by promoting PINK1-mediated mitophagy and regulates sensitivity to lenvatinib. *J. Hematol. Oncol.* 14 (1), 16. doi:10.1186/s13045-020-01029-3

Zhou, J. W., Wang, M., Sun, N. X., Qing, Y., Yin, T. F., Li, C., et al. (2019). Sulforaphane-induced epigenetic regulation of Nrf2 expression by DNA methyltransferase in human Caco-2 cells. *Oncol. Lett.* 18 (3), 2639–2647. doi:10.3892/ol.2019.10569

frontiers | Frontiers in Genetics

TYPE Review
PUBLISHED 01 November 2022
DOI 10.3389/fgene.2022.1027726

OPEN ACCESS

EDITED BY
Zhongxiao Wang,
Biogen Idec, United States

REVIEWED BY
Yingming Sun,
Fujian Medical University, China
Zeming Liu,
Huazhong University of Science and Technology, China

*CORRESPONDENCE
Qing He,
qing.he@whu.edu.cn

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

RECEIVED 25 August 2022
ACCEPTED 10 October 2022
PUBLISHED 01 November 2022

CITATION
Zhang Y and He Q (2022), The role of SELENBP1 and its epigenetic regulation in carcinogenic progression. *Front. Genet.* 13:1027726. doi: 10.3389/fgene.2022.1027726

The role of SELENBP1 and its epigenetic regulation in carcinogenic progression

Yue Zhang and Qing He*

The State Key Laboratory Breeding Base of Basic Science of Stomatology and Key Laboratory for Oral Biomedicine of Ministry of Education, School and Hospital of Stomatology, Wuhan University, Wuhan, China

The initiation and progression of cancer is modulated through diverse genetic and epigenetic modifications. The epigenetic machinery regulates gene expression through intertwined DNA methylation, histone modifications, and miRNAs without affecting their genome sequences. SELENBP1 belongs to selenium-binding proteins and functions as a tumor suppressor. Its expression is significantly downregulated and correlates with carcinogenic progression and poor survival in various cancers. The role of SELENBP1 in carcinogenesis has not been fully elucidated, and its epigenetic regulation remains poorly understood. In this review, we summarize recent findings on the function and regulatory mechanisms of SELENBP1 during carcinogenic progression, with an emphasis on epigenetic mechanisms. We also discuss the potential cancer treatment targeting epigenetic modification of SELENBP1, either alone or in combination with selenium-containing compounds or dietary selenium.

KEYWORDS

SELENBP1, carcinogenesis, epigenetic modification, post-translational modification, therapeutical target

Introduction

Cancer incidence and mortality rate are increasing at an alarming rate worldwide, with an estimated 19.3 million new cases and almost 10.0 million deaths having occurred in 2020 (Sung et al., 2021). Both genetic and epigenetic modifications contribute to carcinogenic progression. Epigenetics refers to heritable changes in the chromatin structure and gene expression during post-transcriptional and translational stages without DNA sequence alterations, including DNA methylation, histone modification, and RNA-based mechanism, which lead to silencing or enhanced expression of the gene or protein (Haig, 2004; Waddington et al., 2012; Zhang et al., 2021; Hussain et al., 2022; Zaib, 2022). DNA methylation that occurs in cytosine–guanine (CpG) islands of the gene promoter regions is associated with gene silencing (Jones and Baylin, 2002; Feinberg, 2018). Histone post-translational modifications (PTMs) include methylation, phosphorylation, acetylation, ubiquitination, glycosylation, and chromatin remodeling, and also could be regulated by non-coding RNAs such as microRNAs (miRNAs), siRNAs, and long-non-coding RNAs (lncRNAs). (Hussain et al., 2022; Zhang et al., 2021; Millan-

Zambrano et al., 2022). Epigenetic changes are closely associated with regulation of tumor suppressors and/or oncogenes *via* DNA hypermethylation, histone modification, and non-coding RNAs (Jones and Baylin, 2002; Hussain et al., 2022).

The study of epigenetics in cancer epidemiology is emerging in the recent 2 decades, and its role in oncogene/tumor suppressor regulation and cancer progression is being emphasized increasingly. Epigenetic changes have opened the way to innovative diagnostic and treatment strategies for a variety of cancers in clinics (Laird, 2003). Epigenetic drugs targeting to inhibit epigenetic modifiers, such as DNA methyltransferase and histone deacetylase (DNMTi and HDACi), have been approved to treat different malignant cancers and showed promising outcomes (Hussain et al., 2022).

Structural and biochemical characteristics of SELENBP1

Mammalian selenium-containing proteins can be divided into three groups: specific selenium-binding proteins, specific incorporation proteins (also called selenoproteins), and non-specific incorporation proteins (Behne and Kyriakopoulos, 2001; Kryukov et al., 2003). The human SELENBP1 (also known as Sbp1, human 56 kDa selenium-binding protein/hSP56, EHMTO, HEL-S-134P, LPSB, MTO, and SBP56 or hSBP) belonging to the specific selenium-binding protein is located at chromosome 1q21–22. The neighbor-joining phylogenetic tree of SELENBP1 has been reported in a previous study (Pol et al., 2018). Its mouse homolog is the SP56 gene (also known as SLP-56, Lp56, Lpsb, MTO, or SBP56) (Bansal et al., 1990; Lanfear et al., 1993; Chang et al., 1997; Porat et al., 2000). The gene sequences of mouse and human SELENBP1 are shown in Figure 1 (A, B) using GSDS2.0 (Hu et al., 2015). The degree of homology between mouse and human SELENBP1 is 86% (Flemetakis et al., 2002), which means many mechanisms apply to both. Selenbp2, as a highly homologous isoform of Selenbp1 differing by only 14 residues, exists in mice but not in humans (Lanfear et al., 1993). Both Selenbp1 and Selenbp2 mRNA levels are downregulated in liver cell lines but remain high in diethylnitrosamine (DEN)-induced mouse liver tumors *in vivo* (Lanfear et al., 1993). This may be due to the redundant role of each other. However, the expressions or functions of Selenbp2 in cancers are scarcely published yet. Both mouse and human SELENBP1 are ubiquitously expressed in the colon, lung, and 17 other tissues (Fagerberg et al., 2014; Yue et al., 2014). Selenium (Se), as the sixth main group of the periodic table of elements, displays both metallic and non-metallic properties (Minich, 2022) and has been widely reported to possess antitumor effects and exert its inhibitory effects (Bansal et al., 1990). SELENBP1 is also implied in cancer prevention. The downregulation and anti-carcinogenic effects of SELENBP1 are assessed in numerous cancers, including kidney/lung/thyroid/stomach/esophagus/liver/breast/prostate/colon/pancreatic/head and neck/skin/bladder/uterine/nerve and ovary cancer. SELENBP1 is highly expressed in the brain (Pol et al., 2018) and is also associated with inflammatory/degenerative central nervous system (CNS) diseases (neuromyelitis optical spectrum and Alzheimer's disease) and schizophrenia (Scz) (Elkjaer et al., 2021; Seelig et al., 2021). In addition, SELENBP1 also plays a role in multiple sclerosis (MS) subtypes, corona virus disease 2019 (COVID-19), extraoral halitosis, and so on (Adam et al., 2011; Pol et al., 2018; Stukalov et al., 2021).

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database was used to identify the interaction partners of human SELENBP1, as shown in Figure 2. One of the interactors of SELENBP1 is ubiquitin-specific peptidase 33 (USP33), also known as von Hippel–Lindau protein (pVHL)-interacting deubiquitinating enzyme 1 (VDU1), which is a deubiquitinating enzyme. The interaction between USP33 and SELENBP1 was detected by a two-hybrid assay in human prostate cancer cells (Jeong et al., 2009), indicating SELENBP1 may play a role in epigenetics regulation. The other mammalian selenium-containing proteins, such as selenium-binding protein and glutathione peroxidase 1 (GPX1), also play important roles in cancers. SELENBP1 regulates GPX1 expression in hepatocellular carcinoma (HCC) (Huang et al., 2012), breast cancer (Fang et al., 2010), colon cancer (Fang et al., 2010), and skin cancer (Schott et al., 2018). In addition, SELENBP1 was identified to interact with tandem BRCA1 carboxyl-terminal (BRCT) domain–mediated proteins, which is frequently present in proteins involved in the damage response (DDR), and bind to phosphorylated peptides. Defects in this network can lead to cancer, while the mechanism of SELENBP1 in DDR during cancer was still unknown (Woods et al., 2012).

Expressions and functions of SELENBP1 in varieties of cancers

Kidney cancer: SELENBP1 could be a useful prognostic factor in renal cell carcinoma (RCC). Using 139 specimens of primary RCC and 59 specimens of control kidney tissues, the mRNA level of SELENBP1 is significantly downregulated in the RCC tissues than in normal adjacent kidney tissues ($p < 0.001$) and is correlated with pathologic (T-stage and Fuhrman grade), prognostic variables (progression and cancer-specific death) as well as cancer-specific death (log-rank test, $p = 0.014$) (Ha et al., 2014). SELENBP1 is downregulated by hepatocyte nuclear factor 4 alpha (HNF4alpha), a tissue-specific transcription factor in RCC microarray studies (Lucas et al., 2005). However, the exact mechanism of SELENBP1 still needs to be elucidated.

Lung cancer: SELENBP1 is downregulated in basaloid carcinoma, large cell carcinoma, lung squamous cell

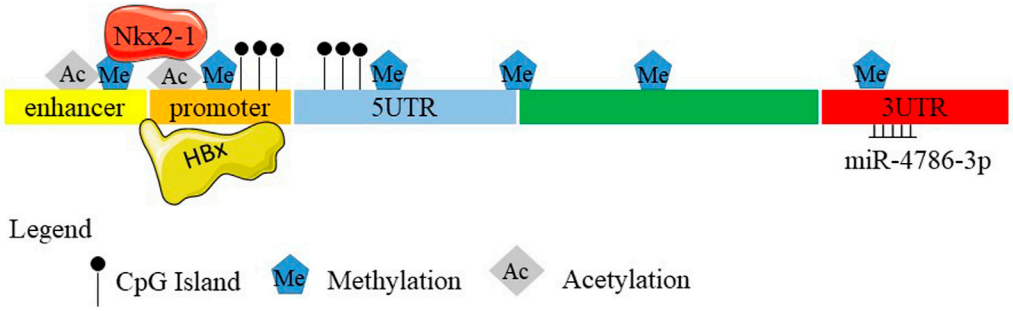


FIGURE 1
Schematic diagram of the *SELENBP1* gene. (A) Nucleic acid sequences of mouse SELENBP1. (B) Nucleic acid sequences of human SELENBP1. Boxes and intervening lines represent exons and introns, respectively. Several CpG sites were found in the 5′-untranslated region and promoter. Methylation occurred at −95 to +90, the 5′-UTR, 3′-UTR, enhancer, promoter as well as in the gene body. The histone acetylation also exists with unknown sites. 5′ flanking promoter is repressed by HBx. Nkx2-1 binds to the enhancer and promoter regions of Selenbp1 through methylation and acetylation. miR-4786-3p targets at the 3′ UTR of SELENBP1.

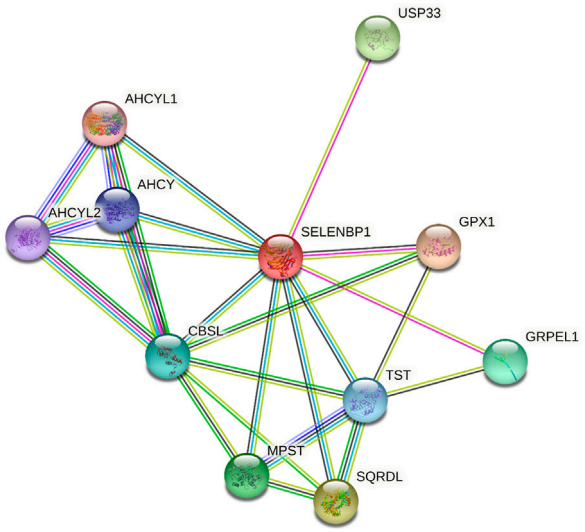


FIGURE 2
Involvement of human SELENBP1 in protein–protein interaction (PPI) databases. SELENBP1, highlighted in red, is shown in the standard human protein–protein association network in STRING 11.5 (https://string-db.org/). The network includes 11 nodes and 25 edges, and the average node degree was 4.55, with a PPI enrichment P-value of 9.81e-05. Nodes represent proteins. Edges represent PPIs.

carcinoma (LSCC), and adenocarcinoma (Li et al., 2004). Using quantitative two-dimensional polyacrylamide gel electrophoresis (2-D PAGE), two isoforms of SELENBP1 are significantly decreased in lung adenocarcinomas at both mRNA and protein levels. Its expression correlates with poor survival, and downregulation of SELENBP1 may increase cell proliferation and decreased differentiation (Chen et al., 2004). The protein level of SELENBP1 is significantly lower in LSCC tissues than in the corresponding normal bronchial epithelium (NBE) tissues ($p = 0.000$) and is associated with higher lymph node metastasis and lower overall survival rate ($p < 0.05$) (Tan et al., 2016). Knockdown of SELENBP1 can increase benzo(a)pyrene (B[a]*p*)-induced human NBE cell transformation to participate in NBE carcinogenesis (Zeng et al., 2013).

Thyroid cancer: Proteomic analysis of papillary thyroid carcinoma (PTC) compared with normal thyroid tissue using difference gel electrophoresis (DIGE), and mass spectrometry confirmed lower expression of SELENBP1 ($p = 0.00097$) (Brown et al., 2006). More studies are needed to examine its function in thyroid tissue and thyroid cancer progression.

Stomach cancer: SELENBP1 was significantly decreased in a proteomic analysis of gastric cancer specimens. Its expression is correlated with differentiation, TNM stage, and lymph node metastasis ($p < 0.05$). It can serve as a potential novel prognostic biomarker (He et al., 2004; Zhang et al., 2011a; Xia et al., 2011), and its level is associated with a poor survival rate in gastric carcinoma (Zhang et al., 2011b). SELENBP1 promotes proliferation, colony formation, and senescence *in vitro* and *in vivo*. In addition, it suppresses tumor growth and metastasis and increases the chemoresistance of gastric cancer cells *via* apoptotic signaling pathways (Zhang C. et al., 2013). The exact mechanism is not known yet.

Esophagus cancer: SELENBP1 is decreased significantly in esophageal adenocarcinoma (EAC) tissues. Downregulation of SELENBP1 in Barrett's esophagus to adenocarcinoma progression could enhance apoptosis, cellular senescence, and cisplatin cytotoxicity in EAC cells (Silvers et al., 2010).

Liver cancer: The protein level of SELENBP1 is decreased in HCC (Raucci et al., 2011). Decreased expression of SELENBP1 could promote tumor invasiveness by increasing GPX1 activity and diminishing hypoxia-inducible protein-

1alpha (HIF-1α, a tumor suppressor) expression in HCC; SELENBP1 could be a novel biomarker for predicting prognosis and guiding personalized therapeutic strategies, especially in patients with advanced HCC (Huang et al., 2012). However, in another study, SELENBP1 mRNA is downregulated in liver cell lines *in vitro* but upregulated in DEN-induced mouse liver tumors compared with the normal tissues *in vivo*, though the function is not known for certain at present. They hypothesized that binding of selenium or acetaminophen metabolites in hepatoxicity might play a role in the processes (Lanfear et al., 1993). Downregulation of SELENBP1 could increase C-X-C motif chemokine receptor 4 (CXCR4) expression and results in epithelial–mesenchymal transition (EMT) of HCC cells (Gao et al., 2018). The 3D model of human protein SELENBP1 was reported to have one cysteine residue, which binds the selenium in HCC and can be used as a therapeutic target (Raucci et al., 2011).

Breast cancer: SELENBP1 remains low in mammary carcinoma (Lanfear et al., 1993). Lower SELENBP1 expression in breast cancer tissues compared to normal control is significantly associated with poor survival ($p < 0.01$). SELENBP1 expression is regulated by estrogen *in vitro*. Supplemental dietary selenium could inhibit cell proliferation in cells expressing high level of SELENBP1 (Zhang S. et al., 2013). SELENBP1 is likely to play roles in modulating selenite-mediated cytotoxicity and the extracellular microenvironment by regulating the levels of extracellular glutathione (GSH) (Wang et al., 2015). The interaction between SELENBP1 and GPX1 is found in MCF-7 breast carcinoma cells (Fang et al., 2010). SELENBP1 can also interact with the retinoid-receptor RARα in the nuclear material, whose interaction is reduced by all trans-retinoic acid (ATRA, a cancer drug) in the breast cancer cell line (Gianni et al., 2019). SELENBP1 interacts with the nuclear receptor estrogen receptor 2 (ESR2, ERβ) to modulate cell proliferation and tumor growth in breast cancer (Giurato et al., 2018).

Prostate cancer: SELENBP1 is observed to be at lower levels in prostate cancer (Yang and Sytkowski, 1998; Ansong et al., 2015). As to energy metabolism, SELENBP1 produces H_2O_2 and H_2S and consequential activation of AMP-activated protein kinase (AMPK), a major regulator of energy homeostasis as well as inhibited oxidative phosphorylation (OXPHOS) in prostate cancer cells. In addition, HNF4 alpha, a novel transcriptional inhibitor of SELENBP1, plays a role in prostate cancer by binding to the SELENBP1 promoter region (Elhodaky et al., 2020). SELENBP1 may influence the plasma selenium levels and may be associated with the risk of advanced prostate cancer (Xie et al., 2016). GPx enzyme activity is inversely correlated with SELENBP1 levels in prostate cancer tissue (Jerome-Morais et al., 2012). SELENBP1 is found to coprecipitate with the androgen receptor (AR) in prostate tumor cells (Veldscholte et al., 1992). In prostate cancer, expressed prostatic secretions (EPS), proximal fluids of the prostate, can be utilized for diagnostic and prognostic assays. SELENBP1 is found in the EPS according to a shotgun proteomics (Principe et al., 2013).

Colon cancer: SELENBP1 is downregulated in the proteomic analysis of colorectal cancer (CRC) (Wang et al., 2012). Both mRNA and protein levels of SELENBP1 are downregulated in CRC, which are correlated with the degree of differentiation, and its levels are higher in benign polyps than in CRC tissues (Kim et al., 2006; Li et al., 2008; Wang et al., 2014). SELENBP1 induces H_2O_2-mediated apoptosis in colon cancer cells and inhibits cancer cell migration *in vitro* and tumor growth *in vivo* (Pohl et al., 2009). SELENBP1 can also inhibit CRC through lipid/glucose metabolic signaling pathways (Ying et al., 2015a) and participate in mitochondrial function in the HCT116 human colorectal carcinoma cell line through cysteine 57 in SELENBP1 (Ying et al., 2015b). The interaction between SELENBP1 and GPX1 is found in colon-derived HCT116 cells (Fang et al., 2010).

Pancreatic cancer: SELENBP1 is downregulated in pancreatic ductal adenocarcinoma (PDAC) patients with skin rash (SR) treated with erlotinib, an epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor (Caba et al., 2016). SELENBP1 interacted with anterior gradient 2 (AGR2), with the latter promoting phosphorylation of RICTOR (T1135), leading to pancreatic tumor metastasis (Tiemann et al., 2019).

Head and neck cancer: SELENBP1 is downregulated in head and neck squamous cell carcinoma (HNSCC) including nasopharyngeal carcinoma (NPC), laryngeal cancer (LC), oral cancer (OC), tonsil cancer (TC), and hypopharyngeal cancer (HPC), which have no association with tumor T-stage, N-stage, and tumor grade. NPC patients with low expression of SELENBP1 have a poor survival rate. Therefore, SELENBP1 could be a novel biomarker for predicting NPC prognosis (Chen et al., 2016). SELENBP1 downregulation in both mRNA and protein levels is also positively correlated with poor prognosis for oral squamous cell carcinoma (OSCC) patients.

Skin cancer: SELENBP1 concentration remains low in skin carcinoma (Lanfear et al., 1993). SELENBP1 is identified to be downregulated in cutaneous melanoma influenced by glutathione peroxidase 1 (GPX1) to regulate proliferation and tumor microenvironment (Schott et al., 2018). An *in vivo* mouse study shows that HIF-1α can transactivate SELENBP1 during murine skin chemical carcinogenesis (Scortegagna et al., 2009).

Bladder cancer: SELENBP1 is significantly downregulated in bladder cancer (Wang et al., 2020).

Uterine cancer: Decreased protein expression of SELENBP1 is found in uterine leiomyoma than in normal myometrium (Zhang C. et al., 2010).

