frontiers | Research Topics

Epigenetic drugs and therapeutic resistance for epithelial malignancies

Edited by
Zhiqian Zhang, Fangfang Tao and Wanjin Hong

Published in
Frontiers in Pharmacology

ISSN 1664-8714
ISBN 978-2-8325-2513-5
DOI 10.3389/978-2-8325-2513-5

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 drugs and therapeutic resistance for epithelial malignancies

Topic editors

Zhiqian Zhang — Southern University of Science and Technology, China
Fangfang Tao — Zhejiang Chinese Medical University, China
Wanjin Hong — Agency for Science, Technology and Research (A*STAR), Singapore

Citation

Zhang, Z., Tao, F., Hong, W., eds. (2023). *Epigenetic drugs and therapeutic resistance for epithelial malignancies*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-2513-5

Table of contents

frontiers | Frontiers in Pharmacology

TYPE Editorial
PUBLISHED 09 May 2023
DOI 10.3389/fphar.2023.1208518

OPEN ACCESS

EDITED AND REVIEWED BY
Olivier Feron,
Université catholique de Louvain,
Belgium

*CORRESPONDENCE
Zhiqian Zhang,
zhangzq@sustech.edu.cn

RECEIVED 19 April 2023
ACCEPTED 02 May 2023
PUBLISHED 09 May 2023

CITATION
Tao F and Zhang Z (2023), Editorial: Epigenetic drugs and therapeutic resistance for epithelial malignancies. *Front. Pharmacol.* 14:1208518. doi: 10.3389/fphar.2023.1208518

Editorial: Epigenetic drugs and therapeutic resistance for epithelial malignancies

Fangfang Tao[1] and Zhiqian Zhang[2,3]*

[1]Department of Immunology and Microbiology, Basic Medical College, Zhejiang Chinese Medical University, Hangzhou, Zhejiang, China, [2]Department of Human Cell Biology and Genetics, School of Medicine, Southern University of Science and Technology, Shenzhen, China, [3]Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology, Shenzhen, China

KEYWORDS

epigenetic modifications, cancer therapy, epigenetic drugs, therapeutic resistance, epithelial malignancies

Editorial on the Research Topic
Epigenetic drugs and therapeutic resistance for epithelial malignancies

Epigenetic modifications are widely recognized for their crucial role in the development and progression of cancer, particularly in epithelial malignancies. These changes involve modifications to DNA molecules and their associated proteins that can influence gene expression without altering the DNA sequence itself. Given their heritable and reversible nature, epigenetic modifications have become an attractive target for cancer therapy. In recent years, there has been a growing interest in developing epigenetic drugs that can target specific modifications and potentially overcome therapeutic resistance. Many cancers, such as breast cancer, lung cancer, and colorectal cancer, are some of the most commonly diagnosed epithelial malignancies worldwide. Although significant progress has been made in developing targeted therapies, drug resistance remains a significant challenge, which often results in treatment failure and disease progression. Epigenetic modifications, such as nuclear dynamic, DNA methylation, covalent histone modification, histone variants, and non-coding RNA (ncRNA)—including microRNA (miRNA/miR) and long ncRNA (lncRNA)—have all been shown to play a critical role in the development of therapeutic resistance in cancer.

Epigenetic modifications play a significant role in the development of drug resistance in cancer patients. However, drugs that target these modifications, such as DNA methyltransferase inhibitors and histone deacetylase inhibitors, have the potential to reverse them and restore sensitivity to standard therapies (Steele et al., 2009; Vijayaraghavalu and Labhasetwar, 2018; Bao et al., 2020). One example of a successful treatment is the use of 5-Aza-2′-deoxycytidine (5-aza-D) to reverse cisplatin resistance in bladder cancer cells. This effect is attributed to the demethylation of the HOXA9 gene promoter (Xylinas et al., 2016). The use of such drugs is a promising approach to combat drug resistance, particularly in patients with hematological cancer types.

The focus of this Research Topic is on epigenetic events in epithelial malignancies, particularly their development, properties, and mechanistic studies.

First, Epigenetic alterations are known to play a significant role in the development of epithelial malignancies and can also serve as biomarkers to predict their outcome. In this Research Topic, Ye et al. explore the roles of 84 methylation-related genes (MRGs)

modification patterns in prostate cancer (PCa) and tumor microenvironment (TME) diversity, clinicopathological characteristics, and various prognostic regulatory mechanisms. The clinical significance of MRGs is highlighted, offering a new perspective for PCa research and advancing our understanding of TME and immunotherapy. Similarly, Cao et al. investigate the role of gene methylation in Pancreatic adenocarcinoma (PAAD), screening potential anti-cancer small molecule drugs and constructing a prediction model to assess PAAD prognosis. The classification model based on differentially methylated and expressed genes (DMEGs) accurately distinguished between normal and tumor samples, underscoring the potential of epigenetic biomarkers and precision medicine in managing PAAD.

Second, therapeutic resistance poses a significant challenge in cancer treatment, and epigenetic changes have been identified as one of the underlying factors. DNA methylation, histone modifications, and chromatin remodeling are examples of epigenetic alterations that contribute to the resistance of cancer cells to chemotherapy therapy. These modifications can disrupt the expression of genes involved in cell cycle regulation, DNA repair, and apoptosis, ultimately leading to the survival and proliferation of cancer cells. Across these publications, Zhang et al. investigate molecular subtypes of thyroid cancer based on immune cell infiltration, underscoring the role of epigenetic modifications in tumor progression and their potential for immunotherapy. This highlights the significance of considering epigenetic modifications in the development of effective cancer therapies.

The role of lncRNA in epigenetics and therapeutic resistance has garnered increasing attention. LncRNAs regulate gene expression through various mechanisms, including transcriptional and post-transcriptional regulation, chromatin remodeling, and competitive binding with miRNAs. The mechanisms involved in lncRNA-mediated gene expression regulation are complex and diverse. Moreover, lncRNA has been linked to the development of therapeutic resistance. In-depth research on lncRNA is expected to yield new strategies and directions for tumor treatment and prevention. Throughout research, Zhang et al. identified RNF157-AS1 as a key lncRNA associated with both doxorubicin resistance and hepatocellular carcinoma (HCC) prognosis. Furthermore, they developed a four-gene risk model that shows potential for predicting HCC prognosis.

Third, in recent years, high-throughput sequencing technology, public databases, and single-cell sequencing technology have become important tools in studying epigenetics. High-throughput sequencing can quickly and accurately detect changes at multiple levels, such as genome, transcriptome, and epigenome levels, providing valuable insights into epigenetic mechanisms. Public databases, such as TCGA and ICGC, include large-scale datasets of tumor samples, enabling researchers to obtain comprehensive epigenetic information and gain important insights into cancer therapy. Within this Research Topic, multiple studies have utilized data from hepatocellular carcinoma (Cheng et al.; Wang et al.), prostate cancer (Deng et al.), pancreatic cancer (Cao et al.; Cao et al.; Ji et al., 2023), lung adenocarcinoma (Liu et al.; Zhou and Zhao), glioblastoma (Lu et al.), colon adenocarcinoma a (Feng et al.), clear cell renal cell carcinoma (Deng et al.), and associated drug resistance data from GEO, TCGA, ICGC, and GTEx. These databases hold a wealth of valuable information that is yet to be fully explored.

Single-cell sequencing technology is a newly emerging tool that enables high-throughput sequencing of the genome, transcriptome, epigenome, and other information of individual cells with high resolution and accuracy. It allows researchers to study the epigenetic differences between cells, better understand the mechanisms underlying tumor occurrence and development, and develop more precise cancer therapy strategies. In this Research Topic, Deng et al. utilized the single-cell dataset of clear cell renal cell carcinoma to identify 24 cell clusters and marker genes for two different cell types in each cluster. Correlation analysis from the research of Lu et al. revealed that paclitaxel treatment affects neurons and may improve glucose metabolism and modulate immune function in glioblastoma. This study highlights the potential of single-cell sequencing to investigate the pharmacological targets and signaling pathways of paclitaxel in glioblastoma and gain insights into its mechanism of action.

In addition, promising approaches have emerged for overcoming therapeutic resistance in epithelial malignancies, including small molecule drugs, traditional Chinese medicine, and network pharmacology. Small molecule drugs, in particular, have demonstrated great potential in targeting epigenetic modifications, specifically histone acetylation and DNA methylation. By selectively inhibiting the enzymes responsible for these modifications, small molecule drugs can disrupt oncogenic signaling pathways and induce apoptosis in cancer cells. Additionally, combination therapy with small molecule drugs and conventional chemotherapy can enhance treatment efficacy and reduce the likelihood of therapeutic resistance. Traditional Chinese medicine has demonstrated promising results in the treatment of epithelial malignancies. These medicines have multi-target effects that can modulate epigenetic modifications, including histone deacetylation and DNA methylation, inhibiting cancer cell growth and proliferation. Moreover, traditional Chinese medicine injections can improve immune function, enhancing the body's ability to fight cancer. Network pharmacology, a novel approach that combines computational methods with experimental validation, has also shown promise in identifying the molecular mechanisms underlying drug action. It can help to identify novel targets for cancer therapy and provide insights into the mechanisms of therapeutic resistance. In this Research Topic, Cao et al. describe a study on paeoniflorin (PF), an herbal active ingredient with anti-tumor effects. They identified PF targets and performed gene enrichment analysis to determine the biological processes impacted by PF. This study also identified the most relevant genes to PF treatment and validated the identified targets using PANC-1 and Capan-2 cells. The study found that PF may regulate inflammatory factors through the p38 MAPK signal pathway and may have potential as a natural anti-tumor compound for pancreatic cancer. Similarly, Wang et al. identified active ingredients and related genes of traditional Chinese medicine injections for treating hepatocellular carcinoma (HCC), characterizing two HCC subtypes and identifying important genes such as *SPP1* as an oncogene in HCC. The study suggests that traditional Chinese medicine injections can serve as an important adjuvant treatment modality for HCC.

As research on the role of epigenetics in cancer continues to advance, the potential for epigenetic drugs to overcome therapeutic resistance and improve patient outcomes becomes increasingly

clear. Future studies are likely to focus on the development of more specific and effective epigenetic therapies that can target specific modifications and combat drug resistance. Additionally, the integration of high-throughput sequencing technology, single-cell sequencing, and network pharmacology will continue to provide new insights into the underlying mechanisms of tumor development and therapeutic resistance. As personalized medicine and precision oncology continue to advance, the identification of epigenetic biomarkers for predicting patient response to therapy will become increasingly important.

Author contributions

All authors listed have made a direct and intellectual contribution to the work and approved it for publication. ZZ and FT wrote and revised the manuscript. ZZ finalized the last version.

Funding

This study is supported by the National Natural Science Foundation of China Grant (32170568), National Natural Science Foundation of China for Youth (82002997), General Project of Basic Research in Shenzhen (JCYJ20190809161215057), Guangdong Provincial Medical Research Fund (A2022131), Youth Project of Guangdong Basic and Applied Basic Regional Joint Fund (2020A1515110281).

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

Bao, Y., Oguz, G., Lee, W. C., Lee, P. L., Ghosh, K., Li, J., et al. (2020). EZH2-mediated PP2A inactivation confers resistance to HER2-targeted breast cancer therapy. *Nat. Commun.* 11, 5878. doi:10.1038/s41467-020-19704-x

Ji, T., Wu, H., and Cao, T. (2023). Establishment of a 7-gene prognostic signature based on oxidative stress genes for predicting chemotherapy resistance in pancreatic cancer. *Front. Pharmacol.* 14, 861.

Steele, N., Finn, P., Brown, R., and Plumb, J. A. (2009). Combined inhibition of DNA methylation and histone acetylation enhances gene re-expression and drug sensitivity *in vivo*. *Br. J. Cancer* 100, 758–763. doi:10.1038/sj.bjc.6604932

Vijayaraghavalu, S., and Labhasetwar, V. (2018). Nanogel-mediated delivery of a cocktail of epigenetic drugs plus doxorubicin overcomes drug resistance in breast cancer cells. *Drug Deliv. Transl. Res.* 8, 1289–1299. doi:10.1007/s13346-018-0556-y

Xylinas, E., Hassler, M. R., Zhuang, D., Krzywinski, M., Erdem, Z., Robinson, B. D., et al. (2016). An epigenomic approach to improving response to neoadjuvant cisplatin chemotherapy in bladder cancer. *Biomolecules* 6. doi:10.3390/biom6030037

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fphar.2022.1011033

OPEN ACCESS

EDITED BY
Zhi-Qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
JinKu Zhang,
Hebei University, China
Maria Anna Siciliano,
Magna Græcia University of Catanzaro, Italy

*CORRESPONDENCE
Xiao Wu,
xwu@njucm.edu.cn
Yongming Li,
lym-569@njucm.edu.cn

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 03 August 2022
ACCEPTED 22 August 2022
PUBLISHED 26 September 2022

CITATION
Cheng X, Li J, Feng L, Feng S, Wu X and Li Y (2022), The role of hypoxia-related genes in TACE-refractory hepatocellular carcinoma: Exploration of prognosis, immunological characteristics and drug resistance based on onco-multi-OMICS approach. *Front. Pharmacol.* 13:1011033. doi: 10.3389/fphar.2022.1011033

The role of hypoxia-related genes in TACE-refractory hepatocellular carcinoma: Exploration of prognosis, immunological characteristics and drug resistance based on onco-multi-OMICS approach

Xuelian Cheng[1], Jingjing Li[1], Limei Feng[1], Songwei Feng[2,3], Xiao Wu[1]* and Yongming Li[1]*

[1]School of Medicine and Holistic Integrative Medicine, Jiangsu Collaborative Innovation Center of Chinese Medicinal Resources Industrialization, Nanjing University of Chinese Medicine, Nanjing, China, [2]School of Medicine, Southeast University, Nanjing, China, [3]State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing, China

Transcatheter arterial chemoembolization (TACE) is an effective treatment for hepatocellular carcinoma (HCC). During TACE, chemotherapeutic agents are locally infused into the tumor and simultaneously cause hypoxia in tumor cells. Importantly, the poor effect of TACE in some HCC patients has been shown to be related to dysregulated expression of hypoxia-related genes (HRGs). Therefore, we identified 33 HRGs associated with TACE (HRGTs) by differential analysis and characterized the mutational landscape of HRGTs. Among 586 HCC patients, two molecular subtypes reflecting survival status were identified by consistent clustering analysis based on 24 prognosis-associated HRGs. Comparing the transcriptomic difference of the above molecular subtypes, three molecular subtypes that could reflect changes in the immune microenvironment were then identified. Ultimately, four HRGTs (*CTSO*, *MMP1*, *SPP1*, *TPX2*) were identified based on machine learning approachs. Importantly, risk assessment can be performed for each patient by these genes. Based on the parameters of the risk model, we determined that high-risk patients have a more active immune microenvironment, indicating "hot tumor" status. And the Tumor Immune Dysfunction and Exclusion (TIDE), the Cancer Immunome Atlas (TCIA), and Genome of Drug Sensitivity in Cancer (GDSC) databases further demonstrated that high-risk patients have a positive response to immunotherapy and have lower IC50 values for drugs targeting cell cycle, PI3K/mTOR, WNT, and RTK related signaling pathways. Finally, single-cell level analysis revealed significant overexpression of *CTSO*, *MMP1*, *SPP1*, and *TPX2* in malignant cell after PD-L1/CTLA-4 treatment. In conclusion, Onco-Multi-OMICS analysis showed that HRGs are potential biomarkers for patients with refractory TACE, and it provides a novel immunological perspective for developing personalized therapies.

KEYWORDS
TACE, drug resistance, hepatocellular carcinoma, prognosis, hypoxia

Introduction

Hepatocellular carcinoma (HCC) accounts for 85% of all liver cancers (Villanueva., 2019). Despite advances in treatment strategies for HCC, the overall 5-years survival rate for patients with HCC remains below 20% (Zheng et al., 2018). Transcatheter arterial chemoembolization (TACE) is a therapy in which drugs are injected into the arteries supplying HCC tissue (Chang et al., 2020). Some studies suggest that TACE-refractory may lead to poor prognosis in HCC patients. It has been shown that TACE procedures can exacerbate hypoxic states (Liu et al., 2016). However, we still lack a multi-omics data-based perspective on the immunological characteristics of hypoxia-associated gene sets in TACE-refractory patients.

Hypoxia is an intrinsic feature of solid tumors due to the imbalance between tumor cell proliferation rate and vascular nutrient supply (Gray et al., 1953). Previous studies have shown that hypoxia can regulate tumor immune microenvironment (TME), such as promoting the recruitment of innate immune cells, and interfering with the differentiation and function of adaptive immune cells (Feng et al., 2022a). Certain cytokines secreted by malignant tumors, especially in hypoxia condition, may induce angiogenesis and metastasis (Zarogoulidis et al., 2014). A retrospective clinical study has shown that high pre-treatment IL-8 levels are a significant predictor of shorter survival and increased refractoriness of TACE (Kim et al., 2015). Therefore, further studies are needed to investigate the hypoxia-related genes (HRGs) that contribute to TACE refractoriness. Importantly, exploring the relationship between the above genes and drug resistance can lead to the development of new therapeutic strategies.

Nowadays, the study of molecular mechanisms based on Onco-Multi-OMICS approach has become one of the most important tools (Feng et al., 2022b; Zhu et al., 2022). Therefore, we searched for hub HRGs contributing to TACE refractoriness and searched for optimal biological markers by combining transcriptome, single cellome, immunome, and whole-exome. Our study also illustrated the immunological characteristics in different risk group and explored their impact on the response to chemotherapy and immunotherapy. In conclusion, our results were beneficial for the management and treatment of TACE-refractory patients.

Materials and methods

Data collection and pre-processing

The mRNA expression profile data of HCC patients were retrieved from TCGA and GEO databases, and the exclusion criteria was as follows: lack of complete follow-up information, 0 days of survival, and repeated sequencing. Supplementary Table S1 showed treatment details for patients in the GSE14520 cohort before exclusion. Finally, 365 tumor samples from the TCGA-LIHC cohort and 221 tumor samples from the GSE14520 cohort were included. Moreover, to study TACE response, we obtained gene expression profile data from GSE104580, which included 100 TACE-responsive samples and 100 TACE-refractory samples (He et al., 2022). Both somatic mutation data and CNV data were obtained from the TCGA-LIHC cohort, including 371 tumor samples. Notably, 'maftools' package was used to present the mutation status of each gene. We removed the batch effect between RNA-seq and microarray data by using the 'sva' package and made the newly generated gene matrix based on two cohorts as a meta cohort.

Identification of hypoxia-related genes in TACE refractoriness

Differentially expressed genes (DEGs) between different response states were identified by using the 'limma' package in the GSE104580 cohort, $p < 0.05$, with $|\log FC| > 1$ as the threshold. In addition, 1,694 HRGs were extracted from the previous study (Zhang et al., 2020). The above DEGs and HRGs were overlaid to identify the HRGs associated with TACE (HRGTs).

Enrichment analysis

Differential genes between subtypes were analyzed using the 'limma' package (adj. $p < 0.05$, $|\log FC| > 1$). Biological pathways were annotated by using the 'clusterProfiler' package for Gene Ontology (GO), Kyoto Gene and Genome Encyclopedia (KEGG). p-value < 0.05 and q-value < 0.05 were considered as significant enrichment pathways. Differences in biological pathways between subtypes were assessed by using 'GSVA' algorithm. And KEGG geneset (c2. cp. kegg. v7. 0. symbols. Gmt) was used as the reference gene set with FDR <0.05 as the threshold.

Consistent clustering analysis

In the meta cohort, the prognostic value of each HRGTs was determined by using univariate cox regression analysis. Consensus clustering, an unsupervised clustering method, is a common method to classify subtypes based on the CDF slope was smallest. Consistent cluster analysis and principal

component analysis (PCA) were performed to determine whether each subtype was relatively independent of the other subtypes based on prognostic HRGTs ($p < 0.05$) and prognostic DEGs ($p < 0.05$). The number of clusters was determined by using 'conensusClusterPlus' package. A 1000 replicates with pltem = 0.8 were performed to verify the stability of the subtypes. We used Kaplan Meier analysis and log-rank test to assess the overall survival (OS) of HCC patients in different subtypes.

Identification and validation of risk scores

Modeling and validation were performed by TCGA-LIHC cohort and GSE14520 cohort, respectively. The least absolute shrinkage and selection operator (LASSO) (Feng et al., 2022c) model was used to remove redundant genes from HRGTs. Subsequently, multivariate Cox regression analysis was performed to integrate the coefficients and then establish risk score formulas by gene expression values. Univariate and multivariate Cox regression analyses were used to assess the prognostic value of risk scores across the entire dataset and the external validation dataset. Time-dependent subject operating characteristic (ROC) curves were used to compare the predictive accuracy of risk scores with traditional clinicopathological parameters.

Drug sensitivity analysis

Half maximal inhibitory concentration (IC50) was calculated using the 'prophetic' package. Relevant drugs targeting cell cycle, PI3K/mTOR, WNT, and RTK pathways were obtained from the Genome of Drug Sensitivity in Cancer (GDSC) database. Charoentong et al. developed a quantitative scoring scheme called the Immunophenotype Score (IPS) to identify the determinants affecting tumor immunogenicity. IPS is a representative gene associated with immunogenicity calculated using z-score, and our meta cohort's IPS was calculated from the Cancer Immunome Atlas (TCIA) database (Wu et al., 2018). Moreover, Peng et al. designed a computational architecture, Tumor Immune dysfunction and ejection (TIDE) score (Jiang et al., 2018), to integrate the two mechanisms of tumor immune escape. Our meta cohort's TIDE score was calculated from the TIDE database.

Single-cell analysis

The HCC single cell dataset was obtained from GSE125499, and single cell expression profile with annotated cell information were obtained from the Tumor Immune Single Cell Hub (TISCH) database (Sun et al., 2021). Finally, we compared the expression changes of *CTSO*, *MMP1*, *SPP1*, *TPX2* in different cell types.

Immunological analysis

We used different algorithms to estimate the abundance of immune cells in different samples, such as ssGSEA, TIMER, CIBERSORT, QUANTISEQ, MCP-counter, XCELL and EPIC. Then, ESTIMATE algorithm was used to calculate the immune score and stromal score to reflect the TME status.

Statistical analysis

All statistical analyses were performed using the R software (v.4.1.1). Kruskal-wallis test was used for comparison between groups, $\chi 2$ test was used for association between covariates, and Kaplan-Meier method was utilized to compare survival differences between groups. More detailed statistical methods for transcriptome data processing are covered in the above section (Ye et al., 2022). $p < 0.05$ was considered statistically significant.

Results

Landscape of HRGTs in HCC

A total of 274 DEGs were first identified from the GSE104580 cohort (Figures 1A,C) and overlapped with existing HRGs genes in the database. Finally, 33 HRGTs were identified (Figure 1B). The above genes may play a key role in TACE refractoriness. HRGTs were mutated in 34 of 371 samples with a frequency of 9.16%, most of which had a low mutation frequency (Figure 1D). In addition, Copy number variation (CNV) were prevalent in HRGTs. ORG1 focused on copy number amplification, while CNV deletion frequency was common in CDC20 (Supplementary Figure S1A). The location of HRGTs on the chromosome (Figure 1E). GSE14520, TCGA-LIHC were included in a meta cohort using the 'combat' algorithm. The network of HRGTs specifically described the combined gene interactions and their prognostic significance for patients (Figure 2A). Cox regression analysis identified 24 HRGTs were indicative of prognostic significance for HCC patients (Figure 2B).

Identification of molecular subtypes based on HRGTs

The classification was optimal when the k value = 2 (Figure 2C). Two different subtypes were finally identified, with 326 cases in subtype A and 260 cases in subtype B. The heat map showed the

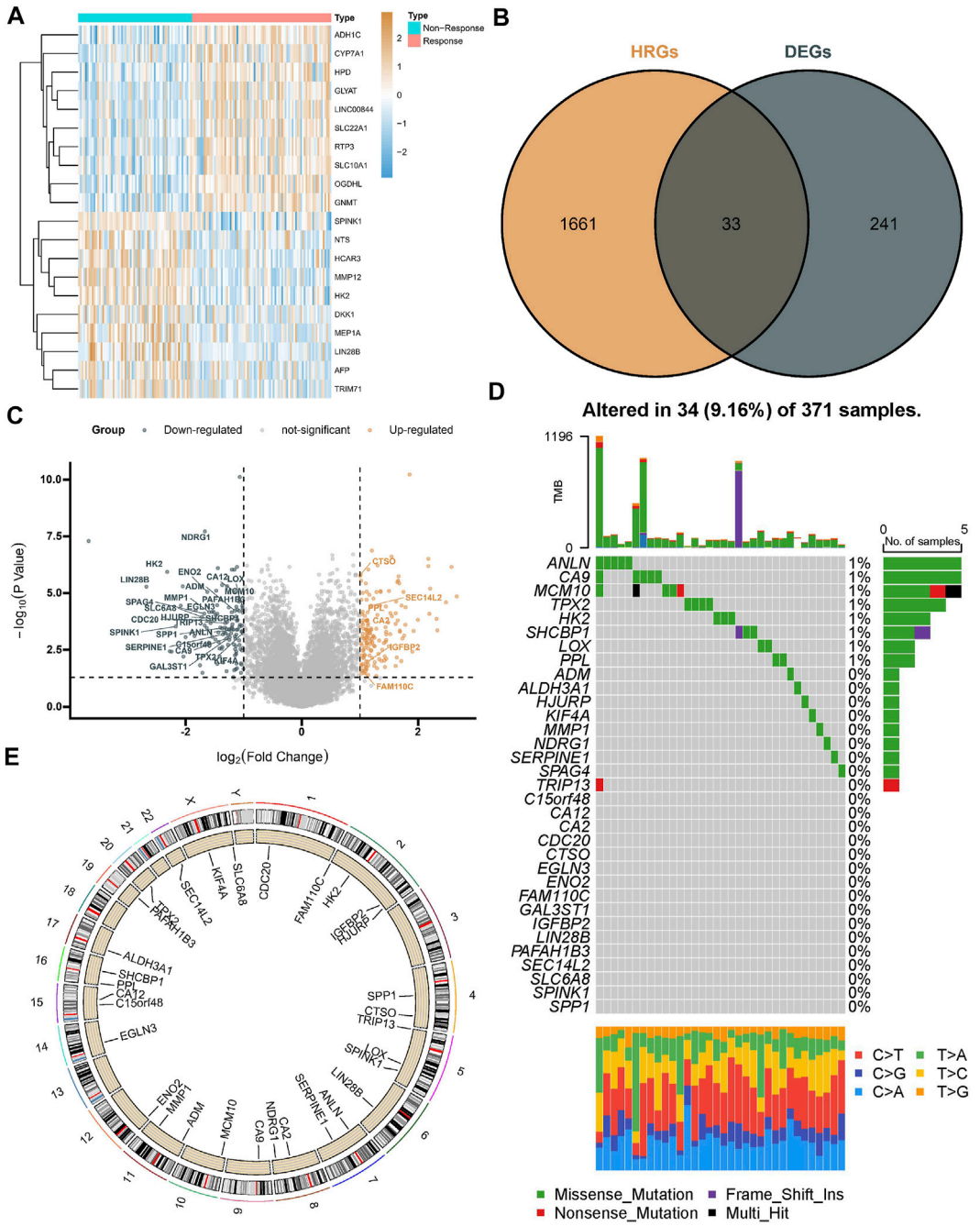


FIGURE 1
Landscape of HRGTs in HCC. **(A)** The heat map showed a total of DEGs were identified from the GSE104580 cohort. **(B)** The venn plot showed DEGs overlapped with existing HRGs in the database. **(C)** The volcano plot showed dysregulation status of DEGs. **(D)** Mutation landscape of HRGTs in 371 samples. **(E)** The location of HRGTs on the human chromosome.

distribution of the clinical features of the different subtypes, with most genes significantly overexpressed in subtype A (Figure 2D). PCA plot revealed that the two molecular subtypes had a relative discrete features (Figure 2E). Prognostic analysis revealed a significant survival disadvantage in the subtype B (Figure 2F).

Immune microenvironment and biological pathways in molecular subtypes

The ESTIMATE algorithm reveals that subtype A has a higher immune score, while the stromal score was significantly

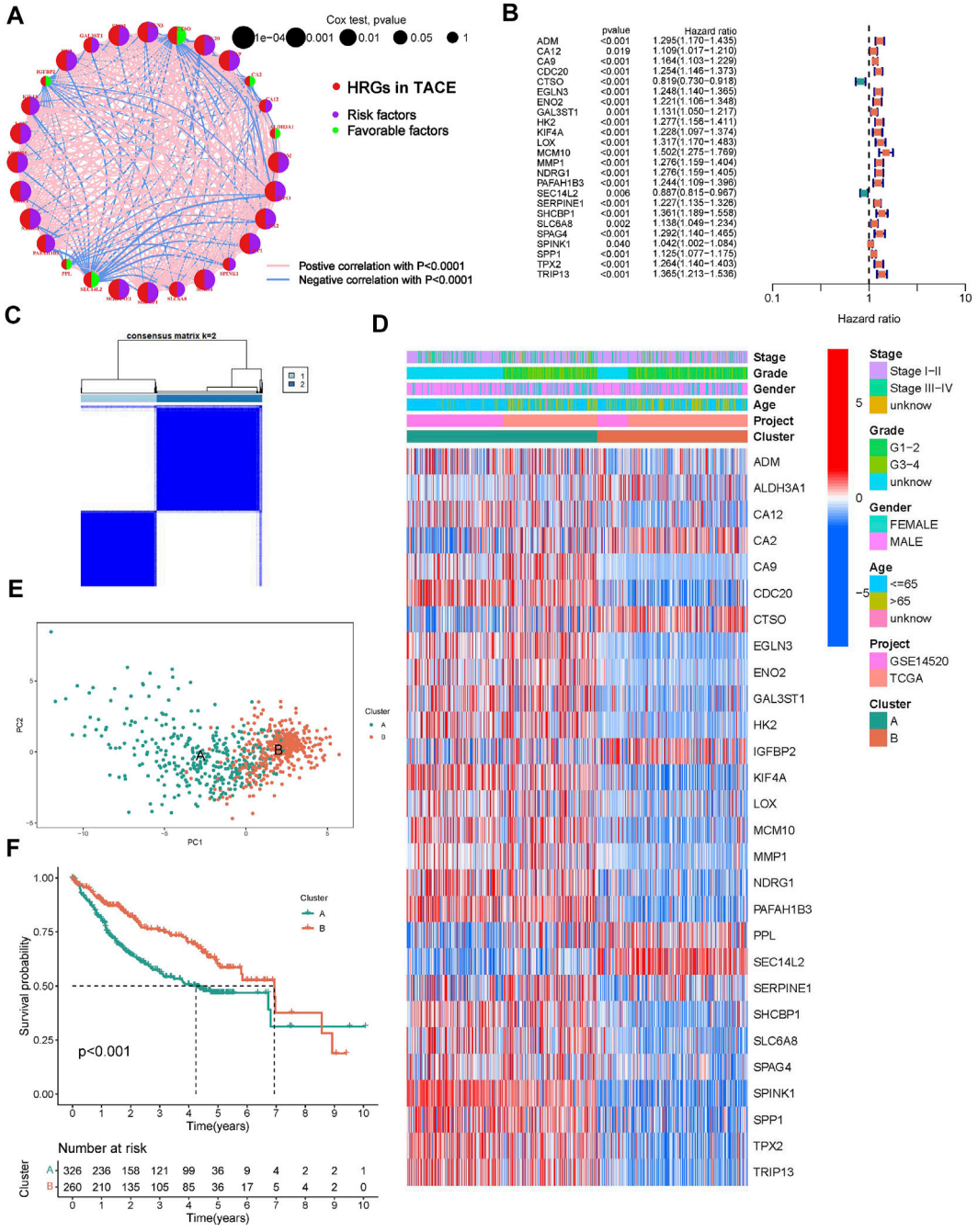


FIGURE 2
Molecular subtypes based on HRGTs. **(A)** The network of HRGTs described the combined interactions and prognostic significance. **(B)** A forest plot showed 24 HRGTs were indicative of prognostic significance. **(C)** The classification was optimal when the k value = 2. **(D)** Heat map of distribution of clinicopathological characteristics and molecular subtypes based on HRGTs. **(E)** PCA plot revealed that the two molecular subtypes had a relative discrete features. **(F)** Kaplan-Meier analysis of overall survival time in different molecular subtypes.

downregulated compared to subtype B (Figure 3A). In addition, ssGSEA analysis demonstrated the TME status in different molecular subtypes. We discovered that subtype A is probably exhibiting hot tumor characteristics. This was due to a significant rise in activated $CD4^+$ T cells, which may have a more active TME (Figure 3B). In addition, we showed the expression of HLA as well as ICI mRNA in different subtypes. Interestingly, the subtype A had higher mRNA expression in most HLAs, such as HLA-A, HLA-B,

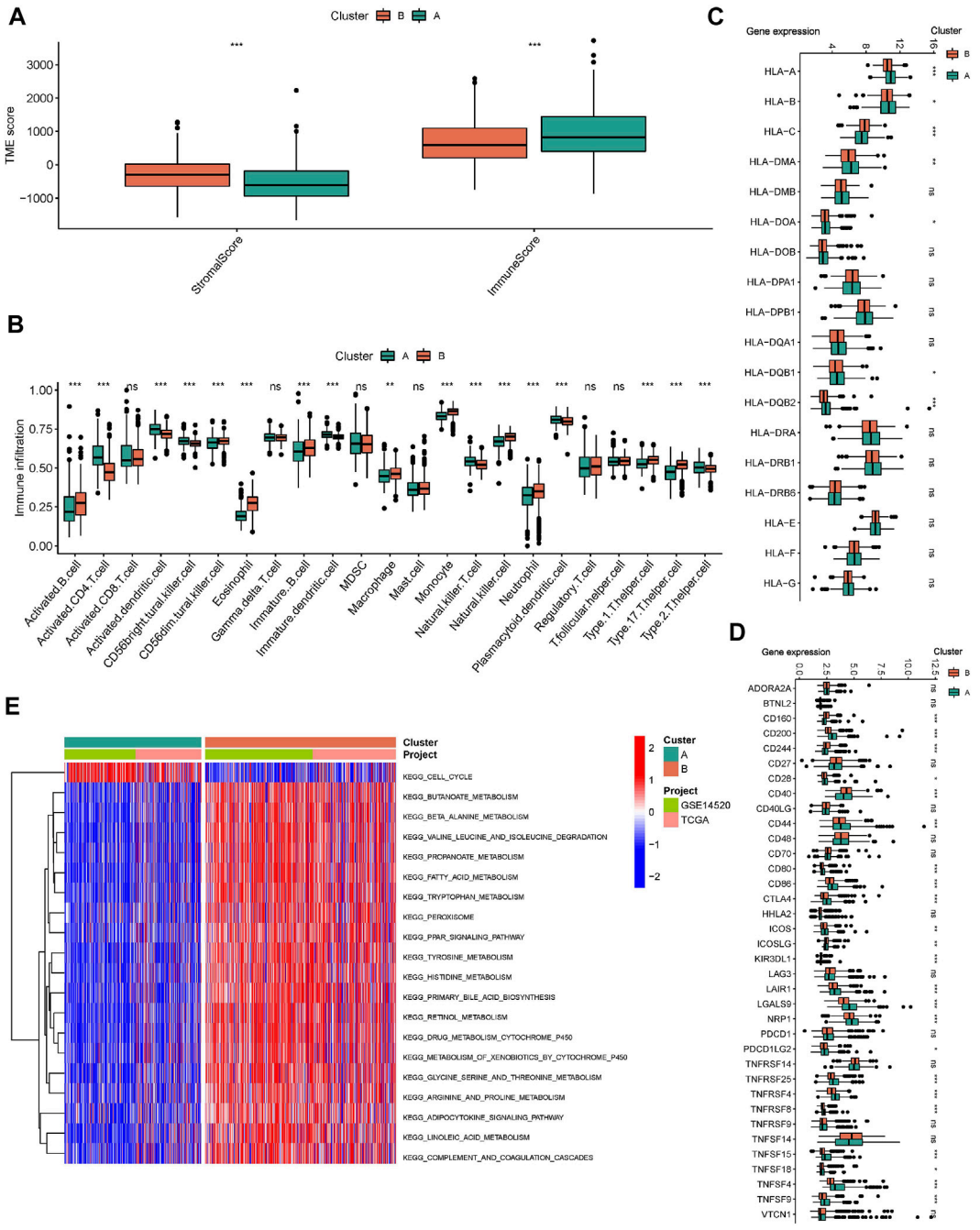


FIGURE 3
Immune microenvironment and biological pathways in molecular subtypes. **(A)** The Box plot of TME score in different molecular subtypes. **(B)** The Box plot of immune cells score based on ssGSEA in different molecular subtypes. The Box plot of mRNA expression of HLA **(C)** and ICIs **(D)** in different molecular subtypes. **(E)** GSVA analysis in different molecular subtypes using KEGG genesets. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

HLA-C, and HLA-DDA (Figure 3C). Similarly, subtype A had higher mRNA expression in most ICIs, such as CTLA4 (Figure 3D). We made a hypothesis that subtype A would benefit more from immunotherapy. To explore the biological behavior between these different subtypes, we performed Gene set variation analysis. Subtype A showed significant enrichment with cell cycle pathways compared to subtype B (Figure 3E). In addition, we performed a differential analysis between the two subtypes. It was found that the major enrichment pathways of the 496 DEGs identified (Figure 4A) may be associated with biological

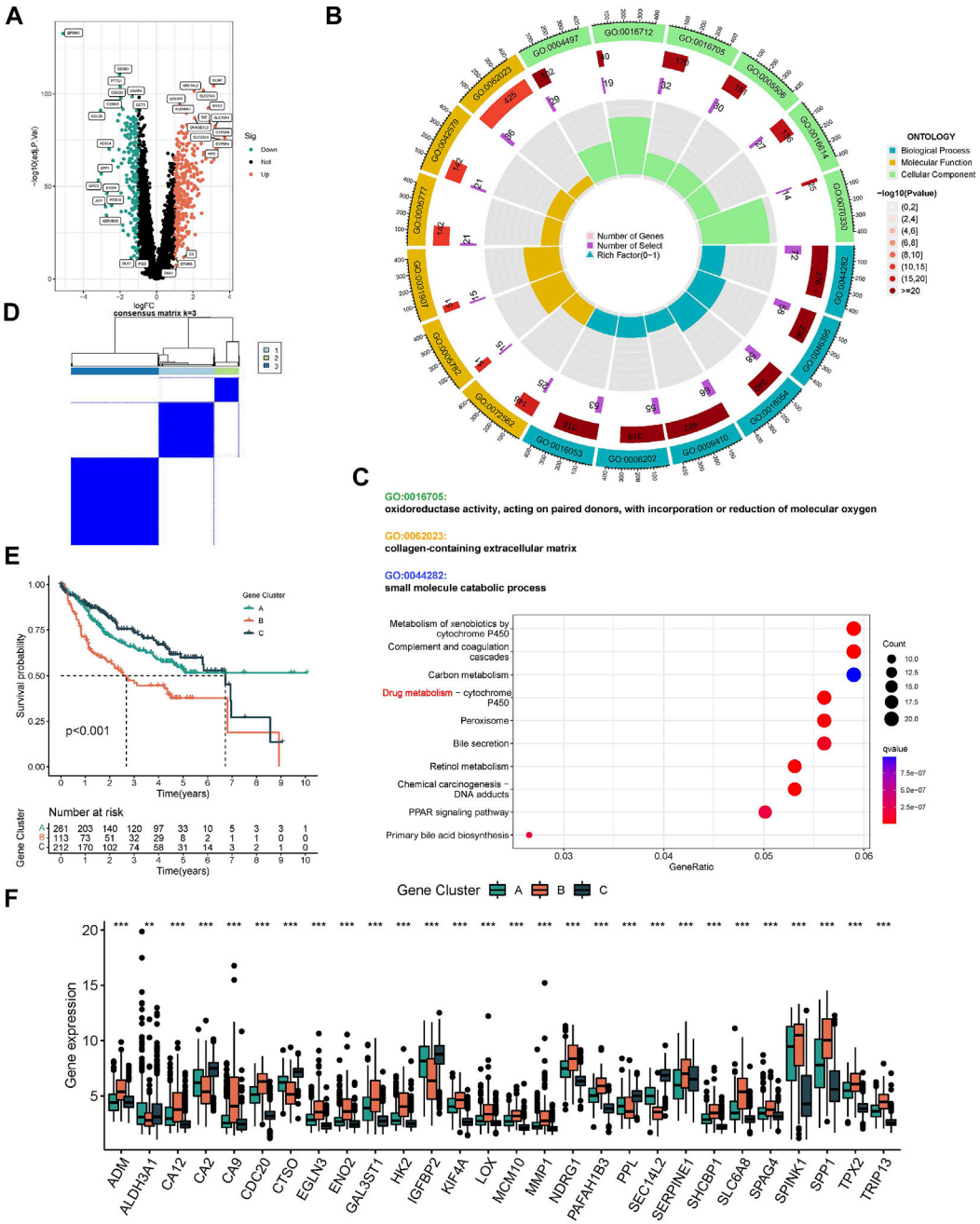


FIGURE 4
Molecular subtypes based on DEGs. **(A)** The Volcano plot showed DEGs in different subtypes. Heat map of unsupervised clustering analysis. **(B)** The circle plot of GO enrichment analysis based on DEGs. **(C)** The bubble plot of KEGG enrichment analysis based on DEGs. **(D)** The classification was optimal when the k value = 3. **(E)** Kaplan-Meier analysis of overall survival time in different molecular subtypes based on DEGs. **(F)** The Box plot of mRNA expression of HRGTs in different molecular subtypes. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

processes related to oxidative stress, extracellular genes and drug metabolism (Figures 4B,C). Finally, we identified three different regulatory subtypes based on the above DEGs (Figure 4D). Among them, subtype B had the worst prognosis, while subtype C had the best prognosis (Figure 4E). And the above 33 HRGTs were differentially expressed in different subtypes (Figure 4F).