Nerve cancer: SELENBP1 is found at the growing tips of the neurites in SH-SY5Y neuroblastoma cells/T98G glioma cells involved in the initial sequential events in rapid cell outgrowth; however, the exact expression change is still unknown (Miyaguchi, 2004).

Malignant pleural mesothelioma (MPM): Gene expression analyses from MPM patients show that SELENBP1 might be associated with the risk of having MPM (Pass et al., 2004).

Uveal melanoma tumor: SELENBP1 expression is induced in uveal melanomas that subsequently develop distant metastases compared with those that do not (Linge et al., 2012).

Ovary cancer: In a membrane proteome profiling analysis, SELENBP1 is decreased in ovary cancer through the selenium/androgen pathway (Huang et al., 2006). In addition, high anti-SELENBP1 levels are also observed in patients with serous ovary cancer, which suggests SELENBP1 participates in an autoimmune process during ovary cancer development (Yu-Rice et al., 2017). Downregulation of SELENBP1 also increases epithelial proliferation and papillary complexity in tumorigenesis of ovarian serous borderline tumor, micropapillary serous borderline tumor, and low-grade serous carcinoma (Zhang P. et al., 2010). SELENBP1 is identified as one of the candidate stromal epithelial cross-talk genes by analyzing common single-nucleotide polymorphisms (SNPs) for their association between the risk of serous ovarian cancer and telomerase reverse transcriptase (*TERT*), a cancer susceptibility "hot-spot" (Johnatty et al., 2010).

SELENBP1 also displays high expression in other tissues, such as the fetal and adult heart, spleen, thymus, and other tissues (Pol et al., 2018); however, no studies have been reported whether SELENBP1 also participates in their carcinogenesis.

Epigenetic mechanisms of SELENBP1

The expression of SELENBP1 is downregulated in almost all cancers, indicating some suppressing functions of SELENBP1 in cancers. Epigenetic changes are likely to account for reduction of SELENBP1 expression.

Epigenetic modification of SELENBP1 through promoter methylation in cancer has been demonstrated in many studies: DNA methylation at CpG islands in promoter regions is regulated by the DNA methyltransferase (DNMT) enzyme. This phenomenon often occurs at an early stage and is a common mechanism of gene silencing during carcinogenesis (Jones and Baylin, 2002; Laird, 2003; Nazemalhosseini et al., 2013; Huss+6ain et al., 2022). Several CpG sites were found in the 5′-untranslated region of SELENBP1 (Pohl et al., 2009; Silvers et al., 2010). Gene hypermethylation in the "epidermal differentiation complex" is located within 700 kb of the SELENBP1 locus (Marenholz et al., 2001; Elder and Zhao, 2002). So, it is possible that the SELENBP1 promoter may be methylated in tumors and results in low expression levels of SELENBP1. Deubiquitinating enzymes (Dubs) function to remove covalently attached ubiquitin from proteins to control substrate activity and/or abundance. SELENBP1 is the bait of USP15 (ubiquitin-specific proteases 15, a Dub) in a global proteomic analysis of Dubs and their associated protein complexes in 293T cells, indicating SELENBP1 levels might also be regulated by USP15 (Sowa et al., 2009). PLEKHA4 sequester E3 ubiquitin ligase adapter controls disheveled polyubiquitination at the plasma membrane to tune sensitivities of cells. SELENBP1 interacts with PLEKHA4 in HEK293 cells, which further directs the crucial function of epigenetic modification in modulating SELENBP1 expression (Shami et al., 2019).

Lung cancer: Nkx2-1 is a transcription factor that suppresses malignant progression of lung adenocarcinoma. SELENBP1 is regulated by Nkx2-1 in lung adenocarcinoma in both the human lung adenocarcinoma and mouse lung cancer model. Nkx2-1 binds to the promoter regions of SELENBP1, which are associated with H3K4me3 and H3K27ac; Nkx2-1 binds to the enhancer regions at H3K4me1 and H3K27ac, which collectively indicate that Nkx2-1 likely regulates the expression of SELENBP1 through epigenetics regulation. In addition, they function in a positive feedback loop during the suppression of malignant progression of lung adenocarcinoma (Caswell et al., 2018).

Esophageal adenocarcinoma: Several CpG sites are found near the predicted promoter region of SELENBP1, where hypermethylation occurred at −95 to +90 and in the 5′-untranslated region of SELENBP1 (Silvers et al., 2010). A histone deacetylase inhibitor trichostatin A (TSA)/valproic acid (VPA) alone and/or the demethylating agent 5-Aza could increase SELENBP1 mRNA expression but not the protein level in EAC cell line Flo-1 cells. In addition, demethylating and acetylase-enhancing agents increased sensitivity of Flo-1 cells to apoptosis with cisplatin, which all indicate that inducible SELENBP1 expression may be regulated at the epigenetic level and play an important role during drug resistance in EAC (Silvers et al., 2010). However, the histone acetylation sites remain unknown.

Liver cancer: Human hepatitis B virus (HBV) is a leading cause of HCC (Beasley et al., 1981). SELENBP1 is decreased in HCC cells expressing the HBV X protein (HBx), and the SELENBP1 promoter is repressed by HBx. In addition, the stepwise deletion of 5′ flanking promoter sequences resulted in a gradual decrease in basal promoter activity and inhibition of SELENBP1 expression by HBx (Lee et al., 2020), which indicates the potential epigenetic regulation of SELENBP1 by HBx; however, whether the promoter hypermethylation exists has not been confirmed yet.

Prostate cancer: SELENBP1 could induce phosphorylation of the p53 tumor suppressor at serine 15 to suppress carcinogenesis (Ansong et al., 2015). In human prostate cancer cells, SELENBP1 interacts with von Hippel–Lindau protein (pVHL)-interacting deubiquitinating enzyme 1 (VDU1), and VUD1 incorporates selenium into SELENBP1. These findings imply the role of SELENBP1 in ubiquitination/de-ubiquitination-mediated protein degradation pathways dependent of selenium. However, the specific mechanism has

not been elucidated (Jeong et al., 2009). SELENBP1 also interacted with VDU2, a closely related isoform of VDU1 (Jeong et al., 2014). VDU2 can interact with HIF-1α, a tumor suppressor, to deubiquitinate and stabilize HIF-1α (Li et al., 2005). SELENBP1 inhibits HIF-1α protein levels and downregulates its stabilization in prostate cells without affecting the mRNA levels of HIF-1α, which indicates SELENBP1 may affect epigenetic regulation of other genes. Whether HIF-1α is regulated by the interaction between SELENBP1 with VDU 1 and/or 2 is unknown (Jeong et al., 2014). SELENBP1 can downregulate HIF-1α protein levels without altering mRNA levels in the human lung carcinoma cell line as well (Jeong et al., 2014). The downregulation of HIF-1α by SELENBP1 also exists in liver cancer (Huang et al., 2012); however, the mechanism is not quite clear yet. SELENBP1 serves as the target gene of HIF-1α in skin squamous carcinogenesis (Scortegagna et al., 2009). SELENBP1 interacts with embryonic ectoderm development (EED), a core component of polycomb group proteins in the nuclear material of VCaP (vertebral cancer of the prostate) cells, a prostate cancer cell line (Cao et al., 2014). EED is a core component of polycomb repressive complexes 2 (PRC2) and is critical for PRC2 to methylate histone H3 at lysine 27 (H3K27me3). Then, EED recruits polycomb repressive complexes 1 (PRC1) directly to the tri-methylated H3K27 loci and enhances PRC1-mediated H2A ubiquitin E3 ligase activity, indicating a potential role for SELENBP1 as an epigenetic exchange factor in prostate cancer (Cao et al., 2014).

Colon cancer: In colon cancer, the promoter of SELENBP1 was methylated in both human colon tissues and cell lines (HCT116, SW480, Caco-2, and HT-29 cells), while it was mostly unmethylated in LS174T cells (Pohl et al., 2009). Treating HCT116 with 5-Aza-dC (5′-Aza-2′-deoxycytidine), a DNA methylation inhibitor (DNMTi), could decrease methylation of the SELENBP1 promoter, increase SELENBP1 promoter activity, and rescue SELENBP1 mRNA and protein expression levels (Pohl et al., 2009). However, very little methylation in the SELENBP1 promoter was detected in the colon cancer cell line SW480 in another study (Silvers et al., 2010). Treating SW480, SW620, and HT-29 cells with 5-Aza-dC alone did not significantly alter SELENBP1 protein and mRNA levels (Wang et al., 2014). However, TSA alone or in combination with 5-Aza-dC could upregulate SELENBP1 expression in the SW480 and SW620 cells, indicating its regulation by histone modification, while no significant change was observed in HT-29 cells (Wang et al., 2014). There are no consistent results about the DNA methylation in colon cancer. NACHT, LRR, and PYD domain-containing protein 7 (NLRP7) are potential biomarkers of CRC. NLRP7 is deubiquitinated by ubiquitin-specific protease 10 (USP10), leading to increased NLRP7 protein stability (Li et al., 2021). SELENBP1 physically interacts with NLRP7, indicating the potential impact of SELENBP1 in epigenetic regulation; however, the specific mechanism remains unknown (Li et al., 2021).

Head and neck cancer: miRNA belongs to the non-coding RNAs (Hussain et al., 2022). SELENBP1 can be epigenetically modified through miRNA in cancer. SELENBP1 downregulation in OSCC is induced by miR-4786-3p binding to the 3′ untranslated region (UTR) of SELENBP1. Nuclear factor erythroid 2-related factor 2 (NRF2, an oncogenic transcription factor) is a downstream responder upon SELENBP1 downregulation. SELENBP1 reduces NRF2 protein levels by promoting its polyubiquitination and degradation. Kelch-like ECH-associated protein 1 (KEAP1) binds to NRF2 to promote ubiquitin–proteasomal degradation of NRF2 (Itoh et al., 1999; Adam et al., 2011). SELENBP1 also acts as a transcriptional factor to induce KEAP1 transcription by binding the KEAP1 promoter (Zeng et al., 2021). Targeting at the miR-4786-3p–SELENBP1–KEAP1–NRF2 signaling axis may enhance the efficacy of chemotherapy for OSCC (Zeng et al., 2021).

Bladder cancer: By using the MEXPRESS tool (http://mexpress.be/), SELENBP1 expression in bladder cancer was found to relate to DNA hypermethylation in its promoter region located close to the transcription start site (TSS) and the first exon. Gene body methylation is negatively associated with SELENBP1 expression; however, DNA methylation in its 3′UTR was positively associated with SELENBP1 expression, which highlights that multiple DNA methylation sites may be differentially involved in regulating SELENBP1 expression. DNA hypermethylation, especially in the gene body, accounts for the reduction of SELENBP1 expression in bladder cancer. While p53-responsive elements are located in − 1,394 bp and − 2,285 bp regions of cyclin-dependent kinase inhibitor 1A (CDKN1A; also known as p21) promoter, SELENBP1-responsive elements are located within approximately ~ 1,300 bp to ~ 200 bp region of the p21 upstream promoter, and SELENBP1 upregulates p21 expression through a p53-independent mechanism transcriptionally and through phosphorylation attenuation of c-Jun and STAT1, leading to the G0/G1 phase cell cycle arrest, thereby inducing attenuation of cancer cell growth (Wang et al., 2020).

The epigenetic regulation of SELENBP1 during carcinogenesis is listed in Table 1 and described in Figure 3. In esophageal adenocarcinoma and colon cancer, histone deacetylase inhibitors or demethylating agents could target SELENBP1 expression (Pohl et al., 2009; Silvers et al., 2010). This leads to the usage of DNA methyltransferase and histone deacetylase to treat cancers targeting SELENBP1.

Selenium and treatment strategies

Se has been widely known to play a role in tumor prevention (Bansal et al., 1990). Studies regarding effects of Se on DNA methylation, DNMT expression or activity, and histone acetylation have been reported in various tissues and cells (Speckmann and Grune, 2015; Jablonska and Reszka, 2017).

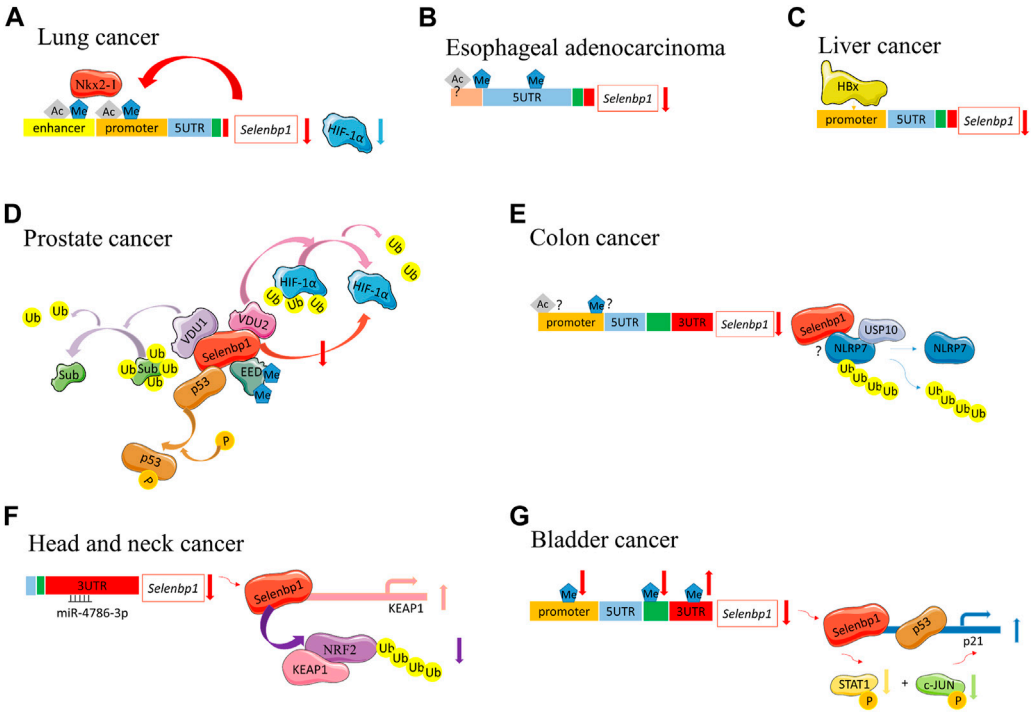


FIGURE 3
Different SELENBP1 epigenetic functions during carcinogenesis. **(A)** Lung cancer: Nkx2-1 binds to the promoter and enhancer regions of SELENBP1 through methylation and acetylation. They also function in a positive feedback loop. SELENBP1 can downregulate HIF-1α protein levels. **(B)** Esophageal adenocarcinoma, methylation occurred at −95 to +90 as well as in the 5′-untranslated region of SELENBP1. The histone acetylation also exists with unknown sites. **(C)** Liver cancer: SELENBP1 is decreased in HCC and the SELENBP1 5′ flanking promoter is repressed by HBx. **(D)** Prostate cancer: SELENBP1 inducts phosphorylation of p53. SELENBP1 interacted with VDU1 and VDU2. VDU2 can deubiquitinate HIF-1α. Sub, potential substrates of VDU1. SELENBP1 inhibits HIF-1α protein levels. SELENBP1 interacts with EED. EED is critical for downstream methylation and ubiquitinoylation. **(E)** Colon cancer: whether SELENBP1 promoter hypermethylation or histone acetylation exists in human colon cancer cells remain controversial. NLRP7 is deubiquitinated by USP10 in colon cancer. SELENBP1 interacted with NLRP7. The specific mechanism between SELENBP1 and NLRP7 remains unclear. **(F)** Head and neck cancer: downregulation of SELENBP1 is induced by miR-4786-3p targeting at the 3′ UTR of SELENBP1. SELENBP1 reduces NRF2 protein levels by promoting its polyubiquitination and degradation. SELENBP1 induces KEAP1 transcription by binding to KEAP1 promoter. **(G)** Bladder cancer: SELENBP1 expression was inversely associated with DNA methylation in the promoter and gene body, but positively correlates with DNA methylation in the 3′-UTR region. SELENBP1-responsive elements located in the upper region of the p21 upstream promoter than p53's to increase p21 expression. SELENBP1-mediated transcriptional induction of p21 protein through combined phosphorylation suppression of c-Jun and STAT1. Ub, ubiquitin; P, phosphorylation; Ac, acetylation.

TABLE 1 List of SELENBP1 regulated by epigenetic modifications during carcinogenesis.

Cancer type	Epigenetic modification	Reference
Lung cancer	Histone modification (H3K4me3, H3K4me1, and H3K27ac)	Caswell et al. (2018)
Esophageal adenocarcinoma (EAC)	Promoter hypermethylation	Silvers et al. (2010)
	Histone modification (acetylation)	
Liver cancer	Potential promoter modification	Lee et al. (2020)
Prostate cancer	Inducting phosphorylation of the p53 (serine 15)	Ansong et al. (2015)
	Histone modification (ubiquitination)	Jeong et al. (2009)
	Potential promoter modification	Jeong et al. (2014)
	Potential promoter modification	Elhodaky et al. (2020)
	Potential histone modification (H3K27me3 and ubiquitination)	Cao et al. (2014)
Colon cancer	Promoter hypermethylation	Pohl et al. (2009)
	Histone modification (acetylation) and promoter hypermethylation	Wang et al. (2014)
	Potential histone modification (ubiquitination)	Li et al. (2021)
Head and neck cancer	miR-4786-3p and histone modification (ubiquitination)	Zeng et al. (2021)
Bladder cancer	DNA hypermethylation and histone modification (phosphorylation)	Wang et al. (2020)

TABLE 2 List of differentially expressed SELENBP1 during carcinogenesis and their functions.

Cancer type	Variation trend	Level	Functional classification	Roles in cancer	Reference
Kidney cancer	Downregulation	mRNA	Potential prognostic factor	Specific mechanism remains unknown	Ha et al. (2014)
		mRNA		Regulated by HNF4alpha	Lucas et al. (2005)
Lung cancer	Downregulation	Protein	Decreased inactivation of carcinogens	Positive feedback loop with Nkx2-1	Caswell et al. (2018)
		Protein	Potential prognostic factor of LSCC	Specific mechanism remains unknown	Tan et al. (2016)
		mRNA and protein	Poor survival	Proliferation and differentiation	Chen et al. (2004)
		mRNA and protein	Potential biomarker for early detection of LSCC	Cell transformation	Zeng et al. (2013)
Thyroid cancer	Downregulation	Protein	Unknown	Unknown	Brown et al. (2006)
Stomach cancer	Downregulation	Protein	Potential suppression target, impact on drug efficacy, and toxicity	Inhibit proliferation and migration	Zhang et al. (2013a)
		Protein	Diagnostic marker	Unknown	(Zhang et al., 2011a; Xia et al., 2011)
Esophageal adenocarcinoma (EAC)	Downregulation	Protein	Predictor of response to chemoprevention or chemosensitization	Epigenetic and posttranscriptional mechanisms	Silvers et al. (2010)
Liver cancer	Downregulation	Protein	Prognosis biomarker	Increasing GPX1 activity and diminishing HIF-1α expression	Huang et al. (2012)
				EMT	Gao et al. (2018)
	Upregulation in vivo		Potential biomarkers and therapeutic targets	Epigenetic regulation	Lee et al. (2020)
	Downregulation in vitro		Unknown	Selenium or acetaminophen metabolites	Lanfear et al. (1993)
Breast cancer	Downregulation		Potential biomarker predicting survival and effectiveness of selenium supplementation	Regulated *via* estrogen	(Lanfear et al., 1993; Zhang et al., 2013b)
				Environment	Wang et al. (2015)
				GPX1	Fang et al. (2010)
				RARα	Gianni et al. (2019)
				ESR2	Giurato et al. (2018)
Prostate cancer	Downregulation		Distinguishing indolent from aggressive disease	Inducting phosphorylation of the p53	Ansong et al. (2015)
				Epigenetic regulation	Jeong et al. (2009)
			Candidate anti-oncogene product	Inhibited HIF-1α protein levels	Jeong et al. (2014)
				Energy metabolism	Elhodaky et al. (2020)
			Potential biomarker of tumor progression	Selenium levels in plasma	Xie et al. (2016)
				GPx enzyme activity	Jerome-Morais et al. (2012)
			Potential biomarker of tumor progression	AR	Veldscholte et al. (1992)
				Potential epigenetic regulation	Cao et al. (2014)
Colon cancer	Downregulation	Protein	Potential pharmacological target	Epigenetic regulation	(Pohl et al., 2009; Wang et al., 2014)
				Epithelial differentiation	(Li et al., 2008)
				Lipid/glucose metabolic signaling pathways;	Ying et al. (2015a)
		Protein	Positive prognostic factor	Mitochondrial function	Fang et al. (2010)
				GPX1	Li et al. (2021)
				Potential epigenetic regulation	
Pancreatic cancer	Downregulation		Unknown	Unknown	Caba et al., (2016)

(Continued on following page)

TABLE 2 (*Continued*) List of differentially expressed SELENBP1 during carcinogenesis and their functions.

Cancer type	Variation trend	Level	Functional classification	Roles in cancer	Reference
Head and neck cancer	Downregulation		Novel prognostic biomarker	Unknown Epigenetic regulation, transcriptional factors, and miRNA in oral cancer	Chen et al. (2016) Zeng et al. (2021)
Skin cancer	Downregulation		Unknown	Proliferation and tumor microenvironment HIF-1α	(Lanfear et al., 1993; Schott et al., 2018) Scortegagna et al. (2009)
Bladder cancer	Downregulation		Potential prognostic biomarker and therapeutic target	Epigenetic regulation	Wang et al., (2020)
Uterine cancer	Downregulation	Protein	Unknown	Unknown	Zhang et al. (2010a)
Nerve cancer	Unknown	/	Unknown	Cell outgrowth	Wang et al. (2015)
Malignant pleural mesothelioma	Unknown	/	Risk	Proliferation and transformation	Pass et al. (2004)
Uveal melanoma tumor	Upregulation		Metastatic prediction	Unknown	Linge et al. (2012)
Ovary cancer	Downregulation	Protein	Potential prognostic indicator	Selenium/androgen pathways Epithelial proliferation and papillary complexity	Huang et al. (2006) Zhang et al. (2010b)
			Risk	Stromal epithelial cross-talk	Johnatty et al. (2010)

High Se exposure leads to inhibition of DNMT expression/activity in prostate cancer (Xiang et al., 2008; Lee et al., 2009), breast carcinoma (de Miranda et al., 2014), and human colon cancer (Uthus et al., 2011). In addition, Se causes demethylation in LNCaP prostate cancer cells (Xiang et al., 2008) and human colon cancer Caco-2 cells (Uthus et al., 2011) and HDAC inhibition in prostate (Xiang et al., 2008; Lee et al., 2009), lymphoma (Kassam et al., 2011), and skin melanoma (Gowda et al., 2012). The epigenome changes influenced by Se may provide potential disease therapies and prevention strategies in cancers (Jablonska and Reszka, 2017).

Since SELENBP1 belongs to specific selenium-binding protein and its potential regulation in epigenetics during carcinogenesis, combined treatment targeting SELENBP1 and selenium might be a potential treatment strategy. In prostate cancer, SELENBP1 may play a role in controlling plasma selenium levels (Xie et al., 2016). For diseases such as prostate cancer, both SELENBP1 and selenium levels should be further examined for better treatment outcomes. The 3D model of human protein SELENBP1 with four cysteine residues was reported to have an alpha–beta structure characterized by four short alpha-helices, one 310 helix (residues 207–210), and 30 antiparallel beta-strands. Only one cysteine (Cys57) is able to bind the selenium in hepatocellular carcinoma, which may serve as a treatment target too (Raucci et al., 2011). The selenium-containing HDAC inhibitor or organic selenium compound

TABLE 3 List of Se regulated by epigenetic modifications during carcinogenesis.

Cancer Type	Cell type	Epigenetic modifications	Reference
Prostate cancer	LNCaP cells	Demethylation; DNMT1 inhibition;	Xiang et al. (2008)
	LNCaP cells	HDAC inhibition	Lee et al. (2009)
Breast cancer	MCF-7	DNMT1 inhibition	de Miranda et al. (2014)
Colon cancer	Human colonic epithelial Caco-2 cell	Demethylation	Uthus et al. (2011)
Diffuse large B-cell lymphoma	Lymphoma cell lines	HDAC inhibition	Kassam et al. (2011)
Cervical carcinoma	Hela cells	HDAC inhibition	Desai et al. (2010)
Skin melanoma	Hela cells and melanoma cell line	HDAC inhibition	Gowda et al. (2012)

becomes the novel inhibitor for cancer treatment (Lee et al., 2009; Desai et al., 2010; Kassam et al., 2011), which merits favorable outcomes *in vitro* and *in vivo*.

There is no evidence of a dose–response relation between selenium status and cancer risk (Vinceti et al., 2018). Whether supplemental dietary selenium should be considered as an adjuvant therapy for cancers still lacks certainty of evidence. Epigenetic drugs such as DNMTi and HDACi targeting SELENBP1 need further investigation before approval to treat different cancers clinically. In addition, microRNA has been reported for years as a potential target in cancers, such as lung, breast, prostate, colon, and liver carcinomas (Szczepanek et al., 2022). In OSCC, SELENBP1 is downregulated by miR-4786-3p through binding to the 3′UTR of SELENBP1 (Zeng et al., 2021). The antagomir of miR-4786-3p implies a potential treatment target of SELENBP1 in OSCC. Since the the number of studies about microRNAs' role in regulating SELENBP1 in other cancers is still limited, more studies should be performed to discover the role of microRNAs on SELENBP1 and develop potential treatment strategies.

Conclusion and future prospects

The expression of SELENBP1 is significantly downregulated in various cancers, which means SELENBP1 plays an important role in cancer progression. Although the molecular aspect of epigenetic changes of SELENBP1 in cancer remains largely unexplored and has not been used clinically in the diagnosis and treatments of patients with cancer yet, the prospects of novel discoveries and potential application of SELENBP1 are promising. Ongoing studies and tremendous progresses in the field of epigenetics have been made through transcriptome analysis and *in vitro*/*in vivo* experiments to unravel the potential epigenetic markers for future cancer therapeutics and prognosis. So, with the limited published articles, we summarize recent findings on the function and regulatory mechanisms of SELENBP1 during carcinogenic progression, with emphasis on epigenetic mechanisms.

There are still some crucial questions that should be addressed but can be easily neglected. 1) Whether the epigenetics change of SELENBP1 is one of the most important mechanisms for its expression and function during carcinogenesis. Many other mechanisms, such as metabolic signaling pathways (Ying et al., 2015a; Elhodaky et al., 2020), estrogen (Lanfear et al., 1993; Zhang C. et al., 2013), also effect the role of SELENBP1 in carcinogenesis. 2) If targeting only one specific epigenetic change at individual loci of SELENBP1 is enough to rescue its levels and activity during carcinogenesis, such as DNA methylation and protein ubiquitination. 3) Could the regulatory role of SELENBP1 epigenetic modification lead to therapy or prognosis? Further investigation is needed to identify novel epigenetic modifications of SELENBP1 through new technologies and dissect out the fundamental roles of these modifications on SELENBP1 expression and activities. Pharmacologic inhibition of epigenetic modifications could be key to reducing carcinogenic progression. Using specific inhibitors to target SELENBP1 might serve as therapeutic targets and lead to the desired clinical outcome (Table 2, Table 3) (Waddington, 1942; Porat et al., 2000; Haig, 2004; Feinberg, 2018; Millan-Zambrano et al., 2022).

Author contributions

YZ: investigation, validation, analysis, writing–original draft preparation, writing—review and editing, and visualization. QH: conceptualization, funding acquisition, writing—review and editing, and visualization. All authors approved the manuscript and agree to be accountable for the work.

Funding

This work was supported by the National Natural Science Foundation of China (82072483), the Natural Science Foundation of Hubei Province (ZRMS2020000955), and the Key Research and Development Program of Hubei Province (2022BCA052).

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

References

Adam, J., Hatipoglu, E., O'Flaherty, L., Ternette, N., Sahgal, N., Lockstone, H., et al. (2011). Renal cyst formation in fh1-deficient mice is independent of the hif/phd pathway: Roles for fumarate in KEAP1 succination and Nrf2 signaling. *Cancer Cell* 20 (4), 524–537. doi:10.1016/j.ccr.2011.09.006

Ansong, E., Ying, Q., Ekoue, D. N., Deaton, R., Hall, A. R., Kajdacsy-Balla, A., et al. (2015). Evidence that selenium binding protein 1 is a tumor suppressor in prostate cancer. *PLoS One* 10 (5), e0127295. doi:10.1371/journal.pone.0127295

Bansal, M. P., Mukhopadhyay, T., Scott, J., Cook, R. G., Mukhopadhyay, R., and MeDina, D. (1990). DNA sequencing of a mouse liver protein that binds selenium: Implications for selenium's mechanism of action in cancer prevention. *Carcinogenesis* 11 (11), 2071–2073. doi:10.1093/carcin/11.11.2071

Beasley, R. P., Hwang, L. Y., Lin, C. C., and Chien, C. S. (1981). Hepatocellular carcinoma and Hepatitis B virus. A prospective study of 22 707 men in Taiwan. *Lancet* 2 (8256), 1129–1133. doi:10.1016/s0140-6736(81)90585-7

Behne, D., and Kyriakopoulos, A. (2001). Mammalian selenium-containing proteins. *Annu. Rev. Nutr.* 21, 453–473. doi:10.1146/annurev.nutr.21.1.453

Brown, L. M., Helmke, S. M., Hunsucker, S. W., Netea-Maier, R. T., Chiang, S. A., Heinz, D. E., et al. (2006). Quantitative and qualitative differences in protein expression between papillary thyroid carcinoma and normal thyroid tissue. *Mol. Carcinog.* 45 (8), 613–626. doi:10.1002/mc.20193

Caba, O., Irigoyen, A., Jimenez-Luna, C., Benavides, M., Ortuno, F. M., Gallego, J., et al. (2016). Identification of gene expression profiling associated with erlotinib-related skin toxicity in pancreatic adenocarcinoma patients. *Toxicol. Appl. Pharmacol.* 311, 113–116. doi:10.1016/j.taap.2016.10.003

Cao, Q., Wang, X., Zhao, M., Yang, R., Malik, R., Qiao, Y., et al. (2014). The central role of EED in the orchestration of polycomb group complexes. *Nat. Commun.* 5, 3127. doi:10.1038/ncomms4127

Caswell, D. R., Chuang, C. H., Ma, R. K., Winters, I. P., Snyder, E. L., and Winslow, M. M. (2018). Tumor suppressor activity of SELENBP1, a direct nkx2-1 target, in lung adenocarcinoma. *Mol. Cancer Res.* 16 (11), 1737–1749. doi:10.1158/1541-7786.MCR-18-0392

Chang, P. W., Tsui, S. K., Liew, C., Lee, C. C., Waye, M. M., and Fung, K. P. (1997). Isolation, characterization, and chromosomal mapping of a novel cDNA clone encoding human selenium binding protein. *J. Cell. Biochem.* 64 (2), 217–224. doi:10.1002/(sici)1097-4644(199702)64:2<217::aid-jcb5>3.0.co;2-#

Chen, G., Wang, H., Miller, C. T., Thomas, D. G., Gharib, T. G., Misek, D. E., et al. (2004). Reduced selenium-binding protein 1 expression is associated with poor outcome in lung adenocarcinomas. *J. Pathol.* 202 (3), 321–329. doi:10.1002/path.1524

Chen, F., Chen, C., Qu, Y., Xiang, H., Ai, Q., Yang, F., et al. (2016). Selenium-binding protein 1 in head and neck cancer is low-expression and associates with the prognosis of nasopharyngeal carcinoma. *Med. Baltim.* 95 (35), e4592. doi:10.1097/MD.0000000000004592

de Miranda, J. X., Andrade, F. O., Conti, A., Dagli, M. L. Z., Moreno, F. S., and Ong, T. P. (2014). Effects of selenium compounds on proliferation and epigenetic marks of breast cancer cells. *J. Trace Elem. Med. Biol.* 28 (4), 486–491. doi:10.1016/j.jtemb.2014.06.017

Desai, D., Salli, U., Vrana, K. E., and Amin, S. (2010). SelSA, selenium analogs of SAHA as potent histone deacetylase inhibitors. *Bioorg. Med. Chem. Lett.* 20 (6), 2044–2047. doi:10.1016/j.bmcl.2009.07.068

Elder, J. T., and Zhao, X. (2002). Evidence for local control of gene expression in the epidermal differentiation complex. *Exp. Dermatol.* 11 (5), 406–412. doi:10.1034/j.1600-0625.2002.110503.x

Elhodaky, M., Hong, L. K., Kadkol, S., and Diamond, A. M. (2020). Selenium-binding protein 1 alters energy metabolism in prostate cancer cells. *Prostate* 80 (12), 962–976. doi:10.1002/pros.24028

Elkjaer, M. L., Nawrocki, A., Kacprowski, T., Lassen, P., Simonsen, A. H., Marignier, R., et al. (2021). CSF proteome in multiple sclerosis subtypes related to brain lesion transcriptomes. *Sci. Rep.* 11 (1), 4132. doi:10.1038/s41598-021-83591-5

Fagerberg, L., Hallstrom, B. M., Oksvold, P., Kampf, C., Djureinovic, D., Odeberg, J., et al. (2014). Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol. Cell. Proteomics* 13 (2), 397–406. doi:10.1074/mcp.M113.035600

Fang, W., Goldberg, M. L., Pohl, N. M., Bi, X., Tong, C., Xiong, B., et al. (2010). Functional and physical interaction between the selenium-binding protein 1 (SBP1) and the glutathione peroxidase 1 selenoprotein. *Carcinogenesis* 31 (8), 1360–1366. doi:10.1093/carcin/bgq114

Feinberg, A. P. (2018). The key role of epigenetics in human disease prevention and mitigation. *N. Engl. J. Med.* 378 (14), 1323–1334. doi:10.1056/NEJMra1402513

Flemetakis, E., Agalou, A., Kavroulakis, N., Dimou, M., Martsikovskaya, A., Slater, A., et al. (2002). Lotus japonicus gene Ljsbp is highly conserved among plants and animals and encodes a homologue to the mammalian selenium-binding proteins. *Mol. Plant. Microbe. Interact.* 15 (4), 313–322. doi:10.1094/MPMI.2002.15.4.313

Gao, P. T., Ding, G. Y., Yang, X., Dong, R. Z., Hu, B., Zhu, X. D., et al. (2018). Invasive potential of hepatocellular carcinoma is enhanced by loss of selenium-binding protein 1 and subsequent upregulation of CXCR4. *Am. J. Cancer Res.* 8 (6), 1040–1049.

Gianni, M., Terao, M., Kurosaki, M., Paroni, G., Brunelli, L., Pastorelli, R., et al. (2019). S100A3 a partner protein regulating the stability/activity of RARα and PML-RARα in cellular models of breast/lung cancer and acute myeloid leukemia. *Oncogene* 38 (14), 2482–2500. doi:10.1038/s41388-018-0599-z

Giurato, G., Nassa, G., Salvati, A., Alexandrova, E., Rizzo, F., Nyman, T. A., et al. (2018). Quantitative mapping of RNA-mediated nuclear estrogen receptor beta interactome in human breast cancer cells. *Sci. Data* 5, 180031. doi:10.1038/sdata.2018.31

Gowda, R., Madhunapantula, S. V., Desai, D., Amin, S., and Robertson, G. P. (2012). Selenium-containing histone deacetylase inhibitors for melanoma management. *Cancer Biol. Ther.* 13 (9), 756–765. doi:10.4161/cbt.20558

Ha, Y. S., Lee, G. T., Kim, Y. H., Kwon, S. Y., Choi, S. H., Kim, T. H., et al. (2014). Decreased selenium-binding protein 1 mRNA expression is associated with poor prognosis in renal cell carcinoma. *World J. Surg. Oncol.* 12, 288. doi:10.1186/1477-7819-12-288

Haig, D. (2004). The (dual) origin of epigenetics. *Cold Spring Harb. Symp. Quant. Biol.* 69, 67–70. doi:10.1101/sqb.2004.69.67

He, Q. Y., Cheung, Y. H., Leung, S. Y., Yuen, S. T., Chu, K. M., and Chiu, J. F. (2004). Diverse proteomic alterations in gastric adenocarcinoma. *Proteomics* 4 (10), 3276–3287. doi:10.1002/pmic.200300916

Hu, B., Jin, J., Guo, A. Y., Zhang, H., Luo, J., and Gao, G. (2015). Gsds 2.0: An upgraded gene feature visualization server. *Bioinformatics* 31 (8), 1296–1297. doi:10.1093/bioinformatics/btu817

Huang, K. C., Park, D. C., Ng, S. K., Lee, J. Y., Ni, X., Ng, W. C., et al. (2006). Selenium binding protein 1 in ovarian cancer. *Int. J. Cancer* 118 (10), 2433–2440. doi:10.1002/ijc.21671

Huang, C., Ding, G., Gu, C., Zhou, J., Kuang, M., Ji, Y., et al. (2012). Decreased selenium-binding protein 1 enhances glutathione peroxidase 1 activity and downregulates HIF-1α to promote hepatocellular carcinoma invasiveness. *Clin. Cancer Res.* 18 (11), 3042–3053. doi:10.1158/1078-0432.CCR-12-0183

Hussain, S., Tulsyan, S., Dar, S. A., Sisodiya, S., Abiha, U., Kumar, R., et al. (2022). Role of epigenetics in carcinogenesis: Recent advancements in anticancer therapy. *Semin. Cancer Biol.* 83, 441–451. (Hussain et al., 2022). doi:10.1016/j.semcancer.2021.06.023

Itoh, K., Wakabayashi, N., Katoh, Y., Ishii, T., Igarashi, K., Engel, J. D., et al. (1999). Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev.* 13 (1), 76–86. doi:10.1101/gad.13.1.76

Jablonska, E., and Reszka, E. (2017). Selenium and epigenetics in cancer: Focus on DNA methylation. *Adv. Cancer Res.* 136, 193–234. doi:10.1016/bs.acr.2017.07.002

Jeong, J. Y., Wang, Y., and Sytkowski, A. J. (2009). Human selenium binding protein-1 (hSP56) interacts with VDU1 in a selenium-dependent manner. *Biochem. Biophys. Res. Commun.* 379 (2), 583–588. doi:10.1016/j.bbrc.2008.12.110

Jeong, J. Y., Zhou, J. R., Gao, C., Feldman, L., and Sytkowski, A. J. (2014). Human selenium binding protein-1 (hSP56) is a negative regulator of HIF-1α and suppresses the malignant characteristics of prostate cancer cells. *BMB Rep.* 47 (7), 411–416. doi:10.5483/bmbrep.2014.47.7.104

Jerome-Morais, A., Wright, M. E., Liu, R., Yang, W., Jackson, M. I., Combs, G. F., et al. (2012). Inverse association between glutathione peroxidase activity and both selenium-binding protein 1 levels and Gleason score in human prostate tissue. *Prostate* 72 (9), 1006–1012. doi:10.1002/pros.21506

Johnatty, S. E., Beesley, J., Chen, X., Macgregor, S., Duffy, D. L., Spurdle, A. B., et al. (2010). Evaluation of candidate stromal epithelial cross-talk genes identifies association between risk of serous ovarian cancer and TERT, a cancer susceptibility "hot-spot. *PLoS Genet.* 6 (7), e1001016. doi:10.1371/journal.pgen.1001016

Jones, P. A., and Baylin, S. B. (2002). The fundamental role of epigenetic events in cancer. *Nat. Rev. Genet.* 3 (6), 415–428. doi:10.1038/nrg816

Kassam, S., Goenaga-Infante, H., Maharaj, L., Hiley, C. T., Juliger, S., and Joel, S. P. (2011). Methylseleninic acid inhibits HDAC activity in diffuse large B-cell

lymphoma cell lines. *Cancer Chemother. Pharmacol.* 68 (3), 815–821. doi:10.1007/s00280-011-1649-1

Kim, H., Kang, H. J., You, K. T., Kim, S. H., Lee, K. Y., Kim, T. I., et al. (2006). Suppression of human selenium-binding protein 1 is a late event in colorectal carcinogenesis and is associated with poor survival. *Proteomics* 6 (11), 3466–3476. doi:10.1002/pmic.200500629

Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R., et al. (2003). Characterization of mammalian selenoproteomes. *Science* 300 (5624), 1439–1443. doi:10.1126/science.1083516

Laird, P. W. (2003). The power and the promise of DNA methylation markers. *Nat. Rev. Cancer* 3 (4), 253–266. doi:10.1038/nrc1045

Lanfear, J., Fleming, J., Walker, M., and Harrison, P. (1993). Different patterns of regulation of the genes encoding the closely related 56 kDa selenium- and acetaminophen-binding proteins in normal tissues and during carcinogenesis. *Carcinogenesis* 14 (3), 335–340. doi:10.1093/carcin/14.3.335