Identification of risk score in HCC

TCGA-LIHC was used as a training cohort with overall survival (OS) as the outcome. The LASSO model was used to remove redundant genes (Figures 5A,B). The coefficients of each gene were obtained by multifactorial Cox regression analysis. A

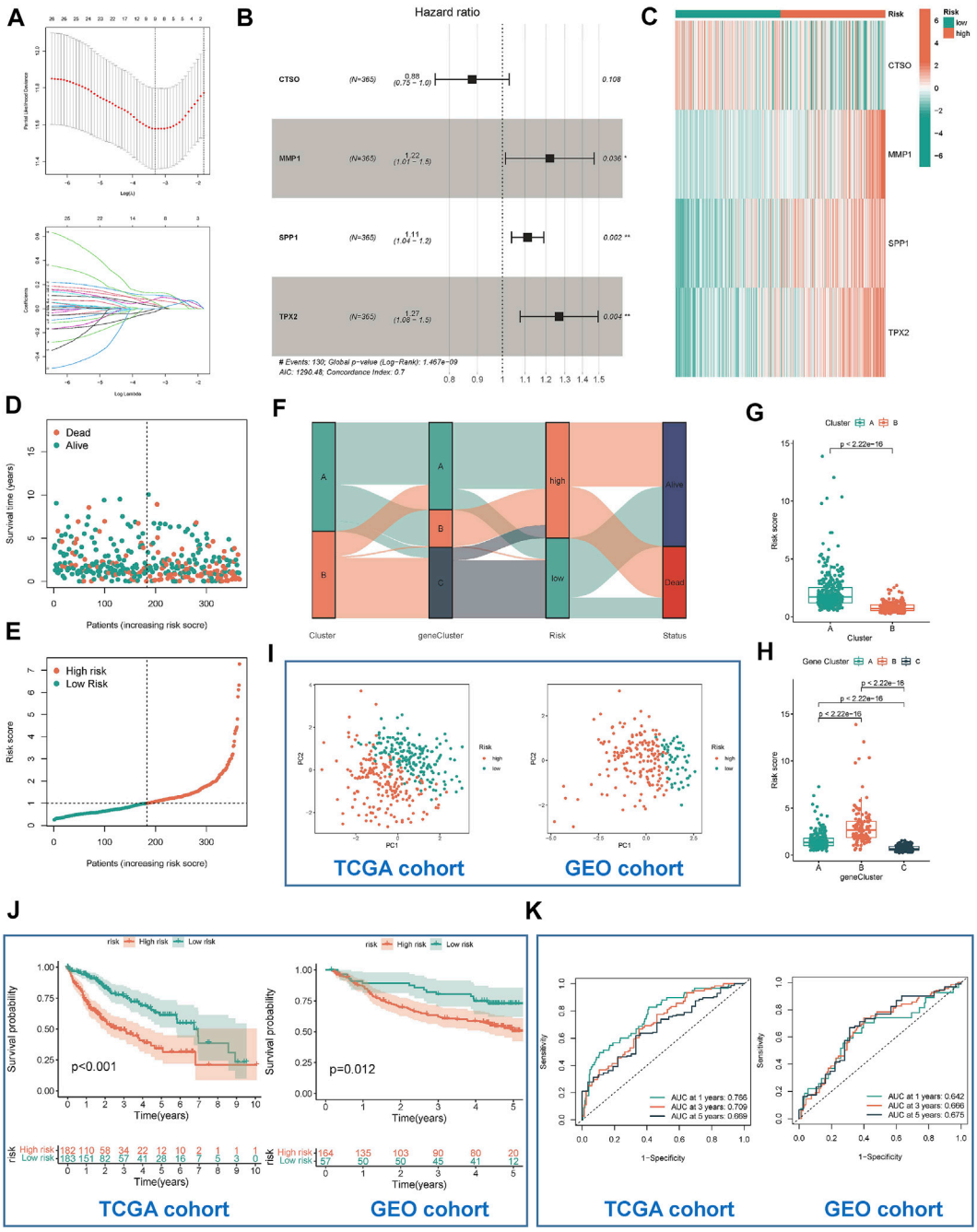


FIGURE 5
Identification and validation of risk model. **(A)** Determination of the number of regulators using LASSO analysis. **(B)** Forest plot of multivariate Cox regression analysis. **(C)** The heat map showed four HRGTs expression in different risk group (TCGA cohort). Risk status plot **(D)** and the survival distribution plot **(E)** demonstrated the poorer prognosis of HCC patients with higher risk score. **(F)** Sankey diagram based on different subtypes. **(G)** Differences in risk scores between the two molecular subtypes based on HRGTs. **(H)** Differences in risk scores between the three molecular subtypes based on DEGs. PCA plot **(I)**, Kaplan-Meier analysis **(J)**, ROC curve of 1,3,5 years **(K)** of different risk groups in the TCGA and GEO cohort. *p < 0.05, **p < 0.01, ***p < 0.001.

final signature containing 4 HRGTs was obtained. The formula of each patient was as follows: riskscore = (-0.1277 × expression level of *CTSO*) + (0.1995 × expression level of *MMP1*) + (0.1061 × expression level of *SPP1*) + (0.2385 × expression level of *TPX2*). Using the median value of risk scores in the TCGA cohort, we identified two risk groups for HCC patients:

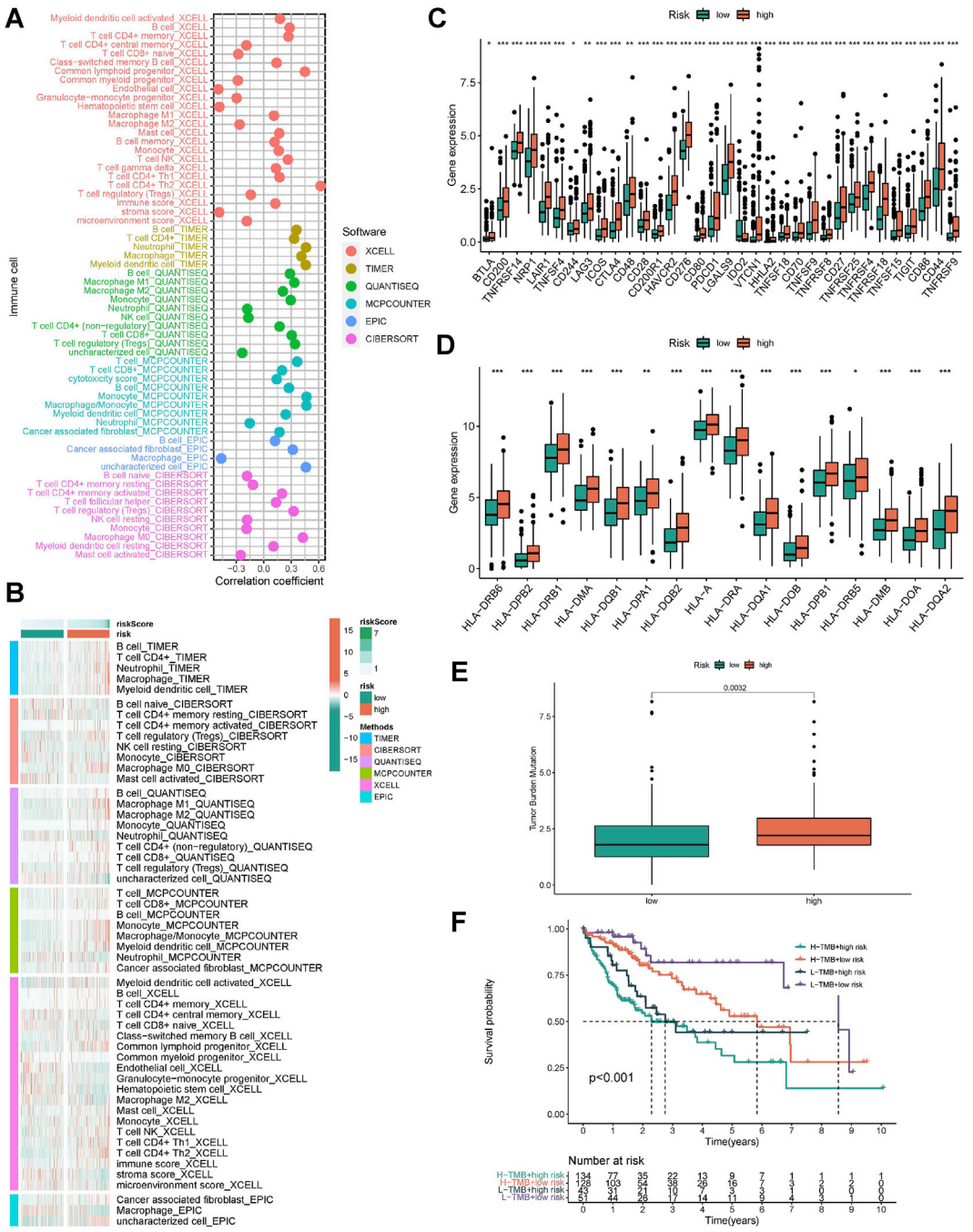


FIGURE 6
Characteristics of immune microenvironment in different risk groups. **(A)** The heat map showed correlation between risk score and immune function score. **(B)** The heat map showed differences in immune function of different risk groups. **(C)** The box plot showed differences in immune checkpoint mRNA expression of different risk groups. **(D)** The box plot showed differences in HLA mRNA expression of different risk groups. **(E)** The box plot showed differences in TMB score of different risk groups. **(F)** Survival analysis by combining TMB score and risk score. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

high-risk group, and low-risk group in all cohorts. Where *MMP1*, *SPP1*, and TPX2 were significantly highly expressed in the high-risk group, while *CTSO* was significantly highly expressed in the low-risk group (Figure 5C). Among them, the risk status plot and the survival distribution plot demonstrated the poorer prognosis of HCC patients with higher risk score (Figures 5D,E). The

results of the sankey plot showed a strong association between risk subtypes and molecular subtypes. And most patients in the subtype A and low risk group were in alive status (Figures 5F–H). PCA also showed genomic heterogeneity between the two risk groups (Figure 5I). Survival analysis and ROC curves indicated (Figures 5J,K) that risk score had a good prognostic value in both the TCGA-LIHC cohort and the GSE14520 cohort, and that survival was suboptimal in patients with both high-risk subtypes. Moreover, we performed correlation analysis between hub genes and m6A methylation regulators, and interestingly, it was positively correlated with most of the regulators except IGFBP1, IGFBP2, and IGF2BP1 (Supplementary Figure S2).

Association of risk subtypes with immune microenvironment

We simultaneously applied different algorithms such as TIMER, CIBERSORT, QUANTISEQ, MCP-counter, XCELL and EPIC to estimate the immune cell infiltration status in each samples. Correlation analysis showed that as the risk score increased, the infiltration score of killer immune cells, such as CD4+ T and CD8+ T cell, also increased (Figure 6A). Similarly, there were differences in the distribution of immune cells in the different risk groups. The high-risk group had a more active TME (Figure 6B). In HLA and ICI analysis, the corresponding mRNA expression was higher in high-risk subtypes (Figures 6C,D). Based on whole-exome sequencing data, patients with both high- and low-risk subtypes did not show significant differences in Top mutated genes, which were *TP53*, *CTNNB1*, and *TTN* (Supplementary Figures S1B,C). Considering the importance of tumor mutational burden (TMB) for immunotherapy, we performed a statistical analysis of the TMB differences between the high- and low-risk groups. It was demonstrated that high-risk group had higher TMB, which suggested that they might have a better response to immunotherapy (Figure 6E). Importantly, when low-TMB and low-riskscore are combined, patients will have the best survival advantage (Figure 6F). In addition, we validated our risk signature in immunotherapy cohort (IMvigor210), and the results were consistent with the above findings, namely, high-risk patients had poor survival outcomes, and high-risk patients were more likely to achieve complete remission (CR), as shown in Supplementary Figure S3.

Risk subtypes could reflect drug resistance in HCC patients

We predicted the drug sensitivity of HCC patients in the meta cohort by utilizing the 'pRRophetic' algorithm and a ridge regression model. The results showed that targeting cell cycle (CGP-60474, GW843682x, BI-2536, and CGP-082996) (Figure 7A), PI3K/mTOR signaling (JW-7-52-1, MK-2206, and A-443654) (Figure 7B), WNT signaling (CHIR-99021) (Figure 7C), and RTK signaling (Sunitinib and PHA-665752) (Figure 7D) were significantly more effective in high-risk group than in low-risk group. The TIDE score showed that the effectiveness of immunotherapy may be better in high-risk patients (Figure 7E). In addition, the IPS results also demonstrated that the high-risk group seems to have more immunogenic phenotypes (Figure 7F).

HRGTs in single cell levels

We annotated the GSE125499 single cell expression profile file based on the TISCH database, and t-SNE plot demonstrated the subpopulation of different cells (Figure 8A). In addition, the violin plot demonstrated the expression of *CTSO*, *MMP1*, *SPP1*, and *TPX2* in different cell types, with *SPP1* being more highly expressed in hepatic progenitor (Figure 8B). Interestingly, after PD-L1/CTLA-4 treatment, *CTSO*, *MMP1*, *SPP1*, and *TPX2* were significantly up-regulated in tumor cells (Figure 8C). Finally, we showed the changes in the proportion of different cell types before and after immunotherapy (Figure 8D). The above data suggest to us that four HRGTs involved in risk signature may have a role in reflecting the response to immunotherapy.

Discussion

Primary liver cancer is one of the sixth most common cancers worldwide and is a common tumor of the digestive system with high aggressiveness and poor prognosis (Choi et al., 2020). Since HCC is not sensitive to conventional radiotherapy, surgery becomes the main treatment method. TACE is the treatment of choice for intermediate-stage hepatocellular carcinoma (Morise et al., 2014). In recent years, however, TACE refractoriness has become a thorny issue and has received increasing attention. This is because TACE accompanied by tumor ischemia plays a dual role in the treatment of HCC. Initially, TACE induces tumor necrosis by blocking the vasculature from the hepatic artery to the HCC. However, TACE also stimulates angiogenesis by inducing hypoxia thereby promoting tumor recurrence and metastasis (Kenji et al., 1997). Tumor angiogenesis and invasiveness by TACE have been found to be mediated by hypoxic signaling, which has been effectively inhibited by antiangiogenic therapies (Liu et al., 2016). However, there are few studies related to HRGs associated with TACE. Although a series of studies have identified predictors or models associated with TACE refractoriness, no studies explore the relevance of HGRTs to the immune microenvironment, prognosis and drug resistance.

Onco-Multi-OMICS approach have been commonly used to discover potential biomarkers (Chakraborty et al., 2018). To date,

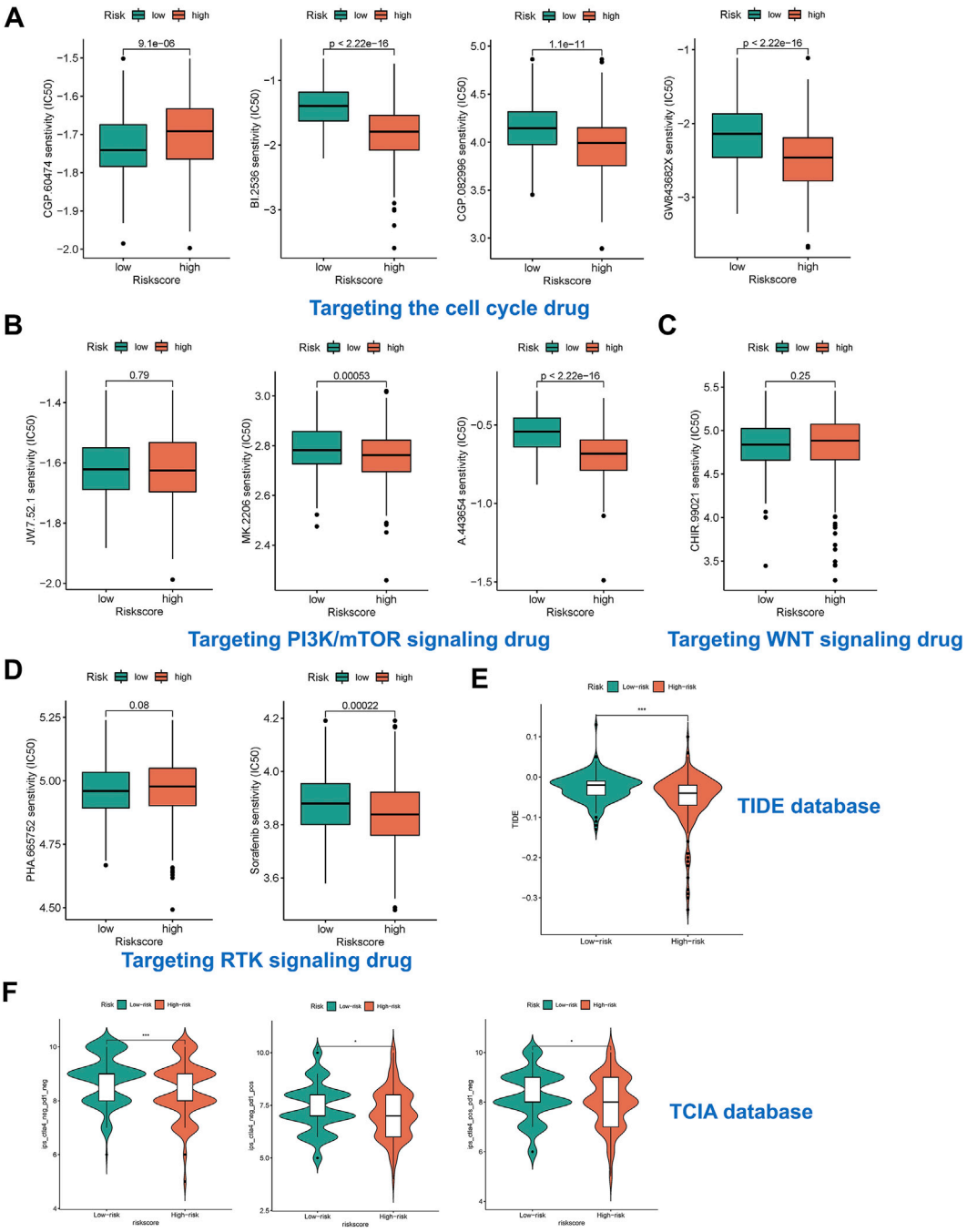


FIGURE 7
Risk subtypes could reflect drug resistance. **(A)** The box plot of targeting cell cycle drug in different risk groups, including CGP-60474, GW843682x, BI-2536, and CGP-082996. **(B)** The box plot of targeting PI3K/mTOR signaling drug in different risk groups, including JW-7-52-1, MK-2206, and A-443654. **(C)** The box plot of targeting WNT signaling drug in different risk groups, including CHIR-99021. **(D)** The box plot of targeting RTK signaling drug in different risk groups, including Sunitinib and PHA-665752. **(E)** The box plot of TIDE score in different risk groups. **(F)** The box plot of IPS in different risk groups.

few studies have constructed prognostic models based on combinations of multiple HRGs in TACE-refractory HCC. Importantly, genetic features and clinical characteristics have performed unsatisfactorily in predicting survival outcomes for TACE-refractory patients. Therefore, we aimed to explore a new HRGTs-based risk stratification and propose potential therapeutic targets for HCC patients. Tumor hypoxia promotes the growth of tumor cells and their transformation

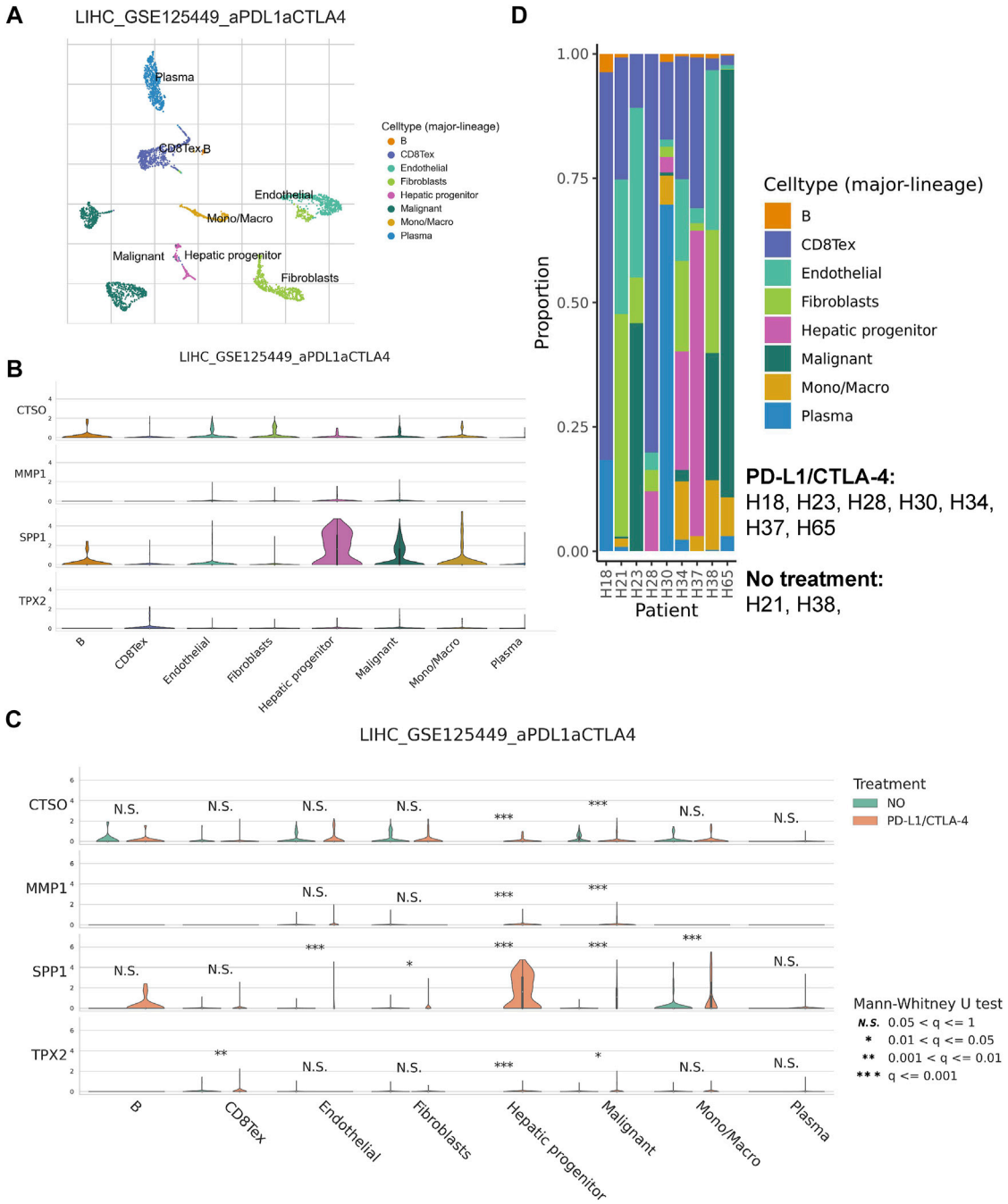


FIGURE 8
HRGTs in single cell levels. **(A)** t-SNE plot demonstrated the subpopulation of different cells. **(B)** The violin plot demonstrated the expression of *CTSO*, *MMP1*, *SPP1*, and *TPX2* in different cell types. **(C)** The violin plot demonstrated the expression of *CTSO*, *MMP1*, *SPP1*, and *TPX2* in different treatment groups. **(D)** The histogram showed the changes in the proportion of different cell types before and after immunotherapys.

to a malignant phenotype. The exploration of hypoxia has opened new perspectives for HCC. Hypoxia is a typical hallmark of TME in almost all solid tumors, caused by rapid and uncontrolled tumor proliferation and inadequate blood supply (Graham and Unger, 2018). Under hypoxic conditions, HIFs bind to transcriptional co-activators and hypoxia response elements to increase the expression of a cascade of target genes, thereby regulating various biological processes, including proliferation, metabolism, angiogenesis, migration and invasion. In addition, hypoxia increases the resistance of

tumor cells to chemotherapy, radiotherapy and even immunotherapy (Li et al., 2004). It can inactivate effector cytokine production by inhibiting T cell proliferation and function. Therefore, it is important to fully understand the effects of hypoxia on TACE. In our study, a total of 274 DEGs were first identified in the GSE104580 cohort and overlapped with existing HRGs genes in the database. Finally, identifying 33 HRGTs that may have played a key role in TACE refractoriness. Patients were classified into different subtypes according to the expression of prognostic HRGTs and DEGs. The ESTIMATE algorithm showed that subtype A had a higher immune score, and subtype A had higher mRNA expression of most HLAs and ICIs.

ICIs therapy has been shown to be a highly effective agent for the treatment of HCC(Graham and Unger, 2018). However, it is unclear how to identify those who may benefit most from ICIs therapy. Hypoxia promotes tumor progression in different ways, including proliferation, metabolism, angiogenesis and migration, and improves resistance to ICIs(Bao and Wong, 2021). In addition, many factors, especially in TME, can influence the effectiveness of ICIs(Zhang and Zhang, 2020). The TME is a complex and integral component of cancer, containing tumor cells, stromal cells, inflammatory cells, fibroblasts, metabolites and cytokines. To investigate the value of risk subtypes in TME status and immunotherapy, multiple algorithms were used simultaneously in the immune cell analysis to estimate the immune cell infiltration score in different samples. Correlation analysis showed that as the risk score increased, the infiltration fraction of killer immune cells, such as CD4+ T and CD8+ T cell, also increased (Li et al., 2021). And the high-risk group had more active TME. Corresponding mRNA expression was higher in the high-risk subtype in HLA and ICIs analysis. For drug resistance, our study suggested that the high-risk group may have a better response to immunotherapy. We used the pRRophetic algorithm to predict drug sensitivity of HCC patients in different risk groups. The results showed that drugs targeting cell cycle, PI3K/mTOR signaling, WNT signaling, and RTK signaling were more effective in high-risk patients. Importantly, the IPS results demonstrated that the high-risk group seems to have more immunogenic phenotypes.

For the four HRGTs involved in the risk signature, we found that all of them were associated with tumor immunity. Secretory phosphorylated protein 1 (*SPP1*) is a secreted multifunctional phosphorylated protein that specifically binds and activates matrix metalloproteinases (MMPs) in cancer (Chen et al., 2019a). Its main biological functions are involved in immune response, biomineralization and tissue remodeling. *SPP1* has also been implicated in cell growth, proliferation, migration, apoptosis and chemotaxis. Previous studies have demonstrated that *SPP1* is overexpressed in a variety of cancers and can be used to predict chemotherapy prognosis, such as ovarian cancer (Zeng et al., 2018), glioblastoma (Kijewska et al., 2017), HCC (Liu et al., 2022) and gastric cancer (Chen et al., 2018a). *MMP1* is a member of a family of zinc-dependent endopeptidases involved in wound healing, inflammation, cancer and angiogenic remodeling of the extracellular matrix (ECM) (Chen et al., 2019b). It has been shown to be closely associated with migration and invasion in many cancers. mmp1 promotes cell cycle acceleration in cancer cells by activating the cdc25a/CDK4-cyclin D1 and p21/cdc2-cyclin B1 complexes (Yu et al., 2021). A newly discovered mechanism of *MMP1* in tumor promotion is by activating *PAR1* to cleave downstream oncogenic signaling pathways (Huang et al., 2018). This is expected to be a promising strategy to address the TACE refractoriness. *TPX2* has been identified as an oncogenic factor in a variety of cancers. For example, upregulated expression of *TPX2* enhances breast cancer metastasis by mediating *MMP2* and *MMP9* expression (Tan et al., 2019). In addition, *TPX2* can inhibit cell proliferation and enhance apoptosis by blocking the PI3k/AKT/p21 pathway and activating the p53 pathway in breast cancer (Chen et al., 2018b). It has been shown that *TPX2* is highly expressed in HCC tissues. *TPX2* expression is associated with the infiltration status of immune cells in HCC involving B, $CD4^+$T and $CD8^+$ T cells, neutrophils, macrophages and DCs(Zhu et al., 2020). In addition, CDK5-mediated stabilization of *TPX2* promotes HCC tumorigenesis (Wang et al., 2019). Clearly, these studies suggest that *TPX2* is an unfavorable marker for HCC and holds promise as a therapeutic target for TACE refractoriness. *CTSO* is a cysteine protease that has been shown to have both extracellular and intracellular functions. This class of proteases mediates intracellular protein catabolism and selectively activates extracellular protein degradation, macrophage function and bone resorption (Shi et al., 1995). The role in cancer therapeutic resistance is an emerging area of interest.

In our study, different hypoxic patterns present different biological processes, signaling pathways and immune features. Based on the parameters of the risk model, we determined that high-risk patients have a more active immune microenvironment, and HRGs are potential biomarkers for TACE-refractory patients. Especially, it may be an independent prognostic factor for HCC patients. However, our study has some limitations. Firstly, we should use advanced artificial intelligence models rather than traditional machine learning models such as Random Forest (RF) or LASSO models. However, for clinical applications, machine learning models with coefficients may be more helpful to clinicians. The clinician can calculate the survival risk of each patient from the mRNA expression and coefficient, however, more advanced deep learning models are a 'black box'. Moreover, due to the limitation of laboratory conditions, we have no more time to conduct *in vivo* or vitro experiments, and we will validate the mechanism of four hub genes in TACE-refractory patients in the future. In conclusion, our study will provide a novel

immunological perspective for the development of treatment options for TACE-refractory HCC.

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

SF and XC conceived and designed the study. SF was responsible for materials. XC, JL, and LF drafted the article. XW and YL revised the article critically. All authors had final approval of the submitted versions.

Funding

This work was partially supported by funding from the National Natural Science Foundation of China (No. 81672944), Natural Science Foundation of Jiangsu Province (Nos. BK20211296 and BK20200843), the Priority Academic Program Development of Jiangsu Higher Education Institutions (Integration of Chinese and Western Medicine), and Innovation Team of Six Talent Peaks Project in Jiangsu Q12 Province (No. WSN-062).

Acknowledgments

We are grateful to Feng from southeast university for bioinformatics analysis, and computer resource from state key laboratory of bioelectronics, school of biological science and medical engineering, southeast university.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.1011033/full#supplementary-material

SUPPLEMENTARY TABLE S1
Treatment details for HCC patients in GSE14520.