Lee, J. I., Nian, H., Cooper, A. J., Sinha, R., Dai, J., Bisson, W. H., et al. (2009). Alpha-keto acid metabolites of naturally occurring organoselenium compounds as inhibitors of histone deacetylase in human prostate cancer cells. *Cancer Prev. Res.* 2 (7), 683–693. doi:10.1158/1940-6207.CAPR-09-0047

Lee, Y. M., Kim, S., Park, R. Y., and Kim, Y. S. (2020). Hepatitis B virus-X downregulates expression of selenium binding protein 1. *Viruses* 12 (5), E565. doi:10.3390/v12050565

Li, L. S., Kim, H., Rhee, H., Kim, S. H., Shin, D. H., Chung, K. Y., et al. (2004). Proteomic analysis distinguishes basaloid carcinoma as a distinct subtype of nonsmall cell lung carcinoma. *Proteomics* 4 (11), 3394–3400. doi:10.1002/pmic.200400901

Li, Z., Wang, D., Messing, E. M., and Wu, G. (2005). VHL protein-interacting deubiquitinating enzyme 2 deubiquitinates and stabilizes HIF-1alpha. *EMBO Rep.* 6 (4), 373–378. doi:10.1038/sj.embor.7400377

Li, T., Yang, W., Li, M., Byun, D. S., Tong, C., Nasser, S., et al. (2008). Expression of selenium-binding protein 1 characterizes intestinal cell maturation and predicts survival for patients with colorectal cancer. *Mol. Nutr. Food Res.* 52 (11), 1289–1299. doi:10.1002/mnfr.200700331

Li, B., Qi, Z. P., He, D. L., Chen, Z. H., Liu, J. Y., Wong, M. W., et al. (2021). NLRP7 deubiquitination by USP10 promotes tumor progression and tumor-associated macrophage polarization in colorectal cancer. *J. Exp. Clin. Cancer Res.* 40 (1), 126. doi:10.1186/s13046-021-01920-y

Linge, A., Kennedy, S., O'Flynn, D., Beatty, S., Moriarty, P., Henry, M., et al. (2012). Differential expression of fourteen proteins between uveal melanoma from patients who subsequently developed distant metastases versus those who did Not. *Invest.. Ophthalmol. Vis. Sci.* 53 (8), 4634–4643. doi:10.1167/iovs.11-9019

Lucas, B., Grigo, K., Erdmann, S., Lausen, J., Klein-Hitpass, L., and Ryffel, G. U. (2005). HNF4alpha reduces proliferation of kidney cells and affects genes deregulated in renal cell carcinoma. *Oncogene* 24 (42), 6418–6431. doi:10.1038/sj.onc.1208794

Marenholz, I., Zirra, M., Fischer, D. F., BaCkendorf, C., Ziegler, A., and Mischke, D. (2001). Identification of human epidermal differentiation complex (EDC)-encoded genes by subtractive hybridization of entire YACs to a gridded keratinocyte cDNA library. *Genome Res.* 11 (3), 341–355. doi:10.1101/gr.114801

Millan-Zambrano, G., Burton, A., Bannister, A. J., and Schneider, R. (2022). Histone post-translational modifications - cause and consequence of genome function. *Nat. Rev. Genet.* 23, 563–580. doi:10.1038/s41576-022-00468-7

Minich, W. B. (2022). Selenium metabolism and biosynthesis of selenoproteins in the human body. *Biochemistry.* 87 (1), S168–S102. doi:10.1134/S0006297922140139

Miyaguchi, K. (2004). Localization of selenium-binding protein at the tips of rapidly extending protrusions. *Histochem. Cell Biol.* 121 (5), 371–376. doi:10.1007/s00418-004-0623-y

Nazemalhosseini, M. E., Kuppen, P. J., Aghdaei, H. A., and Zali, M. R. (2013). The CpG island methylator phenotype (CIMP) in colorectal cancer. *Gastroenterol. Hepatol. Bed Bench* 6 (3), 120–128.

Pass, H. I., Liu, Z., Wali, A., Bueno, R., Land, S., Lott, D., et al. (2004). Gene expression profiles predict survival and progression of pleural mesothelioma. *Clin. Cancer Res.* 10 (3), 849–859. doi:10.1158/1078-0432.ccr-0607-3

Pohl, N. M., Tong, C., Fang, W., Bi, X., Li, T., and Yang, W. (2009). Transcriptional regulation and biological functions of selenium-binding protein 1 in colorectal cancer *in vitro* and in nude mouse xenografts. *PLoS One* 4 (11), e7774. doi:10.1371/journal.pone.0007774

Pol, A., Renkema, G. H., Tangerman, A., Winkel, E. G., Engelke, U. F., de Brouwer, A. P. M., et al. (2018). Mutations in SELENBP1, encoding a novel human methanethiol oxidase, cause extraoral halitosis. *Nat. Genet.* 50 (1), 120–129. doi:10.1038/s41588-017-0006-7

Porat, A., Sagiv, Y., and Elazar, Z. (2000). A 56-kDa selenium-binding protein participates in intra-Golgi protein transport. *J. Biol. Chem.* 275 (19), 14457–14465. doi:10.1074/jbc.275.19.14457

Principe, S., Jones, E. E., Kim, Y., Sinha, A., Nyalwidhe, J. O., Brooks, J., et al. (2013). In-depth proteomic analyses of exosomes isolated from expressed prostatic secretions in urine. *Proteomics* 13 (10-11), 1667–1671. doi:10.1002/pmic.201200561

Raucci, R., Colonna, G., Guerriero, E., Capone, F., Accardo, M., Castello, G., et al. (2011). Structural and functional studies of the human selenium binding protein-1 and its involvement in hepatocellular carcinoma. *Biochim. Biophys. Acta* 1814 (4), 513–522. doi:10.1016/j.bbapap.2011.02.006

Schott, M., de Jel, M. M., Engelmann, J. C., Renner, P., Geissler, E. K., Bosserhoff, A. K., et al. (2018). Selenium-binding protein 1 is down-regulated in malignant melanoma. *Oncotarget* 9 (12), 10445–10456. doi:10.18632/oncotarget.23853

Scortegagna, M., Martin, R. J., Kladney, R. D., Neumann, R. G., and Arbeit, J. M. (2009). Hypoxia-inducible factor-1alpha suppresses squamous carcinogenic progression and epithelial-mesenchymal transition. *Cancer Res.* 69 (6), 2638–2646. doi:10.1158/0008-5472.CAN-08-3643

Seelig, J., Heller, R. A., Haubruck, P., Sun, Q., Georg Klingenberg, J., Hackler, J., et al. (2021). Selenium-binding protein 1 (SELENBP1) as biomarker for adverse clinical outcome after traumatic spinal cord injury. *Front. Neurosci.* 15, 680240. doi:10.3389/fnins.2021.680240

Shami, S. A., Batrouni, A. G., Kim, D., Punyala, A., Cao, W., Han, C., et al. (2019). PLEKHA4/kramer attenuates dishevelled ubiquitination to modulate wnt and planar cell polarity signaling. *Cell Rep.* 27 (7), 2157–2170. doi:10.1016/j.celrep.2019.04.060

Silvers, A. L., Lin, L., Bass, A. J., Chen, G., Wang, Z., Thomas, D. G., et al. (2010). Decreased selenium-binding protein 1 in esophageal adenocarcinoma results from posttranscriptional and epigenetic regulation and affects chemosensitivity. *Clin. Cancer Res.* 16 (7), 2009–2021. doi:10.1158/1078-0432.CCR-09-2801

Sowa, M. E., Bennett, E. J., Gygi, S. P., and Harper, J. W. (2009). Defining the human deubiquitinating enzyme interaction landscape. *Cell* 138 (2), 389–403. doi:10.1016/j.cell.2009.04.042

Speckmann, B., and Grune, T. (2015). Epigenetic effects of selenium and their implications for health. *Epigenetics* 10 (3), 179–190. doi:10.1080/15592294.2015.1013792

Stukalov, A., Girault, V., Grass, V., Karayel, O., Bergant, V., Urban, C., et al. (2021). Multilevel proteomics reveals host perturbations by SARS-CoV-2 and SARS-CoV. *Nature* 594 (7862), 246–252. doi:10.1038/s41586-021-03493-4

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 (3), 209–249. doi:10.3322/caac.21660

Szczepanek, J., Skorupa, M., and Tretyn, A. (2022). MicroRNA as a potential therapeutic molecule in cancer. *Cells* 11 (6), 1008. doi:10.3390/cells11061008

Tan, X., Liao, L., Wan, Y. P., Li, M. X., Chen, S. H., Mo, W. J., et al. (2016). Downregulation of selenium-binding protein 1 is associated with poor prognosis in lung squamous cell carcinoma. *World J. Surg. Oncol.* 14, 70. doi:10.1186/s12957-016-0832-6

Tiemann, K., Garri, C., Lee, S. B., Malihi, P. D., Park, M., Alvarez, R. M., et al. (2019). Loss of ER retention motif of AGR2 can impact mTORC signaling and promote cancer metastasis. *Oncogene* 38 (16), 3003–3018. doi:10.1038/s41388-018-0638-9

Uthus, E., Begaye, A., Ross, S., and Zeng, H. (2011). The von Hippel-Lindau (VHL) tumor-suppressor gene is down-regulated by selenium deficiency in Caco-2 cells and rat colon mucosa. *Biol. Trace Elem. Res.* 142 (2), 223–231. doi:10.1007/s12011-010-8764-4

Veldscholte, J., Berrevoets, C. A., Brinkmann, A. O., Grootegoed, J. A., and MuldEr, E. (1992). Anti-androgens and the mutated androgen receptor of LNCaP cells: Differential effects on binding affinity, heat-shock protein interaction, and transcription activation. *Biochemistry* 31 (8), 2393–2399. doi:10.1021/bi00123a026

Vinceti, M., Filippini, T., Del, G. C., Dennert, G., Zwahlen, M., Brinkman, M., et al. (2018). Selenium for preventing cancer. *Cochrane Database Syst. Rev.* 1, CD005195. doi:10.1002/14651858.CD005195.pub4

Waddington, C. H. (2012). The epigenotype. *Int. J. Epidemiol.* 41 (1), 10–13. doi:10.1093/ije/dyr184

Wang, N., Chen, Y., Han, Y., Zhao, Y., Liu, Y., Guo, K., et al. (2012). Proteomic analysis shows down-regulations of cytoplasmic carbonic anhydrases, CAI and CAII, are early events of colorectal carcinogenesis but are not correlated with lymph node metastasis. *Tumori* 98 (6), 783–791. doi:10.1700/1217.13504

Wang, N., Chen, Y., Yang, X., and Jiang, Y. (2014). Selenium-binding protein 1 is associated with the degree of colorectal cancer differentiation and is regulated by histone modification. *Oncol. Rep.* 31 (6), 2506–2514. doi:10.3892/or.2014.3141

Wang, Y., Fang, W., Huang, Y., Hu, F., Ying, Q., Yang, W., et al. (2015). Reduction of selenium-binding protein 1 sensitizes cancer cells to selenite *via* elevating extracellular glutathione: A novel mechanism of cancer-specific cytotoxicity of selenite. *Free Radic. Biol. Med.* 79, 186–196. doi:10.1016/j.freeradbiomed.2014.11.015

Wang, Y., Zhu, W., Chen, X., Wei, G., Jiang, G., and Zhang, G. (2020). Selenium-binding protein 1 transcriptionally activates p21 expression *via* p53-independent mechanism and its frequent reduction associates with poor prognosis in bladder cancer. *J. Transl. Med.* 18 (1), 17. doi:10.1186/s12967-020-02211-4

Woods, N. T., Mesquita, R. D., Sweet, M., Carvalho, M. A., Li, X., Liu, Y., et al. (2012). Charting the landscape of tandem BRCT domain-mediated protein interactions. *Sci. Signal.* 5 (242), s6. doi:10.1126/scisignal.2002255

Xia, Y. J., Ma, Y. Y., He, X. J., Wang, H. J., Ye, Z. Y., and Tao, H. Q. (2011). Suppression of selenium-binding protein 1 in gastric cancer is associated with poor survival. *Hum. Pathol.* 42 (11), 1620–1628. doi:10.1016/j.humpath.2011.01.008

Xiang, N., Zhao, R., Song, G., and Zhong, W. (2008). Selenite reactivates silenced genes by modifying DNA methylation and histones in prostate cancer cells. *Carcinogenesis* 29 (11), 2175–2181. doi:10.1093/carcin/bgn179

Xie, W., Yang, M., Chan, J., Sun, T., Mucci, L. A., Penney, K. L., et al. (2016). Association of genetic variations of selenoprotein genes, plasma selenium levels, and prostate cancer aggressiveness at diagnosis. *Prostate* 76 (7), 691–699. doi:10.1002/pros.23160

Yang, M., and Sytkowski, A. J. (1998). Differential expression and androgen regulation of the human selenium-binding protein gene hSP56 in prostate cancer cells. *Cancer Res.* 58 (14), 3150–3153.

Ying, Q., Ansong, E., Diamond, A. M., and Yang, W. (2015a). A critical role for cysteine 57 in the biological functions of selenium binding protein-1. *Int. J. Mol. Sci.* 16 (11), 27599–27608. doi:10.3390/ijms161126043

Ying, Q., Ansong, E., Diamond, A. M., Lu, Z., Yang, W., and Bie, X. (2015b). Quantitative proteomic analysis reveals that anti-cancer effects of selenium-binding protein 1 *in vivo* are associated with metabolic pathways. *PLoS One* 10 (5), e0126285. doi:10.1371/journal.pone.0126285

Yu-Rice, Y., Edassery, S. L., Urban, N., Hellstrom, I., Hellstrom, K. E., Deng, Y., et al. (2017). Selenium-Binding Protein 1 (SBP1) autoantibodies in ovarian disorders and ovarian cancer. *Reproduction* 153 (3), 277–284. doi:10.1530/REP-16-0265

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., et al. (2014). A comparative encyclopedia of DNA elements in the mouse genome. *Nature* 515 (7527), 355–364. doi:10.1038/nature13992

Zaib, S., and Khan, I. (2022). Rana N and khan I: Histone modifications and their role in epigenetics of cancer. *Curr. Med. Chem.* 29 (14), 2399–2411. doi:10.2174/0929867328666211108105214

Zeng, G. Q., Yi, H., Zhang, P. F., Li, X. H., Hu, R., Li, M. Y., et al. (2013). The function and significance of SELENBP1 downregulation in human bronchial epithelial carcinogenic process. *PLoS One* 8 (8), e71865. doi:10.1371/journal.pone.0071865

Zeng, H., Zhao, X., and Tang, C. (2021). Downregulation of SELENBP1 enhances oral squamous cell carcinoma chemoresistance through KEAP1-NRF2 signaling. *Cancer Chemother. Pharmacol.* 88 (2), 223–233. doi:10.1007/s00280-021-04284-4

Zhang, C., Wang, Y. E., Zhang, P., Liu, F., Sung, C. J., Steinhoff, M. M., et al. (2010a). Progressive loss of selenium-binding protein 1 expression correlates with increasing epithelial proliferation and papillary complexity in ovarian serous borderline tumor and low-grade serous carcinoma. *Hum. Pathol.* 41 (2), 255–261. doi:10.1016/j.humpath.2009.07.019

Zhang, P., Zhang, C., Wang, X., Liu, F., Sung, C. J., Quddus, M. R., et al. (2010b). The expression of selenium-binding protein 1 is decreased in uterine leiomyoma. *Diagn. Pathol.* 5, 80. doi:10.1186/1746-1596-5-80

Zhang, J., Dong, W. G., and Lin, J. (2011a). Reduced selenium-binding protein 1 is associated with poor survival rate in gastric carcinoma. *Med. Oncol.* 28 (2), 481–487. doi:10.1007/s12032-010-9482-7

Zhang, J., Zhan, N., and Dong, W. G. (2011b). Altered expression of selenium-binding protein 1 in gastric carcinoma and precursor lesions. *Med. Oncol.* 28 (4), 951–957. doi:10.1007/s12032-010-9564-6

Zhang, C., Xu, W., Pan, W., Wang, N., Li, G., Fan, X., et al. (2013a). Selenium-binding protein 1 may decrease gastric cellular proliferation and migration. *Int. J. Oncol.* 42 (5), 1620–1629. doi:10.3892/ijo.2013.1850

Zhang, S., Li, F., Younes, M., Liu, H., Chen, C., and Yao, Q. (2013b). Reduced selenium-binding protein 1 in breast cancer correlates with poor survival and resistance to the anti-proliferative effects of selenium. *PLoS One* 8 (5), e63702. doi:10.1371/journal.pone.0063702

Zhang, Y., Sun, Z., Jia, J., Du, T., Zhang, N., Tang, Y., et al. (2021). Overview of histone modification. *Adv. Exp. Med. Biol.* 1283, 1–16. doi:10.1007/978-981-15-8104-5_1

Glossary

RCC Renal cell carcinoma

LSCC Lung squamous cell carcinoma

NBE Bronchial epithelium

PTC Papillary thyroid carcinoma

DIGE Difference gel electrophoresis

HCC Hepatocellular carcinoma

VDU1 Hippel–Lindau protein (pVHL)-interacting deubiquitinating enzyme 1

HIF-1α Hypoxia-inducible protein-1alpha

OXPHOS Oxidative phosphorylation

AMPK AMP-activated protein kinase

CRC Colorectal cancer

TSA Trichostatin A

VPA Valproic acid

5-Aza-dC 5′-Aza-2′-deoxycytidine

2-D PAGE Two-dimensional polyacrylamide gel electrophoresis

Se Selenium

hSP56 Human 56 kDa selenium-binding protein

DEN Diethylnitrosamine

OSCC Oral squamous cell carcinoma

UTR Untranslated region

NRF2 Nuclear factor erythroid 2-related factor 2

KEAP1 Kelch-like ECH-associated protein 1

GPX1 Glutathione peroxidase 1

HNSCC Head and neck squamous cell carcinoma

NPC Nasopharyngeal carcinoma

LC Laryngeal cancer

OC Oral cancer

TC Tonsil cancer

HPC Hypopharyngeal cancer

EAC Esophageal adenocarcinoma

USP33 Ubiquitin-Specific Peptidase 33

STRING Search Tool for the Retrieval of Interacting Genes/Proteins

PPI Protein–protein interaction

CXCR4 C-X-C motif chemokine receptor 4

EMT Epithelial–mesenchymal transition

B[a]P Benzo(a)pyrene

GSH Glutathione

PDAC Pancreatic ductal adenocarcinoma

SR Skin rash

HNF4alpha Hepatocyte nuclear factor 4 alpha

HBx Hepatitis B virus-X

CpG Cytosine–guanine

MPM Malignant pleural mesothelioma

AR Androgen receptor

ATRA All trans-retinoic acid

NLRP7 NACHT, LRR, and PYD domain-containing protein 7

TERT Telomerase reverse transcriptase

BRCT BRCA1 carboxyl-terminal

DDR Damage response

AGR2 Anterior gradient 2

Dubs Deubiquitinating enzymes

USP15 Ubiquitin-specific proteases 15

EPS Expressed prostatic secretions

ESR2 Estrogen receptor 2

EED Embryonic ectoderm development

PRC2 Polycomb repressive complexes 2

H3K27me3 Histone H3 at lysine 27

PRC1 Polycomb repressive complexes 1

VCaP Vertebral Cancer of the prostate

DNMTi DNA methylation inhibitor

DNMTs DNA methyltransferase

HDACi Histone deacetylase inhibitor

MDS Myelodysplastic syndromes

AML Acute myeloid leukemia

PTMs Post-translational modifications

miRNAs MicroRNAs

lncRNAs Long-non-coding RNAs

MS Multiple sclerosis

Scz Schizophrenia

COVID-19 Corona virus disease 2019

CNS Central nervous system

TSS Transcription start site

CDKN1A Cyclin-dependent kinase inhibitor 1A

USP10 Ubiquitin-specific protease 10

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 13 February 2023
DOI 10.3389/fgene.2023.1032683

OPEN ACCESS

EDITED BY
Mehdi Pirooznia,
Johnson & Johnson, United States

REVIEWED BY
Naoko Hattori,
Hoshi University, Japan
Dong Liu,
Huazhong University of Science and Technology, China

*CORRESPONDENCE
Yan Gong,
yan.gong@whu.edu.cn
Conghua Xie,
chxie_65@whu.edu.cn

†These authors have contributed equally to this work

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

RECEIVED 31 August 2022
ACCEPTED 24 January 2023
PUBLISHED 13 February 2023

CITATION
Liu X, Chen J, Li J, Zeng Z, Jiang X, Gao Y, Huang Z, Wu Q, Gong Y and Xie C (2023), Integrated analysis reveals common DNA methylation patterns of alcohol-associated cancers: A pan-cancer analysis.
Front. Genet. 14:1032683.
doi: 10.3389/fgene.2023.1032683

Integrated analysis reveals common DNA methylation patterns of alcohol-associated cancers: A pan-cancer analysis

Xingyu Liu[1†], Jiarui Chen[1†], Jiali Li[1], Zihang Zeng[1], Xueping Jiang[1], Yanping Gao[1], Zhengrong Huang[2,3], Qiuji Wu[1,4,5], Yan Gong[2,3]* and Conghua Xie[1,4,5]*

[1]Department of Radiation and Medical Oncology, Zhongnan Hospital of Wuhan University, Wuhan, China, [2]Department of Biological Repositories, Zhongnan Hospital of Wuhan University, Wuhan, China, [3]Tumor Precision Diagnosis and Treatment Technology and Translational Medicine, Hubei Engineering Research Center, Zhongnan Hospital of Wuhan University, Wuhan, China, [4]Hubei Key Laboratory of Tumor Biological Behaviors, Zhongnan Hospital of Wuhan University, Wuhan, China, [5]Hubei Cancer Clinical Study Center, Zhongnan Hospital of Wuhan University, Wuhan, China

Background: The role of alcohol in carcinogenesis has received increasing attention in recent years. Evidence shows its impacts on various aspects, including epigenetics alteration. The DNA methylation patterns underlying alcohol-associated cancers are not fully understood.