References

Bao, M. H., and Wong, C. C. (2021). Hypoxia, metabolic reprogramming, and drug resistance in liver cancer. *Cells* 10 (7), 1715. PMID: 34359884; PMCID: PMC8304710. doi:10.3390/cells10071715

Chakraborty, S., Hosen, M. I., Ahmed, M., and Shekhar, H. U. (2018). Onco-multi-OMICS approach: A new frontier in cancer research. *Biomed. Res. Int.* 2018, 9836256. PMID: 30402498; PMCID: PMC6192166. doi:10.1155/2018/9836256

Chang, Y., Jeong, S. W., Young Jang, J., and Jae Kim, Y. (2020). Recent updates of transarterial chemoembolilzation in hepatocellular carcinoma. *Int. J. Mol. Sci.* 21 (21), 8165. PMID: 33142892; PMCID: PMC7662786. doi:10.3390/ijms21218165

Chen, L. Z., He, C. Y., Su, X., Peng, J. L., Chen, D. L., Ye, Z., et al. (2018). SPP1 rs4754 and its epistatic interactions with SPARC polymorphisms in gastric cancer susceptibility. *Gene* 640, 43–50. Epub 2017 Sep 28. PMID: 28962925. doi:10.1016/j.gene.2017.09.053

Chen, M., Zhang, H., Zhang, G., Zhong, A., Ma, Q., Kai, J., et al. (2018). Targeting TPX2 suppresses proliferation and promotes apoptosis via repression of the PI3k/AKT/P21 signaling pathway and activation of p53 pathway in breast cancer. *Biochem. Biophys. Res. Commun.* 507 (1-4), 74–82. Epub 2018 Nov 16. PMID: 30454896. doi:10.1016/j.bbrc.2018.10.164

Chen, X., Xiong, D., Ye, L., Yang, H., Mei, S., Wu, J., et al. (2019). SPP1 inhibition improves the cisplatin chemo-sensitivity of cervical cancer cell lines. *Cancer Chemother. Pharmacol.* 83 (4), 603–613. Epub 2019 Jan 9. PMID: 30627777. doi:10.1007/s00280-018-3759-5

Chen, Y., Peng, S., Cen, H., Lin, Y., Huang, C., Chen, Y., et al. (2019). MicroRNA hsa-miR-623 directly suppresses MMP1 and attenuates IL-8-induced metastasis in pancreatic cancer. *Int. J. Oncol.* 55 (1), 142–156. PMID: 31115512; PMCID: PMC6561617. doi:10.3892/ijo.2019.4803

Choi, J., Lee, D., Shim, J. H., Kim, K. M., Lim, Y. S., Lee, Y. S., et al. (2020). Evaluation of transarterial chemoembolization refractoriness in patients with hepatocellular carcinoma. *PLoS One* 15 (3), e0229696. PMID: 32130270; PMCID: PMC7055892. doi:10.1371/journal.pone.0229696

Feng, S., Xia, T., Ge, Y., Zhang, K., Ji, X., Luo, S., et al. (2022). Computed tomography imaging-based radiogenomics analysis reveals hypoxia patterns and immunological characteristics in ovarian cancer. *Front. Immunol.* 13, 868067. PMID: 35418998; PMCID: PMC8995567. doi:10.3389/fimmu.2022.868067

Feng, S., Xu, Y., Dai, Z., Yin, H., Zhang, K., and Shen, Y. (2022). Integrative analysis from multicenter studies identifies a WGCNA-derived cancer-associated fibroblast signature for ovarian cancer. *Front. Immunol.* 13, 951582. PMID: 35874760; PMCID: PMC9304893. doi:10.3389/fimmu.2022.951582

Feng, S., Yin, H., Zhang, K., Shan, M., Ji, X., Luo, S., et al. (2022). Integrated clinical characteristics and omics analysis identifies a ferroptosis and iron-metabolism-related lncRNA signature for predicting prognosis and therapeutic responses in ovarian cancer. *J. Ovarian Res.* 15 (1), 10. PMID: 35057848; PMCID: PMC8772079. doi:10.1186/s13048-022-00944-y

Graham, K., and Unger, E. (2018). Overcoming tumor hypoxia as a barrier to radiotherapy, chemotherapy and immunotherapy in cancer treatment. *Int. J. Nanomedicine* 13, 6049–6058. PMID: 30323592; PMCID: PMC6177375. doi:10.2147/IJN.S140462

Gray, L., Conger, A., Ebert, M., Hornsey, S., and Scott, O. (1953). The concentration of oxygen dissolved in tissues at the time of irradiation as a

factor in radiotherapy. *Br. J. Radiol.* 26 (312), 638–648. doi:10.1259/0007-1285-26-312-638

He, Q., Yang, J., and Jin, Y. (2022). Development and validation of TACE refractoriness-related diagnostic and prognostic scores and characterization of tumor microenvironment infiltration in hepatocellular carcinoma. *Front. Immunol.* 13, 869993. PMID: 35493518; PMCID: PMC9043752. doi:10.3389/fimmu.2022.869993

Huang, C., Li, Y., Guo, Y., Zhang, Z., Lian, G., Chen, Y., et al. (2018). MMP1/PAR1/SP/NK1R paracrine loop modulates early perineural invasion of pancreatic cancer cells. *Theranostics* 8 (11), 3074–3086. PMID: 29896303; PMCID: PMC5996366. doi:10.7150/thno.24281

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. Epub 2018 Aug 20. PMID: 30127393; PMCID: PMC6487502. doi:10.1038/s41591-018-0136-1

Kenji, J., Hyodo, I., Tanimizu, M., Tanada, M., Nishikawa, Y., Hosokawa, Y., et al. (1997). Total necrosis of hepatocellular carcinoma with a combination therapy of arterial infusion of chemotherapeutic lipiodol and transcatheter arterial embolization: Report of 14 cases. *Semin. Oncol* 24 (2), S6-S–71. PMID: 9151920.

Kijewska, M., Kocyk, M., Kloss, M., Stepniak, K., Korwek, Z., Polakowska, R., et al. (2017). The embryonic type of SPP1 transcriptional regulation is re-activated in glioblastoma. *Oncotarget* 8 (10), 16340–16355. PMID: 28030801; PMCID: PMC5369967. doi:10.18632/oncotarget.14092

Kim, S. S., Cho, H. J., Won, J. H., Bae, J. I., Kang, D. R., Lee, J. D., et al. (2015). Interleukin-8 level as a prognostic marker in patients with Hepatitis B virus-associated hepatocellular carcinoma treated with transarterial chemoembolization. *Cytokine* 76 (2), 449–457. Epub 2015 Jul 8. PMID: 26163999. doi:10.1016/j.cyto.2015.07.001

Li, C., Teixeira, A. F., Zhu, H. J., and Ten Dijke, P. (2021). Cancer associated-fibroblast-derived exosomes in cancer progression. *Mol. Cancer* 20 (1), 154. PMID: 34852849; PMCID: PMC863844. doi:10.1186/s12943-021-01463-y

Li, X., Feng, G. S., Zheng, C. S., Zhuo, C. K., and Liu, X. (2004). Expression of plasma vascular endothelial growth factor in patients with hepatocellular carcinoma and effect of transcatheter arterial chemoembolization therapy on plasma vascular endothelial growth factor level. *World J. Gastroenterol.* 10 (19), 2878–2882. PMID: 15334691; PMCID: PMC4572123. doi:10.3748/wjg.v10.i19.2878

Liu, K., Min, X. L., Peng, J., Yang, K., Yang, L., and Zhang, X. M. (2016). The changes of HIF-1α and VEGF expression after TACE in patients with hepatocellular carcinoma. *J. Clin. Med. Res.* 8 (4), 297–302. PMID: 26985249; PMCID: PMC4780492. doi:10.14740/jocmr2496w

Liu, L., Zhang, R., Deng, J., Dai, X., Zhu, X., Fu, Q., et al. (2022). Construction of TME and Identification of crosstalk between malignant cells and macrophages by SPP1 in hepatocellular carcinoma. *Cancer Immunol. Immunother.* 71 (1), 121–136. Epub 2021 May 24. PMID: 34028567. doi:10.1007/s00262-021-02967-8

Morise, Z., Kawabe, N., Tomishige, H., Nagata, H., Kawase, J., Arakawa, S., et al. (2014). Recent advances in the surgical treatment of hepatocellular carcinoma. *World J. Gastroenterol.* 20 (39), 14381–14392. PMID: 25339825; PMCID: PMC4202367. doi:10.3748/wjg.v20.i39.14381

Shi, G. P., Chapman, H. A., Bhairi, S. M., DeLeeuw, C., Reddy, V. Y., and Weiss, S. J. (1995). Molecular cloning of human cathepsin O, a novel endoproteinase and homologue of rabbit OC2. *FEBS Lett.* 357 (2), 129–134. PMID: 7805878. doi:10.1016/0014-5793(94)01349-6

Sun, D., Wang, J., Han, Y., Dong, X., Ge, J., Zheng, R., et al. (2021). TISCH: A comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. *Nucleic Acids Res.* 49 (D1), D1420–D1430. PMID: 33179754; PMCID: PMC7778907. doi:10.1093/nar/gkaa1020

Tan, G. Z., Li, M., Tan, X., Shi, M. L., and Mou, K. (2019). MiR-491 suppresses migration and invasion via directly targeting TPX2 in breast cancer. *Eur. Rev. Med. Pharmacol. Sci.* 23 (22), 9996–10004. doi:10.26355/eurrev_201911_19566

Villanueva, A. (2019). Hepatocellular carcinoma. *N. Engl. J. Med.* 380 (15), 1450–1462. doi:10.1056/NEJMra1713263

Wang, F., Zhao, W., Gao, Y., Zhou, J., Li, H., Zhang, G., et al. (2019). CDK5-mediated phosphorylation and stabilization of TPX2 promotes hepatocellular tumorigenesis. *J. Exp. Clin. Cancer Res.* 38 (1), 286. PMID: 31272499; PMCID: PMC6610961. doi:10.1186/s13046-019-1297-6

Wu, J., Zhao, W., Zhou, B., Su, Z., Gu, X., Zhou, Z., et al. (2018). TSNAdb: A database for tumor-specific neoantigens from immunogenomics data analysis. *Genomics Proteomics Bioinforma.* 16 (4), 276–282. Epub 2018 Sep 15. PMID: 30223042; PMCID: PMC6203688. doi:10.1016/j.gpb.2018.06.003

Ye, F., Liang, Y., Cheng, Z., Liu, Y., Hu, J., Li, W., et al. (2022). Immunological characteristics of alternative splicing profiles related to prognosis in bladder cancer. *Front. Immunol.* 13, 911902. PMID: 35769470; PMCID: PMC9234272. doi:10.3389/fimmu.2022.911902

Yu, J., Xu, Z., Guo, J., Yang, K., Zheng, J., and Sun, X. (2021). Tumor-associated macrophages (TAMs) depend on MMP1 for their cancer-promoting role. *Cell Death Discov.* 7 (1), 343. PMID: 34753916; PMCID: PMC8578434. doi:10.1038/s41420-021-00730-7

Zarogoulidis, P., Katsikogianni, F., Tsiouda, T., Sakkas, A., Katsikogiannis, N., and Zarogoulidis, K. (2014). Interleukin-8 and interleukin-17 for cancer. *Cancer Invest.* 32 (5), 197–205. Epub 2014 Mar 26. PMID: 24669909. doi:10.3109/07357907.2014.898156

Zeng, B., Zhou, M., Wu, H., and Xiong, Z. (2018). SPP1 promotes ovarian cancer progression via Integrin β1/FAK/AKT signaling pathway. *Onco. Targets. Ther.* 11, 1333–1343. PMID: 29559792; PMCID: PMC5856063. doi:10.2147/OTT.S154215

Zhang, B., Tang, B., Gao, J., Li, J., Kong, L., and Qin, L. (2020). A hypoxia-related signature for clinically predicting diagnosis, prognosis and immune microenvironment of hepatocellular carcinoma patients. *J. Transl. Med.* 18 (1), 342. PMID: 32887635; PMCID: PMC7487492. doi:10.1186/s12967-020-02492-9

Zhang, Y., and Zhang, Z. (2020). The history and advances in cancer immunotherapy: Understanding the characteristics of tumor-infiltrating immune cells and their therapeutic implications. *Cell Mol. Immunol.* 17 (8), 807–821. Epub 2020 Jul 1. PMID: 32612154; PMCID: PMC7395159. doi:10.1038/s41423-020-0488-6

Zheng, R., Qu, C., Zhang, S., Zeng, H., Sun, K., Gu, X., et al. (2018). Liver cancer incidence and mortality in China: Temporal trends and projections to 2030. *Chin. J. Cancer Res.* 30 (6), 571–579. PMID: 30700925; PMCID: PMC6328503. doi:10.21147/j.issn.1000-9604.2018.06.01

Zhu, H., Liu, J., Feng, J., Zhang, Q., Bian, T., Li, X., et al. (2020). Overexpression of TPX2 predicts poor clinical outcome and is associated with immune infiltration in hepatic cell cancer. *Med. Baltim.* 99 (49), e23554. PMID: 33285774; PMCID: PMC7717782. doi:10.1097/MD.0000000000023554

Zhu, Y., Feng, S., Song, Z., Wang, Z., and Chen, G. (2022). Identification of immunological characteristics and immune subtypes based on single-sample gene set enrichment analysis algorithm in lower-grade glioma. *Front. Genet.* 13, 894865. PMID: 35646050; PMCID: PMC9136245. doi:10.3389/fgene.2022.894865

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 12 October 2022
DOI 10.3389/fphar.2022.1030766

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Congxiao Wu,
Shenzhen Qianhai Taikang hospital, China
Shilong Liu,
Harbin Medical University Cancer Hospital, China

*CORRESPONDENCE
Hanchao Zhang,
zhanghanchao@cdu.edu.cn

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 29 August 2022
ACCEPTED 28 September 2022
PUBLISHED 12 October 2022

CITATION
Ye X, Wang R, Yu X, Wang Z, Hu H and Zhang H (2022), $m^6A/ m^1A /m^5C/m^7G$-related methylation modification patterns and immune characterization in prostate cancer.
Front. Pharmacol. 13:1030766.
doi: 10.3389/fphar.2022.1030766

$m^6A/ m^1A /m^5C/m^7G$-related methylation modification patterns and immune characterization in prostate cancer

Xin Ye[1], Ruyi Wang[2], Xiaoqian Yu[3], Zili Wang[2], Haifeng Hu[2] and Hanchao Zhang[2,4]*

[1]Department of Urology, Institute of Urology, West China Hospital of Sichuan University, Chengdu, China, [2]Department of Urology, The Affilated Hospital and Clinical Medical College of Chengdu University, Chengdu, China, [3]Molecular Medicine Research Center and National Clinical Research Center for Geriatrics, West China Hospital, and State Key Laboratory of Biotherapy, Sichuan University, Chengdu, China, [4]Medical College of Soochow University, Suzhou, China

Methylation has a close relationship with immune reactions, metastasis, and cancer cell growth. Additionally, RNA methylation-related proteins have emerged as potential cancer therapeutic targets. The connection between the tumor microenvironment (TME) and methylation-related genes (MRGs) remains unclear. We explored the expression patterns of the MRGs in the genome and transcriptional fields of 796 prostate cancer (PCa) samples using two separate data sets. We identified a relationship between patient clinicopathological characteristics, prognosis, TME cell infiltrating qualities, and different MRG changes, as well as the identification of two distinct molecular groupings. Then, we formed an MRGs model to predict overall survival (OS), and we tested the accuracy of the model in patients with PCa. In addition, we developed a very accurate nomogram to improve the MRG model's clinical applicability. The low-risk group had fewer tumor mutational burden (TMB), greater tumor immune dysfunction and exclusion (TIDE) ratings, fewer mutant genes, and better OS prospects. We discuss how MGRs may affect the prognosis, clinically important traits, TME, and immunotherapy responsiveness in PCa. In order to get a better understanding of MRGs in PCa, we could further explore the prognosis and create more effective immunotherapy regimens to open new avenues.

KEYWORDS

Prostate cancer, Methylation modification, Tumor microenvironment, Molecular subtype, Prognostic model

Introduction

Prostate cancer (PCa) is the most frequent cancer diagnosis in men. Notably, PCa is the second most common neoplasm in senior men and the fifth leading cause of cancer-related mortality globally, accounting for 15% of all new tumor-related cases (Vietri et al., 2021). Most instances progress slowly and pose no danger to life. However, despite recent improvements, PCa still poses a serious medical challenge for the men affected. Therefore, finding novel prognostic indicators is essential for creating efficient treatment plans and enhancing PCa patients' prognoses (Zhao et al., 2020).

Numerous biological processes, including cell differentiation, sex determination, stress response, and others, are known to be impacted by RNA methylation and its connected downstream signaling cascades (Menezo et al., 2020). RNA modification disorders have been linked to a wide range of cancers, including PCa (Haruehanroengra et al., 2020). As the third layer of epigenetics, more than 170 RNA modifications have been identified (Haruehanroengra et al., 2020). N6-methyladenosine (m^6A), 5-methylcytosine (m^5C), N1-methyladenosin (m^1A), N7-methylguanosine (m^7G) are post-transcriptional modifications, which are abundant in most eukaryotic mRNAs and involved in almost all stages of the RNA life cycle, including RNA transcription, translation and degradation. They are found in mRNA, lncRNA, and miRNA. Additionally, it is essential for the growth and development of numerous immune system illnesses, including cancers and a wide range of other human pathogenic activities (Dai et al., 2021). The evidence for RNA modification pathways being dysregulated in human malignancies is growing, and these pathways may provide excellent targets for cancer therapy (Barbieri and Kouzarides, 2020).

Fluctuations in RNA methylation in cancer are known as promising targets for developing useful diagnostic, prognostic and predictive biomarkers (Koch et al., 2018). It is also exciting to note that methylation has been connected to antitumor immunity in cancer immunotherapy (B. Yang et al., 2021). Besides necroptosis, methylation is also an important cellular response that controls the initiation, progression, and metastasis of cancer. Nevertheless, the role of some methylation regulators in the prognosis and possible molecular mechanisms of PCa is not well understood (B. Yang et al., 2021). Studying methylation landscapes can help predict the prognosis of PCa, according to Wen-Juan Li et al. (W. J. Li et al., 2021). A study identified 8 methylation-based biomarkers (cg04633600, cg05219445, cg05796128, cg10834205, cg16736826, cg23523811, cg23881697, cg24755931) which were useful for aggressively detecting PCa (Pu et al., 2021). To increase PCa cell survival and docetaxel resistance, SPOP mutations will upregulate the formation of stress particles (Shi et al., 2019). An invasive tumor is more likely to form in PCa with TP53 mutation (Maxwell et al., 2022). all of which are strongly methylation-deregulated and closely linked to prognosis. There are a few studies on the relationship between methylation and PCa, so we need to further study the fact that it plays a significant role in carcinogenesis and anticancer mechanisms.

Immunological checkpoint blocking, or immunotherapy (ICB, PD-1/L1 and CTLA-4), has shown astounding clinical success in a small minority of patients with long-term responses (Kalbasi et al., 2020). However, a large number of patients received little to no therapeutic benefit, which falls far short of satisfying a clinical need (M. Zhang et al., 2021a). It has only ever been assumed that the multi-step process of tumor formation alters the genetic and epigenetic makeup of tumor cells. But a large number of studies have shown us that the tumor microenvironment (TME) also has a significant role in the growth of the tumor (Vitale et al., 2019). Direct and indirect interactions between TME components can induce changes in biological behaviors such as immune tolerance (M. Zhang et al., 2021a). The MRG risk score for PCa was shown by Zhipeng Xu et al. colleagues to strongly correlate with immune infiltration (Xu et al., 2022). The decreased effectiveness of checkpoint inhibitors (CPIs) in advanced prostate cancer compared to other tumor types is likely largely due to an immunosuppressive tumor microenvironment (TME) and impaired cellular immunity (Bansal et al., 2021). The complexity and variability of the TME landscape should therefore be thoroughly parsed to identify various tumor immune phenotypes, which would also enhance the ability to predict and direct immunotherapeutic responsiveness (Hinshaw et al., 2019; Song et al., 2021). The discovery of very accurate biomarkers to gauge patients' reactions to immunotherapy will aid in the search for novel therapeutic targets (Ehrlich, 2019).

We are now able to fully examine the transcriptome, metabolome, proteome, and genome in order to investigate the biomarkers and carcinogenesis framework for the therapy and prognosis of cancer when we explore the rapid advancement of science and the development of the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases. We sought to determine MRG expression in PCa, prognostic importance, and putative regulatory axis. Our results may provide more information on the molecular processes and prognostic biomarkers of PCa.

Materials and methods

Data sources

From the TCGA (TCGA-PCa) and GEO (GSE65858 and GSE116918) databases, RNA-seq and clinicopathological data for PCa were retrieved (Supplementary Table S1). RNA-seq for PCa was converted to Transcripts Per Kilobase Million (TPM) values as previously mentioned and was taken into consideration to be equivalent to those for microarrays. After integrating two

datasets (TCGA-PCa and GSE65858), batch effects were eliminated using the "Combat" method. The subsequent analyses included 796 PCa patients because we excluded data from people whose OS information was lacking or less than 30 days.

Consensus clustering analysis of MRGs

These 84 MRGs are shown in Supplementary Table S2's details. Using "ConsensusClusterPlus", consensus unsupervised clustering analysis was utilized to divide patients into distinct molecular subgroups based on MRG expression. The following criteria were used to group these items: First, there was a fluid and progressive growth in the cumulative distribution function curve. Second, there was no tiny sample size in any group. Thirdly, although there was a drop in the inter-group correlation, the intra-group correlation rose. Gene set variation analysis (GSVA) was carried out to study MRG variation in biological processes.

Correlation between clinical features and prognosis molecular subtypes

Age, gender, TNM stage, and clinical stage were some of the patient's features. And to assess the two clusters identified by consensus clustering's clinical value, we looked at the connections between molecular subtypes, clinical features, and prognosis. In addition, Kaplan-Meier curves, generated by the "survival" and "survminer" R programs, were used to compare OS among different subtypes.

Relationship of molecular subtypes with TME

Additionally, the CIBERSORT algorithm was used to calculate the scores of 22 different human immune cell types for each PCa sample (Hao et al., 2019). We used single-sample Gene Set Enrichment Analysis (ssGSEA) to explore the levels of immune cell infiltration (Hwang et al., 2021). DEG identification and functional annotation of DEGs with the "limma" package in R, DEGs were discovered with a p-value of 0. 05 and |logFC| of 0. 585. We use the "clusterprofiler" package in R to perform functional enrichment analyses on the DEGs, allowing us to have a better analysis of the hidden functions of the methylation clusters in DEGs and discriminate between the enriched pathways and gene functions that go along with them.

Construction of the prognostic risk model

We used unsupervised clustering to classify patients into different subtypes (gene cluster A and gene cluster B) for further study. All patients with PCA were randomly divided into training and testing groups with a ratio of 0.7:0.3 to establish a prognostic model. The DEGs were used in univariate Cox regression analysis in order to identify the DEGs associated with PCa's OS. We employ the following procedures to calculate the risk score: Risk score is equal to (expi * coefi), where expi and coefi are the expression and risk coefficients of each gene, respectively. To lessen the possibility of over-fitting using prognostic DRGs, the LASSO Cox regression technique was temporarily used. In the two groups, the expression levels of genes connected to MRGs were examined. We divided patients into high-and low-risk score groups by the median of risk scores, and Kaplan-Meier analysis and receiver-operating characteristic (ROC) curves were used to assess the accuracy of risk scores. GSE116918 was applied as an external testing set to validate the model.

Construction of a nomogram scoring system

We use the nomogram calibration plot to plot the forecast value between 3-, 5-, and 8-year survival events and virtual observations. A variable in a nomogram scoring system that combines risk scores and clinical characteristics has a score, and the total score is the sum of all the individual scores (Iasonos et al., 2008).

Mutation, immunotherapy response and drug susceptibility analysis

It is investigated how the genes in the model relate to the 22 immune cells. The ESTIMATE algorithm was applied to assess the immune and stromal scores of each sample. The TCGA database generates mutation annotation formats to identify somatic mutations in various PCa sample groups. We determine the tumor burden mutation (TBM) score for each PCa patient across the two categories. We investigated the associations between tumor immune dysfunction and exclusion (TIDE) and different groups. We created the half-maximal semi-inhibitory concentration (IC50) values of a pRRophetic package of anti-tumor medications for PCa in order to examine the difference in the treatment impact of commonly used anti-tumor agents between the two groups.

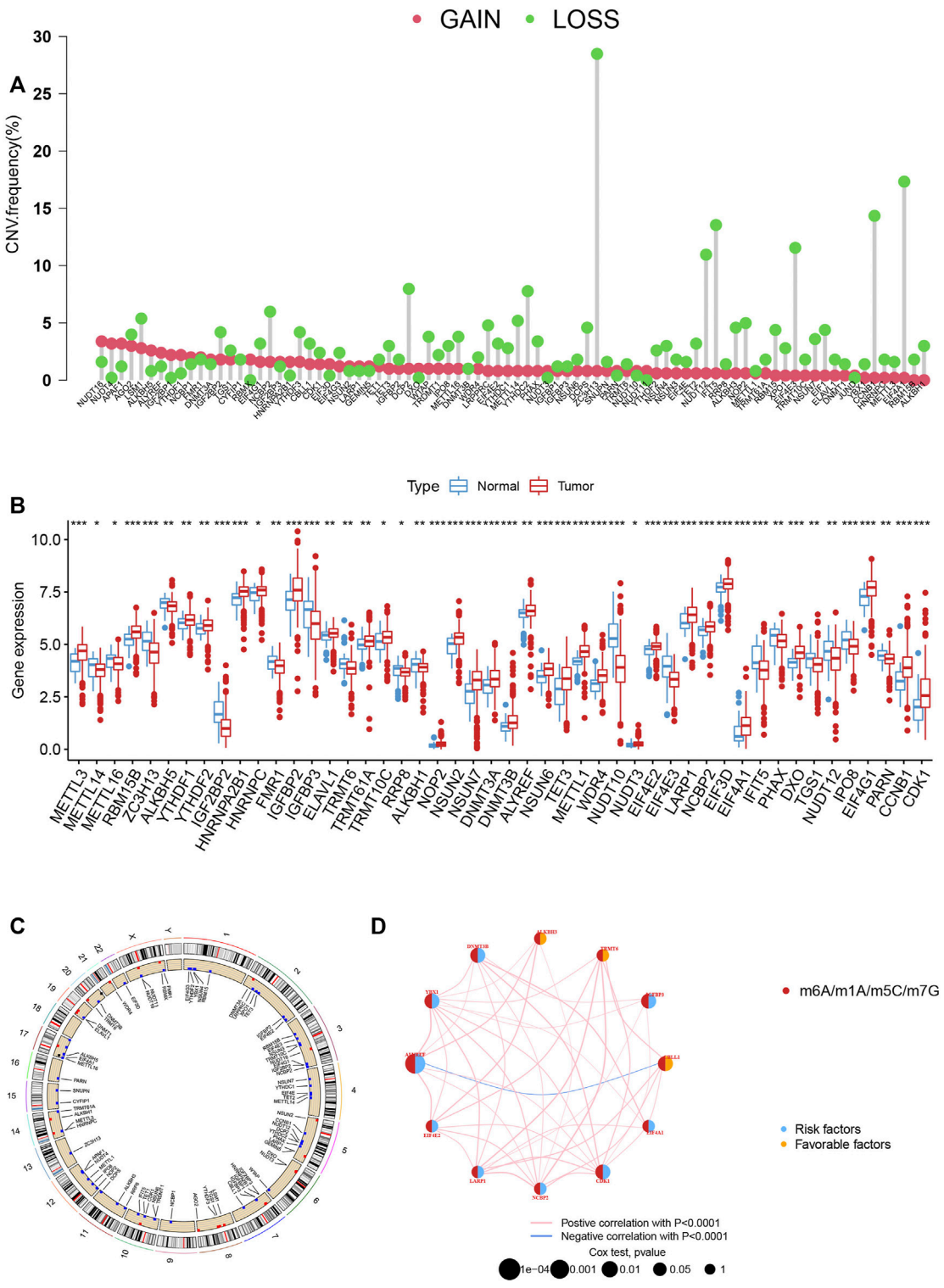


FIGURE 1
(A) The CNV of 84 MRGs. **(B)** Expression distributions of differentially expressed MRGs between normal and PCa tissues. **(C)** The positions of the CNV alterations on their respective chromosomes for these MRGs. **(D)** The overall group of MRG interactions, regulatory factor connectivity and value of prognosis in PCa patients was identified in the network.

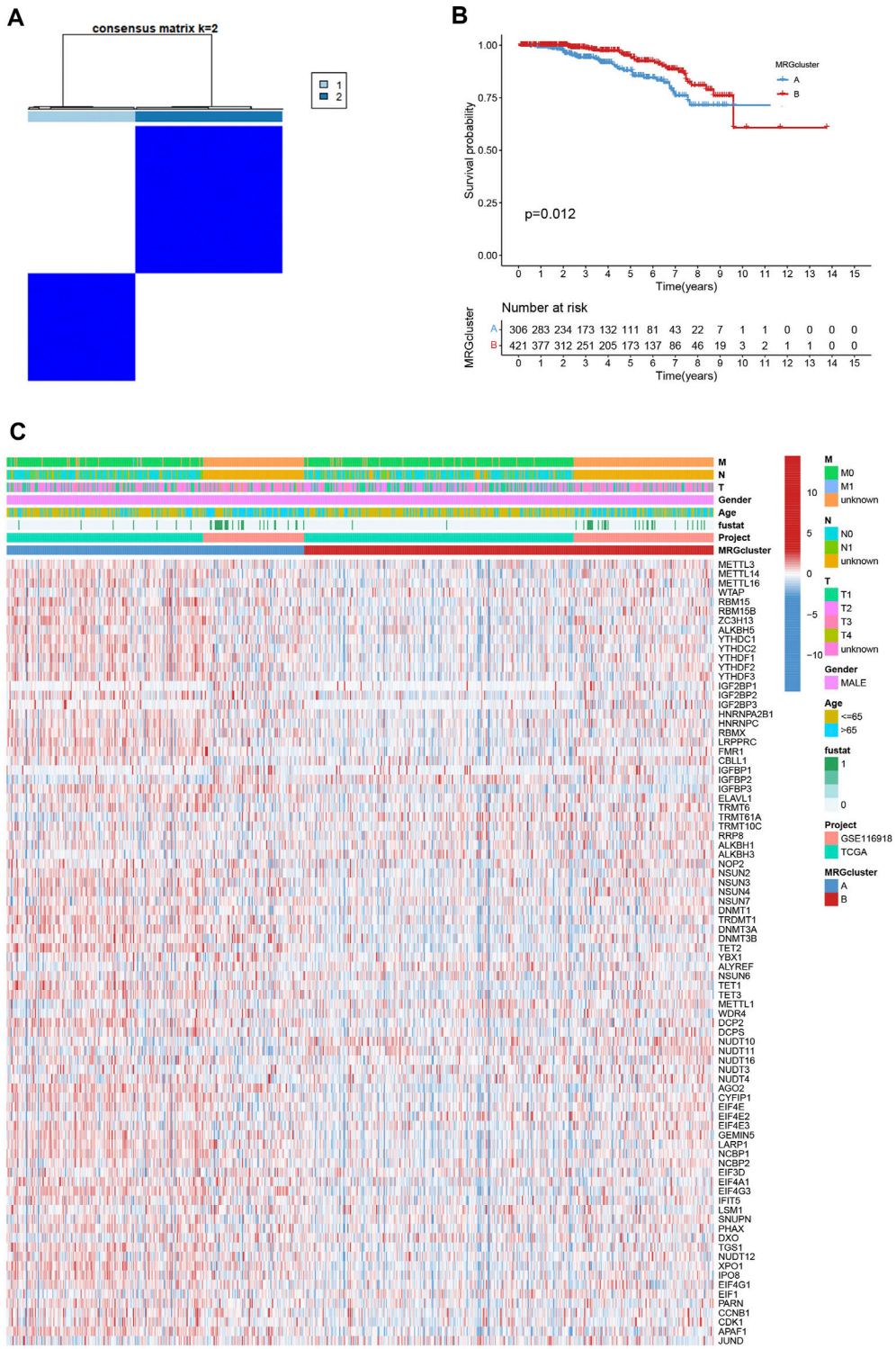


FIGURE 2
(A) Consensus matrix heatmap defining two MRG clusters (*k* = 2). **(B)** Kaplan-Meier curves indicated a shorter OS in patients with MRG cluster A than in patients with MRG cluster **(B) (C)** Differences in clinical features and MRG expression levels between the two MRG subtypes.

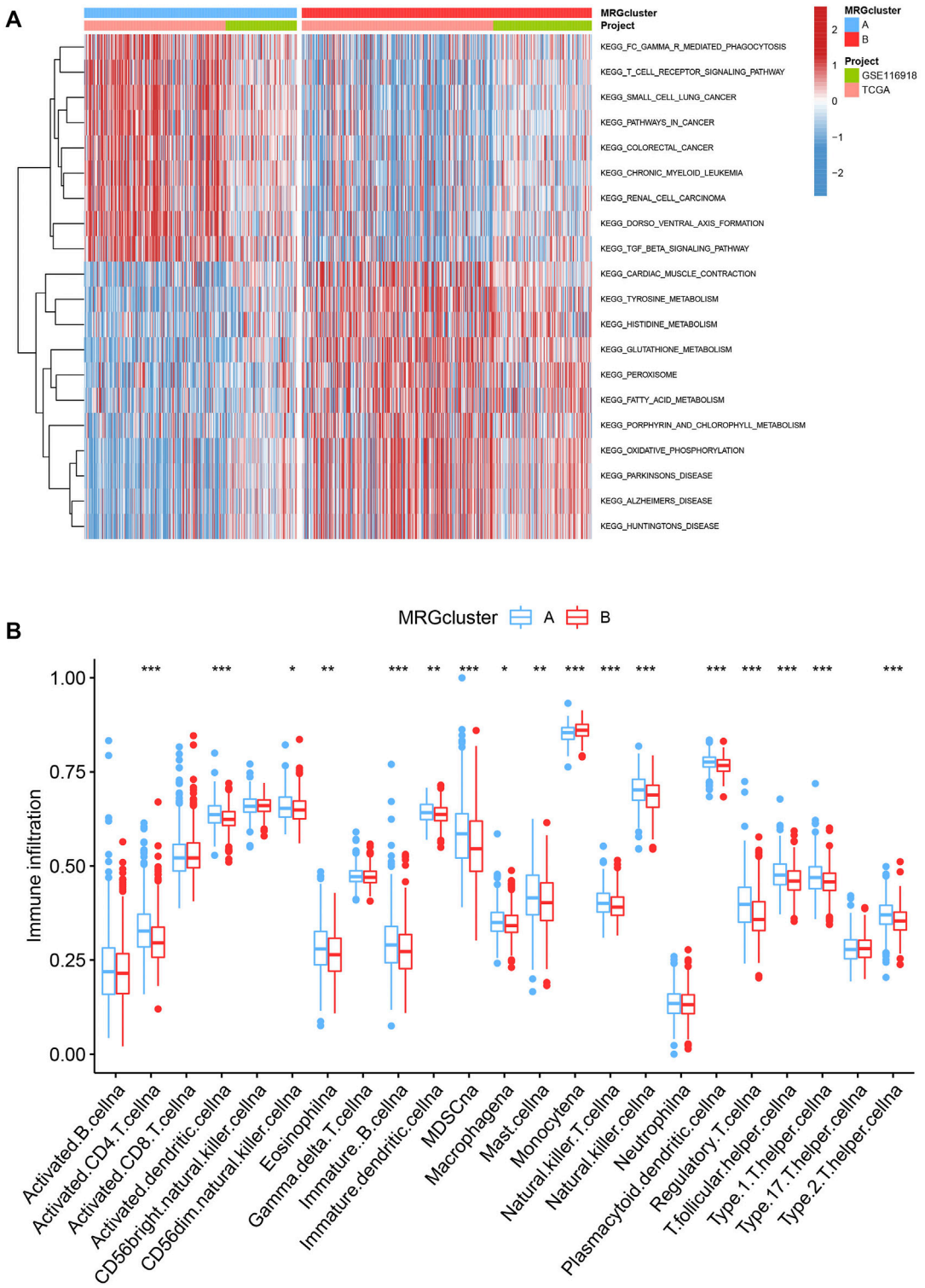


FIGURE 3
(A) Heatmap of GSVA enrichment analysis results. **(B)** Significant differences occurred among the two subtypes in the infiltration of some immune cells.

Results

Genetic and transcriptional alterations of MRGs in PCa

According to the analysis, we could see significant differences in the potential function of MRGs in PCa carcinogenesis with the expression levels and genetic landscape of MRGs between PCa and control samples. In this investigation, 84 MRGs were examined (Supplementary Table S2). We then looked into somatic copy number variation in the 84 MRGs and discovered that there were a number of common copy number alterations, including increases in general copy number variation (CNV) in NUDT16, NUDT4, APAF1, AGO2, LSM1, and ALKBH5, and decreases in CNV in ZC3H13, ELF4A1, CCNB1, IFIT5, ELF4E3, and NUDT12 (Figure 1A). MRGs with CNV loss were expressed at lower levels, such as ZC3H13, IFIT5, ELF4E3 and NUDT12 in PCa samples, when compared to those in normal PCa samples (Figure 1B), hinting that the mRNA expression of MRGs might be regulated by CNV. Figure 1C shows the locations of CNV alterations on their respective chromosomes in MRGs. DNA methylation factors could modulate gene expression (Nishiyama and Nakaanishi., 2021).

Identification of methylation-related subtypes

We picked 796 patients (TCGA and GSE116918) to explore the expression pattern of MRG involved in tumorigenesis for further analysis. The 12 prognostic MRGs were recognized by univariate Cox analysis. The prognostic MRG interactions, regulatory factor connectivity and value of methylation in PCa patients were identified in the methylation network (Figure 1D). Based on the 84 MRGs' expression profiles, we used a consensus clustering approach to classify the PCa patients. We classified the entire cohort as the best choice for MRG cluster A and B based on $k = 2$ (Figure 2A and Supplementary Figure S1). Patients in MRG Group B had a better OS, as hinted by the Kaplan-Meier curves ($p = 0.012$; Figure 2B). Furthermore, we demonstrate that MRG expression and clinical pathology characteristics are significantly different (Figure 2C).

Evaluation of TME

GSVA enrichment analysis showed that MRG cluster B and MRG cluster A were significantly different. One was in fc gamma r mediated phagocytosis, T cell receptor signaling pathway, small cell lung cancer, and pathways in cancer, while another was in huntingtons disease, alzheimers disease, parkinsons disease, and oxidsative phosphorylation (Figure 3A). We examine the relationships between the 22 human immune cell subsets and the two subtypes of each PCa sample by using the CIBERSORT method. There were important variations between the two subtypes in terms of the invasion of certain immune cells. (Figure 3B).

Classification of gene clusters

To investigate the underlying biological behaviour of each focal flash pattern, the R package "limma" was used to recognize 74 DEGs linked to MRG subtypes. These were then subjected to functional enrichment analysis (Figures 4A,B). These DEGs were widely distributed in biological processes and were associated with immunity (Figure 4A and Supplementary Table S3). KEGG analysis revealed an enrichment of immunological and cancer-related pathways, demonstrating the significance of methylation in the immune control of the TME (Figure 4B and Supplementary Table S4). By using univariate Cox regression analysis, 32 prognostic DEGs related to OS time were chosen from 74 DEGs ($p < 0.05$; Supplementary Table S5). In order to validate these regulatory mechanisms, consensus clustering techniques were utilized to share patients into two gene categories based on prognostic genes (Figure 5A and Supplementary Figure S2). According to Kaplan-Meier curves ($p < 0.001$; Figure 5B), patients with gene cluster B had the highest OS, which is obviously better than that of cluster A. The two gene subtypes' MRG expression showed significant variety, which was in line with our predictions (Figure 5C.) Additionally, a comparison of the clinicopathological characteristics of several gene subtypes revealed a substantial difference between clinical aspects and gene expression (Figure 5D).

Construction and validation of the prognostic risk model

We randomly grouped the patients into training and testing groups in a ratio of 0. 7: 0. 3 ("caret package" in R). To further narrow down the best prognostic signature, the prognostic DEGs were run through LASSO and multivariate Cox analysis (Figures 6A–C). The risk model was built using the following steps: risk score = (0.315* COL1A1) + (0.243* ASPN) + (-0.333* PHYHD1) + (-0.134* PCGEM1). A Sankey diagram was used to illustrate the relationship between the MRG cluster, gene cluster, risk groups, and survival status (Figure 6D). The risk score distributions for the two categories are shown in Figures 6E,F. We found that the expression of MRGs varied considerably between groups (Figure 6G).

The Kaplan-Meier analysis, expression profiles, pattern of survival status, and distribution of risk scores are shared in Figures 7A–C, which hints that patients in the low-risk category will live longer. The model's high sensitivity and

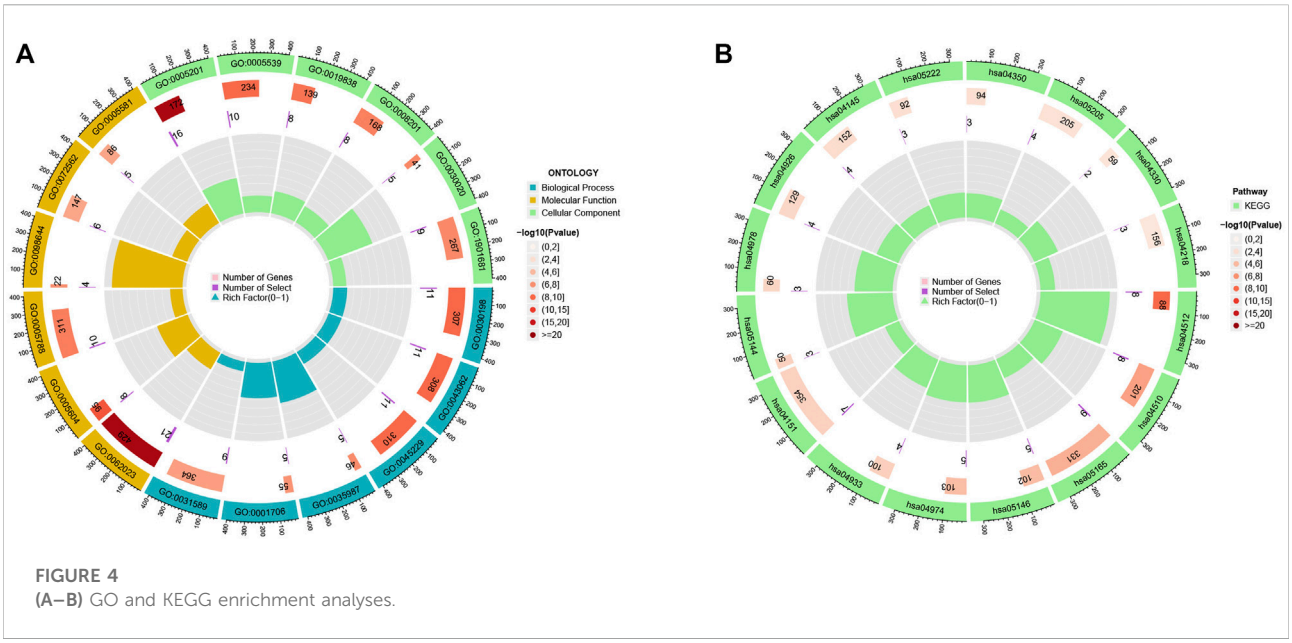


FIGURE 4
(A–B) GO and KEGG enrichment analyses.

specificity for predicting survival were demonstrated by the ROC curves, and the overall set's 8-year AUC value was 0. 759 (Figure 7D). In Supplementary Figures S3–S5, which provide the above analysis for the training, testing and external testing sets, the model's dependability is shown. Figure 7E was the nomogram that included the model and clinical characteristics.

Evaluation of TME

We also looked at the relationship between the number of immune cells and the four genes in the proposed model, and found that the majority of immune cells are obviously related to the four genes (Figure 8A). The low-risk score group was strongly correlated with a low immunological score, while the high-risk score group was linked to a high stromal score (Figure 8B).

Mutation, immunotherapy response and drug susceptibility analysis

We examined how the TCGA-PCa cohort's various risk score groups differed in the somatic mutation distribution. The top 10 mutant genes in the high- and low-risk categories were SPOP, TTN, TP53, KMT2D, FOXA1, MUC16, SYNE1, KMT2C, LRP1B, and SPTA1 (Figures 8C,D). Patients in the low-risk score group had considerably higher frequencies of SPOP mutations compared to those in the high-risk score group. Further, high TBM was connected with poor OS ($p < 0.001$; Figure 8E). The high-risk score group had lower TIDE scores, indicating that they might have responded better to immunotherapy (Figure 8F). Furthermore, by examining the IC50 of regularly used anticancer medicines, we found a significant difference between the two patient groups' susceptibility to the treatments. (Supplementary Figure S6).

Discussion

In vitro and *in vivo* tumor growth, invasion, migration, and the epithelial-mesenchymal transition of cancer cells are all influenced by dynamic RNA methylation and modification events, such as m^6A, m^1A, m^5C and m^7G (X. Y. Li et al., 2022; Traube et al., 2017). In addition to playing essential roles in various cancers and anticancer effects, modification events can also be used as prognostic indicators (Mahmoud and Ali, 2019). There are still several unanswered questions regarding the overall effect and the features of TME penetration adjusted by the effects of numerous MRGs (M. Li et al., 2021).

We identified two distinct molecular subgroups using 84 MRGs. And patients with subtype B had a better OS. The features of the TME varied obviously across the two subtypes. Variations in mRNA transcriptomes between different methylation subtypes were strongly linked with biological pathways involved in MRG and the immune system (Gu et al., 2021; X. Y. Li et al., 2022). We determined two gene subtypes relied on the DEGs between the two methylation subtypes. According to the data, MRGs may be utilized to predict PCa's clinical prognosis and responsiveness to treatment (Zhang et al., 2020). As a result, we discovered and validated the accurate prognostic

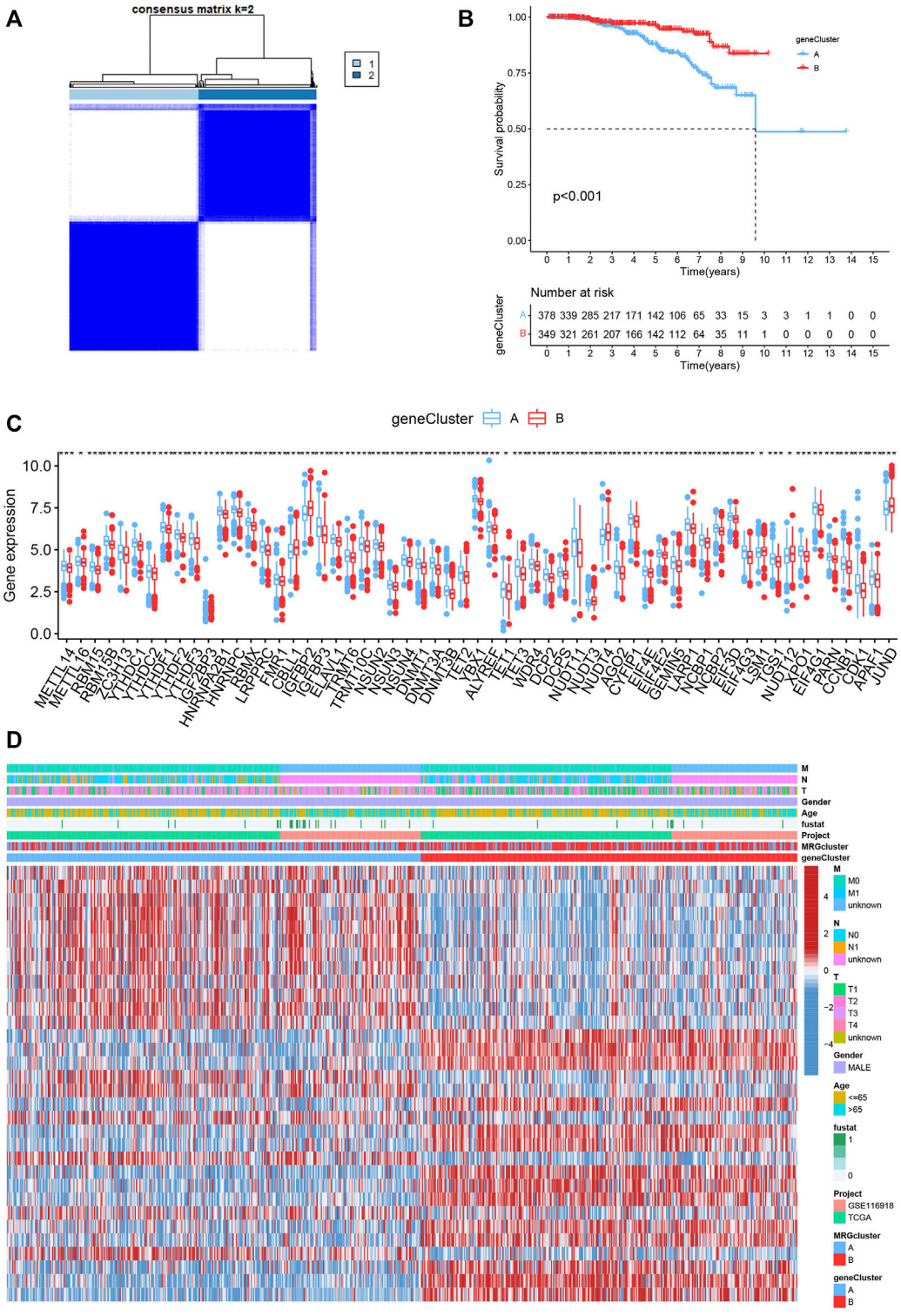


FIGURE 5
(A) Consensus matrix heatmap defining two MRG clusters (*k* = 2). **(B)** Kaplan-Meier curves indicated that patients with gene cluster B had higher OS. **(C)** The expression levels of MRGs in the two gene subtypes. **(D)** Differences in clinical features and MRG expression levels between the gene subtypes.

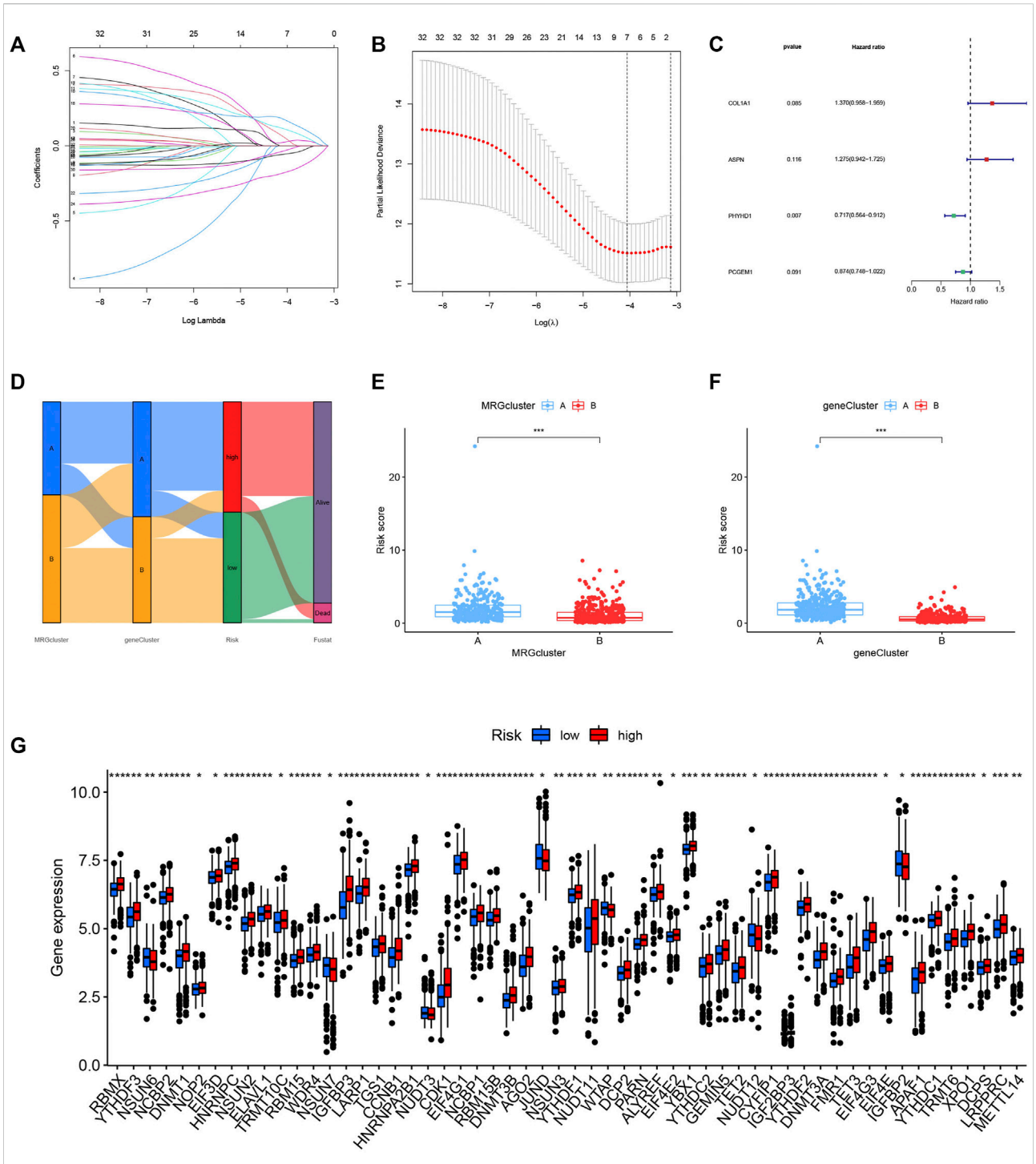


FIGURE 6
(A–C) The model was constructed by LASSO and multivariate Cox regression analysis. **(D)** The relationship between MRG cluster, gene cluster, risk groups, and survival status was visualized in a Sankey diagram. **(E–F)** The distribution of risk scores for the two subgroups. **(G)** The expression of ICIs-related genes was significantly different between groups.

MRG-score. Higher and lower MRG-scores were seen in immune activation- and inhibition-driven PCa patterns, respectively. Finally, we combined the risk score and tumor stage to produce a quantitative nomogram, which dramatically improved performance and made it simpler to utilize the risk score (Jeong et al., 2020).

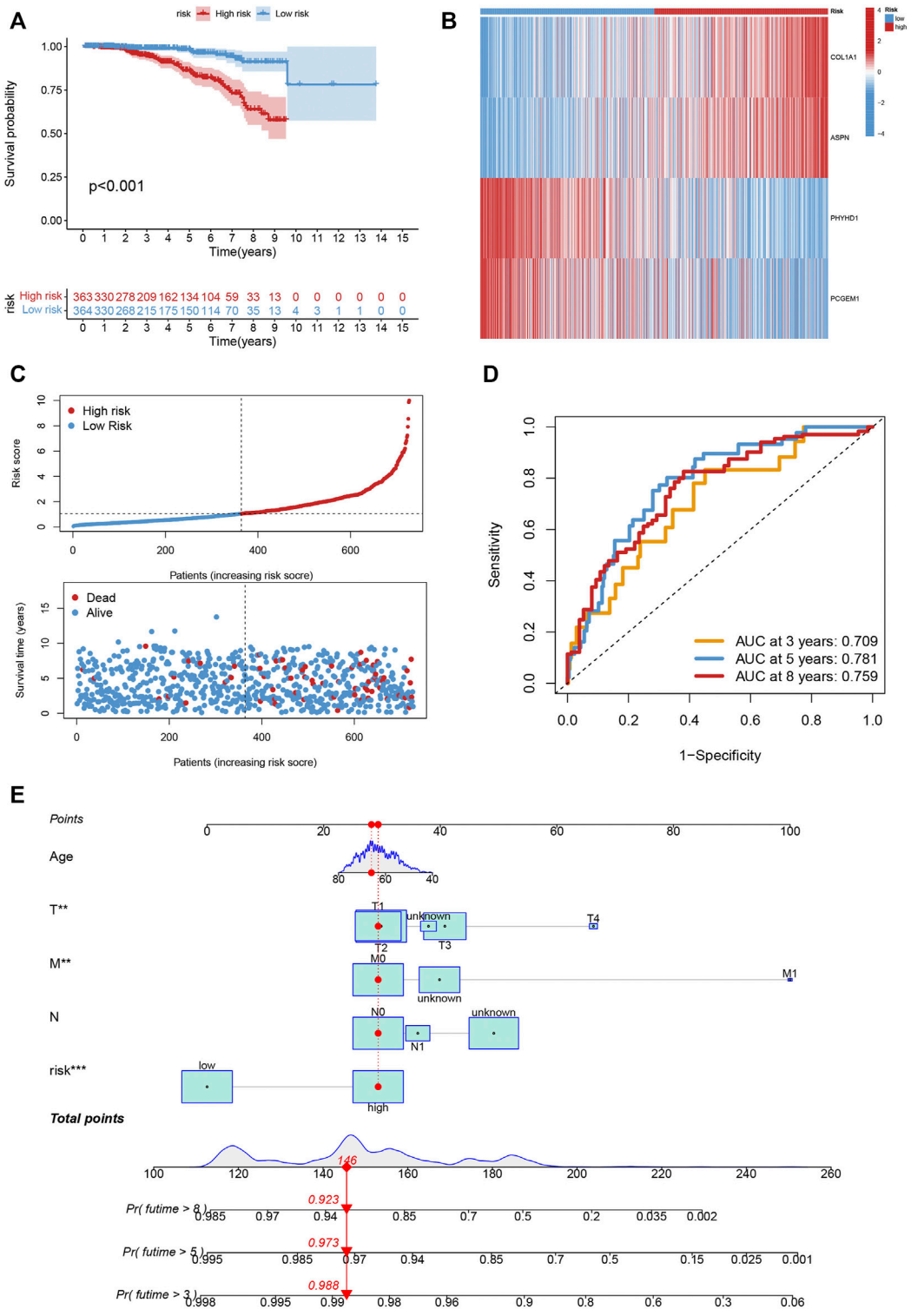


FIGURE 7
(A–C) The Kaplan-Meier analysis, expression profiles, pattern of survival status, and the distribution of risk scores in the entire cohort. **(D)** The ROC curves for the 3, 5, and 8-year AUC values in the entire cohort. **(E)** The nomogram containing the model and clinical features was reliable and sensitive for predicting survival in patients with PCa.

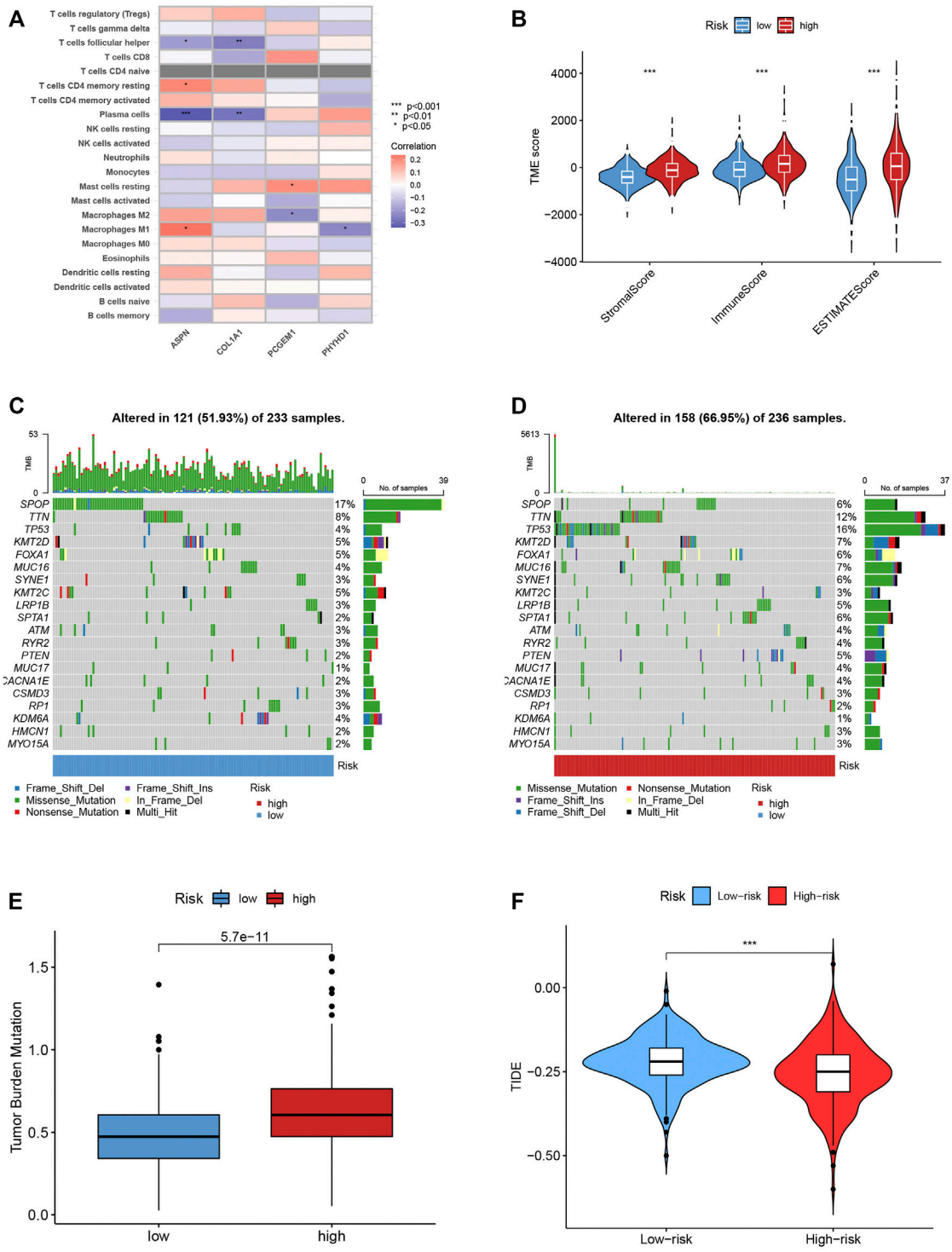


FIGURE 8
(A) The connection between the number of immune cells and the 4 genes in the model. **(B)** The high-risk scores were linked to a low stromal score, and the low-risk scores was highly correlated with a high immune score. **(C–D)** In the high- and low-risk groups, the top 10 mutant genes were SPOP, TTN, TP53, KMT2D, FOXA1, MUC16, SYNE1, KMT2C, LRP1B and SPTA1. **(E)** TBM score between different groups. **(F)** TIDE scores were lower in the high-risk score group, suggesting that the high-risk score group was more responsive to immunotherapy.

A growing body of research has established that MRG alteration played a significant role in the post-transcriptional modification of gene expression, which was strongly associated with tumor formation, maintenance, progression, and prognosis, thanks to advancements in detecting technology. As reported, high TET3 expression (m^5C-related gene) was related to poor prognosis of PCa (Yu et al., 2022). According to certain research, m^6A alteration significantly influences the stability of mRNA, which in turn contributes to PCa development (Du C et al., 2020). PCa bone metastases were related to high m^6A levels of NEAT1-1, and m^6A levels of NEAT1-1 were a reliable indicator of ultimate death (Wen et al., 2020). In recent years, m^7G has been thought to be actively implicated in cancer-related translation problems. The m^7G-score has been shown to be an independent measure of BCR-free survival in patients with PCa (Xin et al., 2022). Additionally, recent research has shown that RNA modification regulators may serve as biomarkers for cancer diagnosis and prognosis surveillance (Haruehanroengra et al., 2020). Nevertheless, a thorough examination of the prognostic significance and functional annotation of MRGs regulators in PCa is still lacking.

PCa patients' prognoses are poor. There were significant differences between patient subgroups in terms of TME, immunological checkpoints, CSC index, prognosis, mutation, and therapy susceptibility after standard therapy because of high levels of checkpoints, lymphocytes that infiltrate tumors, and tumor neoantigens (D. Li et al., 2022). Despite recent developments in immunotherapy, patients with PCa still experience heterogeneity in their results, underlining the important role of TME in the growth and development of PCa tumors (Yu et al., 2022). Immune cells, including granulocytes, lymphocytes, and macrophages, are important biological components of TME. These cells participate in a variety of immunological responses and behaviors, such as the inflammatory response that tumors trigger to help them survive (Schmitt and Greten., 2021). Additional data points to the TME having a significant impact on cancer development, progression, and therapeutic resistance (Cao et al., 2021; Martínez-Reyes and Chandel., 2021). Immune inhibition-driven methylation (subtype A) was associated with a higher risk score, whereas immune activation-driven methylation (subtype B) was related to a lower risk score. We discovered that the relative richness of 22 immune cells as well as the two molecular subtypes' differences in risk scores and TME traits were significantly different.

Various kinds of T cells are crucial components of the immune defense against PCa (K. Yang and Kaliies., 2021). Higher densities indicated a positive prognosis as tumor-infiltrating T cell densities in PCa samples were higher than those in normal tissues (Yu et al., 2022). The enhanced infiltration of activated memory $CD4^+$ and $CD8^+$ T cells as well as gamma delta T cells was seen in the subtype B and low risk score groups, indicating that they favourably contribute to the progression of PCa. A worse prognosis was associated with Treg infiltration, which blocks the immune system's anti-cancer response (Oh and Fong., 2021). This is in line with our observation that patients in the high-risk group and those with subtype B had more Tregs in the TME than those in the low-risk group. Recently, it was shown that B cells aid in the immune response (Fridman et al., 2022; Zhang et al., 2022).

Petitprez et al. believed that in soft-tissue sarcomas, the response to PD-1 inhibition was positively linked with B cell enrichment (Petitprez et al., 2020). Patients who responded to immune checkpoint blockade showed considerably higher levels of the B cell-related genes than those who did not, according to Helmink et al. (Fridman et al., 2022). Additionally, in PCa, tumor-infiltrating B lymphocytes were linked to a good prognosis (Horii et al., 2021). Patients with significant B cell infiltration in their metastatic PCa had prolonged overall survival and a significantly lower risk of the disease coming back (Engelhard et al., 2021). The results of this study demonstrated that B cells are not only incidental contributors to anti-cancer immunotherapy; rather, they present a novel immunotherapy target and may be a potent cancer-fighting tool. In our study, we found subtype B had considerably fewer naive B cells and higher MRG-score, which were associated with poorer overall survival (Franchina et al., 2018).

In this study, the expression levels of a part of immune cells were found to be obviously different in the risk model of MRGs. The stromal score, CD4 memory resting T cells, CD4 memory activated T cells, follicular helper T cells, M0 macrophages, M1 macrophages, and resting mast cells were linked with the risk score. This implies that PCa immune cell infiltration is related to the risk model created using MRGs (He et al., 2022). Our study shows that differentially expressed ASPN,COL1A1, PCGEM1 and PHYHD1 was associated with immune infiltration. The high-risk score group was related to a high stromal score, and the low-risk score group was closely associated with a high immune score. Pu Zhang et al. showed that while ASPN is overexpressed in PCa, a bad prognosis is predicted by excessively high ASPN expression and low expression of other genes, ASPN is independently associated with overall survival (OS) of patients (P. Zhang et al., 2021b). High expression of COL1A1 can predict the prognosis of cancer and is a reliable biomarker and therapeutic target (Ma et al., 2019; Geng et al., 2021). And many studies have shown that the high expression of PCGEM1 and PHYHD1 can promote the value-added migration and invasion of cancer, affecting prognosis (Jiang et al., 2019; Zhang et al., 2019; Liu et al., 2022). Our study

identified the involvement of MRGs and constructed a risk model for PCa. However, this must be confirmed using additional clinical PCa tissue samples and cell experiments. MRGs are generally involved in the occurrence and development of PCa. An independent risk factor for a bad prognosis in PCa patients and a high-risk score is related to patient outcome (Chong et al., 2021). The risk score is associated with PCa stromal score and levels of CD4 memory resting T cells, M0 macrophages, M1 macrophages, resting mast cells, CD4 memory activated T cells, and follicular helper T cells (Xu et al., 2021).

The investigation suffered from a variety of flaws. First and foremost, the samples applied in our investigation were collected retrospectively, all the outcomes were obtained using only data from public databases, and validation in a separate clinical patient cohort is still lacking despite the use of external datasets for validation. Next, surgery, neoadjuvant chemotherapy, and chemoradiotherapy, which may have affected how well the immune response and methylation condition performed.

Conclusion

Here, we disclosure the roles of MRGs modification patterns in the PCa and TME diversity, clinicopathological characteristics and a wide range of prognostic regulatory mechanisms. Next, the therapeutic obligations of MRGs in immunotherapy and commonly used antineoplastic drugs are explained by us. These discoveries emphasize the key clinical significance of MRGs, which offer a new view into the field of PCa research and promote the understanding of TME and immunotherapy in the future.

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

HZ conceived and designed the manuscript. XYe, RW and XYu collected and analyzed the data. ZW and HH checked the article.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.1030766/full#supplementary-material

SUPPLEMENTARY TABLE S1
The details of clinical characteristics.

SUPPLEMENTARY TABLE S2
The list of MRGs.

SUPPLEMENTARY TABLE S3
The details of GO analysis.

SUPPLEMENTARY TABLE S4
The details of KEGG analysis.

SUPPLEMENTARY TABLE S5
32 prognostic DEGs by univariate Cox analysis.

References

Bansal, D., Reimers, M. A., Knoche, E. M., and Pachynski, R. K. (2021). Immunotherapy and immunotherapy combinations in metastatic Castration-Resistant prostate cancer. *Cancers (Basel)* 13, 334. doi:10.3390/cancers13020334

Barbieri, I., and Kouzarides, T. (2020). Role of RNA modifications in cancer. *Nat. Rev. Cancer* 20, 303–322. doi:10.1038/s41568-020-0253-2

Cao, S., Lin, C., Li, X., Liang, Y., and Saw, P. E. (2021). TME-Responsive multistage nanoplatform for siRNA delivery and effective cancer therapy. *Int. J. Nanomedicine* 16, 5909–5921. doi:10.2147/IJN.S322901

Chong, W., Shang, L., Liu, J., Fang, Z., Du, F., Wu, H., et al. (2021). M(6)A regulator-based methylation modification patterns characterized by distinct tumor microenvironment immune profiles in colon cancer. *Theranostics* 11, 2201–2217. doi:10.7150/thno.52717

Dai, X., Ren, T., Zhang, Y., and Nan, N. (2021). Methylation multiplicity and its clinical values in cancer. *Expert Rev. Mol. Med.* 23, e2. doi:10.1017/erm.2021.4

Du, C., Lv, C., Feng, Y., and Yu, S. (2020). Activation of the KDM5A/miRNA-495/YTHDF2/m6A-MOB3B axis facilitates prostate cancer progression. *J. Exp. Clin. Cancer Res.* 39, 223. doi:10.1186/s13046-020-01735-3

Ehrlich, M. (2019). DNA hypermethylation in disease: Mechanisms and clinical relevance. *Epigenetics* 14, 1141–1163. doi:10.1080/15592294.2019.1638701

Engelhard, V., Conejo-Garcia, J. R., Ahmed, R., Nelson, B. H., Willard-Gallo, K., Bruno, T. C., et al. (2021). B cells and cancer. *Cancer Cell* 39, 1293–1296. doi:10.1016/j.ccell.2021.09.007

Franchina, D. G., Grusdat, M., and Brenner, D. (2018). B-Cell metabolic remodeling and cancer. *Trends Cancer* 4, 138–150. doi:10.1016/j.trecan.2017.12.006

Fridman, W. H., Meylan, M., Petitprez, F., Sun, C. M., Italiano, A., and Sautès-Fridman, C. (2022). B cells and tertiary lymphoid structures as determinants of tumour immune contexture and clinical outcome. *Nat. Rev. Clin. Oncol.* 19, 441–457. doi:10.1038/s41571-022-00619-z

Geng, Q., Shen, Z., Li, L., and Zhao, J. (2021). COL1A1 is a prognostic biomarker and correlated with immune infiltrates in lung cancer. *PeerJ* 9, e11145. doi:10.7717/peerj.11145

Gu, Y., Wu, X., Zhang, J., Fang, Y., Pan, Y., Shu, Y., et al. (2021). The evolving landscape of N(6)-methyladenosine modification in the tumor microenvironment. *Mol. Ther.* 29, 1703–1715. doi:10.1016/j.ymthe.2021.04.009

Hao, Y., Yan, M., Heath, B. R., Lei, Y. L., and Xie, Y. (2019). Fast and robust deconvolution of tumor infiltrating lymphocyte from expression profiles using least trimmed squares. *PLoS Comput. Biol.* 15, e1006976. doi:10.1371/journal.pcbi.1006976

Haruehanroengra, P., Zheng, Y. Y., Zhou, Y., Huang, Y., and Sheng, J. (2020). RNA modifications and cancer. *RNA Biol.* 17, 1560–1575. doi:10.1080/15476286.2020.1722449

He, R., Man, C., Huang, J., He, L., Wang, X., Lang, Y., et al. (2022). Identification of RNA Methylation-Related lncRNAs signature for predicting hot and cold tumors and prognosis in colon cancer. *Front. Genet.* 13, 870945. doi:10.3389/fgene.2022.870945

Hinshaw, D. C., and Shevde, L. A. (2019). The tumor microenvironment innately modulates cancer progression. *Cancer Res.* 79, 4557–4566. doi:10.1158/0008-5472.CAN-18-3962

Horii, M., and Matsushita, T. (2021). Regulatory B cells and T cell regulation in cancer. *J. Mol. Biol.* 433, 166685. doi:10.1016/j.jmb.2020.10.019

Hwang, B. O., Park, S. Y., Cho, E. S., Zhang, X., Lee, S. K., Ahn, H. J., et al. (2021). Platelet CLEC2-Podoplanin axis as a promising target for oral cancer treatment. *Front. Immunol.* 12, 807600. doi:10.3389/fimmu.2021.807600

Iasonos, A., Schrag, D., Raj, G. V., and Panageas, K. S. (2008). How to build and interpret a nomogram for cancer prognosis. *J. Clin. Oncol.* 26, 1364–1370. doi:10.1200/JCO.2007.12.9791

Jeong, S. H., Kim, R. B., Park, S. Y., Park, J., Jung, E. J., Ju, Y. T., et al. (2020). Nomogram for predicting gastric cancer recurrence using biomarker gene expression. *Eur. J. Surg. Oncol.* 46, 195–201. doi:10.1016/j.ejso.2019.09.143

Jiang, H., Guo, S., Zhao, Y., Wang, Y., Piao, H. Y., Wu, Y., et al. (2019). Circulating long non-coding RNA PCGEM1 as a novel biomarker for gastric cancer diagnosis. *Pathol. Res. Pract.* 215, 152569. doi:10.1016/j.prp.2019.152569

Kalbasi, A., and Ribas, A. (2020). Tumour-intrinsic resistance to immune checkpoint blockade. *Nat. Rev. Immunol.* 20, 25–39. doi:10.1038/s41577-019-0218-4

Koch, A., Joosten, S. C., Feng, Z., de Ruijter, T. C., Draht, M. X., Melotte, V., et al. (2018). Analysis of DNA methylation in cancer: Location revisited. *Nat. Rev. Clin. Oncol.* 15, 459–466. doi:10.1038/s41571-018-0004-4

Li, D., Li, K., Zhang, W., Yang, K. W., Mu, D. A., Jiang, G. J., et al. (2022). The m6A/m5C/m1A regulated gene signature predicts the prognosis and correlates with the immune status of hepatocellular carcinoma. *Front. Immunol.* 13, 918140. doi:10.3389/fimmu.2022.918140