Methods: We investigated the aberrant DNA methylation patterns in four alcohol-associated cancers based on the Illumina HumanMethylation450 BeadChip. Pearson coefficient correlations were identified between differential methylated CpG probes and annotated genes. Transcriptional factor motifs were enriched and clustered using MEME Suite, and a regulatory network was constructed.

Results: In each cancer, differential methylated probes (DMPs) were identified, and 172 hypermethylated and 21 hypomethylated pan-cancer DMPs (PDMPs) were examined further. Annotated genes significantly regulated by PDMPs were investigated and enriched in transcriptional misregulation in cancers. The CpG island *chr19:58220189–58220517* was hypermethylated in all four cancers and silenced in the transcription factor *ZNF154*. Various biological effects were exerted by 33 hypermethylated and seven hypomethylated transcriptional factor motifs grouped into five clusters. Eleven pan-cancer DMPs were identified to be associated with clinical outcomes in the four alcohol-associated cancers, which might provide a potential point of view for clinical outcome prediction.

Conclusion: This study provides an integrated insight into DNA methylation patterns in alcohol-associated cancers and reveals the corresponding features, influences, and potential mechanisms.

Abbreviations: DMPs, differential methylated probes; PDMPs, pan-cancer differential methylated probes; ESCA, esophageal carcinoma; HNSCC, head and neck squamous cell carcinoma; LIHC, liver hepatocellular carcinoma; PAAD, pancreatic adenocarcinoma; ZNF154, zinc finger protein 154; TF, transcription factor; FC, fold change; PCCs, Pearson coefficient correlations; KEGG, Kyoto Encyclopedia of Genes and Genomes; MPI, methylation prognosis index; AUC, area under the curve; HeP, hypermethylated promotor; HeB, hypermethylated body; HoP, hypomethylated promotor; HoB, hypomethylated body; PATZ1, POZ/BTB and AT-hook-containing zinc finger protein 1.

KEYWORDS

alcohol consumption, alcohol-associated cancer, pan-cancer differentially methylated probe, CpG island, transcription factor motifs

1 Introduction

A strong association has been established between alcohol consumption and carcinogenesis in multiple cancers, including head and neck squamous carcinoma, esophageal squamous carcinoma, hepatocellular carcinoma, breast cancer, and colorectal cancer (Scoccianti et al., 2015). Alcohol consumption is also suspected to associate with pancreatic and lung cancers (Boffetta and Hashibe, 2006). Previous studies have demonstrated a dose-dependent effect of alcohol consumption on survival in head, neck (Lee et al., 2019), and colorectal cancer patients (Cai et al., 2014) and have also demonstrated that alcohol affected tumor development during gene-encoding enzyme alterations for alcohol metabolism, folate metabolism, DNA repair, and oxidative stress (Boffetta and Hashibe, 2006; Seitz and Stickel, 2007). Relationships between alcohol and cancers deserved extensive investigations.

DNA methylation alterations are one of the most frequently identified events in various malignant tumors (Kodach et al., 2010; Lee et al., 2013; Dong et al., 2019) and play important roles in cancer initialization, progression, and recurrence (Ehrlich, 2006; Pellacani et al., 2014; Tahara and Arisawa, 2015; Chen et al., 2016; Das et al., 2019). A group of enzymes known as DNA methyltransferases (DNMTs) are the key regulators that catalyze the methylation process. DNMT1 functions as the maintenance DNMT completing the methylation of the partially methylated DNA, while DNMT3A and DNMT3B catalyze *de novo* methylation. The abnormal expression of DNMTs can silence numerous tumor suppressor genes and affect important biological functions in multiple cancers (Sharma et al., 2010; Fattahi et al., 2018; Zhang et al., 2020). Accumulating evidence has shown that alcohol consumption may cause epigenetic changes, particularly abnormal DNA methylation, which could be important contributory factors to alcohol-induced carcinogenesis (Schernhammer et al., 2010; Boycott et al., 2022; Zhao et al., 2022; Zhou et al., 2022). Alcohol could regulate the expression of DNMTs in mouse models (Mukhopadhyay et al., 2013; Miozzo et al., 2018) and affect DNA methylation-associated pathways *via* reprogramming S-adenosylmethionine (SAMe) metabolism or disturbing the intake of vitamins B6 and B12 (Voigt, 2005; Varela-Rey et al., 2013; Na and Lee, 2017). According to several methylome-wide analyses, a robust association was established between the methylation of specific sites and the clinical outcomes of cancer patients with alcohol consumption (Villanueva et al., 2015; Zhou et al., 2019). Nonetheless, most studies focused on a limited number of CpG sites or genes in a single cancer type, and few analyses rely on the common patterns of alcohol-induced methylome changes. The underlying molecular mechanisms of these aberrant alterations remain elusive, and further explorations are required to uncover their possible effects on tumor biological behaviors.

In this study, we investigated whether a common methylation pattern existed in alcohol-associated neoplasms. The methylome, transcriptome, and clinical information on four alcohol-associated tumors, including esophageal carcinoma (ESCA), head and neck squamous cell carcinoma (HNSCC), liver hepatocellular carcinoma (LIHC), and pancreatic adenocarcinoma (PAAD), was downloaded from The Cancer Genome Atlas (TCGA) portal. The differentially methylated probes (DMPs) were identified between tumors developed due to alcohol consumption and normal tissues. A total of 172 hypermethylated and 21 hypomethylated pan-cancer DMPs (PDMPs) were extracted. The PDMP-annotated genes were investigated and mostly enriched in transcriptional misregulation of cancer pathways. A CpG island *chr19:58220189–58220517* was significantly hypermethylated in all four cancer types and strongly correlated with the downregulation of transcription regulator zinc finger protein 154 (ZNF154), which could serve as a potential epigenetic therapeutic biomarker. Transcription factor (TF) motifs enriched in PDMP regions were clustered into five groups and were significantly associated with multiple tumor hallmarks, on which a TF network was constructed. Furthermore, our studies also identified 11 hypermethylated PDMPs that may be used to predict the overall survival of alcohol-associated cancer patients.

2 Methods

2.1 Data acquisition

The clinical and transcriptome data on ESCA, HNSCC, LIHC, and PAAD were obtained from TCGA data portal (https://portal.gdc.cancer.gov/). The methylome data were downloaded from the UCSC Xena browser (https://xenabrowser.net/datapages/), which was based on the Illumina HumanMethylation450 BeadChip. GSE123781 is a dataset of oral squamous cell carcinoma patients under alcohol consumption documented in clinical information (Nemeth et al., 2019). The methylation data and clinical data were downloaded from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/).

2.2 Data preprocessing

We extracted 129 ESCA, 352 HNSCC, 118 LIHC, and 102 PAAD patients with a clear alcohol consumption history. After samples with no methylation data or gene expression data were excluded, 102 ESCA, 328 HNSCC, 117 LIHC, and 101 PAAD patients were finally applied for further investigation.

Several criteria were used to filter out inappropriate DNA methylation data in this study, including 1) probes having a 'Not Applicable' (NA) beta value in more than 50% samples; 2) probes containing no greater than three beads in a minimum of 5% samples; 3) probes with a detection p-value ≥ 0.01; 4) multiple-location-hitting probes; 5) probes that are non-CpG; 6) probes associated with single nucleotide polymorphism (SNP) sites; and 7) sex chromosome-specific probes. The k-nearest neighbors (KNNs) method was then applied to estimate the missing value, followed by type II probe

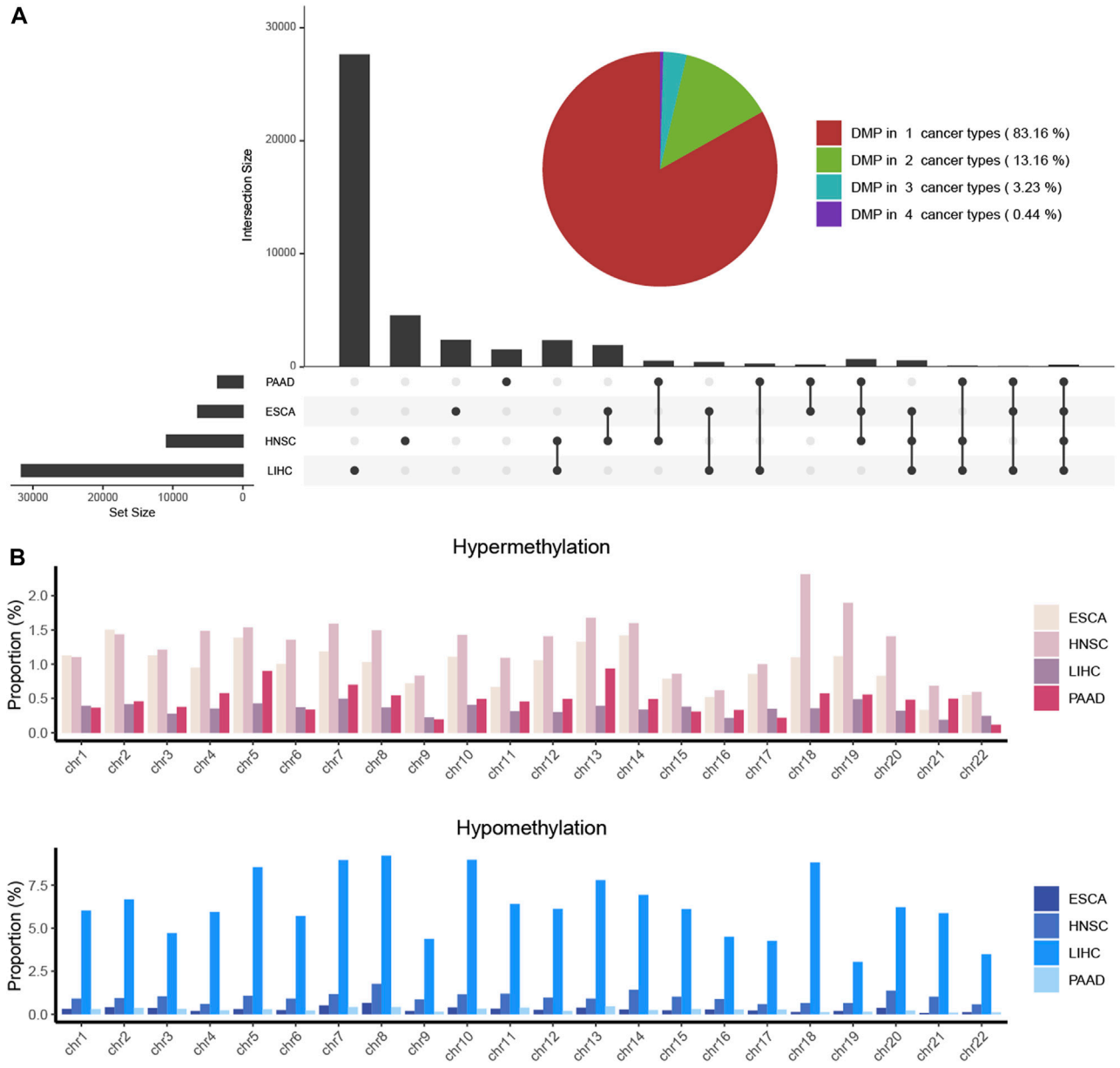


FIGURE 1
DMPs were investigated in four alcohol-associated cancers. **(A)** Combined plot showed the number of DMPs in each cancer type and that of multiple cancer types. The number of the same DMPs between different cancer types was calculated and shown as the intersection size. **(B)**. Proportion of hypermethylated and hypomethylated DMPs located in chromosomes of each cancer type. The proportion was determined as the ratio of the number of identified DMPs to the total number of CpG probes on each chromosome.

normalization using the BMIQ method (Fortin et al., 2017). Using the *ChAMP* package, the aforementioned steps were performed to preprocess raw methylation data (Morris et al., 2014).

2.3 DMP and PDMP definitions

The Limma method in the ChAMP package was applied to identify DMPs in each cancer. The threshold of a significant probe is the fold change (FC) of a beta value > 1.2 and p-value < 0.05. A DMP with a higher methylation level (beta value) in the tumor sample was defined as hypermethylated DMP, while a DMP with a lower methylation level was defined as hypomethylated DMP. A probe is defined as PDMP if it can be identified as DMP in all four cancer types and has the same methylation status (hypermethylated or hypomethylated).

2.4 Regional methylation status assessment

Using probe beta values for one cancer type in a specific region, we investigated the methylation status of various chromosomes and functional CpG regions. The formula is as follows:

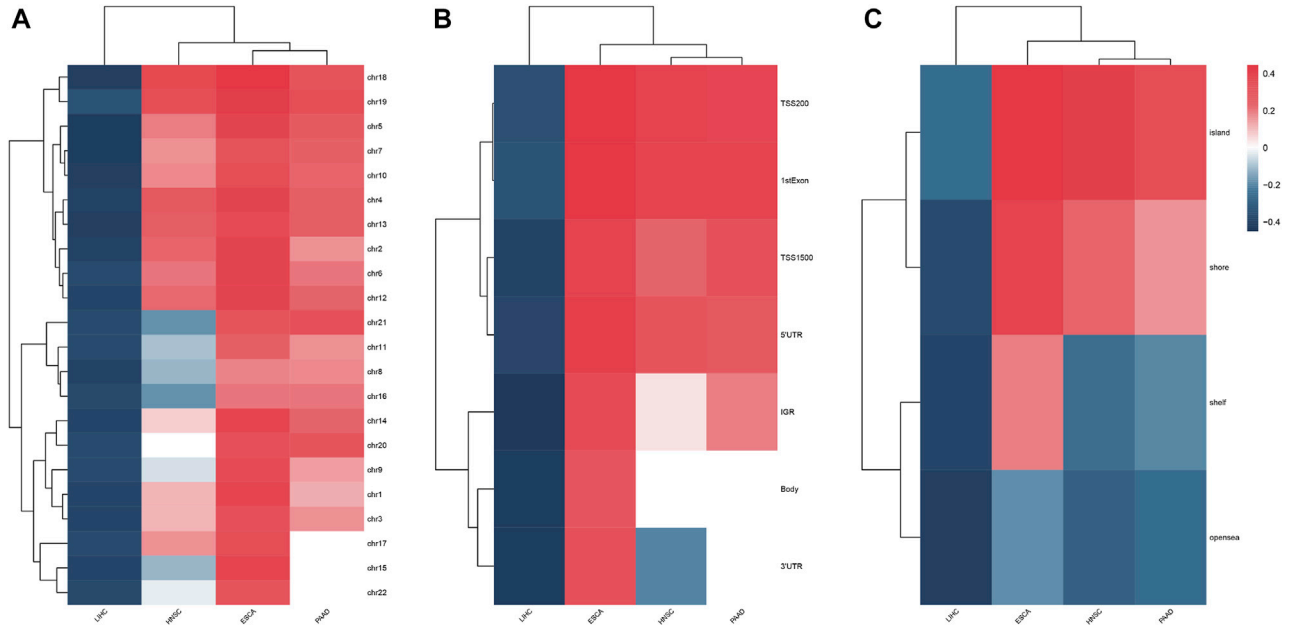


FIGURE 2
Heatmap showed the methylation status of chromosomes **(A)**, feature regions **(B)**, and CpG regions **(C)** in four alcohol-associated cancers.

$$MS = median\left(\sum M_{i,T}^{p,r}\right) - median\left(\sum M_{j,N}^{p,r}\right).$$

For a given cancer type, the median beta value M of probe p in region r of all tumor samples T or normal samples N was calculated. A hypermethylated region was defined as having a p-value < 0.05 and MS > 0, while a hypomethylated region was defined as having a p-value < 0.05 and MS < 0.

2.5 Correlation between PDMPs and gene expression

We calculated the Pearson coefficient correlations (PCCs) between PDMPs and their annotated genes. A significant correlation was defined as having |PCC| > 0.2 and p-value < 0.05. Because the positions of CpG probes in different regions may have a different effect on genes (Yang et al., 2015), probe regions were divided into two groups: a) the promotor region, consisting of *TSS200*, *TSS1500*, *5′UTR*, and *1stExon*, and b) the body region, consisting of *Body*, *3′UTR*, and *IGR*. Upstream and downstream genes that may potentially be regulated by enhancers located on specific probes were investigated using the ELMER package (Silva et al., 2019). A significant correlation was defined as having |PCC| > 0.1 and p-value < 0.05.

2.6 Motif enrichment analysis

As a measurement of TF-binding motifs enriched in PDMP regions, we generated sequences of 500 bp between 250 bp upstream and 250 bp downstream of each probe and divided them into hypermethylated and hypomethylated groups, followed by motif enrichment analysis using the AME function on the MEME website (http://meme-suite.org/tools/ame) (Buske et al., 2010; Bailey et al., 2009; Kulakovskiy et al., 2018). The TF-enriched motif was obtained with a E-value < 10^{-5}, which was subsequently clustered using STAMP (http://www.benoslab.pitt.edu/stamp/index.php) (Mahony and Benos, 2007). *clusterProfiler* was used to explore pathways enriched in the Kyoto Encyclopedia of Genes and Genomes (KEGG) (Yu et al., 2012).

2.7 Construction of the TF-regulated network

The TRRUST database (https://www.grnpedia.org/trrust/) is a practical website for TF-gene interaction prediction, collecting 8,444 TF-target regulatory relationships derived from 11,237 experiment-based articles (Han et al., 2018), which was used to screen and filter high-reliability TF-gene pairs. Subsequently, 40 TFs and their target genes were selected to construct a TF-regulated network in each cancer based on the adjacent matrix created by the weighted method (Langfelder and Horvath, 2008). Each cancer regulatory network's top 10% weighted edges, which represent TF-gene interaction intensity, were combined into one pan-cancer regulatory network (Yu, 2020). Genes were clustered based on the greedy algorithm (Girvan and Newman, 2002).