Li, M., Zha, X., and Wang, S. (2021). The role of N6-methyladenosine mRNA in the tumor microenvironment. *Biochim. Biophys. Acta. Rev. Cancer* 1875, 188522. doi:10.1016/j.bbcan.2021.188522

Li, W. J., He, Y. H., Yang, J. J., Hu, G. S., Lin, Y. A., Ran, T., et al. (2021). Profiling PRMT methylome reveals roles of hnRNPA1 arginine methylation in RNA splicing and cell growth. *Nat. Commun.* 12, 1946. doi:10.1038/s41467-021-21963-1

Li, X. Y., Wang, S. L., Chen, D. H., Liu, H., You, J. X., Su, L. X., et al. (2022). Construction and validation of a m7G-Related Gene-Based prognostic model for gastric cancer. *Front. Oncol.* 12, 861412. doi:10.3389/fonc.2022.861412

Liu, H., He, X., Li, T., Qu, Y., Xu, L., Hou, Y., et al. (2022). PCGEM1 promotes proliferation, migration and invasion in prostate cancer by sponging miR-506 to upregulate TRIAP1. *BMC Urol.* 22, 14. doi:10.1186/s12894-022-00969-x

Ma, H. P., Chang, H. L., Bamodu, O. A., Yadav, V. K., Huang, T. Y., Wu, A., et al. (2019). Collagen 1A1 (COL1A1) is a reliable biomarker and putative therapeutic target for hepatocellular carcinogenesis and metastasis. *Cancers (Basel)* 11, E786. doi:10.3390/cancers11060786

Mahmoud, A. M., and Ali, M. M. (2019). Methyl donor micronutrients that modify DNA methylation and cancer outcome. *Nutrients* 11, E608. doi:10.3390/nu11030608

Martínez-Reyes, I., and Chandel, N. S. (2021). Cancer metabolism: Looking forward. *Nat. Rev. Cancer* 21, 669–680. doi:10.1038/s41568-021-00378-6

Maxwell, K. N., Cheng, H. H., Powers, J., Gulati, R., Ledet, E. M., Morrison, C., et al. (2022). Inherited TP53 variants and risk of prostate cancer. *Eur. Urol.* 81, 243–250. doi:10.1016/j.eururo.2021.10.036

Menezo, Y., Clement, P., Clement, A., and Elder, K. (2020). Methylation: An ineluctable biochemical and physiological process essential to the transmission of life. *Int. J. Mol. Sci.* 21, E9311. doi:10.3390/ijms21239311

Nishiyama, A., and Nakanishi, M. (2021). Navigating the DNA methylation landscape of cancer. *Trends Genet.* 37, 1012–1027. doi:10.1016/j.tig.2021.05.002

Oh, D. Y., and Fong, L. (2021). Cytotoxic CD4(+) T cells in cancer: Expanding the immune effector toolbox. *Immunity* 54, 2701–2711. doi:10.1016/j.immuni.2021.11.015

Petitprez, F., Meylan, M., de Reyniès, A., Sautès-Fridman, C., and Fridman, W. H. (2020). The tumor microenvironment in the response to immune checkpoint blockade therapies. *Front. Immunol.* 11, 784. doi:10.3389/fimmu.2020.00784

Pu, Y., Li, C., Yuan, H., and Wang, X. (2021). Identification of prostate cancer specific methylation biomarkers from a multi-cancer analysis. *BMC Bioinforma.* 22, 492. doi:10.1186/s12859-021-04416-w

Schmitt, M., and Greten, F. R. (2021). The inflammatory pathogenesis of colorectal cancer. *Nat. Rev. Immunol.* 21, 653–667. doi:10.1038/s41577-021-00534-x

Shi, Q., Zhu, Y., Ma, J., Chang, K., Ding, D., Bai, Y., et al. (2019). Prostate Cancer-associated SPOP mutations enhance cancer cell survival and docetaxel resistance by upregulating Caprin1-dependent stress granule assembly. *Mol. Cancer* 18, 170. doi:10.1186/s12943-019-1096-x

Song, W., Ren, J., Xiang, R., Kong, C., and Fu, T. (2021). Identification of pyroptosis-related subtypes, the development of a prognosis model, and characterization of tumor microenvironment infiltration in colorectal cancer. *Oncoimmunology* 10, 1987636. doi:10.1080/2162402X.2021.1987636

Traube, F. R., and Carell, T. (2017). The chemistries and consequences of DNA and RNA methylation and demethylation. *RNA Biol.* 14, 1099–1107. doi:10.1080/15476286.2017.1318241

Vietri, M. T., D'Elia, G., Caliendo, G., Resse, M., Casamassimi, A., Passariello, L., et al. (2021). Hereditary prostate cancer: Genes related, target therapy and prevention. *Int. J. Mol. Sci.* 22, 3753. doi:10.3390/ijms22073753

Vitale, I., Manic, G., Coussens, L. M., Kroemer, G., and Galluzzi, L. (2019). Macrophages and metabolism in the tumor microenvironment. *Cell Metab.* 30, 36–50. doi:10.1016/j.cmet.2019.06.001

Wen, S., Wei, Y., Zen, C., Xiong, W., Niu, Y., and Zhao, Y. (2020). Long non-coding RNA NEAT1 promotes bone metastasis of prostate cancer through N6-methyladenosine. *Mol. Cancer* 19, 171. doi:10.1186/s12943-020-01293-4

Xin, S., Deng, Y., Mao, J., Wang, T., Liu, J., Wang, S., et al. (2022). Characterization of 7-Methylguanosine identified biochemical recurrence and tumor immune microenvironment in prostate cancer. *Front. Oncol.* 12, 900203. doi:10.3389/fonc.2022.900203

Xu, B., Lu, M., Yan, L., Ge, M., Ren, Y., Wang, R., et al. (2021). A Pan-Cancer analysis of predictive methylation signatures of response to cancer immunotherapy. *Front. Immunol.* 12, 796647. doi:10.3389/fimmu.2021.796647

Xu, Z., Chen, S., Zhang, Y., Liu, R., and Chen, M. (2022). Roles of m5C RNA modification patterns in biochemical recurrence and tumor microenvironment characterization of prostate adenocarcinoma. *Front. Immunol.* 13, 869759. doi:10.3389/fimmu.2022.869759

Yang, B., Wang, J. Q., Tan, Y., Yuan, R., Chen, Z. S., and Zou, C. (2021). RNA methylation and cancer treatment. *Pharmacol. Res.* 174, 105937. doi:10.1016/j.phrs.2021.105937

Yang, K., and Kallies, A. (2021). Tissue-specific differentiation of CD8(+) resident memory T cells. *Trends Immunol.* 42, 876–890. doi:10.1016/j.it.2021.08.002

Yu, G., Bao, J., Zhan, M., Wang, J., Li, X., Gu, X., et al. (2022). Comprehensive analysis of m5C methylation regulatory genes and tumor microenvironment in prostate cancer. *Front. Immunol.* 13, 914577. doi:10.3389/fimmu.2022.914577

Zhang, B., Wu, Q., Li, B., Wang, D., Wang, L., and Zhou, Y. L. (2020). M(6)A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol. Cancer* 19, 53. doi:10.1186/s12943-020-01170-0

Zhang, M., Song, J., Yuan, W., Zhang, W., and Sun, Z. (2021a). Roles of RNA methylation on tumor immunity and clinical implications. *Front. Immunol.* 12, 641507. doi:10.3389/fimmu.2021.641507

Zhang, Q., Zheng, J., and Liu, L. (2019). The long noncoding RNA PCGEM1 promotes cell proliferation, migration and invasion via targeting the miR-182/FBXW11 axis in cervical cancer. *Cancer Cell Int.* 19, 304. doi:10.1186/s12935-019-1030-8

Zhang, Q., Zhu, Z., Guan, J., and Zheng, C. (2022). Identification and assessment of Necroptosis-Related genes in clinical prognosis and immune cells in diffuse large B-Cell lymphoma. *Front. Oncol.* 12, 904614. doi:10.3389/fonc.2022.904614

Zhao, S. G., Chen, W. S., Li, H., Foye, A., Zhang, M., Sjöström, M., et al. (2020). The DNA methylation landscape of advanced prostate cancer. *Nat. Genet.* 52, 778–789. doi:10.1038/s41588-020-0648-8

Zhang, P., Qian, B., Liu, Z., Wang, D., Lv, F., Xing, Y., et al. (2021b). Identification of novel biomarkers of prostate cancer through integrated analysis. *Transl. Androl. Urol.* 10, 3239–3254. doi:10.21037/tau-21-401

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 20 October 2022
DOI 10.3389/fphar.2022.1032282

OPEN ACCESS

EDITED BY
Zhi-Qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Jinxiao Li,
Huazhong University of Science and Technology, China
Fangdie Ye,
Fudan University, China

*CORRESPONDENCE
Xiaofeng Rong,
cyrxf@163.com
Chaoxu Wu,
wucx_cq@163.com

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

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 30 August 2022
ACCEPTED 03 October 2022
PUBLISHED 20 October 2022

CITATION
Cao C, Zhao W, Chen X, Shen B, Wang T, Wu C and Rong X (2022), Deciphering the action mechanism of paeoniflorin in suppressing pancreatic cancer: A network pharmacology study and experimental validation.
Front. Pharmacol. 13:1032282.
doi: 10.3389/fphar.2022.1032282

Deciphering the action mechanism of paeoniflorin in suppressing pancreatic cancer: A network pharmacology study and experimental validation

Chunhao Cao[1†], Wenting Zhao[2†], Xianglin Chen[2], Bin Shen[3], Teng Wang[3], Chaoxu Wu[1*] and Xiaofeng Rong[1*]

[1]Department of Integrated Traditional Chinese and Western Medicine, The First Affiliate Hospital of Chongqing Medical University, Chongqing, China, [2]Hubei University of Chinese Medicine, Wuhan, China, [3]Chongqing Medical University, Chongqing, China

Background: Paeoniflorin (PF) is the main active component of Chinese herbaceous peony that has been shown to have an anti-tumor effect. However, there are few studies on the prevention and treatment of pancreatic cancer with PF.

Methods: We gathered Microarray data pertaining to paeoniflorin intervention in pancreatic cancer by utilizing the GEO database (GSE97124). Then, the DEGs were filtered by the 33R program. RNA-seq data of pancreatic cancer and normal tissue samples were taken from the TCGA and GTEx databases, respectively, and the WGCNA technique was utilized to examine the pancreatic cancer-specific genes. Paeoniflorin target genes for the treatment of pancreatic cancer were determined based on the overlap between DEGs and WGCNA. GO and KEGG enrichment analyses were then performed on paeoniflorin target genes to discover which biological processes were impacted. Using the 3 hierarchical methods included in the Cytohubba plugin, we re-screened the hub genes in the target genes to find the genes most relevant to paeoniflorin treatment. The overall survival effects of hub genes were confirmed using the TCGA database. Finally, the paeoniflorin targets identified by the network pharmacology analysis were validated using PANC-1 and Capan-2 cells.

Results: We identified 148 main potential PF targets, and gene enrichment analysis suggested that the aforementioned targets play a crucial role in the regulation of MAPK, PI3K-AKT, and other pathways. The further screening of the prospective targets resulted in the identification of 39 hub genes. Using the TCGA database, it was determined that around 33.33% of the hub gene's high expression was linked with a bad prognosis. Finally, we demonstrated that PF inhibits IL-6 and IL-10 expression and p38 phosphorylation in pancreatic cancer cells, thereby reducing inflammation.

Conclusion: PF may regulate inflammatory factors mainly through the p38 MAPK signal pathway. These findings provide theoretical and

experimental evidence suggesting the PF as a promising natural source of anti-tumor compounds for pancreatic cancer.

KEYWORDS

paeoniflorin, pancreatic cancer, network pharmacology, p38 MAPK signal pathway, WGCNA

1 Introduction

Over the previous decade, the yearly number of pancreatic cancer diagnoses has increased from 43,140 to 60,430 (Jemal et al., 2010; Siegel et al., 2021). It is now the third highest cause of cancer mortality in the United States, behind lung cancer and breast cancer (Grossberg et al., 2020). Despite racial disparities in the risk of gastrointestinal disease, studies from numerous nations on various continents and with diverse ethnic structures indicate that the prevalence of pancreatic cancer continues to increase due to the aging of the global population (Jia et al., 2018; Nipp et al., 2018; Klein, 2021). Moreover, the mortality rate of pancreatic cancer is still high compared with other malignancies, and the ensuing social impact cannot be underestimated. This may be owing to the subtle start, quick development, and early metastasis of pancreatic cancer and the fact that the majority of patients have lost the window of opportunity for surgery at the time of diagnosis (Strobel et al., 2019). In addition, pancreatic cancer surgery is challenging and needs negative margins under the microscope. Only a minority of patients benefit from surgery (Kamisawa et al., 2016; Kang et al., 2016). Therefore, the great majority of patients with pancreatic cancer must receive radiation treatment or chemotherapy.

In recent years, novel targeted medicines and immunotherapy, including Erlotinib, Everolimus, and Olaparix, have offered patients some hope (Wong and Lemoine, 2009; Leroux and Konstantinidou, 2021). However, these medications and treatments continue to have many adverse events. For instance, during the administration of erlotinib, side effects such as dermatitis and diarrhea occurred with a frequency that was not negligible (Rudin et al., 2008). There have also been occasional reports of erlotinib causing interstitial pneumonia and treatment-related fatalities (Wang et al., 2016). These adverse effects have a devastating impact on older pancreatic cancer patients and may even force them to discontinue therapy. In addition, owing to individual variances, pricey targeted medications are ineffective for certain individuals. This puts the therapeutic use of some targeted medications in a dilemma. If pancreatic cancer patients are going to live longer and have a better quality of life, it is still important to find new medicines to treat the disease.

Many herbs have healing properties and have been used to treat diseases in China for thousands of years. However, it should be reminded that hazardous substances and non-pharmaceutical chemicals are also present throughout the herb. Therefore, it is a reasonable choice to investigate natural substances with well-defined chemical structures that are isolated from herbal remedies as medicinal pharmaceuticals. Numerous natural compounds, like paclitaxel, resveratrol, etc., have shown potent anticancer activity. Even at large concentrations, several natural chemicals are well tolerated by patients (Rejhová et al., 2018). Therefore, novel natural chemicals have a promising future in the creation of anti-cancer medications. Paeoniflorin (PF) is the main active component of Radix Paeoniae Alba, Radix Paeoniae Rubra, and Paeonia Suffruticosa Andr, which is a water-soluble monoterpene glycoside (Wu et al., 2010), extracted from the peony in 1963 for the first time (Hu et al., 2013). By blocking the activity of the Notch-1 signaling pathway, Zhang and colleagues have shown that PF reduces the growth and invasion of breast cancer cells (Zhang et al., 2016b). Treatment of colorectal cancer cells with PF leads to downregulation of FoxM1 and inhibits colorectal cancer cell migration (Yue et al., 2018). However, the chemical and pharmacological foundation of PF as a pancreatic cancer inhibitor has not been proven or investigated.

Network pharmacology is a novel inter-disciplinary technique that has assisted several researchers in investigating the pharmacological effects of natural substances and compound herbal remedies. To further elucidate the mechanism of action of PF in the treatment of pancreatic cancer, we used network pharmacology to investigate the impact of PF on pancreatic cancer and confirmed the regulatory link between PF and key signaling pathways *in vitro*. Figure 1 depicts the workflow for this research.

2 Materials and methods

2.1 Predicting the targets genes of the paeoniflorin in pancreatic cancer

GSE97124 (Li et al., 2017) is a dataset related to paeoniflorin treatment of pancreatic cancer from the GEO database (http://ww.ncbinlm.nih.gov/geo/). The dataset for pancreatic cancer was preprocessed and normalized for future analysis using the normalize Quantities function of the limma (Ritchie et al., 2015) package in R (Version 4.1.0). Using the limma, differentially expressed genes (DEGs) were assessed between paeoniflorin-free and paeoniflorin-treated samples. The results are plotted with the ggplot2 package as a volcano map and heatmap (Ito and Murphy, 2013).

Weighted gene correlation networks Analysis (WGCNA) reveals modules of co-expressed genes in complex biological procedures (Langfelder and Horvath, 2008). For WGCNA analysis, we utilized 177 pancreatic cancer tissues from The

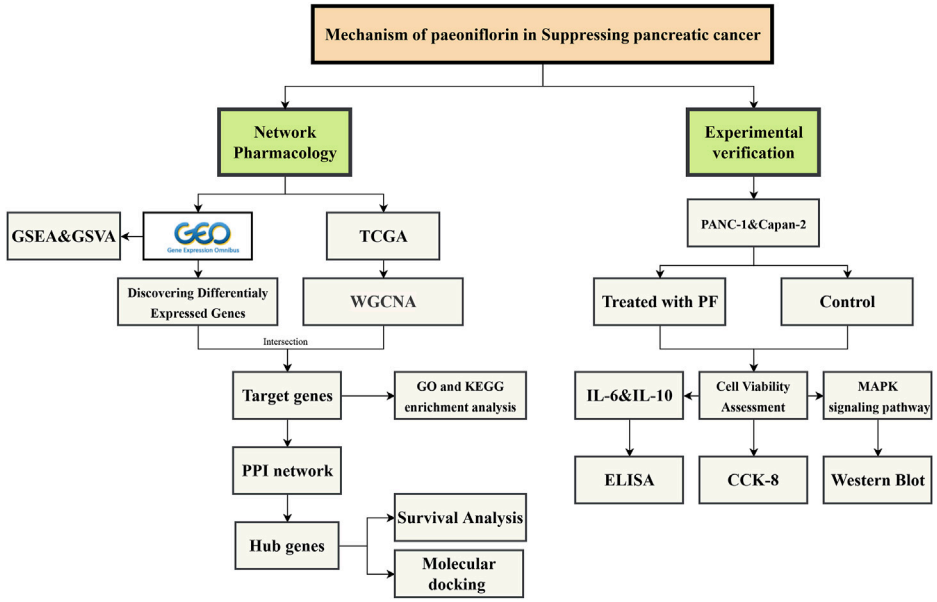


FIGURE 1
The workflow of this study.

Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) and 167 normal tissues from Genotype-Tissue Expression (GTEx, https://commonfund.nih.gov/GTEx). Confirm the soft threshold using the pickSoftThreshold function of the WGCNA package in R with reference to Linbang et al. (Wang et al., 2021). Following the formation of a scale-free network according to the soft threshold, a topology matrix and hierarchical clustering are applied. That meant 60 genes were the bare minimum per module needed for dynamic gene module cleavage. Then, each module's Eigengenes were established. Based on the Eigengenes module, hierarchical clustering was carried out once the correlation between modules was established. A total of six modules were created by once more combining the earlier components. The link between modules and between modules and PF was investigated using Pearson correlation. Significantly connected modules were regarded as crucial PF elements for a subsequent investigation.

Finally, paeoniflorin's target genes for the therapy of pancreatic cancer were determined by looking at the overlap between WGCNA and DEGs screening.

2.2 Gene set enrichment analysis and Gene Set Variation Analysis

Many of the most enriched gene sets defining metabolic activities were uncovered using GSEA (Subramanian et al., 2005). We ran a gene set enrichment analysis on Microarray data collected from the GEO database to investigate the changes induced by paeoniflorin. The gene set "c2. cp.v7.2. symbols.gmt" from the MSigDB database (https://www.gsea-msigdb.org/gsea/msigdb/) was used for the aforementioned procedures (Liberzon et al., 2015). The R package "GSVA" was used to perform Gene Set Variation Analysis (GSVA). An enrichment score (ES) was generated for each sample and pathway as a result of the analysis using a non-parametric unsupervised approach that converted a traditional gene matrix (gene-by-sample) into a gene set-by-sample matrix. Then, the mean values of ES of cells in the two groups of samples were compared using the *t*-test. The MSigDB database's "c2. cp.kegg.v7.5.1. symbols" is one of the target gene sets used in this study. A false discovery rate (FDR) < 0. 25 was considered significant enrichment (Zhang et al., 2016a).

2.3 Gene ontology and kyoto encyclopedia of genes and genomes enrichment analysis

Gene ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis were performed on target genes (Gene Ontology Consortium, 2015; Kanehisa et al., 2019). The "clusterProfiler" package (Yu et al., 2012) of the R program was used to carry out

these gene enrichment analyses. Pathways with $p < 0.01$ were considered statistically significant. Then, the visualized the top 20 or 30 pathways using the "ggplot2" package in R.

2.4 Building a PPI network and Kaplan-Meier curves

The Protein-Protein Interaction Network was built by inserting target genes into the STRING database (Szklarczyk et al., 2017). Change the needed minimum interaction score to 0.9. The PPI network was visualized and displayed by Cytoscape (Shannon et al., 2003). Likewise, the cytoHubba plugin (Shannon et al., 2003) was used to exclude non-central genes from the target genes, substantially reducing the number of target genes. We then used the TCGA database to examine the association between the expression of these hub genes and prognosis. The specific strategy involves downloading the pancreatic cancer dataset from the TCGA database through the TCGAbiolinks package (Colaprico et al., 2016), and then choosing a total of 178 individuals with complete overall survival (OS) and clinical characteristics. Patients were divided into high- and low-expression groups based on the median of the difference in expression of a single hub gene; the difference in overall survival between the two groups was then examined using the Kaplan-Meier curves. The "survminer" package is utilized for data visualization, whereas the "survival" package is utilized for survival data statistical analysis.

2.5 Molecular docking verification

Molecular docking was utilized to anticipate interactions between paeoniflorin and its primary targets. The Protein Data Bank (Berman et al., 2000) contains detailed information and 3D structures, including the original structures of important targets. Paeoniflorin's chemical structure was derived from the PubChem website (Kim et al., 2021). Import the structure of the aforementioned protein into the AutoDock Vina tool (Seeliger and de Groot, 2010), select the default settings, and set the Grid Box to the entire protein molecule. Then, run AutoDock Vina for molecular docking to confirm the binding activity of the target and the compound and obtain the binding energy.

2.6 Experimental validation *in Vitro*

2.6.1 Reagents and materials

Prior to usage, the purity of paeoniflorin (CAS: 23180-57-6) acquired from Sigma-Aldrich (St. Louis, United States) was verified by UPLC-MS to be more than 95%. From Procell Co., Ltd. (Wuhan, China) purchased the human pancreatic cancer cell lines PANC-1 and Capan-2. Gibco supplied the DEME, fetal bovine serum, and penicillin-streptomycin necessary for cell culture (Grand Island, United States). From BOSTER Technology Co., Ltd. (Wuhan, China) bought ELISA kits. DMSO, CCK-8 kit, RIPA buffer, and BCA Protein Assay Kit were procured from Solarbio Technology Co., Ltd. (Beijing, China). Primary antibodies against p38, phospho-p38, p44/42 MAPK (Erk1/2), phospho-p44/42 MAPK (Erk1/2), SAPK/JNK, phospho-SAK/JNK, and β-actin were bought from Abcam and Cell Signaling Technology, both in the United States.

2.6.2 Cell culture and viability assay

Panc-1 cells and Capan-2 were cultured in a 10% FBS and 100 U/mol penicillin-streptomycin solution in an incubator at 37°C, 5% CO_2, and 95% relative humidity. Upon intervention, every cell was in the logarithmic growth phase. PF was dissolved in DMSO to achieve a concentration of 400 μM. 100 ml of PF (400 μM) was then diluted to 200, 100, and 50 μM, respectively, in 100, 300, and 700 ml of DMEM. In six-well plates, cells (about 1×10^6 per well) were seeded and cultured for 24 h with varying doses of PF-treated DEME or DEME without PF. The collection of cells and cell culture media for use in later investigations. CCK-8 cell viability was evaluated in accordance with the manufacturer's instructions. Panc-1 cells were planted at a density of 3×10^4 cells per well in 96-well plates (approximately 100 μl medium per well). After 24 h of treatment with various PF concentrations, 10 μl of CCK-8 solution was applied to the 96-well plate. The plate was incubated at 37°C for 4 h. Use a microplate reader (Bio-Rad, United States) to figure out the optical density of each well at 450 nm.

2.6.3 Measurement of IL-6 and IL-10

Using an ELISA kit and the manufacturer's instructions, the IL-6 and IL-10 levels in cell culture supernatants were assessed.

2.6.4 Western blot analysis

After 48 h of treatment with various doses of PF (0, 50, 100, and 200 μM), PANC-1 cells were harvested. PANC-1 cells were lysed in RIPA buffer with a protease and phosphatase inhibitor cocktail. Upon completion of the lysis, the protein concentration was determined with the BCA kit. 30 mg of protein was loaded per lane on 10% SDS-PAGE gels and transferred to 0.45 m PVDF membranes (Millipore). After 60 min of blocking with 5% skim milk, the membrane was incubated at 4°C overnight with the primary antibody (1:1000) indicated in 2.6.1. After three washes, the membrane was re-incubated with HRP-conjugated secondary antibody IgG (1:2000) at room temperature for 1 h in the dark. Signals were recognized by autoradiography after applying ECL (FluorChem E, Proteinsimple, United States). The densitometric findings were examined using ImageJ (National Institutes of Health, United States).

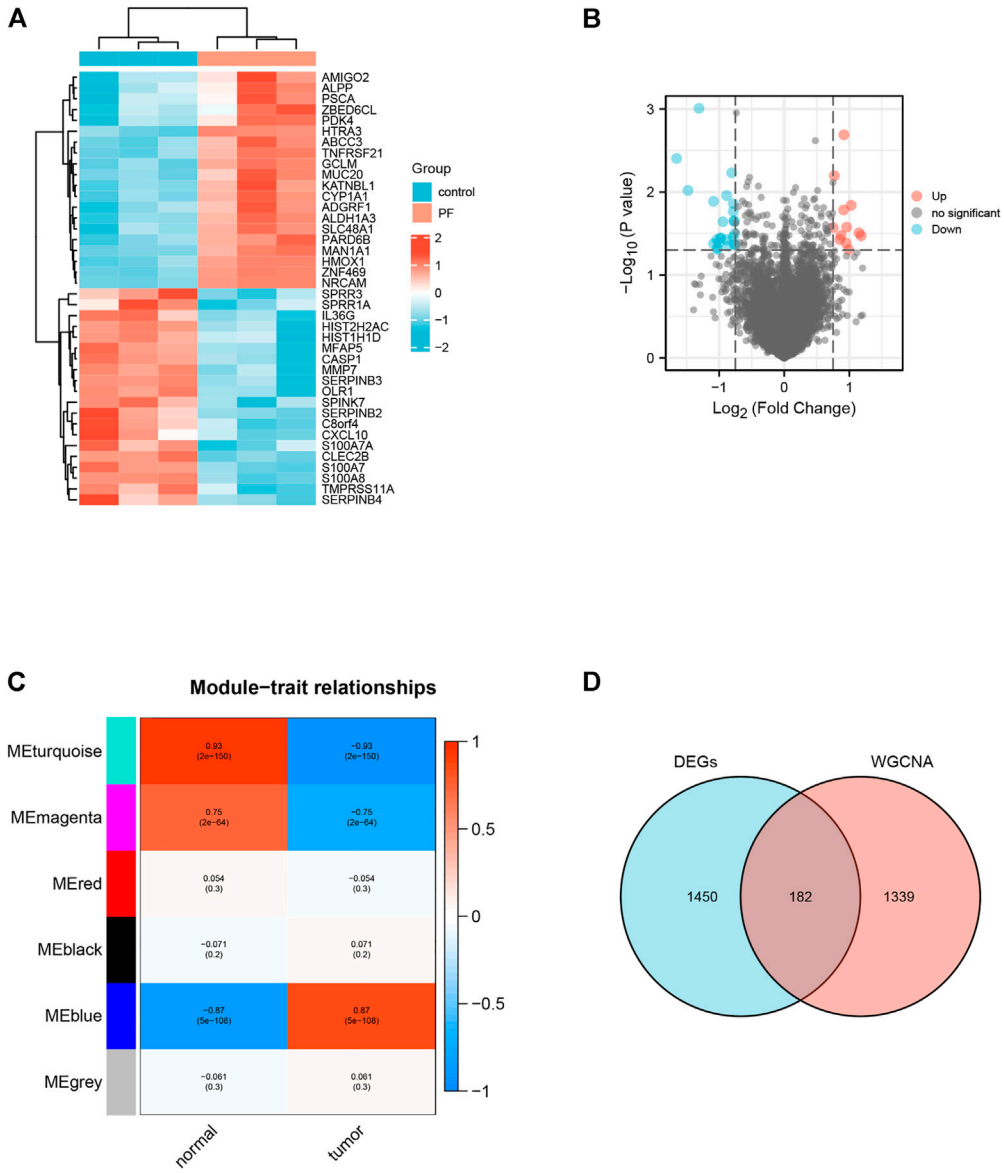


FIGURE 2
Identification of Target genes for paeoniflorin treatment of pancreatic cancer. **(A)** Heatmap of the top 40 DEGs associated with Paeoniflorin-treated pancreatic cancer. **(B)** The Volcano plot shows the distribution of genes that are differentially expressed (log2 fold change) compared to a measure of statistical significance (–log10 *p*-value) in the GSE97124 dataset. Genes that were downregulated are shown in blue, whereas genes that were upregulated are shown in red. **(C)** The results show a highly significant correlation between GS and MM in the MEblue module. **(D)** Genes screened by DEGs (pink circles), genes in modules strongly connected to PF in WGCNA analysis (blue circles), and portions of these two classes of genes that overlap (purple circles).

2.7 Statistical analysis

The information was given in the form of means and standard deviations. A student's t-test was used to assess differences between pairs of groups, while one-way analysis of variance (ANOVA) was used to investigate differences between three or more groups. Data analysis and graphing were performed using GraphPad Prism 7 (San Diego, United States).

3 Results

3.1 Identification of Target genes in datasets

We identified 1632 differentially expressed genes (DEGs) in the GSE97124 dataset based on a p value <0.01 (Figures 2A,B). Next, we did an analysis of the TCGA and GTEx data

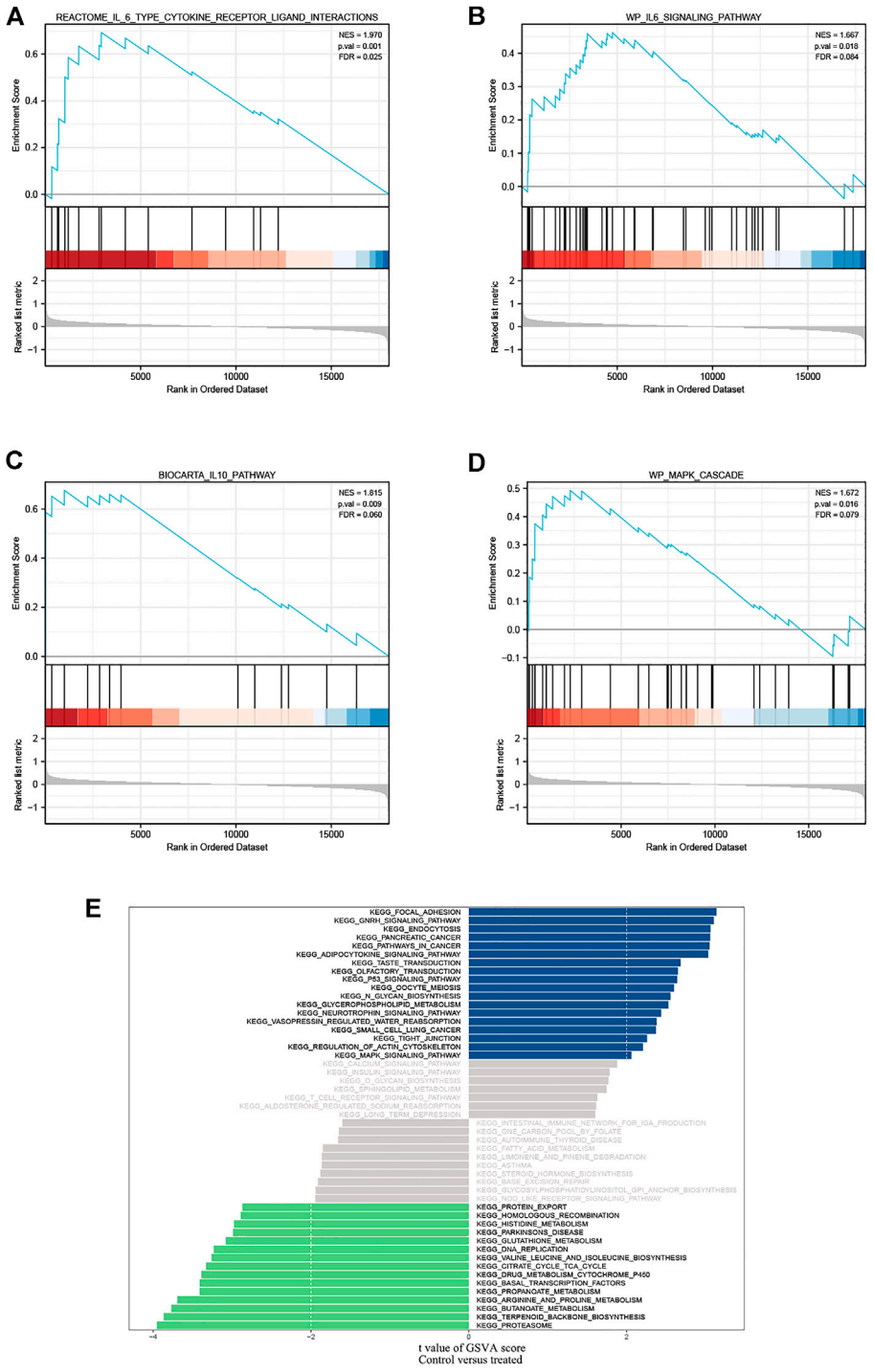


FIGURE 3
GSEA identified IL-6 type cytokine receptor ligand interactions. **(A,B)** IL-6 family signaling pathways; **(C)** IL-10 pathway and **(D)** MAPK cascade as regulatory targets of paeoniflorin in pancreatic cancer. **(E)** Differences in pathway activities scored per sample by GSVA between PF-treated group and control group. Blue represents the pathway with t value greater than 2, and green represents the pathway with t value less than -2. NES, normalized enrichment score; FDR, False discovery rate.

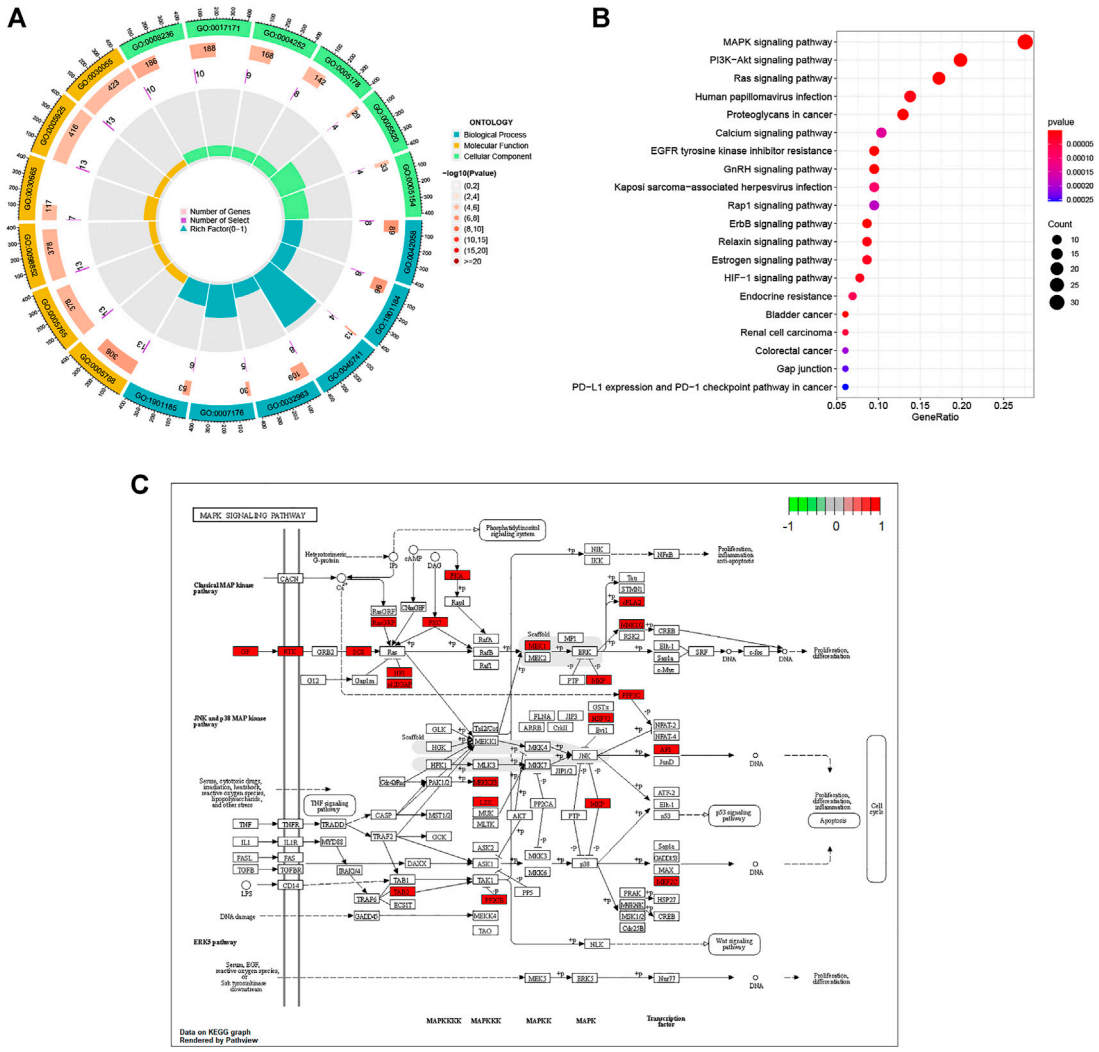


FIGURE 4
Biological Function and Pathway Analysis. **(A)** GO enrichment analysis of target genes. In the figure, terms enriched in the BP category are depicted in blue, terms enriched in the CC category in green, and terms enriched in the MF category in yellow. The *p* value's cutoff value was set at 0.05. *p* value is used to rank terms within a category. **(B)** KEGG enrichment analysis of target gene. **(C)** Distribution of Target genes in the MAPK signaling pathway. The red box in the figure represents Target genes.

using the WGCNA method. The pickSoftThreshold function identifies 9 as the optimal soft threshold (Supplementary Figure S1A). Then, the dynamic tree cutting algorithm identifies six modules, which are differentiated by distinct colors and marked on the diagram (Supplementary Figures S1B,C). Module eigengenes (ME) refer to the major components of every module. By evaluating the correlation between the modules' eigengenes, it was evident that the MEblue modules were most strongly related to pancreatic cancer (Supplementary Figure S1D; Figure 2C). Therefore, we considered the genes in the MEblue module as disease-related characteristic gene sets. The MEblue module has 1521 genes, of which 182 (11.97%) overlap with DEGs (Figure 2D). Select as target genes the genes that overlap between DEGs and this module gene.

3.2 GSEA and GSVA

Using the previously mentioned R software package, we investigated significantly enriched pathways for paeoniflorin intervention and control groups. In the Microarray data set, IL-6, IL-10, and MAPK pathways were strongly related to paeoniflorin intervention, indicating that paeoniflorin may play a crucial role *via* these pathways (Figures 3A–D). We also performed a GSVA analysis. The results of a GSVA

analysis were arranged by the magnitude of the absolute t-value and displayed in figure format (Figure 3E). In terms of biological processes such as focal adhesion, gnrh signaling pathway, endocytosis, proteasome, terpenoid backbone biosynthesis, and butanoate metabolism, there were substantial variations between the treatment group and the control group.

3.3 Biological Function and Pathway Analysis

We utilized the "clusterProfiler" package to do GO annotation and KEGG pathway analysis on pancreatic cancer target genes. Figure 4A; Supplementary Table S1 depict the most enriched items across the three categories of biological process (BP), cellular component (CC), and molecular function (MF). Epidermal growth-related pathway, ERBB signaling pathway, collagen metabolic process, ribosome, and regulation of protein tyrosine kinase activity were among the targets in the BP test. The four most frequently occurring GO terms under the category of cellular component were endoplasmic reticulum lumen, lysosomal membrane, lytic vacuole membrane, and clathrin-coated vesicle membrane. The overwhelming majority of CC terms were linked to vesicles or membranes. In MF, the primary targets were serine-type peptidase activity, serine hydrolase activity, serine-type endopeptidase activity, and integrin binding. Based on a *p*-value of 0.05, KEGG analysis identified 50 enriched pathways (Figure 4B, Only the top ten enriched terms are shown in the figure). This suggests that paeoniflorin's regulatory mechanisms include the PI3K-AKT signaling route, the RAS signaling pathway, and the MAPK signaling network. Since the MAPK signaling route was included in both GSEA, GSVA and KEGG enrichment, we visualized the link between important genes and the MAPK signaling pathway (Figure 4C).

3.4 PPI and survival analysis

Import the 184 targets discovered in 3.1 into the STRING platform, set the confidence score to 0.9, and construct a PPI network. Due to the enormous number of nodes, the 182 target genes with the highest node values are displayed on the graph There are 126 nodes in the PPI network and 238 edges (as shown in Figure 5A, the color gradually turns yellow, the greater the possibility of becoming a core protein). The Cytohubba plugin includes the following degree algorithms: MNC, MCC, and DEGREE. We implement these algorithms to filter nodes in the PPI network and designate the intersection nodes as hub genes (Figures 5B–D). The conclusion is represented by a Venn diagram (Figure 5E). These hub genes are predominantly associated with the matrix metalloprotein (MMP) family. The above findings show that paeoniflorin may have a role in the therapy of pancreatic cancer by interfering with these main targets.

In addition, using pancreatic cancer patient information from the TCGA database, we analyzed the effect of these hub genes on overall survival. Four of the twelve genes had clearly separated KM curves. This indicates that low expression of these four genes (MMP1, MMP7, MMP14, and HBEGF) predicts longer overall survival than high expression (Figure 6). These four genes and PF were subsequently considered ligands and receptors, respectively. Lower binding energies suggest a more stable binding configuration between the receptor and ligand. Figure 7 depicts interactions between receptors and their ligands and affinity. The image on the left was captured when the lens was focused further away, while the one on the right was captured when the lens was focused closer.

3.5 PF inhibited proliferation of pancreatic cancer cells

PANC-1 and Capan-2 cells were treated with PF (0, 50, 100, 200, and 300 μM) for 24 h to determine the effects of various PF concentrations on human pancreatic cancer PANC-1 and Capan-2 cells. The findings demonstrated that the presence of PF lowered the number of viable PANC-1 and Capan-2 cells in comparison to the control group (Figure 8). In addition, PF inhibited both types of pancreatic cancer cells in a concentration-dependent manner between 0 and 200 μM. The inhibitory effect did not change appreciably when the concentration of PF was increased to 300 μM from 200 μM. In the subsequent trials, dosages ranging from 0 to 200 μM are used.

3.6 PF alleviates inflammation in PANC-1 cells and Capan-2 cells

The levels of inflammatory factors were measured to establish whether PF may attenuate the inflammatory response in pancreatic cancer. ELISA showed that the levels of IL-6 and IL-10 were lower in the group that was given PF than in the control group (Figure 9).

3.7 PF suppressed the MAPK signal pathway

Firstly, we conducted docking experiments between PF and p38. Figure 10A depicts the final output optimization using AutoDock Vina. Docking data indicated that the absolute value of the affinity of the target protein for PF is 8. The lower the binding energy, the greater the binding activity and the greater the target protein's capacity to bind to PF. Consequently, p38 can tightly bind to PF. Subsequently, we

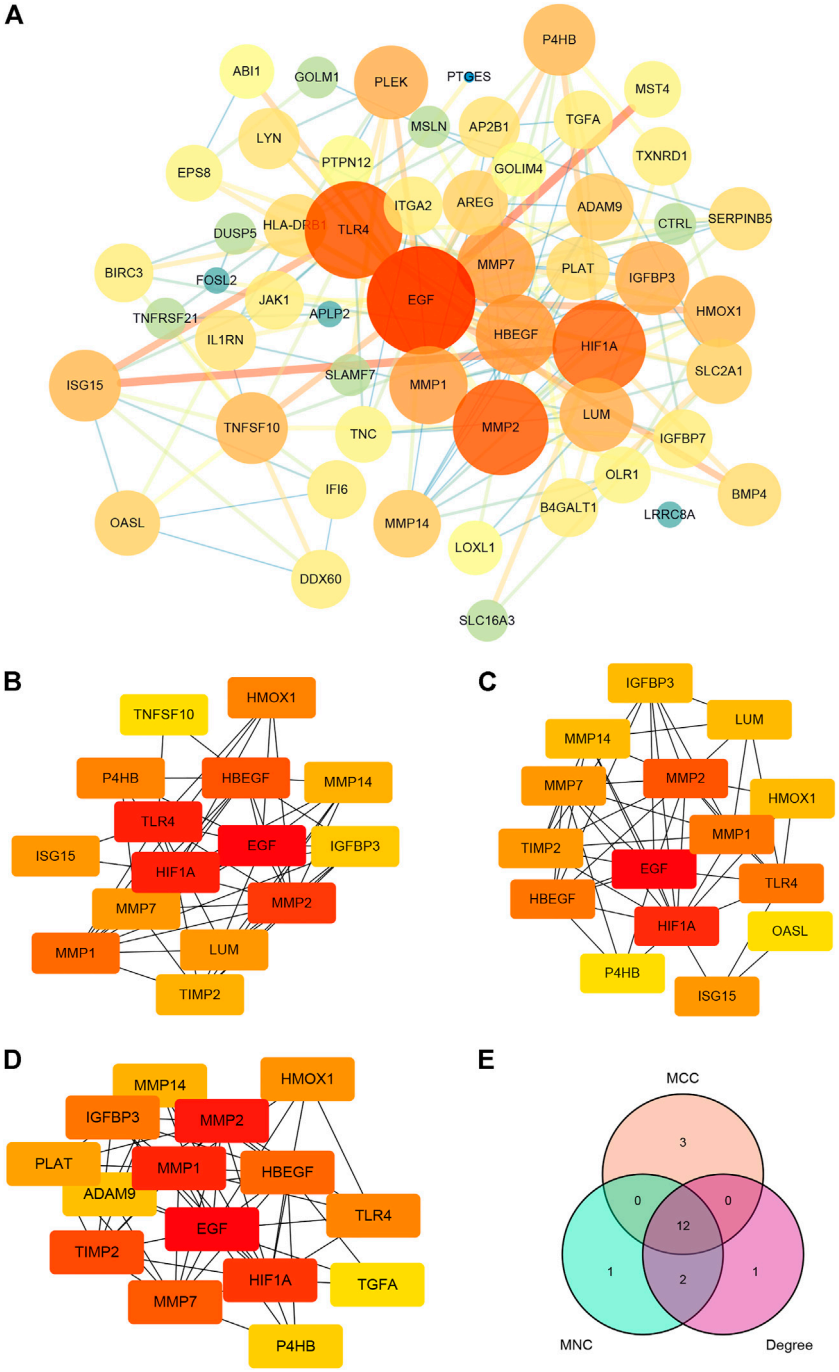


FIGURE 5
PPI network construction and hub genes screening. **(A)** PPI network of paeoniflorin for treating pancreatic cancer. The greater the number of linked nodes, the larger and darker the circle and its color. Three screening techniques for hub genes (implemented by the Cytohubba plugin). Yellow nodes have the lowest correlation strength, whereas red nodes have the greatest. **(B)** MCC; **(C)** MNC; **(D)** DEGREE. **(E)** A Venn diagram analysis of hub gene screening using Cytohubba. Each hue corresponds to a screening algorithm. The core gene is the hub gene, which is the gene present in all five algorithms.

demonstrated the therapy of pancreatic cancer with PF through the MAPK pathway in pancreatic cancer cells. In the range of 0–200 μm, the inhibitory impact of PF on p38 phosphorylation grew progressively, but p38 concentrations did not differ substantially from those of the controls. PF reduced the activation of P38 MAPK, but not ERK and JNK

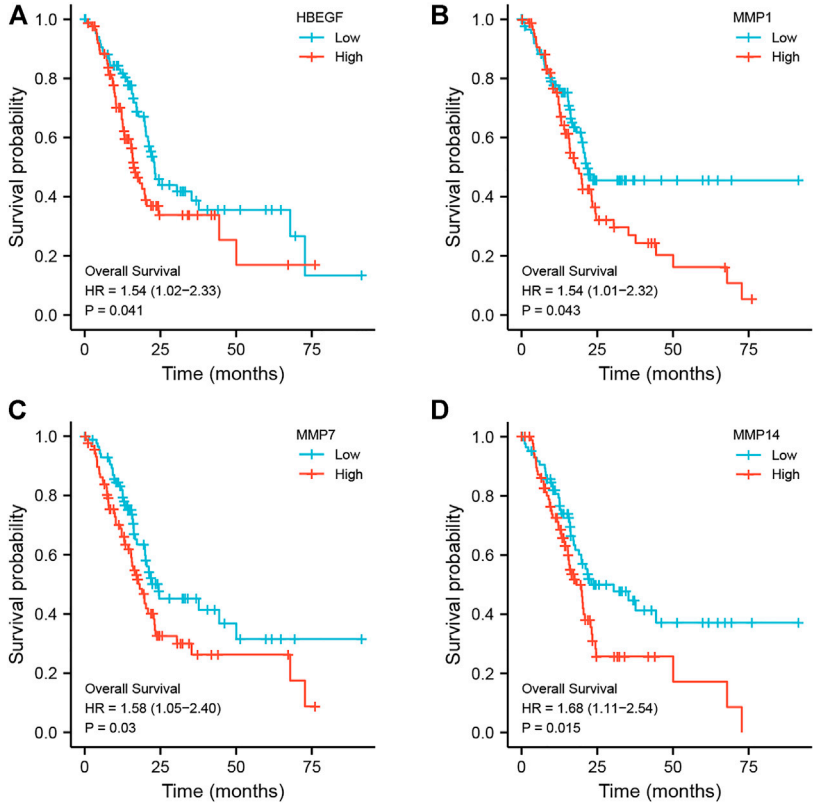


FIGURE 6
Kaplan-Meier curves of hub genes. According to the expression level of each hub gene, it is separated into high expression (red line) and low expression (blue line) groups, and the KM curve of each gene is drawn independently. **(A)** HBEGF; **(B)** MMP1; **(C)** MMP7; **(D)** MMP14.

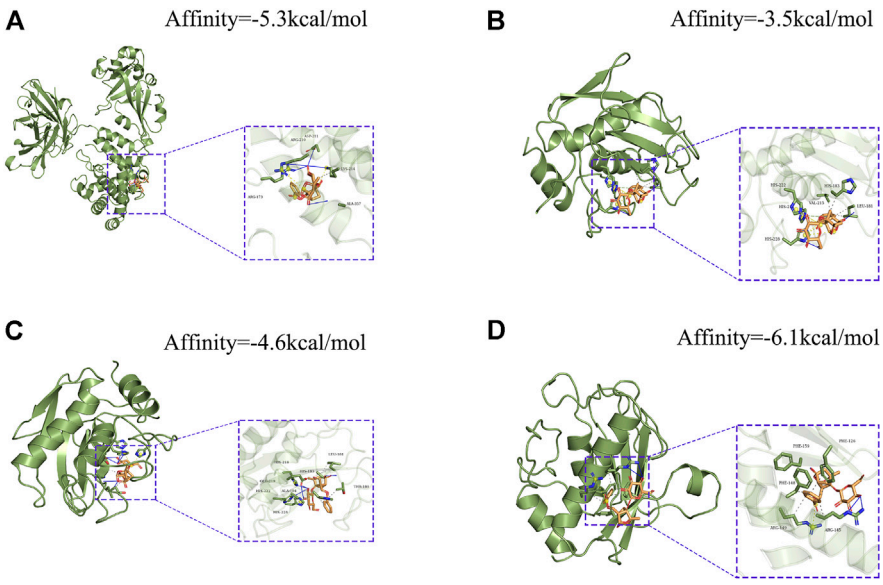


FIGURE 7
Molecular docking models of PF binding to potential targets. **(A)** HBEGF; **(B)** MMP1; **(C)** MMP7; **(D)** MMP14.

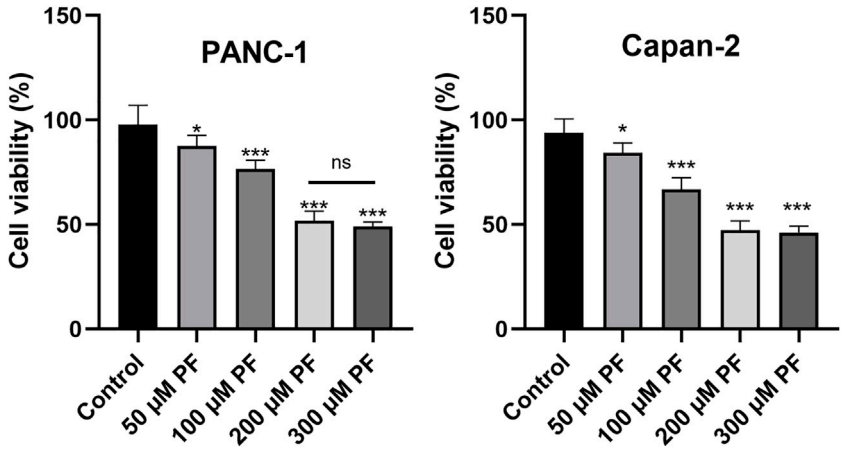


FIGURE 8
Effects of PF intervention on the growth of pancreatic cancer cells. PANC-1 cells and Capan-2 were treated with 50 μM, 100 μM, 200 μM, and 300 μMPF, and the CCK-8 assay was used to evaluate cell viability. The experiments were repeated at least three times. Data were expressed as mean ± SD. *$p < 0.05$; ***$p < 0.001$, *versus* control group.

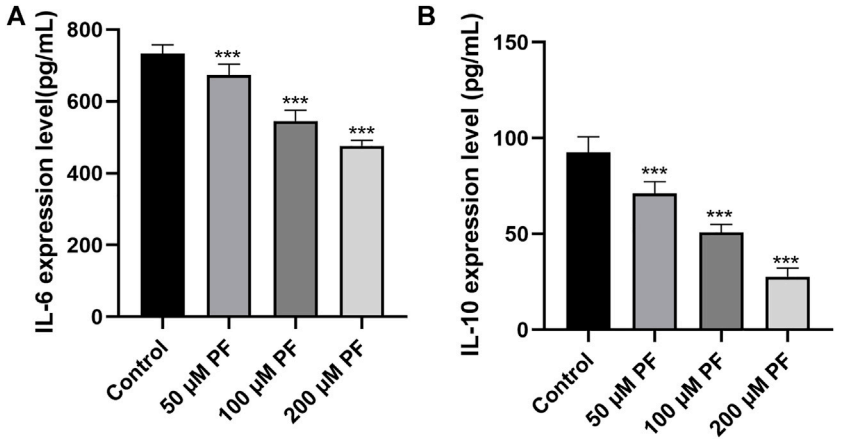


FIGURE 9
The contents of IL-6 and IL-10 in the culture supernatant of pancreatic cancer cells were determined by ELISA. **(A)** IL-6 expression in the culture supernatant of PANC-1 cells. **(B)** IL-10 expression in the culture supernatant of Capan-1 cells. *$p < 0.05$; ***$p < 0.001$, *versus* control group.

phosphorylation, as revealed by a standard western blot (Figures 10B–E).

4 Discussion

Due to the high lethality of pancreatic cancer, it is vital to create new therapies and medications for pancreatic cancer. It should be highlighted that the molecular process underlying the formation and progression of pancreatic cancer is complicated, including numerous proteins or pathways, and therefore a single targeted treatment may not have the desired therapeutic impact (Li et al., 2019). According to previous studies, a number of Chinese medicines and herbal extracts offer unique advantages in the treatment of pancreatic cancer (Triantafillidis et al., 2022). Quercetin, for instance, may block EMT by reducing TGF-β, cause cell death, and reduce the development of pancreatic cancer cells by downregulating c-Myc expression (Asgharian et al., 2021). Apigenin contains antioxidant and anti-inflammatory characteristics and can exert a therapeutic effect on PC cells *via* HIF, VEGF, and GLUT-1 (Ashrafizadeh et al., 2020). The medicinal effects of the chemicals derived from these plants are multitarget and multipathway. Then we set our sights on paeoniflorin, a natural compound with anti-inflammatory

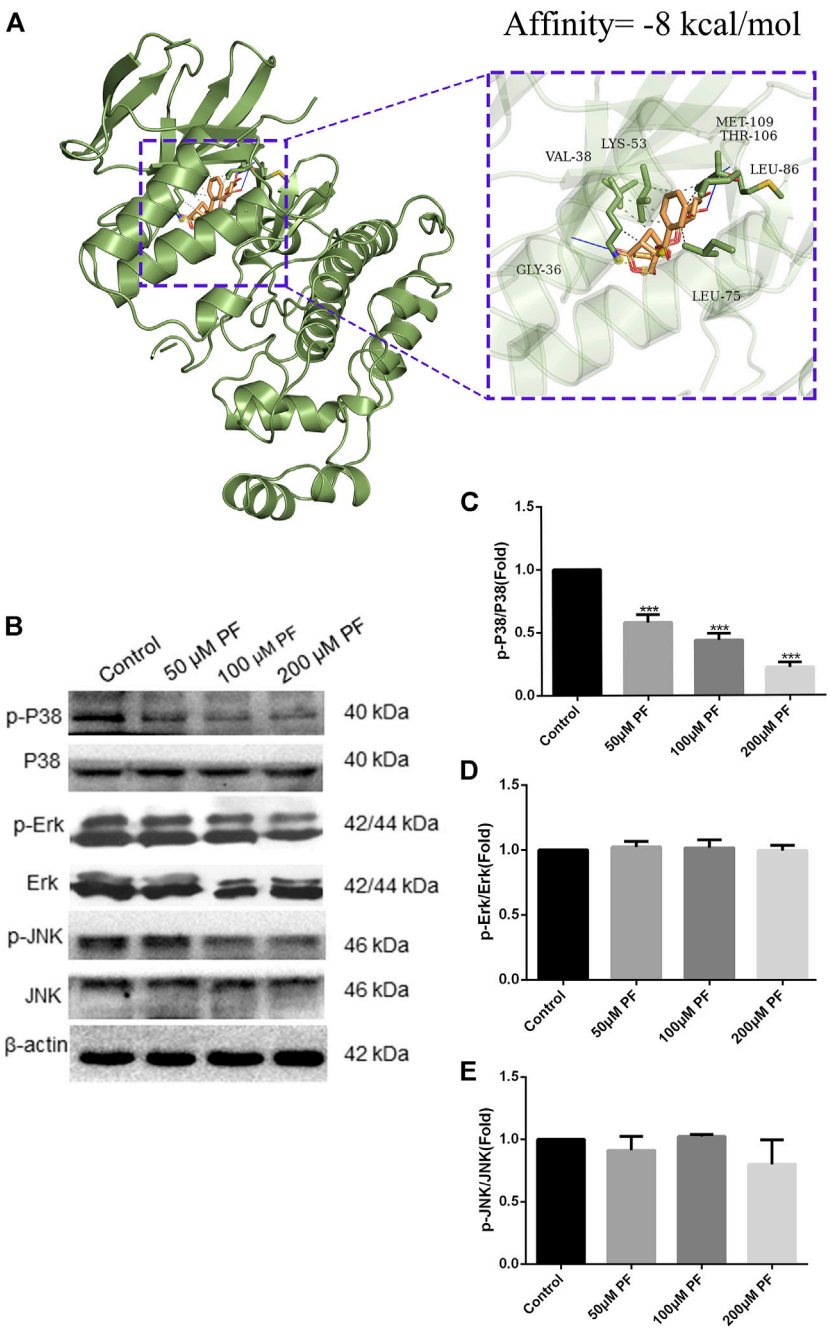


FIGURE 10
PF suppressed MAPK signal pathway in PANC-1. **(A)** PF exhibited good binding activity to p38 as determined by molecular docking. **(B)** Based on the results of Western blot analyses, PF was shown to decrease p38-related phosphorylation in MAPK signaling. This image displays some typical consequences from WB banding. **(C–E)** Each protein's relative expression level was measured using statistical methods: p-ERK, p-JNK, and p-P38. ***$p < 0.001$ vs. control group.

and anti-tumor effects (Zhou et al., 2020; Wang et al., 2022). Paeoniflorin has been demonstrated to inhibit the early phases of EMT induced by TGF-β. This may be accomplished by inhibiting the production of transcription factors Snail and Slug *via* the Smad pathway (Ji et al., 2016). Paeoniflorin impacts the progression of hepatocellular carcinoma, which is also a cancer of the digestive system, by downregulating the 5-HT1D inhibitory Wnt/β-Catenin pathway (Zhou et al., 2021). Moreover, pancreatic cancer growth is inhibited by paeoniflorin, which does so through increasing HTRA3 (Li

et al., 2017). So, we used network pharmacology to look into the possible pharmacological mechanism of paeoniflorin, thinking of it as a possible molecule for treating pancreatic cancer.

Previous network pharmacology is defined by the use of several internet databases to construct a network of multiple links between medications, targets, and disorders in order to study drug pharmacology (Jiao et al., 2021). For paeoniflorin, a substance isolated from plants with a well-defined chemical structure and characteristics, the usage of pharmacological data retrieval targets cannot be used to create an intersecting network. In addition, since certain databases are updated slowly, it is difficult to acquire more targets (Zhuang et al., 2019). To investigate the pharmacological effects of paeoniflorin in more depth, we downloaded the Microarray data of pancreatic cancer from the GEO database and screened its targets using the WGCNA method in this work. The genes in the modules screened by WGCNA were crossed with DEGs, and a total of 182 genes were obtained as target genes. KEGG and GO enrichment analysis were performed on these genes. The findings of the GO analysis suggested that epidermal growth and serine were crucial for PF intervention in pancreatic cancer. Serine promotes protein, amino acid, and glutathione production, which are critical for cell development and survival. Many tumor cells show dependence on exogenous serine and dietary serine (Yang and Vousden, 2016; Tajan et al., 2021). Pancreatic ductal adenocarcinoma patients had lower blood levels of essential and non-essential amino acids. PGAM1 knockdown increases 3-PG accumulation in serine-starved PDAC cells, resulting in increased cell proliferation and tumor formation (Itoyama et al., 2021). As for the traditional EGFR receptor, it is often referred to as ERBB-1. The MAPK pathway robustly stimulates the transcription and release of numerous ERBB ligands in pancreatic cancer (Mendelsohn and Baselga, 2003). Tumor-specific targeted delivery of 5FU using EGFR aptamers as the carrier achieved high target specificity, overcame 5 FU resistance (Mahajan et al., 2021). The KEGG analysis supports this conclusion. According to KEGG analysis, cancer MAPK signaling, PI3K-Akt signaling, and RAS signaling pathways could be regulated by PF. These pathways are intrinsically associated with the inflammatory response. Previous research has demonstrated that inflammation is present at every stage of tumor development (Greten and Grivennikov, 2019). In recent years, several endeavors have been made to discover the processes behind inflammation-induced carcinogenesis (Hausmann et al., 2014). Similarly, the GESA enrichment results were significantly correlated with GO and KEGG analyses, which were both enriched in inflammation-related aspects. Meanwhile, through algorithmic screening, we obtained 12 hub genes. The survival analysis of these genes indicated that the high expression of four of them (HBEGF, MMP1, MMP7, and MMP14) was indicative of a poor prognosis for overall survival. MMP1, MMP7, and MMP14 are all members of the family of Matrix Metalloproteinases. The MMP family drives tumor invasion and the creation of distant metastases, and is recognized to play a crucial role in a number of human malignancies, including pancreatic cancer (Huang et al., 2018). Expression of MMP1 in the early stages of a number of malignancies is correlated with a dismal prognosis. By activating MAPK pathways, RAS oncogenes may play a crucial role in the constitutive production of MMP1 in human pancreatic cancer cells (Huang et al., 2018, 1). MMP7 is involved in the injury response of mucosal epithelia and the degradation of extracellular matrix components, and it has been shown to be overexpressed in pancreatic ductal adenocarcinoma and its precursors, PanIN and intraductal papillary mucinous neoplasms, with MMP7 alterations evident even in intermediate-grade (Kartsonaki et al., 2022). IL-17 stimulates MMP7 expression in prostate cancer to destroy the E-cadherin/β-catenin complex and release -catenin, hence promoting EMT and tumor cell invasion (Zhang et al., 2017). Additionally, exosome-transferred MMP14 is a crucial facilitator of gemcitabine resistance in pancreatic cancer (Li et al., 2022).

Based on the results of CCK-8 and early studies, it was observed that PF intervenes in pancreatic cancer cells in a concentration-dependent way (in the range of 0–200 μm) in this experiment; hence, the 0–200 μm concentration was selected for further investigations. The presence of the IL-10 cytokine in the culture supernatant of PANC-1 cells was barely detectable in the first ELISA tests. This corroborates the findings of Graziella et al. (G et al., 2006). Capan-2 cells that released substantially higher IL-10 were therefore chosen. Then, we subjected PANC-1 and Capan-2 cells to 50, 100, and 200 μm dosages for 24 h and measured the levels of the cytokines IL-10 and IL-6, as well as the expression of MAPK pathway-related proteins. The results suggest that PF can have a therapeutic function by reducing the activation of inflammatory pathways and variables associated with pancreatic cancer. The activation of inflammatory signals can release a large number of cytokines, inflammatory mediators, and free radicals. Genes of the Interleukin 10 (IL-10) family play a dual, contentious function in a variety of cancers. According to a number of studies, an increase in IL-10 levels can dramatically drive tumor proliferation, metastasis, and immune evasion in a range of tumor models, including pancreatic cancer, and hence result in diverse pathologies (Zhuang et al., 2019). Our studies demonstrated that paeoniflorin dose-dependently inhibits IL-10 expression in pancreatic cancer cells. In addition, there is a complicated interaction effect between IL-10 and the MAPK pathway, whereby the activation of the MAPK pathway, which increases the production of IL-10 and TGF by cancer cells by maturing immune suppressive regulatory $CD4^+$ T-cells, results in a complex interaction effect (Hou et al., 2020). The activation of an ERK/JNK/P38 MAPK inflammatory pathway by IL-10 appears to enhance the

progression of chronic pancreatitis to pancreatic cancer. IL-6 can promote EMT by activating numerous molecular pathways of the JAK2/STAT3 and MAPK signaling axis, hence enhancing the metastatic potential of tumors (Abaurrea et al., 2021). In addition, our enrichment analysis revealed that the signaling pathways PI3K-AKT, and MAPK, etc. were related to pancreatic cancer. Since our hub genes are closely related to MAPK signaling pathway. Therefore, we chose MAPK signaling pathway for validation. Phosphorylation activates members of the Mitogen-activated protein kinase (MAPK) family, which includes ERK1/2, JNK, and p38. This activation causes the expression of target genes. Under normal circumstances, the expression of each member of the MAPK family maintains a dynamic balance, thereby maintaining cell proliferation and apoptosis in a balanced state. More and more studies have shown that the occurrence and development of tumors are related to the regulation of MAPK information transmission. In our study, paeoniflorin inhibited p38 but not ERK or JNK to affect pancreatic cancer. Wang et al. showed that p38gamma-MAPK promotes pancreatic cancer by activating PFKFB3 and GLUT2 through the KRAS oncogene signaling and aerobic glycolysis (Wang et al., 2020).

Nonetheless, we must recognize that PF is an unstable, water-soluble monoterpene glycoside that is strongly impacted by strong alkali and high temperatures. Therefore, when it is used in the body, whether it will have a decomposition reaction or *in volve* more complex pathways to play a role requires further research.

5 Conclusion

In conclusion, our research implemented a combination of bioinformatics, network pharmacology, and *in vitro* experiments to analyze the potential PF treatment pathways for pancreatic cancer. PF inhibits the proliferation of pancreatic cancer cells by interfering with the MAKP signaling pathway, IL-10, IL-6, and additional inflammatory factors. We provide some references for the development of PF as a follow-up treatment for pancreatic cancer.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: GSE97124, https://www.ncbi.nlm.nih.gov/geo/; The Cancer Genome Atlas, https://portal.gdc.cancer.gov/; Genotype-Tissue Expression, https://commonfund.nih.gov/GTEx.

Author contributions

CC gathered pertinent information and conducted bioinformatics correlation analysis. WZ and XC completed the *in vitro* experimental validation. CW finished the networkpharmacology analysis. WT and BS participated in the research design. WZ and CW wrote the manuscript. XR planned and directed this study.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.1032282/full#supplementary-material

References

Abaurrea, A., Araujo, A. M., and Caffarel, M. M. (2021). The role of the IL-6 cytokine family in epithelial-mesenchymal plasticity in cancer progression. *Int. J. Mol. Sci.* 22, 8334. doi:10.3390/ijms22158334

Asgharian, P., Tazehkand, A. P., Soofiyani, S. R., Hosseini, K., Martorell, M., Tarhriz, V., et al. (2021). Quercetin impact in pancreatic cancer: An overview on its therapeutic effects. *Oxid. Med. Cell. Longev.* 2021, 4393266. doi:10.1155/2021/4393266

Ashrafizadeh, M., Bakhoda, M. R., Bahmanpour, Z., Ilkhani, K., Zarrabi, A., Makvandi, P., et al. (2020). Apigenin as tumor suppressor in cancers: Biotherapeutic activity, nanodelivery, and mechanisms with emphasis on pancreatic cancer. *Front. Chem.* 8, 829. doi:10.3389/fchem.2020.00829

Bellone, G., Smirne, C., Mauri, F. A., Tonel, E., Carbone, A., Buffolino, A., et al. (2006). Cytokine expression profile in human pancreatic carcinoma cells and in surgical specimens: Implications for survival. *Cancer Immunol. Immunother.* 55, 684–698. doi:10.1007/s00262-005-0047-0

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The protein Data Bank. *Nucleic Acids Res.* 28, 235–242. doi:10.1093/nar/28.1.235

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, e71. doi:10.1093/nar/gkv1507

Gene Ontology Consortium (2015). Gene ontology Consortium: Going forward. *Nucleic Acids Res.* 43, D1049–D1056. doi:10.1093/nar/gku1179

Greten, F. R., and Grivennikov, S. I. (2019). Inflammation and cancer: Triggers, mechanisms and consequences. *Immunity* 51, 27–41. doi:10.1016/j.immuni.2019.06.025

Grossberg, A. J., Chu, L. C., Deig, C. R., Fishman, E. K., Hwang, W. L., Maitra, A., et al. (2020). Multidisciplinary standards of care and recent progress in pancreatic ductal adenocarcinoma. *Ca. Cancer J. Clin.* 70, 375–403. doi:10.3322/caac.21626

Hausmann, S., Kong, B., Michalski, C., Erkan, M., and Friess, H. (2014). The role of inflammation in pancreatic cancer. *Adv. Exp. Med. Biol.* 816, 129–151. doi:10.1007/978-3-0348-0837-8_6

Hou, P., Kapoor, A., Zhang, Q., Li, J., Wu, C.-J., Li, J., et al. (2020). Tumor microenvironment remodeling enables bypass of oncogenic KRAS dependency in pancreatic cancer. *Cancer Discov.* 10, 1058–1077. doi:10.1158/2159-8290.CD-19-0597

Hu, S., Sun, W., Wei, W., Wang, D., Jin, J., Wu, J., et al. (2013). Involvement of the prostaglandin E receptor EP2 in paeoniflorin-induced human hepatoma cell apoptosis. *Anticancer. Drugs* 24, 140–149. doi:10.1097/CAD.0b013e32835a4dac

Huang, C., Li, Y., Guo, Y., Zhang, Z., Lian, G., Chen, Y., et al. (2018). MMP1/PAR1/SP/NK1R paracrine loop modulates early perineural invasion of pancreatic cancer cells. *Theranostics* 8, 3074–3086. doi:10.7150/thno.24281

Ito, K., and Murphy, D. (2013). Application of ggplot2 to pharmacometric graphics. *CPT. Pharmacometrics Syst. Pharmacol.* 2, e79. doi:10.1038/psp.2013.56

Itoyama, R., Yasuda-Yoshihara, N., Kitamura, F., Yasuda, T., Bu, L., Yonemura, A., et al. (2021). Metabolic shift to serine biosynthesis through 3-PG accumulation and PHGDH induction promotes tumor growth in pancreatic cancer. *Cancer Lett.* 523, 29–42. doi:10.1016/j.canlet.2021.09.007

Jemal, A., Siegel, R., Xu, J., and Ward, E. (2010). Cancer statistics, 2010. *Ca. Cancer J. Clin.* 60, 277–300. doi:10.3322/caac.20073

Ji, Y., Dou, Y., Zhao, Q., Zhang, J., Yang, Y., Wang, T., et al. (2016). Paeoniflorin suppresses TGF-β mediated epithelial-mesenchymal transition in pulmonary fibrosis through a Smad-dependent pathway. *Acta Pharmacol. Sin.* 37, 794–804. doi:10.1038/aps.2016.36

Jia, X., Du, P., Wu, K., Xu, Z., Fang, J., Xu, X., et al. (2018). Pancreatic cancer mortality in China: Characteristics and prediction. *Pancreas* 47, 233–237. doi:10.1097/MPA.0000000000000976

Jiao, X., Jin, X., Ma, Y., Yang, Y., Li, J., Liang, L., et al. (2021). A comprehensive application: Molecular docking and network pharmacology for the prediction of bioactive constituents and elucidation of mechanisms of action in component-based Chinese medicine. *Comput. Biol. Chem.* 90, 107402. doi:10.1016/j.compbiolchem.2020.107402

Kamisawa, T., Wood, L. D., Itoi, T., and Takaori, K. (2016). Pancreatic cancer. *Lancet* 388, 73–85. doi:10.1016/S0140-6736(16)00141-0

Kanehisa, M., Sato, Y., Furumichi, M., Morishima, K., and Tanabe, M. (2019). New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* 47, D590–D595. doi:10.1093/nar/gky962