2.8 Survival analysis

To evaluate the prognostic ability of PDMPs for each cancer type, a univariate Cox regression model was applied. Probes with

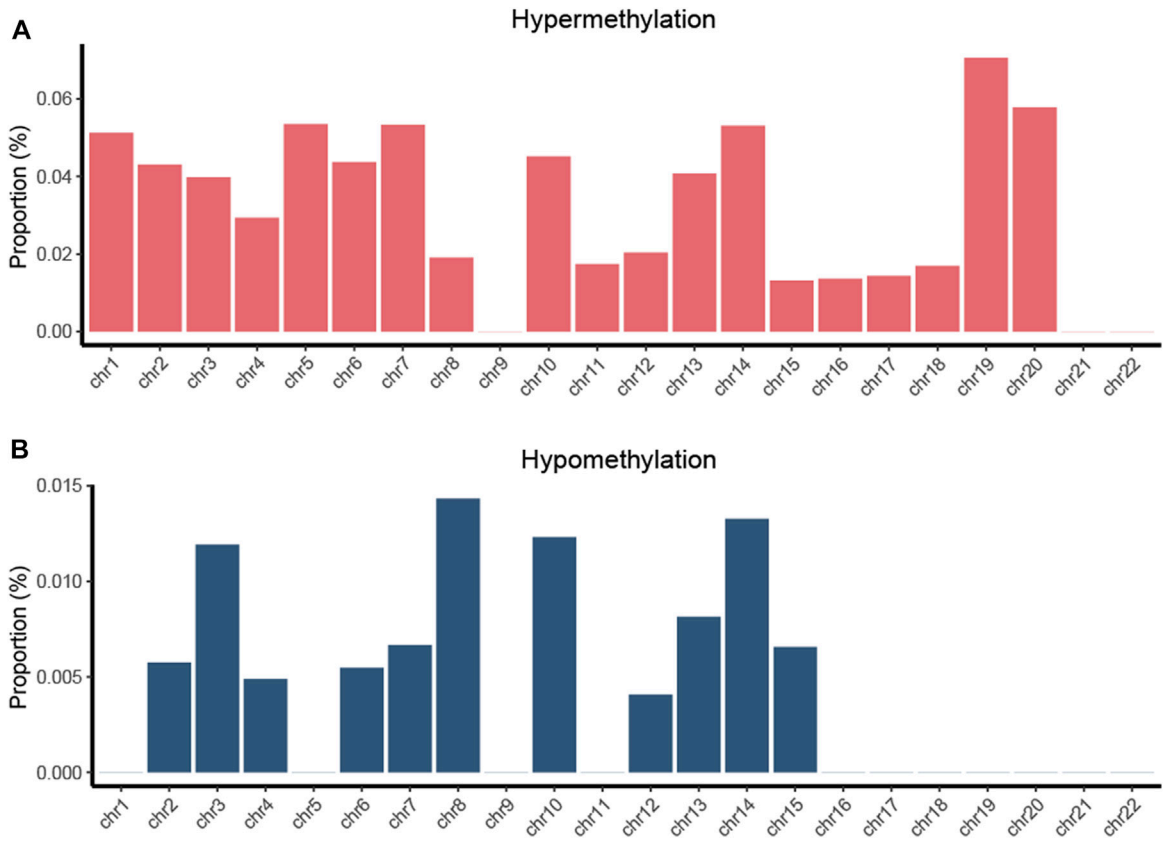


FIGURE 3
Proportion of hypermethylated **(A)** and hypomethylated **(B)** PDMPs located in chromosomes. The proportion was determined as the ratio of the number of identified PDMPs to the total number of CpG probes on each chromosome.

p-value < 0.1 in at least three cancer types were extracted and considered prognosis-related PDMPs, followed by multivariate Cox analysis. For each cancer, we evaluated the methylation prognosis index (MPI) of each sample and divided them into high- and low-risk groups by the following formula:

$$MPI_i = \sum_{p}^{n} \beta_{p,i} I_p.$$

For a given sample *i*, the beta value of prognosis-related PDMP *p* to the index coefficient *p* was summed. To evaluate the survival difference between two groups of patients, Kaplan–Meier survival analysis was performed. AUC was calculated using the *timeROC* package of R software to validate the predictive ability of the MPI model (Blanche et al., 2013).

2.9 Statistical analysis

The statistical significance of continuous variables was estimated using the Wilcoxon rank-sum test. Univariate and multivariate Cox regression analyses were performed to construct the MPI model using the *survival* package (Holleczek and Brenner, 2013). All the aforementioned statistics were analyzed by R software (version 4.0.2).

3 Results

3.1 Characteristics of alcohol-associated DNA methylation patterns

The clinical information on each alcohol-associated aforementioned cancer was downloaded from TCGA database. ESCA, HNSCC, LIHC, and PAAD were included in further studies, and other alcohol-related cancer types, such as lung and colon cancer, were excluded due to the absence of a documented history of alcohol consumption. We identified a set of DMPs in four cancers, with 6,436 in ESCA, 10,891 in HNSCC, 31,620 in LIHC, and 3,592 in PAAD patients (Figure 1A), and the proportion of hypermethylated or hypomethylated DMPs distributed on each normal chromosome to all probes in that region was measured (Figure 1B). Given the number of DMPs in each cancer, DNA methylation alterations in pancreatic cancer might be affected slightly by alcohol consumption, whereas liver cancer was more affected. The methylation status of chromosomes in ESCA and PAAD separately was more hypermethylated or hypomethylated and that of HNSCC and LIHC was in an intermediate status (Figure 2A). Most DMPs tended to be located in non-promotor regions (Supplementary Figure S1A, S1B) and an approximately equal percentage of the CpG island and open-sea regions (Supplementary Figure S1C, S1D). Functional regions were

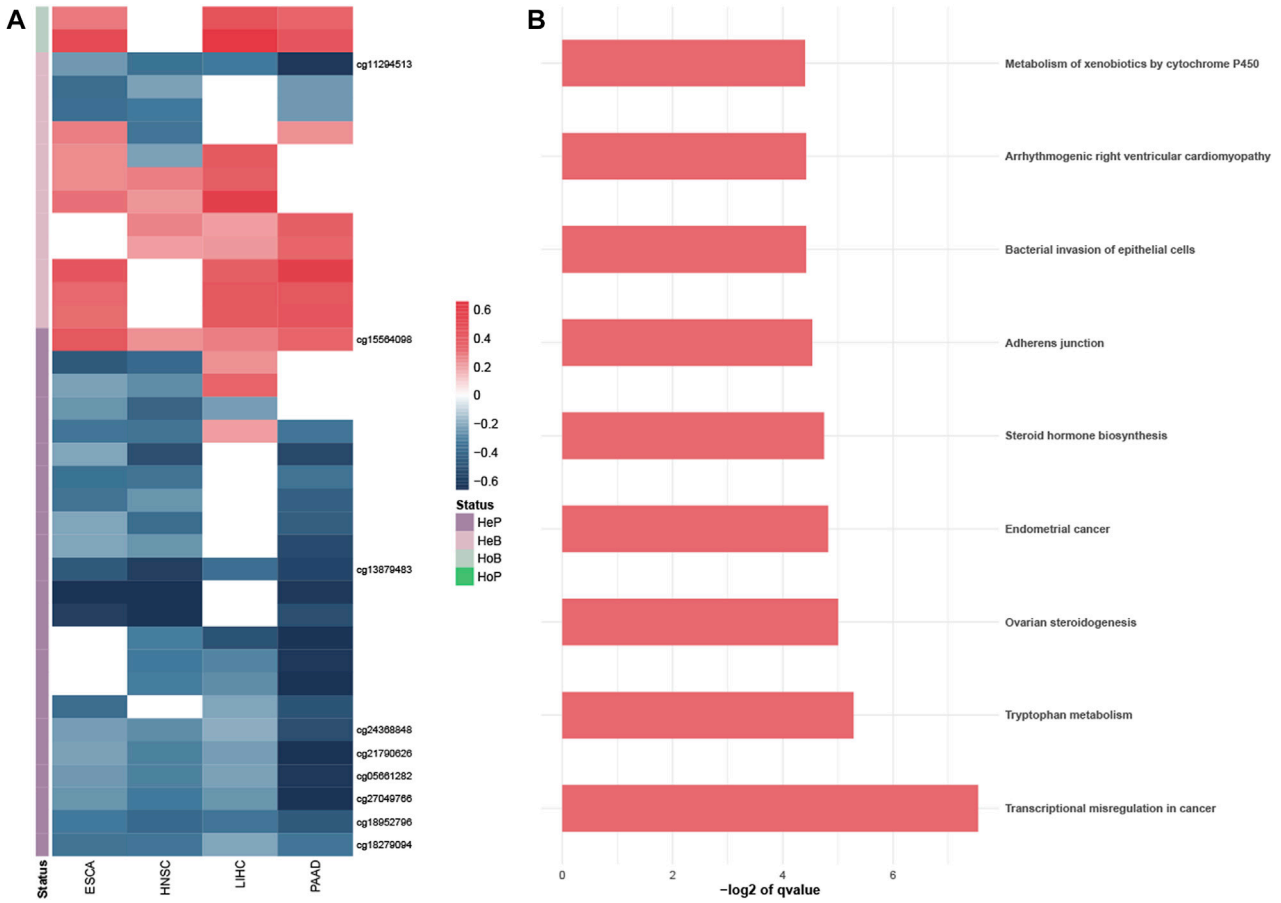


FIGURE 4
Exploration of the PDMP-annotated genes. **(A)** Heatmap revealed the Pearson coefficient correlations between PDMPs and annotated genes. PDMPs were defined into four groups: HeP, hypermethylated probe in the promotor region; HeB, hypermethylated probe in the body region; HoP, hypomethylated probe in the promotor region; HoB, hypomethylated probe in the body region. **(B)**. KEGG enrichment analysis of 33 PDMP-annotated genes.

categorized based on the methylation status into two groups, namely, the promotor (*TSS200*, *TSS1500*, *5′UTR*, and *1stExon*) and body region (*Body*, *3′UTR*, and *IGR*) or CGI (*CpG island*), and non-CGI region (*CpG shore*, *shelf*, and *opensea*) (Figures 2B,C), each of which was annotated with biological functions.

We computed DMPs with the same methylation status in all four types of cancers and obtained 172 hypermethylated and 21 hypomethylated PDMPs (Supplementary Figure S2A, S2B). The highest proportion of hypermethylated PDMPs was found on chromosome 19, while the highest proportion of hypomethylated PDMPs was found on chromosome 8. (Figure 3). Interestingly, there was no PDMP located in chromosomes 9, 21, and 22. One of the possible explanations could be that DMPs in these chromosomes exhibit biological behaviors that are specific to cancer. Compared with distribution of the functional region of DMPs, PDMPs located in promotor or body regions showed an almost equal percentage distribution (46.7% and 53.3%) (Supplementary Figure S3A, S3B). In addition, 72% of PDMPs were located in the CGI region (Supplementary Figure S3C, S3D). We also validated the PDMPs in GSE123781, and 182 of 193 probes were found significantly differentially methylated (Supplementary Table S1). Taken together, alcohol-associated aberrant pan-cancer methylation changes were more likely to exert biological effects in promotor and CGI regions.

Aging was suggested to be associated with the increase of methylation at a global level (Maegawa et al., 2017). To identify whether aging contributed to the alcohol-associated methylation alterations in our study, we investigated the DMPs between the young group (<65 years old) and the old group (≥65 years old). We found that there was only a total of 26 CpG probes with an adjusted p-value < 0.05, with 21 in ESCA, 1 in HNSCC, 4 in LIHC, and none in PAAD patients (Supplementary Table S2). Interestingly, none of these probes reached the DMP-defined threshold of fold change > 1.2. These results suggest that aging has a very slight impact on our findings.

3.2 PDMPs regulated the expression of annotated genes

To better understand the potential biological effect of these PDMPs, PCCs were used to evaluate the correlation between

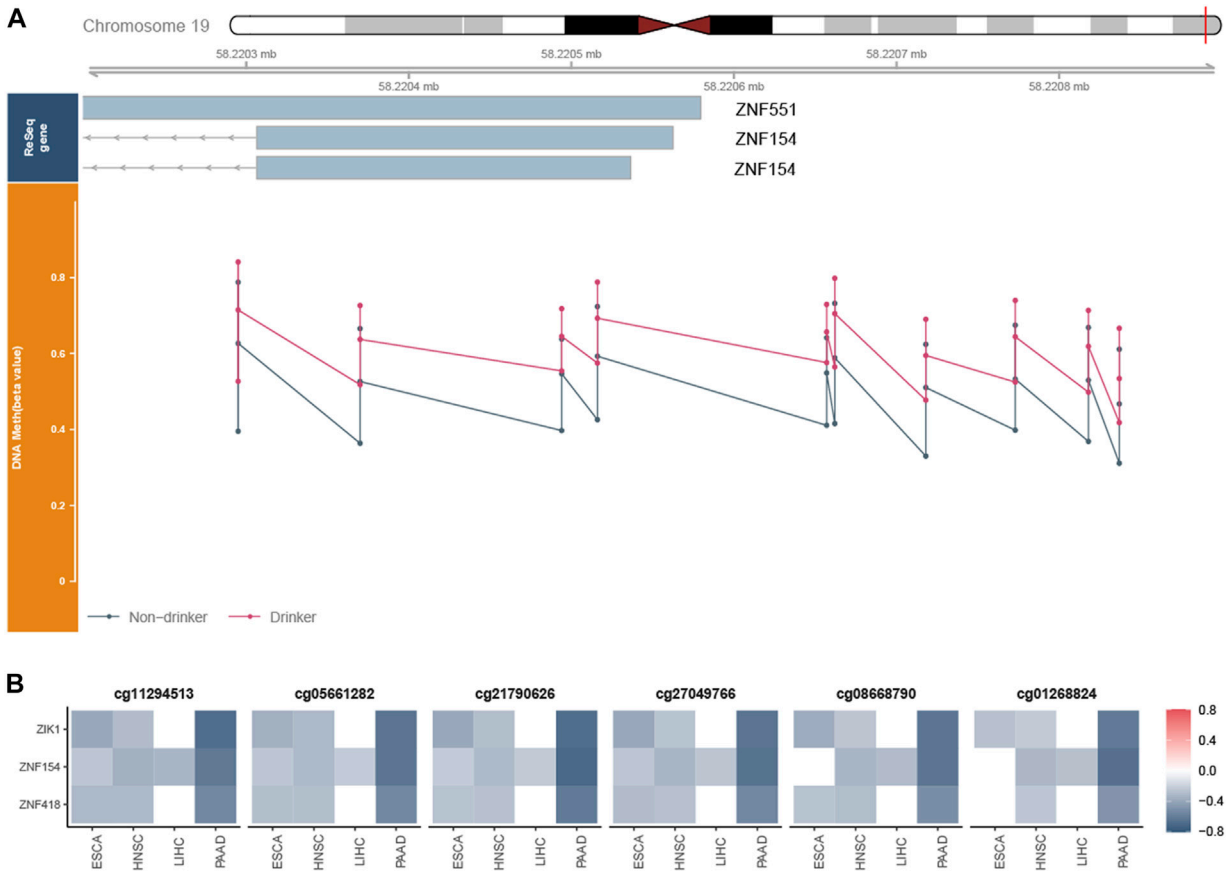


FIGURE 5
Exploration of the CpG island *chr19:58220189–58220517*. **(A)** Whole view of CpG probes in the CpG island *chr19:58220189–58220517* demonstrated its hypermethylation status. The bottom dot plots from left to right represent the beta values of cg1129451, cg0566128, cg21790626, cg27049766, cg03234186, cg08668790, cg12506930, cg26465391, cg01268824, and cg27324426, respectively. *cg03142586* was excluded from further analysis since its beta value was NA in most samples. **(B)** Heatmap revealed the correlation between the beta value of six PDMPs and the expression of *ZNF154*, *ZIK1*, and *ZNF418* in four cancer types.

PDMPs and their annotated downstream genes. Since probes with different relative locations may have a different regulatory effect on expression, 193 PDMPs were defined into four groups: HeP = hypermethylated probe in the promotor region, HeB = hypermethylated probe in the body region, HoP = hypomethylated probe in the promotor region, and HoB = hypomethylated probe in the body region. Thirty-seven PDMP–gene pairs were identified in at least three cancer types, and nine PDMP–gene pairs were found in all cancers (Figure 4A; Supplementary Table S3). These 37 PDMP-annotated genes were enriched in several cancer hallmarks, among which *transcriptional misregulation in cancer* (q-value = $5.35e^{-3}$) stood out as the most significant one (Figure 4B). There was a surprising outcome that four of nine aforementioned PDMPs (*cg01268824*, *cg11294513*, *cg21790626*, and *cg27049766*) belonged to the same CpG island *chr19:58220189-58220517* and correlated with the same gene *ZNF154*. The role of this CpG island required further exploration.

3.3 Hypermethylated CpG island *chr19: 58220189-58220517* silenced downstream gene expression

We checked all probes in *chr19:58220189–58220517* and found a total set of 11 probes, among which six probes were PDMPs (*cg11294513*, *cg05661282*, *cg21790626*, *cg27049766*, *cg08668790*, and *cg01268824*). *cg03142586* was excluded from further analysis since its beta value was NA in most samples. Probes in this region were remarkably hypermethylated (Figure 5A; Supplementary Figure S4), which were validated in GSE123781 (Supplementary Figure S5). It has been shown that exposure to alcohol increases methylation levels at certain loci in normal tissues (Lee et al., 2011). To obtain an overall view of the methylation status of the region, we investigated the beta values of samples exposed or not exposed to alcohol. Significant differences were observed between alcohol-associated tumor samples and non-alcohol-associated tumor samples (Supplementary Figure S6). A strong negative correlation was found between the beta values

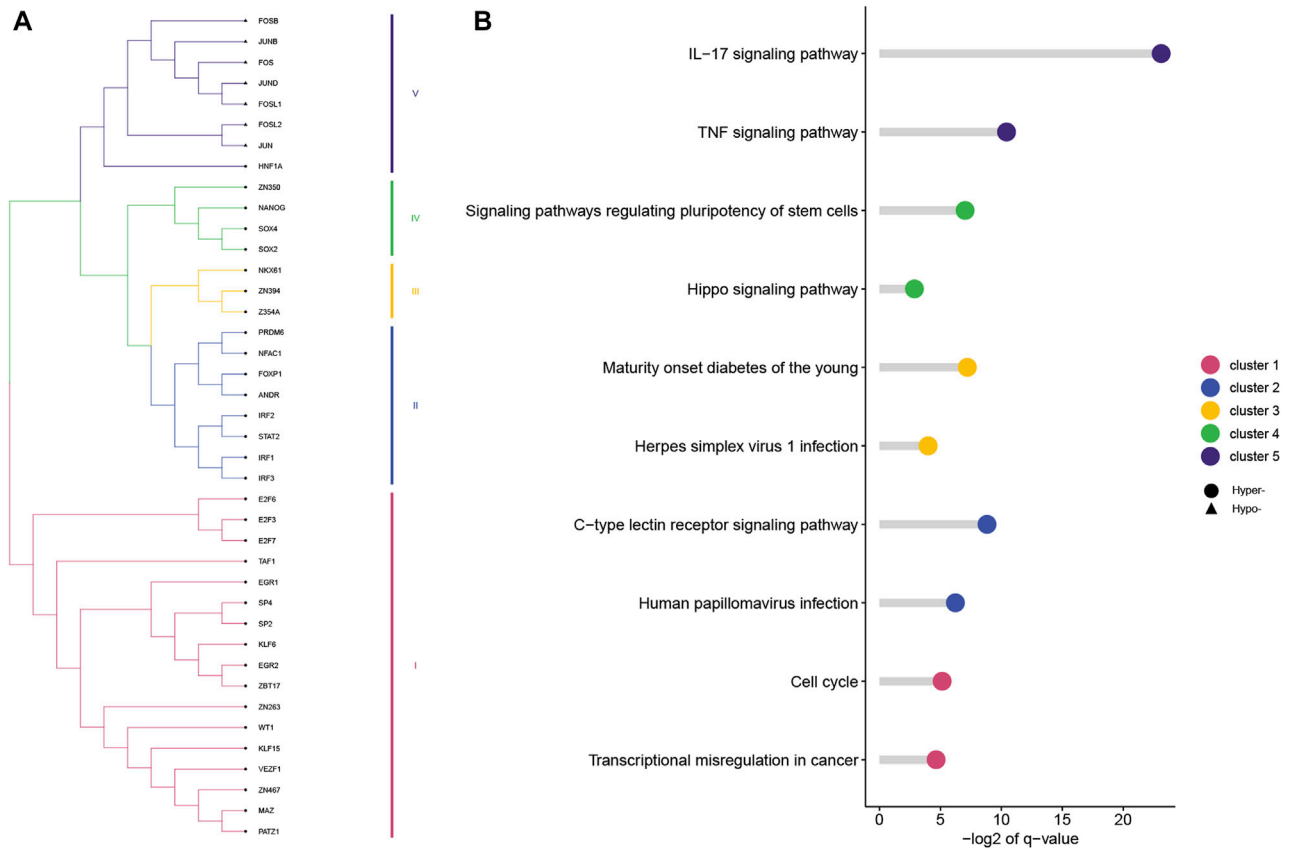


FIGURE 6
TF motifs enriched in hypermethylated and hypomethylated PDMPs. **(A)** TF motifs were divided into five clusters based on the sequence similarity. **(B)** KEGG enrichment analysis of five clusters. The top two enriched pathways of each cluster were visualized.

of these PDMPs and the expression of *ZNF154* (Figure 5B), suggesting a potential silence effect. Recently, the methylome and transcriptome-based algorithm has been developed to identify cancer-specific enhancers (Yao et al., 2015). The CpG probes located on these enhancers possess the potential ability to regulate the expression of upstream and downstream genes. We then applied this method and explored the relationship between these PDMPs and the closest 10 upstream and downstream genes (Supplementary Figure S7). Together with *ZNF154*, zinc finger protein interacting with K protein 1 (*ZIK1*, the seventh closest upstream gene) and *ZNF418* (the ninth closest downstream gene) was also found to have a negative correlation with six PDMPs (Figure 5B), suggesting the hypermethylation status of the CpG island *chr19: 58220189–58220517* could potentially mediate transcription activity of these three genes. All three genes were associated with DNA-binding activity, suggesting that the disorder of DNA recognition might be involved in alcohol-associated tumor behaviors.