Kang, M. J., Jang, J.-Y., and Kim, S.-W. (2016). Surgical resection of pancreatic head cancer: What is the optimal extent of surgery? *Cancer Lett.* 382, 259–265. doi:10.1016/j.canlet.2016.01.042

Kartsonaki, C., Pang, Y., Millwood, I., Yang, L., Guo, Y., Walters, R., et al. (2022). Circulating proteins and risk of pancreatic cancer: A case-subcohort study among Chinese adults. *Int. J. Epidemiol.* 51, 817–829. doi:10.1093/ije/dyab274

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., et al. (2021). PubChem in 2021: New data content and improved web interfaces. *Nucleic Acids Res.* 49, D1388–D1395. doi:10.1093/nar/gkaa971

Klein, A. P. (2021). Pancreatic cancer epidemiology: Understanding the role of lifestyle and inherited risk factors. *Nat. Rev. Gastroenterol. Hepatol.* 18, 493–502. doi:10.1038/s41575-021-00457-x

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

Leroux, C., and Konstantinidou, G. (2021). Targeted therapies for pancreatic cancer: Overview of current treatments and new opportunities for personalized oncology. *Cancers (Basel)* 13, 799. doi:10.3390/cancers13040799

Li, S., Xu, H.-X., Wu, C.-T., Wang, W.-Q., Jin, W., Gao, H.-L., et al. (2019). Angiogenesis in pancreatic cancer: Current research status and clinical implications. *Angiogenesis* 22, 15–36. doi:10.1007/s10456-018-9645-2

Li, X., Li, K., Li, M., Lin, X., Mei, Y., Huang, X., et al. (2022). Chemoresistance transmission via exosome-transferred MMP14 in pancreatic cancer. *Front. Oncol.* 12, 844648. doi:10.3389/fonc.2022.844648

Li, Y., Gong, L., Qi, R., Sun, Q., Xia, X., He, H., et al. (2017). Paeoniflorin suppresses pancreatic cancer cell growth by upregulating HTRA3 expression. *Drug Des. devel. Ther.* 11, 2481–2491. doi:10.2147/DDDT.S134518

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425. doi:10.1016/j.cels.2015.12.004

Mahajan, U. M., Li, Q., Alnatsha, A., Maas, J., Orth, M., Maier, S. H., et al. (2021). Tumor-specific delivery of 5-fluorouracil-incorporated epidermal growth factor receptor-targeted aptamers as an efficient treatment in pancreatic ductal adenocarcinoma models. *Gastroenterology* 161, 996–1010.e1. doi:10.1053/j.gastro.2021.05.055

Mendelsohn, J., and Baselga, J. (2003). Status of epidermal growth factor receptor antagonists in the biology and treatment of cancer. *J. Clin. Oncol.* 21, 2787–2799. doi:10.1200/JCO.2003.01.504

Nipp, R., Tramontano, A. C., Kong, C. Y., Pandharipande, P., Dowling, E. C., Schrag, D., et al. (2018). Disparities in cancer outcomes across age, sex, and race/ethnicity among patients with pancreatic cancer. *Cancer Med.* 7, 525–535. doi:10.1002/cam4.1277

Rejhová, A., Opattová, A., Čumová, A., Slíva, D., Vodička, P., Rejhova, A., et al. (2018). Natural compounds and combination therapy in colorectal cancer treatment. *Eur. J. Med. Chem.* 144, 582–594. doi:10.1016/j.ejmech.2017.12.039

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43, e47. doi:10.1093/nar/gkv007

Rudin, C. M., Liu, W., Desai, A., Karrison, T., Jiang, X., Janisch, L., et al. (2008). Pharmacogenomic and pharmacokinetic determinants of erlotinib toxicity. *J. Clin. Oncol.* 26, 1119–1127. doi:10.1200/JCO.2007.13.1128

Seeliger, D., and de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J. Comput. Aided. Mol. Des.* 24, 417–422. doi:10.1007/s10822-010-9352-6

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, 2498–2504. doi:10.1101/gr.1239303

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer statistics, 2021. *Ca. Cancer J. Clin.* 71, 7–33. doi:10.3322/caac.21654

Strobel, O., Neoptolemos, J., Jäger, D., and Büchler, M. W. (2019). Optimizing the outcomes of pancreatic cancer surgery. *Nat. Rev. Clin. Oncol.* 16, 11–26. doi:10.1038/s41571-018-0112-1

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, 15545–15550. doi:10.1073/pnas.0506580102

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

Tajan, M., Hennequart, M., Cheung, E. C., Zani, F., Hock, A. K., Legrave, N., et al. (2021). Serine synthesis pathway inhibition cooperates with dietary serine and glycine limitation for cancer therapy. *Nat. Commun.* 12, 366. doi:10.1038/s41467-020-20223-y

Triantafillidis, J. K., Triantafyllidi, E., Sideris, M., Pittaras, T., and Papalois, A. E. (2022). Herbals and plants in the treatment of pancreatic cancer: A systematic review of experimental and clinical studies. *Nutrients* 14, 619. doi:10.3390/nu14030619

Wang, F., Qi, X.-M., Wertz, R., Mortensen, M., Hagen, C., Evans, J., et al. (2020). p38γ MAPK is essential for aerobic glycolysis and pancreatic tumorigenesis. *Cancer Res.* 80, 3251–3264. doi:10.1158/0008-5472.CAN-19-3281

Wang, L., He, T., Liu, J., Tai, J., Wang, B., Zhang, L., et al. (2021). Revealing the immune infiltration landscape and identifying diagnostic biomarkers for lumbar disc herniation. *Front. Immunol.* 12, 666355. doi:10.3389/fimmu.2021.666355

Wang, X. Z., Xia, L., Zhang, X. Y., Chen, Q., Li, X., Mou, Y., et al. (2022). The multifaceted mechanisms of Paeoniflorin in the treatment of tumors: State-of-the-Art. *Biomed. Pharmacother.* 149, 112800. doi:10.1016/j.biopha.2022.112800

Wang, Y., Hu, G., Zhang, Q., Tang, N., Guo, J., Liu, L., et al. (2016). Efficacy and safety of gemcitabine plus erlotinib for locally advanced or metastatic pancreatic cancer: A systematic review and meta-analysis. *Drug Des. devel. Ther.* 10, 1961–1972. doi:10.2147/DDDT.S105442

Wong, H. H., and Lemoine, N. R. (2009). Pancreatic cancer: Molecular pathogenesis and new therapeutic targets. *Nat. Rev. Gastroenterol. Hepatol.* 6, 412–422. doi:10.1038/nrgastro.2009.89

Wu, S.-H., Wu, D.-G., and Chen, Y.-W. (2010). Chemical constituents and bioactivities of plants from the genus Paeonia. *Chem. Biodivers.* 7, 90–104. doi:10.1002/cbdv.200800148

Yang, M., and Vousden, K. H. (2016). Serine and one-carbon metabolism in cancer. *Nat. Rev. Cancer* 16, 650–662. doi:10.1038/nrc.2016.81

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, 284–287. doi:10.1089/omi.2011.0118

Yue, M., Li, S., Yan, G., Li, C., and Kang, Z. (2018). Paeoniflorin inhibits cell growth and induces cell cycle arrest through inhibition of FoxM1 in colorectal cancer cells. *Cell Cycle* 17, 240–249. doi:10.1080/15384101.2017.1407892

Zhang, D., Park, D., Zhong, Y., Lu, Y., Rycaj, K., Gong, S., et al. (2016a). Stem cell and neurogenic gene-expression profiles link prostate basal cells to aggressive prostate cancer. *Nat. Commun.* 7, 10798. doi:10.1038/ncomms10798

Zhang, Q., Liu, S., Parajuli, K. R., Zhang, W., Zhang, K., Mo, Z., et al. (2017). Interleukin-17 promotes prostate cancer via MMP7-induced epithelial-to-mesenchymal transition. *Oncogene* 36, 687–699. doi:10.1038/onc.2016.240

Zhang, Q., Yuan, Y., Cui, J., Xiao, T., and Jiang, D. (2016b). Paeoniflorin inhibits proliferation and invasion of breast cancer cells through suppressing Notch-1 signaling pathway. *Biomed. Pharmacother.* 78, 197–203. doi:10.1016/j.biopha.2016.01.019

Zhou, Y.-X., Gong, X.-H., Zhang, H., and Peng, C. (2020). A review on the pharmacokinetics of paeoniflorin and its anti-inflammatory and immunomodulatory effects. *Biomed. Pharmacother.* 130, 110505. doi:10.1016/j.biopha.2020.110505

Zhou, Y., Liu, X., Gao, Y., Tan, R., Wu, Z., Zhong, Q., et al. (2021). Paeoniflorin affects hepatocellular carcinoma progression by inhibiting wnt/β-catenin pathway through downregulation of 5-ht1d. *Curr. Pharm. Biotechnol.* 22, 1246–1253. doi:10.2174/1389201021666201009153808

Zhuang, J., Holay, M., Park, J. H., Fang, R. H., Zhang, J., and Zhang, L. (2019). Nanoparticle delivery of immunostimulatory agents for cancer immunotherapy. *Theranostics* 9, 7826–7848. doi:10.7150/thno.37216

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 10 November 2022
DOI 10.3389/fphar.2022.1015842

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Zhixiang Yu,
Fourth Military Medical University, China
Min Wei,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Wei Zhang,
weizhang_tj@126.com

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal *Frontiers in Pharmacology*

RECEIVED 10 August 2022
ACCEPTED 18 October 2022
PUBLISHED 10 November 2022

CITATION
Zhang Z, Chen W, Luo C and Zhang W (2022), Exploring a four-gene risk model based on doxorubicin resistance-associated lncRNAs in hepatocellular carcinoma.
Front. Pharmacol. 13:1015842.
doi: 10.3389/fphar.2022.1015842

Exploring a four-gene risk model based on doxorubicin resistance-associated lncRNAs in hepatocellular carcinoma

Zunyi Zhang, Weixun Chen, Chu Luo and Wei Zhang*

Hepatic Surgery Center, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Background: Liver cancer is a lethal cancer type among which hepatocellular carcinoma (HCC) is the most common manifestation globally. Drug resistance is a central problem impeding the efficiency of HCC treatment. Long non-coding RNAs reportedly result in drug resistance. This study aimed to identify key lncRNAs associated with doxorubicin resistance and HCC prognosis.

Materials and Methods: HCC samples with gene expression profiles and clinical data were accessed from public databases. We applied differential analysis to identify key lncRNAs that differed between HCC and normal samples and between drug-fast and control samples. We also used univariate Cox regression analysis to screen lncRNAs or genes associated with HCC prognosis. The least absolute shrinkage and selection operator (LASSO) was used to identify the key prognostic genes. Finally, we used receiver operating characteristic analysis to validate the effectiveness of the risk model.

Results: The results of this study revealed RNF157-AS1 as a key lncRNA associated with both doxorubicin resistance and HCC prognosis. Metabolic pathways such as fatty acid metabolism and oxidative phosphorylation were enriched in RNF157-AS1-related genes. LASSO identified four protein-coding genes—*CENPP*, *TSGA10*, *MRPL53*, and *BFSP1*—to construct a risk model. The four-gene risk model effectively classified HCC samples into two risk groups with different overall survival. Finally, we established a nomogram, which showed superior performance in predicting the long-term prognosis of HCC.

Conclusion: RNF157-AS1 may be involved in doxorubicin resistance and may serve as a potential therapeutic target. The four-gene risk model showed potential for the prediction of HCC prognosis.

KEYWORDS

hepatocellular carcinoma, long non-coding RNAs, RNF157-AS1, risk model, lncRNAs

Introduction

Primary liver cancer is the seventh most frequently diagnosed cancer worldwide and shows a high mortality rate among all cancer types (Sung et al., 2021). Asia and Africa are two regions with large numbers of cases (Sung et al., 2021). Hepatocellular carcinoma (HCC) accounts for 85–90% of all cases of liver cancer (McGlynn et al., 2021). Hepatitis B virus (HBV) infection largely contributes to the development of HCC; other factors such as hepatitis C virus (HCV) infection, alcohol use, nonalcoholic fatty hepatitis, and liver cirrhosis can also increase the risk of HCC. Surgical treatment is a recognized strategy for HCC treatment, in which hepatectomy and liver transplantation are two major methods to increase survival time (Zhou et al., 20192020). However, for patients with metastatic HCC, surgical treatment is not effective. Chemotherapy or radiotherapy is the main treatment choice for killing metastatic cancer cells.

Drugs for systemic chemotherapy, including sorafenib, lenvatinib, and oxaliplatin, have demonstrated effectiveness as first-line therapy against late-stage HCC in clinical studies and are included in treatment guidelines (Zhou et al., 20192020). However, the use of single drugs is unsatisfactory and drug resistance is common with the continuous use of these drugs (Lohitesh et al., 2018). For example, only around 30% of patients with HCC can benefit from sorafenib and show only 6 months of drug response (Tang et al., 2020). Therefore, understanding drug resistance is required to optimize these chemotherapeutic drugs. Recent studies suggest the roles of epigenetics, regulated RNAs, cell death, and the tumor microenvironment in the occurrence of drug resistance in HCC (Ding et al., 2018; Liang et al., 2020; Tang et al., 2020).

Long non-coding RNAs (lncRNAs) are involved in multiple biological or pathological processes regulating cell proliferation, cell death, immune response, and others, which also contribute to drug resistance in HCC (Ding et al., 2018). For example, the knockdown of lncARSR inhibited PTEN expression and activated the PI3K/Akt pathway during the treatment of HCC, which was involved in doxorubicin resistance (Li et al., 2017). Tsang and Kwok reported that H19 knockdown suppressed MDR1 expression and increased HepG2 cell sensitivity to doxorubicin (Tsang and Kwok, 2007). Zhou et al. (2017)observed that the knockdown of HOTAIR lncRNA increased cisplatin efficiency on HCC cells by regulating the STAT3/ABCB1 signaling pathway. Therefore, lncRNAs play a critical role in drug resistance and HCC progression, and it is crucial to identify lncRNAs related to doxorubicin resistance to guide HCC treatment. Therefore, in the current study, we aimed to identify key lncRNAs related to drug resistance in HCC based on expression profiles of HCC and to construct a lncRNA-related risk model to predict HCC prognosis.

Materials and methods

Datasets

The Cancer Genome Atlas-Liver Hepatocellular Carcinoma (TCGA-LIHC) dataset, which contains RNA sequencing (RNA-seq) data and clinical information from HCC and paracancerous samples, was downloaded through the TCGA GDC API. The GSE76427 and GSE125180 datasets were downloaded from Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE76427, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE125180). The GSE76427 dataset contains microarray data from HCC samples, while the GSE125180 dataset contains microarray data from three doxorubicin drug-fast samples and three control samples.

RNA-seq and microarray data processing

For RNA-seq data in the TCGA dataset, we included samples with data on overall survival and survival status. The Ensembl ID of each gene was transferred to the gene symbol. We determined the average expression level for genes with multiple gene symbols. After processing, 360 primary HCC samples and 50 paracancerous samples remained.

For microarray data in the GEO datasets, normal samples were removed. Probes were converted to gene symbols using the platform annotation file. Probes corresponding to multiple genes were removed. Samples lacking data on survival time and survival status in the GSE76427 dataset were excluded. Finally, 115 HCC samples from the GSE76427 dataset remained.

Acquisition of lncRNA data

Gene transfer format (GTF) file (v32) was obtained from the GENCODE website (https://www.gencodegenes.org/). The expression data from the TCGA and GSE76427 datasets were converted to mRNA and lncRNA data. Mutual lncRNA data from the two datasets were included in the analyses.

Identification of key lncRNAs related to drug resistance and HCC prognosis

The limma R package (Ritchie et al., 2015) was applied to perform differential analysis and identify differentially expressed lncRNAs (DElncRNAs) between HCC and paracancerous samples in the TCGA dataset, drug-fast and control samples of GSE76427. $p < 0.05$ and |log2(fold change)|>log2(1.2) (Chen et al., 2021) were determined to screen DElncRNAs. The intersection of DElncRNAs in two datasets was included, and univariate Cox regression analysis was further performed on the

screened DElncRNAs, with $p < 0.05$ indicating potential prognostic DElncRNAs.

Pathway analysis

We downloaded the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways from the KEGG website (https://www.genome.jp/kegg/pathway.html). Single-sample gene set enrichment analysis (ssGSEA) was performed for each HCC sample in the TCGA dataset using the ssGSEA algorithm in the GSVA R package (Hänzelmann et al., 2013).

Evaluation of tumor microenvironment and the relationships between DElncRNAs and immune infiltration

We used CIBERSORT (Chen et al., 2018) to estimate the proportions of 22 immune-related cells and the ESTIMATE tool (Yoshihara et al., 2013) to calculate the immune and stromal scores. We applied the Hmisc R package (https://cran.r-project.org/web/packages/Hmisc/index.html) to conduct Pearson correlation analyses to assess the correlations between DElncRNA expression and immune infiltration. WebGestaltR package (Liao et al., 2019) was used to perform enrichment analysis on gene ontology (GO) and KEGG pathways using a false discovery rate (FDR) of <0.05 to screen for enriched GO terms and pathways.

Construction of a risk model related to key DElncRNAs

We first performed Pearson correlation analysis using the "rcorr" function in the Hmisc R package to screen for genes significantly correlated with the key DElncRNAs (|correlation coefficient| > 0.4 and $p < 0.001$). Next, we used the survival R package to perform a univariate Cox regression analysis of the screened genes, with genes with $p < 0.001$ defined as potential prognostic genes. Least absolute shrinkage and selection operator (LASSO) Cox regression analysis was conducted using the glmnet R package (Friedman et al., 2010) to decrease the number of prognostic genes. For each variable (prognostic gene), a trajectory of coefficient variation with the changing lambda value was visualized. With increasing lambda values, the coefficients of prognostic genes approached zero. We applied ten-fold cross-validation to construct the model, which was defined as follows: Risk Score = $\Sigma\beta i \times Expi$, where βi represents the Cox coefficients of the prognostic genes and Expi represents the expression levels. For each sample, a risk score was calculated and converted to a z-score. According to z-score = 0, the samples were divided into high-risk (z-score > 0) and low-risk (z-score < 0) groups. We conducted Kaplan–Meier survival analysis to draw survival plots of the two risk groups. The TimeROC R package (Blanche et al., 2013) was used to examine the efficiency of the risk model in predicting overall survival.

Statistical analysis

The bioinformatics analysis in this study was supported by the Sangerbox tool (http://vip.sangerbox.com/) (Shen et al., 2022). All statistical analyses were conducted in R software (v4.1). Wilcoxon tests were performed for comparisons between groups. Log-rank tests were used in the Cox regression and survival analyses. $p < 0.05$ was defined as statistical significance.

Results

Screening lncRNAs related to both drug sensitivity and HCC prognosis

We first screened for DElncRNAs between the drug-fast and control groups in the GSE125180 dataset by differential analysis. We identified a total of 28 DElncRNAs, including 18 that were upregulated and 10 that were downregulated ($p < 0.01$, |log2(fold change)| > log2(1.2); Supplementary Figure S1A). We then performed the same analysis between HCC and paracancerous samples in the TCGA dataset, which revealed 534 upregulated and 49 downregulated DElncRNAs (Supplementary Figure S1B). The Venn plot showed that three DElncRNAs (HNF4A-AS1, RNF157-AS1, and LINC00488) were found in both datasets (Figure 1A). The expression of the three lncRNAs was upregulated in the drug-fast group compared to that in the control group ($p < 0.05$; Figure 1B). These lncRNAs also displayed differential expression levels between HCC and paracancerous samples ($p < 0.05$; Figure 1C). To explore the relationship between three lncRNAs and HCC prognosis, we classified the HCC samples as having high or low expression based on the median cut-off value of the expression of the three lncRNAs. Survival analysis revealed that the high- and low-expression groups of HNF4A-AS1 and RNF157-AS1 had differential overall survival ($p < 0.05$; Figure 1D). However, univariate Cox regression analysis showed that only RNF157-AS1 expression was independently associated with overall survival ($p = 0.015$, hazard ratio (HR) = 1.231; Figure 1E).

Potential pathways related to RNF157-AS1

As we identified that RNF157-AS1 expression was dysregulated in drug-fast and HCC samples and was also an independent risk factor, we considered RNF157-AS1 as an

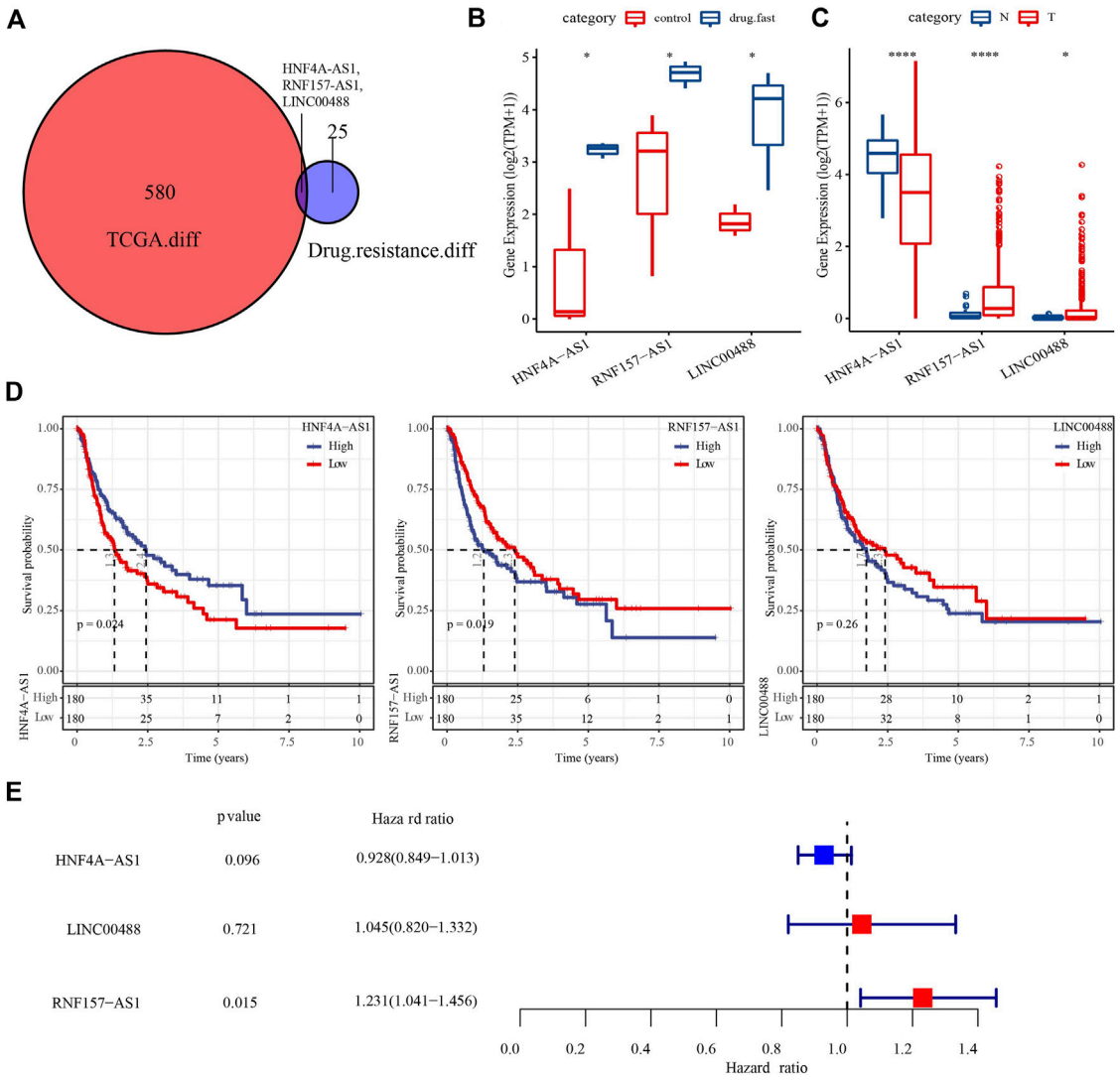


FIGURE 1
Identification of key DElncRNAs related to drug resistance and HCC prognosis. **(A)** Venn plot of DElncRNAs in the TCGA and GSE125180 datasets. **(B)** Expression of three DElncRNAs in control and drug-fast samples from the GSE125180 dataset. Wilcoxon tests were conducted. **(C)** Expression of three DElncRNAs in HCC and paracancerous samples in the TCGA dataset. Wilcoxon tests were conducted. **(D)** Kaplan–Meier survival plots of high- and low-expression of three DElncRNAs. **(E)** Univariate Cox regression analysis of three DElncRNAs. Log-rank tests were conducted. *$p < 0.05$, ****$p < 0.0001$.

important lncRNA involved in HCC development and drug resistance. Therefore, we further explored the functional pathways with which it was correlated. The enrichment score for all hallmark pathways was calculated for each HCC sample. We further performed a correlation analysis between the pathway score and the RNF157-AS1 expression level. We identified 10 functional pathways that were associated with RNF157-AS1 expression (Figure 2A). For example, bile acid metabolism, fatty acid metabolism, and oxidative phosphorylation were negatively associated with RNF157-AS1 expression, while MYC target V1 and WNT-β catenin signaling were positively associated with RNF157-AS1 expression. A heatmap also showed the correlations between the 10 pathways and RNF157-AS1 expression, consistent with Figure 2A (Figure 2B).

Association between RNF157-AS1 and immune infiltration

The characteristics of the tumor microenvironment reflect cancer development and prognosis, as well as the sensitivity to

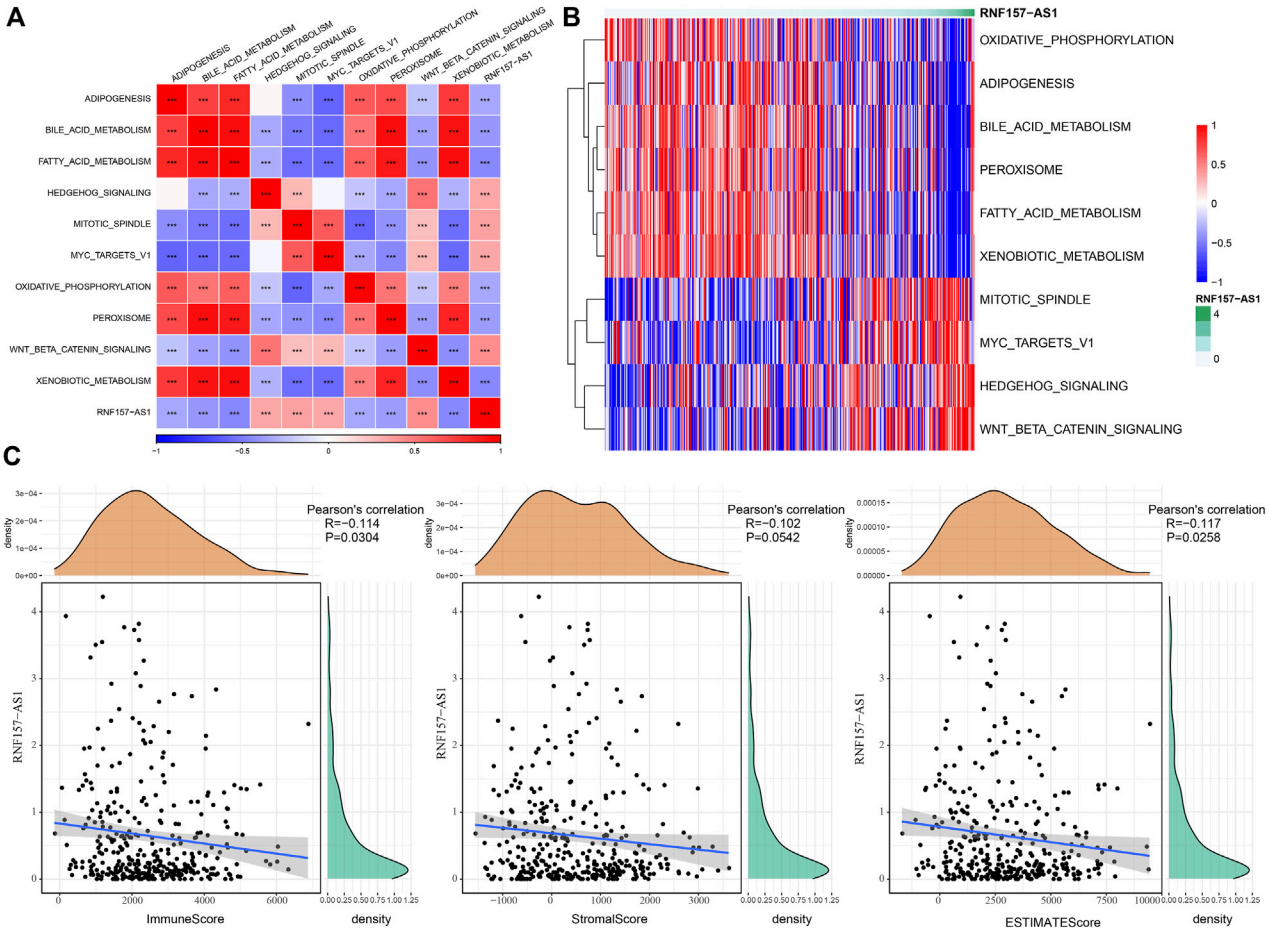


FIGURE 2
Functional pathways related to RNF157-AS1 and the relationship between RNF157-AS1 and immune infiltration. **(A)** Correlation analysis between the enrichment of pathways and RNF157-AS1 expression. Red and blue indicate positive and negative correlations, respectively. **(B)** Heatmap showing the enrichment score of pathways ranked by increasing RNF157-AS1 expression. Darker green: higher expression. Red and blue: high and low expression levels, respectively. **(C)** Pearson correlation analysis of RNF157-AS1 with immune, stromal, and ESTIMATE scores.

cancer treatment. Therefore, we assessed the potential correlation between RNF157-AS1 and the tumor microenvironment in HCC. The CIBERSORT results indicated that M2 macrophages, CD8 T cells, and gamma delta T cells were negatively correlated with RNF157-AS1 expression, while naïve CD4 T cells, M0 macrophages, eosinophils, and neutrophils were positively correlated with RNF157-AS1 expression ($p < 0.05$; Supplementary Figure S2A). The results of the ESTIMATE analysis further demonstrated a significant correlation between RNF157-AS1 expression and immune infiltration (Figure 2C). In addition, we calculated the ssGSEA scores of immune-related pathways and discovered that complement and coagulation cascades, natural killer cell-mediated cytotoxicity, and FC gamma-R-mediated phagocytosis were associated with RNF157-AS1 expression (Supplementary Figure S2B). Together, these results suggested that RNF157-AS1 was involved in immune infiltration regulation through some immune-related pathways.

Identification of prognostic genes related to RNF157-AS1

We performed Pearson correlation analysis to identify protein-coding genes (PCGs) associated with RNF157-AS1 based on their expression levels. The results revealed 1,498 PCGs that were significantly associated with RNF157-AS1 (correlation coefficient $|>0.4$, $p < 0.001$). Functional analysis revealed enrichment of mRNA-related terms, cell cycle, and RNA transport in the 1,498 PCGs (Supplementary Figures S3A–D). Univariate Cox regression analysis to identify prognostic genes identified 166 risk genes that were associated with HCC prognosis (Supplementary Figure S4A). Subsequently,

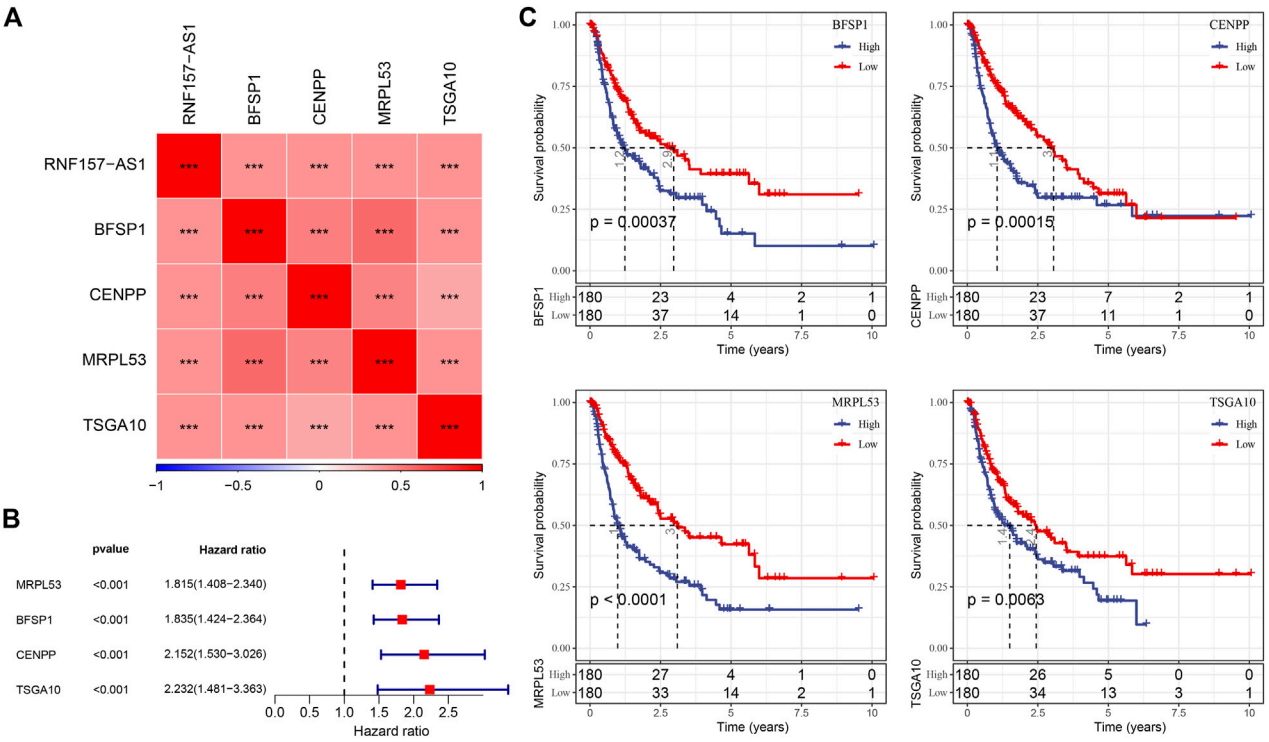


FIGURE 3
Relationship of RNF157-AS1 with four key PCGs and the association between four PCGs and HCC prognosis. **(A)** Pearson correlation analysis between the expression of RNF157-AS1 and four lncRNAs. Red and blue: positive and negative correlations, respectively. **(B)** Univariate Cox regression analysis of four lncRNAs. **(C)** Kaplan–Meier survival plots of high- and low-expression groups in the four lncRNAs. Log-rank tests were conducted. ***$p < 0.001$.

we implemented LASSO regression analysis to determine the key prognostic genes for constructing an optimal risk model. At a lambda value of 0.0708, the model reached the optimal status (Supplementary Figures S4B,C). Finally, four risk genes remained, including *CENPP*, *TSGA10*, *MRPL53*, and *BFSP1* (Supplementary Figure S4D). The expression levels of these four genes were related to RNF157-AS1 expression ($p < 0.001$; Figure 3A). Cox regression analysis showed that the four genes were independent risk factors ($p < 0.001$, HR > 1.8; Figure 3B). Survival analysis showed distinct overall survival between the high- and low-expression groups of all four genes ($p < 0.01$; Figure 3C), indicating that these four genes were highly associated with HCC prognosis. Therefore, we included these genes in the construction of the risk model.

Validation of the four-gene risk model

For each HCC sample, we calculated the risk score and validated the efficiency of the risk model in predicting prognosis by ROC analysis. The results showed that the risk model had a good performance in evaluating 1- and 5-year prognosis, with AUC values of 0.69 and 0.70, respectively, in the TCGA dataset (Figure 4A). HCC samples were classified into two risk groups (high and low risk) based on the cut-off (z-score = 0). The Kaplan–Meier survival plots showed a significant difference in the prognosis of the two risk groups ($p < 0.0001$; Figure 4A). In the GSE76427 dataset, we observed similar results, which suggested the robust performance of the risk model (Figure 4B). Notably, we observed a markedly positive correlation between RNF157-AS1 expression and risk score (R = 0.532; Figure 4C). We then evaluated the distribution of RNF157-AS1 expression levels and risk scores in different clinical features. The RNF157-AS1 expression level was much higher in G3 and G4 compared to that of G1 and G2 ($p < 0.0001$; Figure 4D). Moreover, severe clinical stages had higher risk scores than moderate stages ($p < 0.001$; Figure 4E), suggesting the reliability of the risk model.

Establishing a nomogram based on the four-gene risk model and clinical features

After demonstrating the robustness and reliability of the four-gene risk model related to RNF157-AS1, we next tried to further increase the accuracy of its application in the clinical

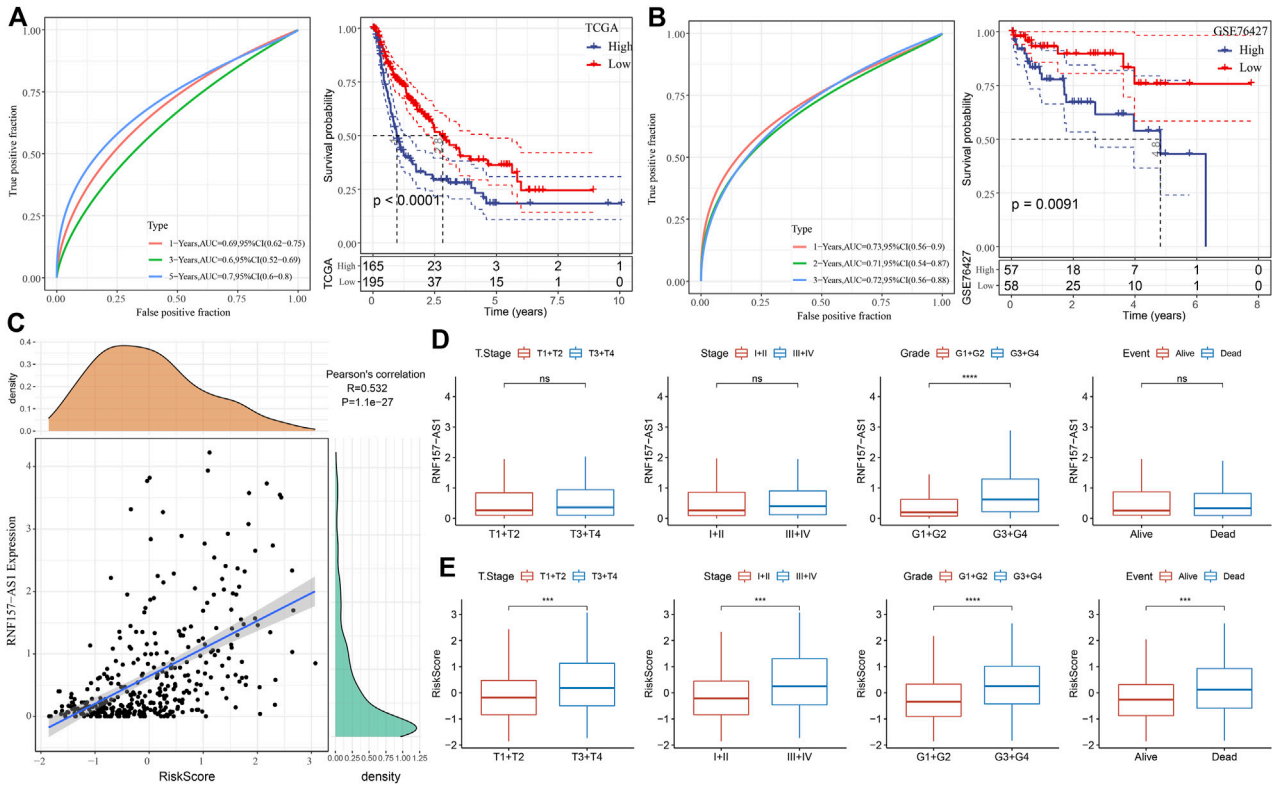


FIGURE 4
Performance of the four-gene risk model. **(A)** ROC analysis of survival plots of the risk model in the TCGA dataset. **(B)** ROC analysis of survival plots of the risk model in the GSE76427 dataset. **(C)** Pearson correlation analysis between RNF157-AS1 expression and risk score. **(D)** RNF157-AS1 expression of HCC samples grouped by different clinical features. **(E)** Distribution of risk scores for different clinical features. Log-rank tests were conducted in **(A, B)**. Wilcoxon tests were performed in **(D, E)**. ns: no significance. ***$p < 0.01$, ****$p < 0.0001$.

setting. Cox regression analysis of the clinical features and risk score showed that only stage and risk score were independent risk factors of HCC prognosis (Figures 5A,B). Therefore, we used these factors to establish a nomogram for predicting 1-, 3-, and 5-year death rates (Figure 5C). The calibration curve showing the predicted overall survival was similar to the observed values (Figure 5D). Decision curve analysis (DCA) further demonstrated that patients could benefit more from the nomogram compared to other indicators (Figure 5E). Finally, ROC analysis revealed the superior performance of the risk score and nomogram for predicting long-term overall survival (Figure 5F).

Discussion

Trans-arterial chemoembolization (TACE), an effective strategy for patients with unresectable HCC, requires the injection of embolizing agents in combination with doxorubicin (Fong and Tanabe, 2014). TACE is a recognized management scheme for patients with intermediate-stage HCC (Bruix and Sherman, 2011). The combination of the embolic effect with doxorubicin suppressed tumor progression and improved the prognosis of patients with HCC in randomized clinical trials (Llovet et al., 2002; Malagari et al., 2010). Moreover, doxorubicin-based TACE has a positive effect on reducing tumor size, which provides conditions benefitting liver transplantation (Fong and Tanabe, 2014). Nevertheless, resistance to doxorubicin markedly limits the treatment efficiency of patients with HCC, with only 27% of tumors treated with TACE showing a complete response and approximately 50% showing no response (Lammer et al., 2010). Consequently, it is of great importance to understand the mechanisms of resistance to doxorubicin in HCC. As a novel gene regulator, lncRNA is closely related to the occurrence, development, and prognosis of human disease, especially cancer. The abnormal expression of some lncRNAs may be related to the overgrowth, apoptosis inhibition, invasion, metastasis, and poor prognosis of HCC cells (Li et al., 2018; Pan et al., 2019; Huang et al., 2020). The present study is the first to comprehensively analyze the potential lncRNAs and genes associated with doxorubicin resistance in HCC.

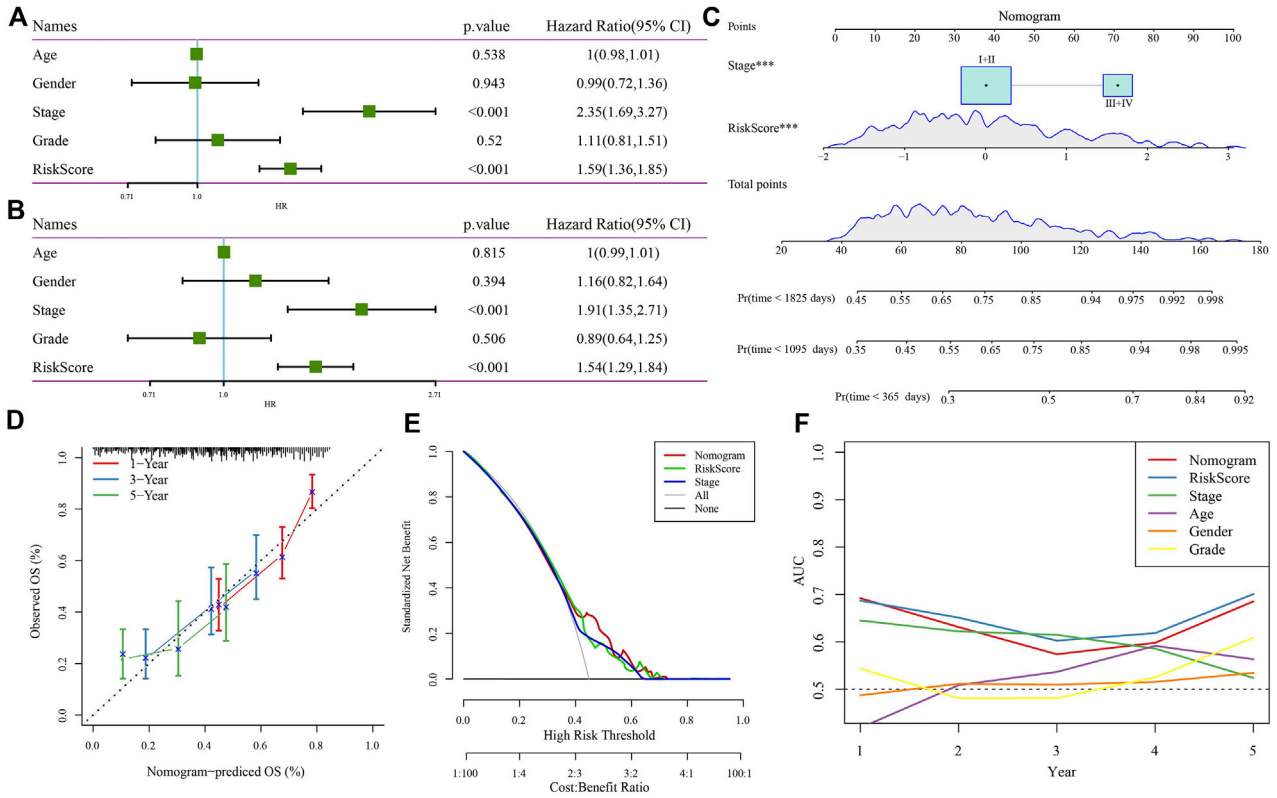


FIGURE 5
Nomogram construction and evaluation. Univariate **(A)** and multivariate **(B)** Cox regression analysis of the risk score, age, sex, stage, and grade. **(C)** Nomogram based on risk score and stage for predicting death rate at 1, 3, and 5 years. **(D)** Calibration curve for predicting the accuracy of the nomogram. **(E)** DCA plot of the nomogram, risk score, and stage. **(F)** AUC of the nomogram, risk score, and clinical features.

In this study, we compared RNA-seq data from HCC samples and doxorubicin-resistant HCC samples to control data to identify key lncRNAs. Differential analysis between HCC and paracancerous samples and between doxorubicin-resistant and control samples revealed three candidate DElncRNAs. However, only RNF157-AS1 was an independent risk factor significantly associated with overall survival in HCC. Functional analysis revealed that some metabolic pathways, including bile acid metabolism, fatty acid metabolism, and oxidative phosphorylation, were negatively correlated with RNF157-AS1 expression. These metabolic pathways are involved in the pathogenesis or immune response of HCC (Wang et al., 2016; Ma et al., 2018; Liu et al., 2020); thus, RNF157-AS1 may regulate HCC progression and induce doxorubicin resistance by interacting with these metabolic pathways.

RNF157-AS1 has not been specifically reported in HCC or other cancer types. Some studies focusing on constructing gene signatures for cancer have identified RNF157-AS1 as a prognostic lncRNA in the signature. For instance, Lin et al. (2021) identified a five-lncRNA signature including RNF157-AS1 for predicting ovarian cancer prognosis. Jiang et al. (2020) reported RNF157-AS1 as a candidate gene related to HBV-based HCC.

To understand the potential mechanism of RNF157-AS1 in HCC development, we identified a series of RNF157-AS1-associated PCGs. Function analysis showed significant enrichment of the mRNA surveillance pathway, RNA transport, and the cell cycle in RNF157-AS1-associated PCGs, implying that RNF157-AS1 may be involved in regulating mRNA-related biological processes and cell proliferation. Based on these RNF157-AS1-associated PCGs, we constructed a risk model containing four prognostic PCGs (*CENPP*, *TSGA10*, *MRPL53*, and *BFSP1*). The four-gene risk model showed reliable performance and a high AUC score in different datasets. The result also showed that the risk score was an independent indicator of overall survival in HCC. In addition, we established a nomogram based on the HCC risk score and stage to increase the accuracy in the prediction of HCC prognosis. Compared to other indicators, the nomogram showed superior performance in the DCA plot, indicating that the nomogram could benefit patients with HCC.

Among these four genes, only *TSGA10* has been reported relatively often in cancer; the other three genes were less often reported. Tanaka et al. (2004) observed TSGA10 overexpression in 4 of 20 patients with HCC, two of whom showed antibodies against recombinant TSGA10 protein. TSGA10 was also suggested as a potential biomarker in cancer tumorigenesis (Mobasheri et al., 2007). Dianatpour et al. (2012) observed upregulated TSGA10 expression in breast cancer cell lines and suggested its important role in breast cancer proliferation and prognosis. However, the association between RNF157-AS1 and these four genes requires further analysis. The four-gene risk model also requires further verification in clinical cases. However, this study has several limitations. For example, the sample size of the included dataset is limited, which may lead to bias in the screening results and biological function analysis. Moreover, the prognostic model requires further experimental verification.

Conclusion

In conclusion, the results of this study identified a key lncRNA (RNF157-AS1) that may contribute to doxorubicin resistance by involving metabolic pathways in HCC. Based on RNF157-AS1-associated PCGs, we constructed a four-gene risk model that reliably predicted HCC prognosis.

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

ZZ designed the study and contributed to the literature research. WC analyzed and interpreted the data. CL wrote the initial draft of the manuscript. WZ reviewed and edited the manuscript. All authors read and approved the manuscript.

Funding

The research was supported by a Project of the Natural Science Foundation of China (81860117).

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/fphar.2022.1015842/full#supplementary-material

References

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

Bruix, J., and Sherman, M. (2011). Management of hepatocellular carcinoma: An update. *Hepatol. Baltim. Md)* 53 (3), 1020–1022. doi:10.1002/hep.24199

Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Chen, J. H., Wang, L. L., Tao, L., Qi, B., Wang, Y., Guo, Y. J., et al. (2021). Identification of MYH6 as the potential gene for human ischaemic cardiomyopathy. *J. Cell. Mol. Med.* 25 (22), 10736–10746. doi:10.1111/jcmm.17015

Dianatpour, M., Mehdipour, P., Nayernia, K., Mobasheri, M. B., Ghafouri-Fard, S., Savad, S., et al. (2012). Expression of testis specific genes TSGA10, TEX101 and ODF3 in breast cancer. *Iran. Red. Crescent Med. J.* 14 (11), 722–726. doi:10.5812/ircmj.3611

Ding, B., Lou, W., Xu, L., and Fan, W. (2018). Non-coding RNA in drug resistance of hepatocellular carcinoma. *Biosci. Rep.* 38 (5), BSR20180915. doi:10.1042/BSR20180915

Fong, Z. V., and Tanabe, K. K. (2014). The clinical management of hepatocellular carcinoma in the United States, europe, and asia: A comprehensive and evidence-based comparison and review. *Cancer* 120 (18), 2824–2838. doi:10.1002/cncr.28730

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33 (1), 1–22. doi:10.18637/jss.v033.i01

Hänzelmann, 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

Huang, Z., Zhou, J. K., Peng, Y., He, W., and Huang, C. (2020). The role of long noncoding RNAs in hepatocellular carcinoma. *Mol. Cancer* 19 (1), 77. doi:10.1186/s12943-020-01188-4

Jiang, D., Deng, J., Dong, C., Ma, X., Xiao, Q., Zhou, B., et al. (2020). Knowledge-based analyses reveal new candidate genes associated with risk of Hepatitis B virus

related hepatocellular carcinoma. *BMC cancer* 20 (1), 403. doi:10.1186/s12885-020-06842-0

Lammer, J., Malagari, K., Vogl, T., Pilleul, F., Denys, A., Watkinson, A., et al. (2010). Prospective randomized study of doxorubicin-eluting-bead embolization in the treatment of hepatocellular carcinoma: Results of the PRECISION V study. *Cardiovasc. Interv. Radiol.* 33 (1), 41–52. doi:10.1007/s00270-009-9711-7

Li, X., Lei, Y., Wu, M., and Li, N. (2018). Regulation of macrophage activation and polarization by HCC-derived exosomal lncRNA TUC339. *Int. J. Mol. Sci.* 19 (10), E2958. doi:10.3390/ijms19102958

Li, Y., Ye, Y., Feng, B., and Qi, Y. (2017). Long noncoding RNA lncARSR promotes doxorubicin resistance in hepatocellular carcinoma via modulating PTEN-PI3K/akt pathway. *J. Cell. Biochem.* 118 (12), 4498–4507. doi:10.1002/jcb.26107

Liang, Y., Liang, Q., Qiao, L., and Xiao, F. (2020). MicroRNAs modulate drug resistance-related mechanisms in hepatocellular carcinoma. *Front. Oncol.* 10, 920. doi:10.3389/fonc.2020.00920

Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z., and Zhang, B. (2019). WebGestalt 2019: Gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* 47 (W1), W199–w205. doi:10.1093/nar/gkz401

Lin, N., Lin, J. Z., Tanaka, Y., Sun, P., and Zhou, X. (2021). Identification and validation of a five-lncRNA signature for predicting survival with targeted drug candidates in ovarian cancer. *Bioengineered* 12 (1), 3263–3274. doi:10.1080/21655979.2021.1946632

Liu, G., Luo, Q., Li, H., Liu, Q., Ju, Y., and Song, G. (2020). Increased oxidative phosphorylation is required for stemness maintenance in liver cancer stem cells from hepatocellular carcinoma cell line HCCLM3 cells. *Int. J. Mol. Sci.* 21 (15), E5276. doi:10.3390/ijms21155276

Llovet, J. M., Real, M. I., Montaña, X., Planas, R., Coll, S., Aponte, J., et al. (2002). Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: A randomised controlled trial. *Lancet (London, Engl.* 359 (9319), 1734–1739. doi:10.1016/S0140-6736(02)08649-X

Lohitesh, K., Chowdhury, R., and Mukherjee, S. (2018). Resistance a major hindrance to chemotherapy in hepatocellular carcinoma: An insight. *Cancer Cell Int.* 18, 44. doi:10.1186/s12935-018-0538-7

Ma, C., Han, M., Heinrich, B., Fu, Q., Zhang, Q., Sandhu, M., et al. (2018). Gut microbiome-mediated bile acid metabolism regulates liver cancer via NKT cells, *Science* 360. eaan5931. doi:10.1126/science.aan5931

Malagari, K., Pomoni, M., Kelekis, A., Pomoni, A., Dourakis, S., Spyridopoulos, T., et al. (2010). Prospective randomized comparison of chemoembolization with doxorubicin-eluting beads and bland embolization with BeadBlock for hepatocellular carcinoma. *Cardiovasc. Interv. Radiol.* 33 (3), 541–551. doi:10.1007/s00270-009-9750-0

McGlynn, K. A., Petrick, J. L., and El-Serag, H. B. (2021). Epidemiology of hepatocellular carcinoma. *Hepatol. Baltim. Md)* 73, 4–13. doi:10.1002/hep.31288

Mobasheri, M. B., Jahanzad, I., Mohagheghi, M. A., Aarabi, M., Farzan, S., and Modarressi, M. H. (2007). Expression of two testis-specific genes, TSGA10 and SYCP3, in different cancers regarding to their pathological features. *Cancer detect. Prev.* 31 (4), 296–302. doi:10.1016/j.cdp.2007.05.002

Pan, W., Li, W., Zhao, J., Huang, Z., Zhao, J., Chen, S., et al. (2019). lncRNA-PDPK2P promotes hepatocellular carcinoma progression through the PDK1/AKT/Caspase 3 pathway. *Mol. Oncol.* 13 (10), 2246–2258. doi:10.1002/1878-0261.12553

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Shen, W., Song, Z., Xiao, Z., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 1:e36. doi:10.1002/imt2.36

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

Tanaka, R., Ono, T., Sato, S., Nakada, T., Koizumi, F., Hasegawa, K., et al. (2004). Over-expression of the testis-specific gene TSGA10 in cancers and its immunogenicity. *Microbiol. Immunol.* 48 (4), 339–345. doi:10.1111/j.1348-0421.2004.tb03515.x

Tang, W., Chen, Z., Zhang, W., Cheng, Y., Zhang, B., Wu, F., et al. (2020). The mechanisms of sorafenib resistance in hepatocellular carcinoma: Theoretical basis and therapeutic aspects. *Signal Transduct. Target. Ther.* 5 (1), 87. doi:10.1038/s41392-020-0187-x

Tsang, W. P., and Kwok, T. T. (2007). Riboregulator H19 induction of MDR1-associated drug resistance in human hepatocellular carcinoma cells. *Oncogene* 26 (33), 4877–4881. doi:10.1038/sj.onc.1210266

Wang, M., Han, J., Xing, H., Zhang, H., Li, Z., Liang, L., et al. (2016). Dysregulated fatty acid metabolism in hepatocellular carcinoma. *Hepat. Oncol.* 3 (4), 241–251. doi:10.2217/hep-2016-0012

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

Zhou, J., Sun, H., Wang, Z., Cong, W., Wang, J., Zeng, M., et al. (2020). Guidelines for the diagnosis and treatment of hepatocellular carcinoma (2019 edition). *Liver cancer* 9 (6), 682–720. doi:10.1159/000509424

Zhou, J. J., Cheng, D., He, X. Y., Meng, Z., Ye, H. L., and Chen, R. F. (2017). Knockdown of long non-coding RNA HOTAIR sensitizes hepatocellular carcinoma cell to cisplatin by suppressing the STAT3/ABCB1 signaling pathway. *Oncol. Lett.* 14 (6), 7986–7992. doi:10.3892/ol.2017.7237

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 21 November 2022
DOI 10.3389/fphar.2022.1076958

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Li Ding,
The Affiliated Hospital of Xuzhou Medical University, China
Xian Shao,
Tianjin Medical University, China

*CORRESPONDENCE
Qun Li,
Garylina@126.com

†These authors have contributed equally to this work.

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 22 October 2022
ACCEPTED 11 November 2022
PUBLISHED 21 November 2022

CITATION
Lu J, Xu F, Rao C, Shen C, Jin J, Zhu Z, Wang C and Li Q (2022), Mechanism of action of paclitaxel for treating glioblastoma based on single-cell RNA sequencing data and network pharmacology.
Front. Pharmacol. 13:1076958.
doi: 10.3389/fphar.2022.1076958

Mechanism of action of paclitaxel for treating glioblastoma based on single-cell RNA sequencing data and network pharmacology

Jianglong Lu†, Fanjie Xu†, Changjun Rao, Chaodong Shen, Jinghao Jin, Zhangzhang Zhu, Chengde Wang and Qun Li*

Department of Neurosurgery First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

Paclitaxel is an herbal active ingredient used in clinical practice that shows anti-tumor effects. However, its biological activity, mechanism, and cancer cell-killing effects remain unknown. Information on the chemical gene interactions of paclitaxel was obtained from the Comparative Toxicogenomics Database, SwishTargetPrediction, Binding DB, and TargetNet databases. Gene expression data were obtained from the GSE4290 dataset. Differential gene analysis, Kyoto Encyclopedia of Genes and Genomes, and Gene Ontology analyses were performed. Gene set enrichment analysis was performed to evaluate disease pathway activation; weighted gene co-expression network analysis with diff analysis was used to identify disease-associated genes, analyze differential genes, and identify drug targets *via* protein-protein interactions. The Molecular Complex Detection (MCODE) analysis of critical subgroup networks was conducted to identify essential genes affected by paclitaxel, assess crucial cluster gene expression differences in glioma *versus* standard samples, and perform receiver operator characteristic mapping. To evaluate the pharmacological targets and signaling pathways of paclitaxel in glioblastoma, the single-cell GSE148196 dataset was acquired from the Gene Expression Omnibus database and preprocessed using Seurat software. Based on the single-cell RNA-sequencing dataset, 24 cell clusters were identified, along with marker genes for the two different cell types in each cluster. Correlation analysis revealed that the mechanism of paclitaxel treatment involves effects on neurons. Paclitaxel may affect glioblastoma by improving glucose metabolism and processes involved in modulating immune function in the body.

KEYWORDS

glioblastoma, paclitaxel, single-cell RNA sequencing, bioinformatics, network pharmacology

1 Introduction

The most common primary malignant brain tumors of the central nervous system are gliomas, which originate from neuroectodermal cells (Jiang et al., 2011a; Wang et al., 2016) and are responsible for 74.6% of malignant tumors and 24.7% of initial brain tumors (Mat Zin and Zulkarnain, 2019). Gliomas are characterized by rapid growth, aggression, relapse after surgery, and a high death rate (Liu et al., 2012). Surgery, chemotherapy, radiation, and other methods are the main treatment options for glioblastoma (Yang et al., 2013). Treatments for glioblastoma have advanced in recent years through the development of chemotherapeutic medicines. Chemical drugs improve outcomes following surgery or radiation therapy and prolong the survival time and tumor-free survival time (Chen et al., 2017). However, the targets of chemotherapeutic drugs are unclear, the drugs do not easily cross the blood-brain barrier, and their effects are insufficient; additionally, effective drugs do not concentrate at the lesion site and do not remain at this site long-term (Johnson and Phillips, 1996). These drugs also show low bioavailability (Talibi et al., 2014). Thus, new treatment options for glioblastoma are needed (Bush et al., 2017).

The anticancer drug paclitaxel is extracted from the bark of the yew tree and targets microtubule proteins (Bastiancich et al., 2019). Paclitaxel accelerates the formation of microtubules from microtubule dimers and prevents their separation, which induces abnormal mechanical reorganization of the microtubules and inhibits normal cell division. This drug also inhibits the effects of other factors on the microtubule system and, together with the stable binding of microtubule proteins, eventually induces apoptosis (Liu et al., 2017a). In clinical applications, paclitaxel has shown good efficacy in treating non-small cell lung cancer (Hoang et al., 2012), breast cancer (Manhas et al., 2022), gastric cancer (Tu et al., 2022), nasopharyngeal cancer (Xia et al., 2022), ovarian cancer (Kong et al., 2023)and cervical cancer (Yasunaga et al., 2022), particularly for drug-resistant tumors (Song et al., 2018; Kawiak et al., 2019). Although the therapeutic efficacy of paclitaxel in glioma has been confirmed, its therapeutic mechanism remains unclear.

In addition, the activity of paclitaxel against brain tumors was disappointing in phase II experiments due to the presence of the blood-tumor barrier (BTB) and/or blood-brain barrier (BBB) (Zhang et al., 2012). In recent years, more and more studies have been devoted to the combined administration to break through the blood-brain barrier and act precisely on gliomas, and p-glycoprotein has been confirmed to be an important obstacle to preventing paclitaxel from entering the brain through studies of paclitaxel crossing the blood-brain barrier *in vitro* and *in vivo* (Fellner et al., 2002; Zhang et al., 2012; Li et al., 2016). One study showed angiopep-2 modified cationic liposomes for effective co-delivery of therapeutic genes encoding human tumor necrosis factor-related apoptosis-inducing ligand (pEGFP-hTRAIL) and paclitaxel to gliomas (Sun et al., 2012). Local delivery of brain-penetrating nanoparticles significantly improved the efficacy of paclitaxel for malignant gliomas and substantially delayed tumor growth (Nance et al., 2014). These studies and methods provide great help for paclitaxel to break through the blood-brain barrier and act as a precise drug-targeted therapy, and also make our study meaningful. This study was conducted to evaluate the specific effects of paclitaxel on glioblastoma and provide a new approach for treating this disease in clinical settings.

In this study, we investigated the mechanism of action of paclitaxel in glioblastoma therapy by using network, pharmacology, and genetics analyses. We determined the crucial role of immune function regulation in the prognosis of patients with glioblastoma. Analysis of transcription data from the Gene Expression Omnibus (GEO) database and corresponding clinical information revealed differentially expressed genes (DEGs). We also explored the correlations between drug- and disease-acting genes and levels of immune function activation, constructed a glioblastoma prediction model, and identified several different genes associated with immune activation as potential biomarkers. The findings were validated using the GEO single-cell dataset. Our findings revealed a crucial role for immunomodulation in treating glioblastoma with paclitaxel, which may act on neuronal cells and improve processes such as glucose metabolism to regulate the body's immune function.

2 Materials and methods

Flowchart was shown in Figure 1.

2.1 Identification of targets of paclitaxel

The SwissTargetPrediction database (https://www.swisstargetprediction.ch/) was used to query paclitaxel for its targets and associated target genes. The Comparative Toxicogenomics Database (CTD; https://ctdbase.org/), BindingDB (http://bindingdb.org/bind/index.jsp), and TargetNet database (https://targetnet.scbdd.com/) were also used to identify potential target genes. Additionally, we used the UniProt database (https://www.uniprot.org/uploadlists/) to query genes corresponding to potential target proteins to screen for active ingredients. Our results were used to locate paclitaxel lactones through Excel searching and sorting. Gene Ontology (GO) was used for functional annotation. The Database for Annotation, Visualization, and Integrated Discovery (DAVID) (https://david.ncifcrf.gov/) was employed for functional annotation, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, and Disease Ontology (DO) functional annotation of the target genes. DAVID integrates biological information and statistical tools to help researchers

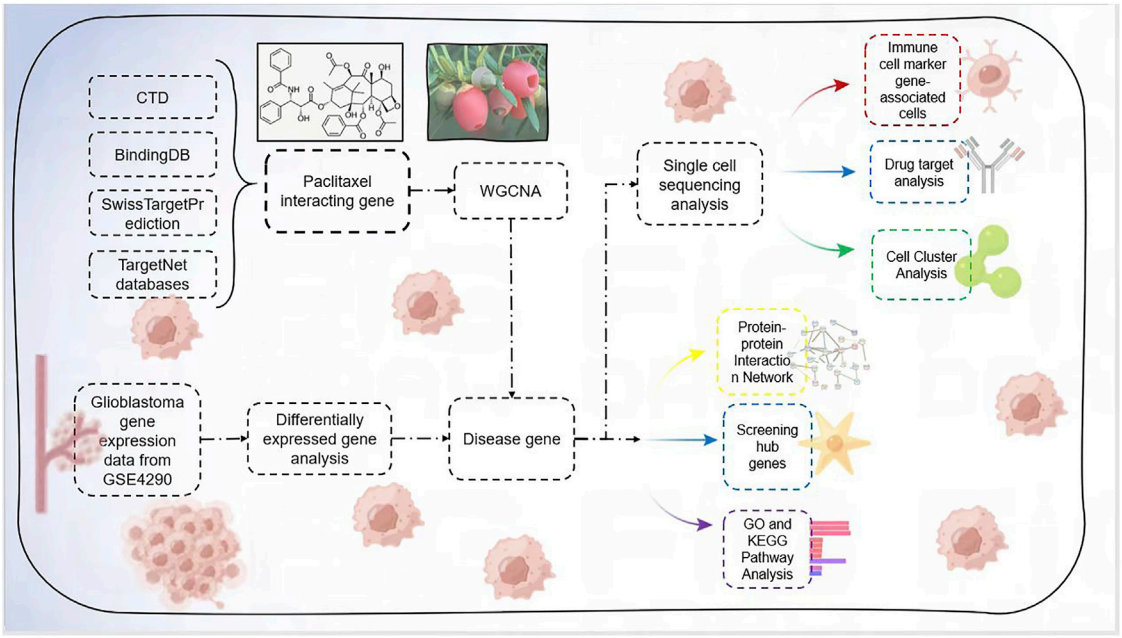


FIGURE 1
Flowchart.

identify gene and protein material. The bioinformatics tool GO analyzes and classifies biological processes into genes, with molecular functions, biological processes, and cells as the three GO components. Molecular data obtained using high-volume experimental techniques can be utilized to investigate signaling pathways, including numerous protein interactions and activities that regulate cellular function and metabolic activity. The ggplot2 tool in R was used to visualize the data, and an adjusted $p < 0.05$ was utilized for barrier testing.

2.2 Identification of DEGs in glioblastoma

77 glioblastoma samples and 23 healthy controls comprised the GSE4290 microarray dataset downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo). To obtain a gene expression matrix for the samples, it was first normalized and integrated. The genomes of glioblastoma samples and healthy controls were analyzed using the R package "limma (version 3.5.1)". *p*-values were adjusted using the Benjamin–Hochberg method. The segmentation criteria were modified to |[log2 fold-change]| >1 and $p < 0.05$. Using the R packages "ggplot2 (version 3.3.2)" and "heatmap (version 0.3.2)," all genes were displayed in a volcano plot. Heatmap (version 0.7.7) was used to show the top 20 DEGs (Ito and Murphy, 2013). Ridge plots were designed, and the defined genomes were analyzed using gene set enrichment analysis (GSEA) to identify significant differences between the two characteristics (Subramanian et al., 2007). The biological pathways and processes involved in the pathogenesis of module membership (MM) were predicted using GSEA (version 3.0, http://www.gsea-msigdb.org/gsea/index.jsp). Hub gene expression values were employed as phenotype files to calculate Pearson correlation coefficients, and the KEGG pathway gene set was used as an enrichment background. The above gene sets were used as background genes for enrichment analysis, and the correlation coefficients of each hub gene with other genes were sorted in descending order as scan sequences. Analyses were performed using the following settings: false discovery rate <0.25, nominal *p*-value < 0.05, |normalized enrichment score| > 1.

2.3 Weighted gene co-expression network analysis of GEO

In the weighted gene co-expression network analysis (WGCNA) package of the R software, 5,000 genes with the highest average expression were selected to construct a weighted gene co-expression network using expression as a screening condition. The screening threshold was set to convert the paired correlation matrix into a neighborhood correlation matrix to ensure that the scale-free network calculated the paired Pearson correlation coefficients between all genes individually. The minimum number of genes per gene

module was set to 30 using the dynamic hybrid shear tree algorithm criterion, and eigenvector values were calculated for each module. The modules were analyzed by clustering, and close modules were combined into a new module. The WGCNA algorithm calculates the module feature correlation to determine the correlation between module genes and disease subgroup phenotypes, and the heat map reflects the strength of the correlation. Individual modules were considered as significantly correlated with the phenotype when $p < 0.05$. The module showing the highest correlation coefficient with glioblastoma was selected as the key module. Pearson's correlation coefficients were calculated for each co-expression module with gene identity values to screen for key genes. Genes with module membership (MM) > 0.8 and gene importance (GS) > 0.65 were selected as key genes. Differential genes were intersected with WGCNA as disease-related genes and imported into DAVID 6.8 for GO and KEGG pathway enrichment analyses. Pathway enrichment analysis was performed to validate the significant gene functional categories ($p < 0.05$).

2.4 Generation of protein-protein interaction networks

Protein-protein interactions (PPIs) were investigated using the cross-targets identified in STRING (version 10.5, https://string-db.org/). The network nodes and edges depict protein and high-binding conversations, respectively. Cytoscape software was used to create and visualize the PPI interaction networks (version 3.6.0). The Molecular Complex Detection (MCODE) algorithm detects dense regions of tightly linked protein or PPI networks and is used to screen for critical subnetworks that contribute to glioblastoma development, derive essential subpopulation genes, and perform GO enrichment analysis.

2.5 Differential expression of crucial subpopulation genes in glioblastoma and normal tissues

Differential expression analysis of crucial subpopulation genes was performed on the GEO dataset using statistical software R4.1.3 (The R Project for Statistical Computing, Vienna, Austria). Differential expression of crucial subpopulation genes between disease and control groups was explored under screening conditions of p-value < 0.05 and |[(log2 fold-change)]| > 1 and visualized as heat maps in R language. Data from GSE4290 were used to construct a disease control model validation set to assess the association of critical genes with glioblastoma in R language software (Robin et al., 2011). Receiver operating characteristic (ROC) curves were plotted, and screened core genes were evaluated by calculating the area under the ROC curve.

2.6 Single-cell RNA sequencing data analysis and identification of glioblastoma-associated genes

The original expression profile dataset (GSE148196) used for analysis was screened using the GEO public database. The dataset consisted of biopsies from four patients with active glioblastoma. Tissues were extracted and then analyzed using expression profiling microarrays on the Illumina NextSeq assay platform. The raw dataset was preprocessed using the Seurat R package to ensure the quality of the results. The total number of molecules within the cell (nCount RNA) and genes detected in each cell (nFeature RNA) were determined, and the number of genes was compared to the number of reads obtained from sequencing of each cell. Widespread mitochondrial genomic contamination in low-quality or dead cells was assessed by calculating the number of reads paired with the mitochondrial genome using a percentage feature set function. Cells were clustered based on the filtered principal components and visually classified using the unified manifold approximation and projection dimensionality reduction technique. Immune cell marker genes with adjusted p-values < 0.05 were screened. Immune cell marker genes were retrieved using the PanglaoDB database and intersected with the corresponding genes for each class group to identify the class group of the immune cells. The results revealed the potential targets of paclitaxel in glioblastoma.

2.7 Statistical analysis

A two-sided p-value of 0.05 was considered to indicate statistically significant results. Rstudio (www.r-project.org; version 4.2.1) was used to sort and observe the data (Packages: limma, edgeR, ggplot2, survminer, survival, RMS, randomForest, pROC, glmnet, heatmap, timeROC, *via* storyline, complot, ConsensusClusterPlus, forest plot, survival rock, beeswarm, edgeR, "TxDb.Hsapiens.UCSC.hg38," "known gene," "cluster profile," "org.Hs.eg.DB," "karyoploteR," "GSVA," "GSEABase," "stringr," "GEOquery," "dplyr," "ComplexHeatmap," and "RColorBrewer").

3 Results

3.1 Target genes of paclitaxel

Using the SwissTargetPrediction, CTD, BindingDB, and TargetNet databases, we identified and retrieved 1,010 target genes associated with paclitaxel lactone (Figure 2A; Supplementary Table S1). We performed GO and DO functional enrichment and KEGG pathway enrichment analyses. The GO biological process category was mainly enriched in regulation of peptidase activity, response to

FIGURE 2
Screening analysis of paclitaxel targets. **(A)** Venn diagram of paclitaxel in the four databases. **(B)** Gene Ontology (GO) enrichment analysis of paclitaxel targets. **(C,D)** DO enrichment analysis of paclitaxel targets. **(E,F)** Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of paclitaxel targets.