3.4 Different pathways are correlated with TF motifs in PDMP regions

Based on 500-bp sequences generated from PDMPs, possible TF-binding motifs were scanned, enriched, and divided into hypermethylated and hypomethylated groups. Thirty-three and seven motifs located in hypermethylated and hypomethylated regions, respectively, were identified, followed by the grouping of these motifs into five clusters with STAMP. (Figure 6A; Supplementary Table S4). POZ/BTB and AT-hook-containing zinc finger protein 1 (*PATZ1*) showed the most significant enrichment in the hypermethylated group (*E*-value = $3.36e^{-17}$), and among hypomethylated groups, *JUN* was identified as the most significant motif (*E*-value = $6.47e^{-17}$). All seven hypomethylated motifs shared a close relatedness and were clustered into one group, suggesting a similar regulatory mechanism among them. The motif clusters exhibited distinct biological pathways, and several pathways appeared to be involved in the development and differentiation of cancers. (Figure 6B). Cluster I was significantly enriched in the *cell cycle* (*q*-value = $2.80e^{-2}$), and Cluster IV was associated with *signaling pathways regulating the pluripotency of stem cells* (*q* = $7.00e^{-3}$). Additionally, Cluster IV was associated with the *IL-17 signaling pathway* (*q*-value = $1.10e^{-7}$) and *TNF signaling pathway* (*q*-value = $7.28e^{-4}$), which is possibly involved in modulation of the immune response. According to the TRRUST database, we further explored the target genes associated with the aforementioned TF motifs and constructed a transcriptional regulatory network, which was recognized as six distinct modules (Figure 7).

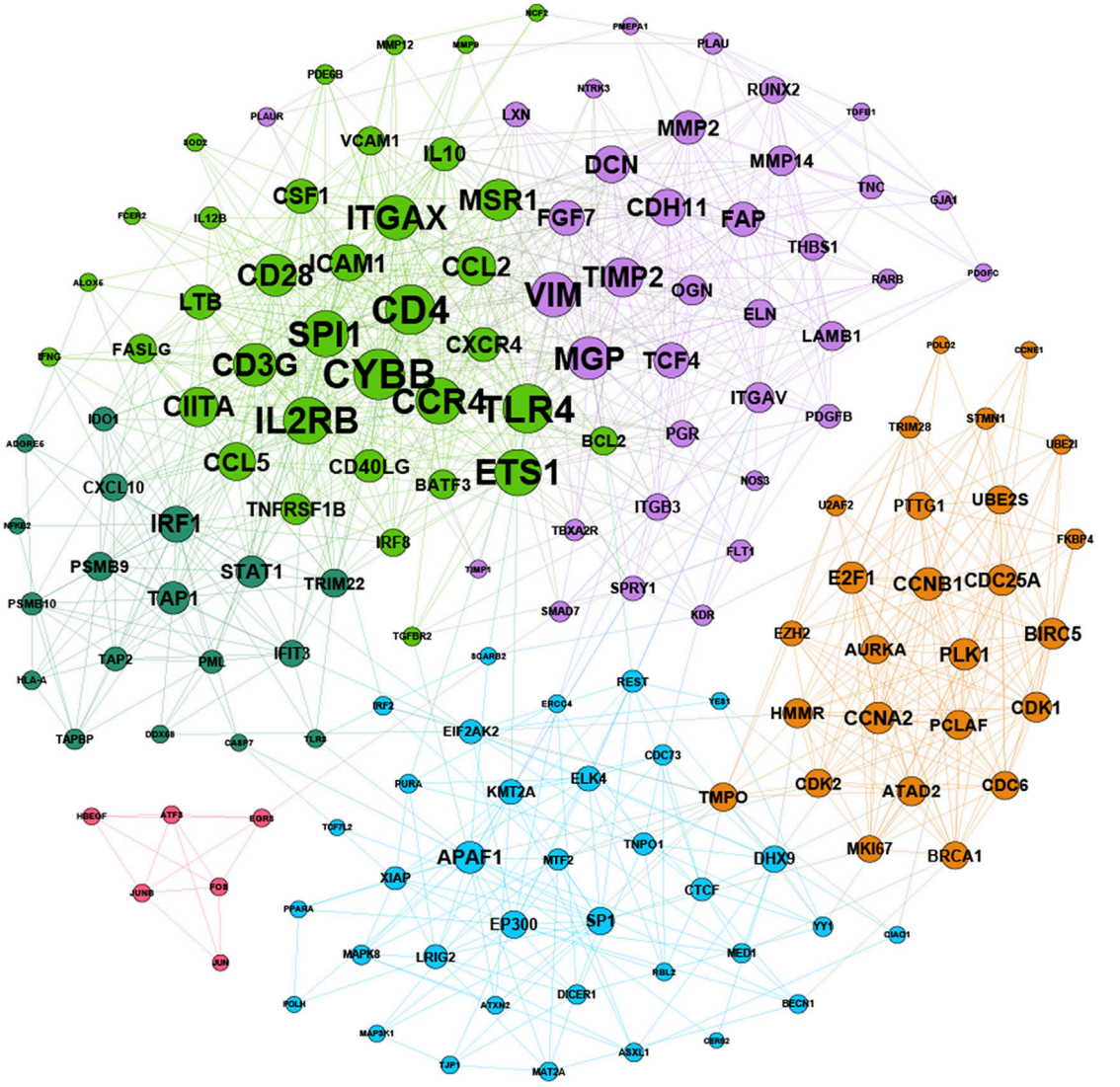


FIGURE 7
Construction of the combined TF-gene regulatory network. For each cancer, a regulatory network was constructed from the adjacent matrix. The top 10% of TF-gene interactions were considered significant and combined into one network.

3.5 The prognostic potential of PDMPs

To evaluate the prognostic value of specific PDMPs, we performed univariate Cox regression analyses in each cancer separately (Figure 8A). We screened 11 PDMPs as they were associated with prognosis in at least three types of cancer, and these PDMPs were all hypermethylated (Figure 8B; Supplementary Table S5). Seven out of 11 PDMPs (63.6%) were found in promoter regions, while 10 out of 11 (90.1%) were found in CGI regions, which confirmed the aforementioned findings that alcohol-associated PDMPs were more likely to exert biological effects on these regions. Patients were classified into low- and high-MPI groups in each cancer based on the MPIs. High-MPI groups were significantly associated with a worse outcome than low-MPI groups by Kaplan–Meier survival analysis (Figure 8C). The ROC curves were obtained, and the AUC of 3-year OS (range from 0.705 to 0.859) and 5-year OS (range from 0.621 to 0.881) was calculated separately, revealing the reliability of MPIs (Figure 8D).

4 Discussion

The purpose of this study was to investigate PDMPs in four cancers that were associated with alcohol consumption and the distribution of PDMPs within these cancers. Promotor and CGI regions were the two main regions that PDMPs were more likely to enrich. Thirty-seven PDMPs were found to show a close regulatory relationship with annotated genes and correlate with transcriptional misregulation. Furthermore, we observed a hypermethylated CpG island *chr19:58220189–58220517*, showing a strong correlation with transcription regulatory genes. Thirty-three hypermethylated and seven hypomethylated TF motifs were clustered into five groups and

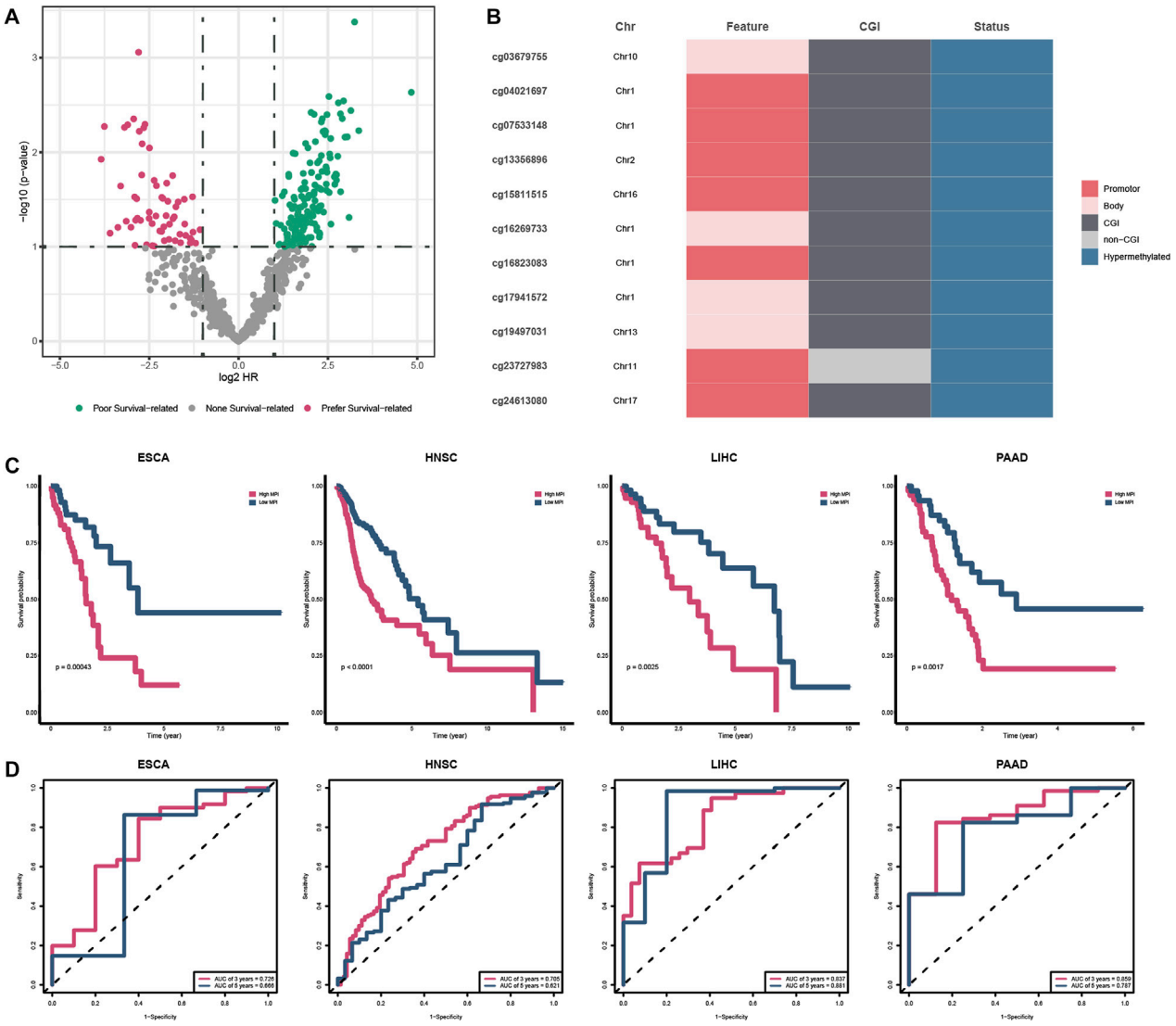


FIGURE 8
Survival analysis of PDMPs. **(A)** Volcano plot showing the survival-related PDMPs in each cancer. **(B)**. Characteristics of 11 survival-related PDMPs. **(C)** Kaplan–Meier analysis of high- and low-MPI group patients in each cancer type. **(D)**. ROC curve validating the predictive ability of MPIs.

exerted various biological effects. Finally, 11 PDMPs were reported to be associated with overall survival, providing valuable prediction of clinical outcomes.

Many DMPs were found in each of the four alcohol-associated cancers, and only a small percentage of DMPs were shared, suggesting the presence of cancer-specific methylation patterns. A study by Fan et al., for instance, revealed that preferred genomic DNA sequence patterns were observed between different DNMTs in hepatocellular carcinoma cells, which might promote to the formation of a hepatocellular carcinoma-specific methylation landscape (Fan et al., 2016). Cancer-specific methylation patterns have not been explored here due to the limited scope of our studies on common patterns of altered DNA methylation, which requires further study.

The genome methylation landscape is wholly poor of CpG and locally rich (in CpG islands). CGIs could be found in more than half of the genes in the vertebrate genome and commonly maintain a low methylated level (Jones, 2012). We reported a hypermethylated CpG island *chr19:58220189–58220517* in all four alcohol-associated cancers, and this finding is consistent with that of Ruike Y et al., who also found that CGI regions tend to be hypermethylated in cancers (Ruike et al., 2010). The expression of *ZNF154* was significantly downregulated by the CpG island *chr19: 58220189–58220517*, which has been partially demonstrated by previous studies. Based on a computational algorithm, we also found that the transcription activity of *ZIK1* and *ZNF418* might be regulated by hypermethylation of the CpG island *chr19: 58220189–58220517*. According to a blood-based diagnostic model, hypermethylation of the *ZNF154* CpG island was identified as a relevant biomarker for detecting circulating solid tumor DNA (Margolin et al., 2016). Epigenetic silencing of *ZNF154* was associated with multiple cancers and could serve as a biomarker

predicting the recurrence of alcohol-associated pancreatic cancer (Yamada et al., 2016; Mishra et al., 2019; Wiesmueller et al., 2019), as well as non-alcohol-associated prostate cancer (Nahar et al., 2018) and bladder cancer (Reinert et al., 2012). Hu et al. showed that *ZNF154* might serve as a tumor suppressor, and overexpression of *ZNF154* significantly inhibited cell migration and invasion *in vivo,* as well as lung nodule formation *in vitro, via* suppressing Wnt/β-catenin signaling pathway activation (Hu et al., 2017). *ZIK1* was identified as a transcriptional repressor blocking a gene promotor bearing Gal4-binding elements *via* interacting with heterogeneous nuclear ribonucleoprotein particle K protein (Denisenko et al., 1996). Silenced *ZIK1* was observed in noncancerous esophageal mucosae (Oka et al., 2009), suggesting low expression of *ZIK1* occurred at the early stages of cancer. Only a few studies documented that *ZNF418* negatively regulated transcription and the MAPK signaling pathway (Li et al., 2008) and that lower expression of *ZNF418* was associated with poorer prognosis in gastric cancer (Hui et al., 2018). As our scope was limited to methylation alterations, the role of associated genes was not identified, which required further investigation.

Thirty-three and seven TF motifs located in hypermethylated and hypomethylated PDMP regions, respectively, were identified as transcription regulators in alcohol-associated cancers, of which *PATZ1* and *JUN* were most notable. *PATZ1*, containing an A–T hook DNA-binding motif, binds to other DNA-binding structures to participate in chromatin modeling and transcription regulation. *PATZ1* is involved in the inhibition of the mesenchymal-to-epithelial transition *via* disturbing the combination between p53 gene and its response elements (Chiappetta et al., 2015; Keskin et al., 2015) and could be an independent prognostic factor in multiple cancers (Guadagno et al., 2017; Zhao et al., 2018; Passariello et al., 2019). *Jun* was involved in a number of biological processes that were triggered by oxidants and toxic stimuli when combined with other members of the Fos family (Angel and Karin, 1991; Shaulian and Karin, 2002). An *in vitro* experiment showed that alcohol promotes the direct recruitment of *c-Jun* to TATA-binding protein, *Brf1*, and tRNA gene promotors, inducing RNA polymerase III-dependent transcription, thus contributing to liver tumor development (Zhong et al., 2011). We established a TF-gene regulatory network in our article that might assist in understanding the regulatory functions and mechanisms of alcohol and carcinogenesis.

Aging may also contribute to methylation changes. Several studies indicated that aging was closely linked to global hypermethylation. In normal breast tissues, age acceleration was observed (Hofstatter et al., 2018).The methylation level of highly variable sites increased from a mean of 3% in the newborn to 20% in the old, and these age-related hypermethylation cases were found to be enriched in the CGI region (Maegawa et al., 2017). A large-scale methylation study revealed that cancer risks and mortality increased with aging epigenetic changes, ranging from 4% to 9% and 2% to 6%, respectively, with 5 years of age acceleration (Dugue et al., 2018). According to Zheng et al., most epigenetic drifts were nonfunctional, while some might randomly affect TF expression or the binding affinity, resulting in an abnormality in tissue homeostasis (Zheng et al., 2016). In light of the complex function and mechanism underpinning aging and tumor epigenetics, the effect of aging on PDMPs identified in this article has been difficult to estimate and remove.

This study only examined a small part of methylation sites. Although the Illumina HumanMethylation450 array provided coverage of 98.9% UCSC RefGenes, the coverage rate of total genome CpG sites was only 2% (Bibikova et al., 2011; Plongthongkum et al., 2014), leading to incomplete exploration of alcohol-associated genome-aberrant CpG patterns. Non-coding RNAs also play significant roles in cancer initiation, progression, and metastasis (Gupta et al., 2010; Li et al., 2015; Ferreira and Esteller, 2018). The function and correlation of CpG sites annotated in these regions were not explored. Due to a lack of alcohol consumption documents in TCGA database, other alcohol-associated cancers such as colorectal and lung cancers were not included in this study.

5 Conclusion

Based on multi-omics data on pan-cancer, this study explored the global DNA methylation alterations of ESCA, HNSCC, LIHC, and PAAD patients with alcohol consumption documents from TCGA database. A total of 193 PDMPs were identified, and the preference patterns of alcohol-associated DNA methylation changes were located in promotor and CGI regions. PDMP-annotated genes were enriched in multiple pathways, especially transcriptional misregulation in cancer, demonstrating that alcohol might contribute to transcriptional disorder by inducing the methylation status of transcription regulators, thus leading to tumor development. A hypermethylated CpG island *chr19:58220189–58220517* was identified, and it regulated the transcription activity of downstream genes, serving as a potential therapeutic biomarker. Five sets of enriched TF motifs were involved in numerous cancer hallmarks including tumorigenesis and immunoregulation, and a TF-gene regulatory network was constructed for a better understanding of potential regulation mechanisms. Additionally, 11 PDMPs were reported to be associated with the overall survival of patients. Estimated MPIs are reliable and provide a potential point of view for clinical outcome prediction.

Data availability statement

Publicly available datasets were analyzed in this study. These data can be found at: TCGA (https://cancergenome.nih.gov) and GEO (https://www.ncbi.nlm.nih.gov/geo/) repositories.

Author contributions

XL and JC carried out most of the analysis and wrote the main manuscript. XL, YGO, and CX designed the study and revised the manuscript. JL and ZZ participated in data acquisition and preprocessing. JC, XJ, and YGA participated in identification of methylation alterations and their influence. XL and ZZ carried out the visualization of figures. ZH and QW helped in drafting the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (81972852 and 81803061), the Key

Research and Development Project of Hubei Province (2020BCA069), the Health Commission of Hubei Province Medical Leading Talent Project, the Young and Middle-Aged Medical Backbone Talents of Wuhan (WHQG201902), the Application Foundation Frontier Project of Wuhan (2020020601012221), and the Translational Medicine and Interdisciplinary Research Joint Fund of Zhongnan Hospital of Wuhan University (ZNJC201922 and ZNJC 202007).

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1032683/full#supplementary-material

References

Angel, P., and Karin, M. (1991). The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. *Biochim. Biophys. Acta* 1072 (2-3), 129–157. doi:10.1016/0304-419x(91)90011-9

Bailey, T. L., Boden, M., Buske, F. A., Frith, M., Grant, C. E., Clementi, L., et al. (2009). Meme suite: Tools for motif discovery and searching. *Nucleic Acids Res.* 37, W202–W208. Web Server issue. doi:10.1093/nar/gkp335

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., et al. (2011). High density DNA methylation array with single CpG site resolution. *Genomics* 98 (4), 288–295. doi:10.1016/j.ygeno.2011.07.007

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

Boffetta, P., and Hashibe, M. (2006). Alcohol and cancer. *Lancet Oncol.* 7 (2), 149–156. doi:10.1016/S1470-2045(06)70577-0

Boycott, C., Beetch, M., Yang, T., Lubecka, K., Ma, Y., Zhang, J., et al. (2022). Epigenetic aberrations of gene expression in a rat model of hepatocellular carcinoma. *Epigenetics* 17 (11), 1513–1534. doi:10.1080/15592294.2022.2069386

Buske, F. A., Boden, M., Bauer, D. C., and Bailey, T. L. (2010). Assigning roles to DNA regulatory motifs using comparative genomics. *Bioinformatics* 26 (7), 860–866. doi:10.1093/bioinformatics/btq049

Cai, S., Li, Y., Ding, Y., Chen, K., and Jin, M. (2014). Alcohol drinking and the risk of colorectal cancer death: A meta-analysis. *Eur. J. Cancer Prev.* 23 (6), 532–539. doi:10.1097/CEJ.0000000000000076

Chen, J., Wang, F., Xu, J., He, Z., Lu, Y., and Wang, Z. (2016). The role of PAQR3 gene promoter hypermethylation in breast cancer and prognosis. *Oncol. Rep.* 36 (3), 1612–1618. doi:10.3892/or.2016.4951

Chiappetta, G., Valentino, T., Vitiello, M., Pasquinelli, R., Monaco, M., Palma, G., et al. (2015). PATZ1 acts as a tumor suppressor in thyroid cancer via targeting p53-dependent genes involved in EMT and cell migration. *Oncotarget* 6 (7), 5310–5323. doi:10.18632/oncotarget.2776

Das, D., Ghosh, S., Maitra, A., Biswas, N. K., Panda, C. K., Roy, B., et al. (2019). Epigenomic dysregulation-mediated alterations of key biological pathways and tumor immune evasion are hallmarks of gingivo-buccal oral cancer. *Clin. Epigenetics* 11 (1), 178. doi:10.1186/s13148-019-0782-2

Denisenko, O. N., O'Neill, B., Ostrowski, J., Van Seuningen, I., and Bomsztyk, K. (1996). Zik1, a transcriptional repressor that interacts with the heterogeneous nuclear ribonucleoprotein particle K protein. *J. Biol. Chem.* 271 (44), 27701–27706. doi:10.1074/jbc.271.44.27701

Dong, S., Li, W., Wang, L., Hu, J., Song, Y., Zhang, B., et al. (2019). Histone-related genes are hypermethylated in lung cancer and hypermethylated HIST1H4F could Serve as a pan-cancer biomarker. *Cancer Res.* 79 (24), 6101–6112. doi:10.1158/0008-5472.CAN-19-1019

Dugue, P. A., Bassett, J. K., Joo, J. E., Jung, C. H., Ming Wong, E., Moreno-Betancur, M., et al. (2018). DNA methylation-based biological aging and cancer risk and survival: Pooled analysis of seven prospective studies. *Int. J. Cancer* 142 (8), 1611–1619. doi:10.1002/ijc.31189

Ehrlich, M. (2006). Cancer-linked DNA hypomethylation and its relationship to hypermethylation. *Curr. Top. Microbiol. Immunol.* 310, 251–274. doi:10.1007/3-540-31181-5_12

Fan, H., Zhao, Z., Cheng, Y., Cui, H., Qiao, F., Wang, L., et al. (2016). Genome-wide profiling of DNA methylation reveals preferred sequences of DNMTs in hepatocellular carcinoma cells. *Tumour Biol.* 37 (1), 877–885. doi:10.1007/s13277-015-3202-z

Fattahi, S., Golpour, M., Amjadi-Moheb, F., Sharifi-Pasandi, M., Khodadadi, P., Pilehchian-Langroudi, M., et al. (2018). DNA methyltransferases and gastric cancer: Insight into targeted therapy. *Epigenomics* 10 (11), 1477–1497. doi:10.2217/epi-2018-0096

Ferreira, H. J., and Esteller, M. (2018). Non-coding RNAs, epigenetics, and cancer: Tying it all together. *Cancer Metastasis Rev.* 37 (1), 55–73. doi:10.1007/s10555-017-9715-8

Fortin, J. P., Triche, T. J., Jr., and Hansen, K. D. (2017). Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. *Bioinformatics* 33 (4), 558–560. doi:10.1093/bioinformatics/btw691

Girvan, M., and Newman, M. E. (2002). Community structure in social and biological networks. *Proc. Natl. Acad. Sci. U. S. A.* 99 (12), 7821–7826. doi:10.1073/pnas.122653799

Guadagno, E., Vitiello, M., Francesca, P., Cali, G., Caponnetto, F., Cesselli, D., et al. (2017). PATZ1 is a new prognostic marker of glioblastoma associated with the stem-like phenotype and enriched in the proneural subtype. *Oncotarget* 8 (35), 59282–59300. doi:10.18632/oncotarget.19546

Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* 464 (7291), 1071–1076. doi:10.1038/nature08975

Han, H., Cho, J. W., Lee, S., Yun, A., Kim, H., Bae, D., et al. (2018). TRRUST v2: An expanded reference database of human and mouse transcriptional regulatory interactions. *Nucleic Acids Res.* 46 (D1), D380–D386. doi:10.1093/nar/gkx1013

Hofstatter, E. W., Horvath, S., Dalela, D., Gupta, P., Chagpar, A. B., Wali, V. B., et al. (2018). Increased epigenetic age in normal breast tissue from luminal breast cancer patients. *Clin. Epigenetics* 10 (1), 112. doi:10.1186/s13148-018-0534-8

Holleczek, B., and Brenner, H. (2013). Model based period analysis of absolute and relative survival with R: Data preparation, model fitting and derivation of survival estimates. *Comput. Methods Programs Biomed.* 110 (2), 192–202. doi:10.1016/j.cmpb.2012.10.004

Hu, Y., Qi, M. F., Xu, Q. L., Kong, X. Y., Cai, R., Chen, Q. Q., et al. (2017). Candidate tumor suppressor ZNF154 suppresses invasion and metastasis in NPC by inhibiting the EMT via Wnt/β-catenin signalling. *Oncotarget* 8 (49), 85749–85758. doi:10.18632/oncotarget.20479

Hui, H. X., Hu, Z. W., Jiang, C., Wu, J., Gao, Y., and Wang, X. W. (2018). ZNF418 overexpression protects against gastric carcinoma and prompts a good prognosis. *Onco Targets Ther.* 11, 2763–2770. doi:10.2147/OTT.S160802

Jones, P. A. (2012). Functions of DNA methylation: Islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* 13 (7), 484–492. doi:10.1038/nrg3230

Keskin, N., Deniz, E., Eryilmaz, J., Un, M., Batur, T., Ersahin, T., et al. (2015). PATZ1 is a DNA damage-responsive transcription factor that inhibits p53 function. *Mol. Cell Biol.* 35 (10), 1741–1753. doi:10.1128/MCB.01475-14

Kodach, L. L., Jacobs, R. J., Heijmans, J., van Noesel, C. J., Langers, A. M., Verspaget, H. W., et al. (2010). The role of EZH2 and DNA methylation in the silencing of the tumour suppressor RUNX3 in colorectal cancer. *Carcinogenesis* 31 (9), 1567–1575. doi:10.1093/carcin/bgq147

Kulakovskiy, I. V., Vorontsov, I. E., Yevshin, I. S., Sharipov, R. N., Fedorova, A. D., Rumynskiy, E. I., et al. (2018). Hocomoco: Towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-seq analysis. *Nucleic Acids Res.* 46 (D1), D252–D259. doi:10.1093/nar/gkx1106

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

Lee, C. H., Wong, T. S., Chan, J. Y., Lu, S. C., Lin, P., Cheng, A. J., et al. (2013). Epigenetic regulation of the X-linked tumour suppressors BEX1 and LDOC1 in oral squamous cell carcinoma. *J. Pathol.* 230 (3), 298–309. doi:10.1002/path.4173

Lee, W. T., Hsiao, J. R., Ou, C. Y., Huang, C. C., Chang, C. C., Tsai, S. T., et al. (2019). The influence of prediagnosis alcohol consumption and the polymorphisms of ethanol-metabolizing genes on the Survival of head and neck cancer patients. *Cancer Epidemiol. Biomarkers Prev.* 28 (2), 248–257. doi:10.1158/1055-9965.EPI-18-0425

Lee, Y. C., Wang, H. P., Wang, C. P., Ko, J. Y., Lee, J. M., Chiu, H. M., et al. (2011). Revisit of field cancerization in squamous cell carcinoma of upper aerodigestive tract: Better risk assessment with epigenetic markers. *Cancer Prev. Res. (Phila)* 4 (12), 1982–1992. doi:10.1158/1940-6207.CAPR-11-0096

Li, Y., Yang, D., Bai, Y., Mo, X., Huang, W., Yuan, W., et al. (2008). ZNF418, a novel human KRAB/C2H2 zinc finger protein, suppresses MAPK signaling pathway. *Mol. Cell Biochem.* 310 (1-2), 141–151. doi:10.1007/s11010-007-9674-4

Li, Y., Zhang, Y., Li, S., Lu, J., Chen, J., Wang, Y., et al. (2015). Genome-wide DNA methylome analysis reveals epigenetically dysregulated non-coding RNAs in human breast cancer. *Sci. Rep.* 5, 8790. doi:10.1038/srep08790

Maegawa, S., Lu, Y., Tahara, T., Lee, J. T., Madzo, J., Liang, S., et al. (2017). Caloric restriction delays age-related methylation drift. *Nat. Commun.* 8 (1), 539. doi:10.1038/s41467-017-00607-3

Mahony, S., and Benos, P. V. (2007). Stamp: A web tool for exploring DNA-binding motif similarities. *Nucleic Acids Res.* 35, W253–W258. Web Server issue). doi:10.1093/nar/gkm272

Margolin, G., Petrykowska, H. M., Jameel, N., Bell, D. W., Young, A. C., and Elnitski, L. (2016). Robust detection of DNA hypermethylation of ZNF154 as a pan-cancer locus with *in silico* modeling for blood-based diagnostic development. *J. Mol. Diagn* 18 (2), 283–298. doi:10.1016/j.jmoldx.2015.11.004

Miozzo, F., Arnould, H., de Thonel, A., Schang, A. L., Saberan-Djoneidi, D., Baudry, A., et al. (2018). Alcohol exposure promotes DNA methyltransferase DNMT3A upregulation through reactive oxygen species-dependent mechanisms. *Cell Stress Chaperones* 23 (1), 115–126. doi:10.1007/s12192-017-0829-2

Mishra, N. K., Southekal, S., and Guda, C. (2019). Survival analysis of multi-omics data identifies potential prognostic markers of pancreatic ductal adenocarcinoma. *Front. Genet.* 10, 624. doi:10.3389/fgene.2019.00624

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

Mukhopadhyay, P., Rezzoug, F., Kaikaus, J., Greene, R. M., and Pisano, M. M. (2013). Alcohol modulates expression of DNA methyltranferases and methyl CpG-/CpG domain-binding proteins in murine embryonic fibroblasts. *Reprod. Toxicol.* 37, 40–48. doi:10.1016/j.reprotox.2013.01.003

Na, H. K., and Lee, J. Y. (2017). Molecular basis of alcohol-related gastric and colon cancer. *Int. J. Mol. Sci.* 18 (6), 1116. doi:10.3390/ijms18061116

Nahar, R., Zhai, W., Zhang, T., Takano, A., Khng, A. J., Lee, Y. Y., et al. (2018). Elucidating the genomic architecture of Asian EGFR-mutant lung adenocarcinoma through multi-region exome sequencing. *Nat. Commun.* 9 (1), 216. doi:10.1038/s41467-017-02584-z

Nemeth, C. G., Rocken, C., Siebert, R., Wiltfang, J., Ammerpohl, O., and Gassling, V. (2019). Recurrent chromosomal and epigenetic alterations in oral squamous cell carcinoma and its putative premalignant condition oral lichen planus. *PLoS One* 14 (4), e0215055. doi:10.1371/journal.pone.0215055

Oka, D., Yamashita, S., Tomioka, T., Nakanishi, Y., Kato, H., Kaminishi, M., et al. (2009). The presence of aberrant DNA methylation in noncancerous esophageal mucosae in association with smoking history: A target for risk diagnosis and prevention of esophageal cancers. *Cancer* 115 (15), 3412–3426. doi:10.1002/cncr.24394

Passariello, A., Errico, M. E., Donofrio, V., Maestrini, M., Zerbato, A., Cerchia, L., et al. (2019). PATZ1 is overexpressed in pediatric glial tumors and correlates with worse event-free Survival in high-grade gliomas. *Cancers (Basel)* 11 (10), 1537. doi:10.3390/cancers11101537

Pellacani, D., Kestoras, D., Droop, A. P., Frame, F. M., Berry, P. A., Lawrence, M. G., et al. (2014). DNA hypermethylation in prostate cancer is a consequence of aberrant epithelial differentiation and hyperproliferation. *Cell Death Differ.* 21 (5), 761–773. doi:10.1038/cdd.2013.202

Plongthongkum, N., Diep, D. H., and Zhang, K. (2014). Advances in the profiling of DNA modifications: Cytosine methylation and beyond. *Nat. Rev. Genet.* 15 (10), 647–661. doi:10.1038/nrg3772

Reinert, T., Borre, M., Christiansen, A., Hermann, G. G., Orntoft, T. F., and Dyrskjot, L. (2012). Diagnosis of bladder cancer recurrence based on urinary levels of EOMES, HOXA9, POU4F2, TWIST1, VIM, and ZNF154 hypermethylation. *PLoS One* 7 (10), e46297. doi:10.1371/journal.pone.0046297

Ruike, Y., Imanaka, Y., Sato, F., Shimizu, K., and Tsujimoto, G. (2010). Genome-wide analysis of aberrant methylation in human breast cancer cells using methyl-DNA immunoprecipitation combined with high-throughput sequencing. *BMC Genomics* 11, 137. doi:10.1186/1471-2164-11-137

Schernhammer, E. S., Giovannucci, E., Kawasaki, T., Rosner, B., Fuchs, C. S., and Ogino, S. (2010). Dietary folate, alcohol and B vitamins in relation to LINE-1 hypomethylation in colon cancer. *Gut* 59 (6), 794–799. doi:10.1136/gut.2009.183707

Scoccianti, C., Cecchini, M., Anderson, A. S., Berrino, F., Boutron-Ruault, M. C., Espina, C., et al. (2015). European code against cancer 4th edition: Alcohol drinking and cancer. *Cancer Epidemiol.* 39 (1), S67–S74. doi:10.1016/j.canep.2015.01.007

Seitz, H. K., and Stickel, F. (2007). Molecular mechanisms of alcohol-mediated carcinogenesis. *Nat. Rev. Cancer* 7 (8), 599–612. doi:10.1038/nrc2191

Sharma, S., Kelly, T. K., and Jones, P. A. (2010). Epigenetics in cancer. *Carcinogenesis* 31 (1), 27–36. doi:10.1093/carcin/bgp220

Shaulian, E., and Karin, M. (2002). AP-1 as a regulator of cell life and death. *Nat. Cell Biol.* 4 (5), E131–E136. doi:10.1038/ncb0502-e131

Silva, T. C., Coetzee, S. G., Gull, N., Yao, L., Hazelett, D. J., Noushmehr, H., et al. (2019). ELMER v.2: An R/bioconductor package to reconstruct gene regulatory networks from DNA methylation and transcriptome profiles. *Bioinformatics* 35 (11), 1974–1977. doi:10.1093/bioinformatics/bty902

Tahara, T., and Arisawa, T. (2015). DNA methylation as a molecular biomarker in gastric cancer. *Epigenomics* 7 (3), 475–486. doi:10.2217/epi.15.4

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 (1), 25–35.

Villanueva, A., Portela, A., Sayols, S., Battiston, C., Hoshida, Y., Mendez-Gonzalez, J., et al. (2015). DNA methylation-based prognosis and epidrivers in hepatocellular carcinoma. *Hepatology* 61 (6), 1945–1956. doi:10.1002/hep.27732

Voigt, M. D. (2005). Alcohol in hepatocellular cancer. *Clin. Liver Dis.* 9 (1), 151–169. doi:10.1016/j.cld.2004.10.003

Wiesmueller, F., Kopke, J., Aust, D., Roy, J., Dahl, A., Pilarsky, C., et al. (2019). Silenced ZNF154 is associated with longer Survival in resectable pancreatic cancer. *Int. J. Mol. Sci.* 20 (21), 5437. doi:10.3390/ijms20215437

Yamada, N., Yasui, K., Dohi, O., Gen, Y., Tomie, A., Kitaichi, T., et al. (2016). Genome-wide DNA methylation analysis in hepatocellular carcinoma. *Oncol. Rep.* 35 (4), 2228–2236. doi:10.3892/or.2016.4619

Yang, X., Shao, X., Gao, L., and Zhang, S. (2015). Systematic DNA methylation analysis of multiple cell lines reveals common and specific patterns within and across tissues of origin. *Hum. Mol. Genet.* 24 (15), 4374–4384. doi:10.1093/hmg/ddv172

Yao, L., Shen, H., Laird, P. W., Farnham, P. J., and Berman, B. P. (2015). Inferring regulatory element landscapes and transcription factor networks from cancer methylomes. *Genome Biol.* 16, 105. doi:10.1186/s13059-015-0668-3

Yu, G. (2020). Using ggtree to visualize data on tree-like Structures. *Curr. Protoc. Bioinforma.* 69 (1), e96. doi:10.1002/cpbi.96

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, J., Yang, C., Wu, C., Cui, W., and Wang, L. (2020). DNA methyltransferases in cancer: Biology, paradox, aberrations, and targeted therapy. *Cancers (Basel)* 12 (8), 2123. doi:10.3390/cancers12082123

Zhao, C., Yan, M., Li, C., and Feng, Z. (2018). POZ/BTB and AT-hook-containing zinc finger protein 1 (PATZ1) Suppresses progression of ovarian cancer and Serves as an independent prognosis factor. *Med. Sci. Monit.* 24, 4262–4270. doi:10.12659/MSM.908766

Zhao, R., Shi, H., Yin, J., Sun, Z., and Xu, Y. (2022). Promoter Specific methylation of SSTR4 is associated with alcohol dependence in han Chinese males. *Front. Genet.* 13, 915513. doi:10.3389/fgene.2022.915513

Zheng, S. C., Widschwendter, M., and Teschendorff, A. E. (2016). Epigenetic drift, epigenetic clocks and cancer risk. *Epigenomics* 8 (5), 705–719. doi:10.2217/epi-2015-0017

Zhong, S., Machida, K., Tsukamoto, H., and Johnson, D. L. (2011). Alcohol induces RNA polymerase III-dependent transcription through c-Jun by co-regulating TATA-binding protein (TBP) and Brf1 expression. *J. Biol. Chem.* 286 (4), 2393–2401. doi:10.1074/jbc.M110.192955

Zhou, H., Sun, H., Liu, X., Chen, J., Zhang, L., Lin, S., et al. (2019). Combined effect between WT1 methylation and *Helicobacter pylori* infection, smoking, and alcohol consumption on the risk of gastric cancer. *Helicobacter* 24 (5), e12650. doi:10.1111/hel.12650

Zhou, X., Yu, L., Wang, L., Xiao, J., Sun, J., Zhou, Y., et al. (2022). Alcohol consumption, blood DNA methylation and breast cancer: A mendelian randomisation study. *Eur. J. Epidemiol.* 37 (7), 701–712. doi:10.1007/s10654-022-00886-1

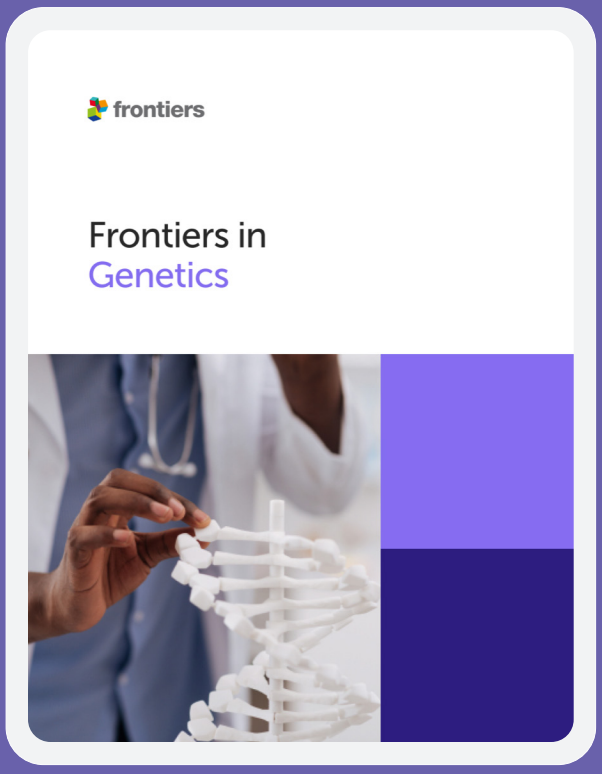
Frontiers in Genetics

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

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact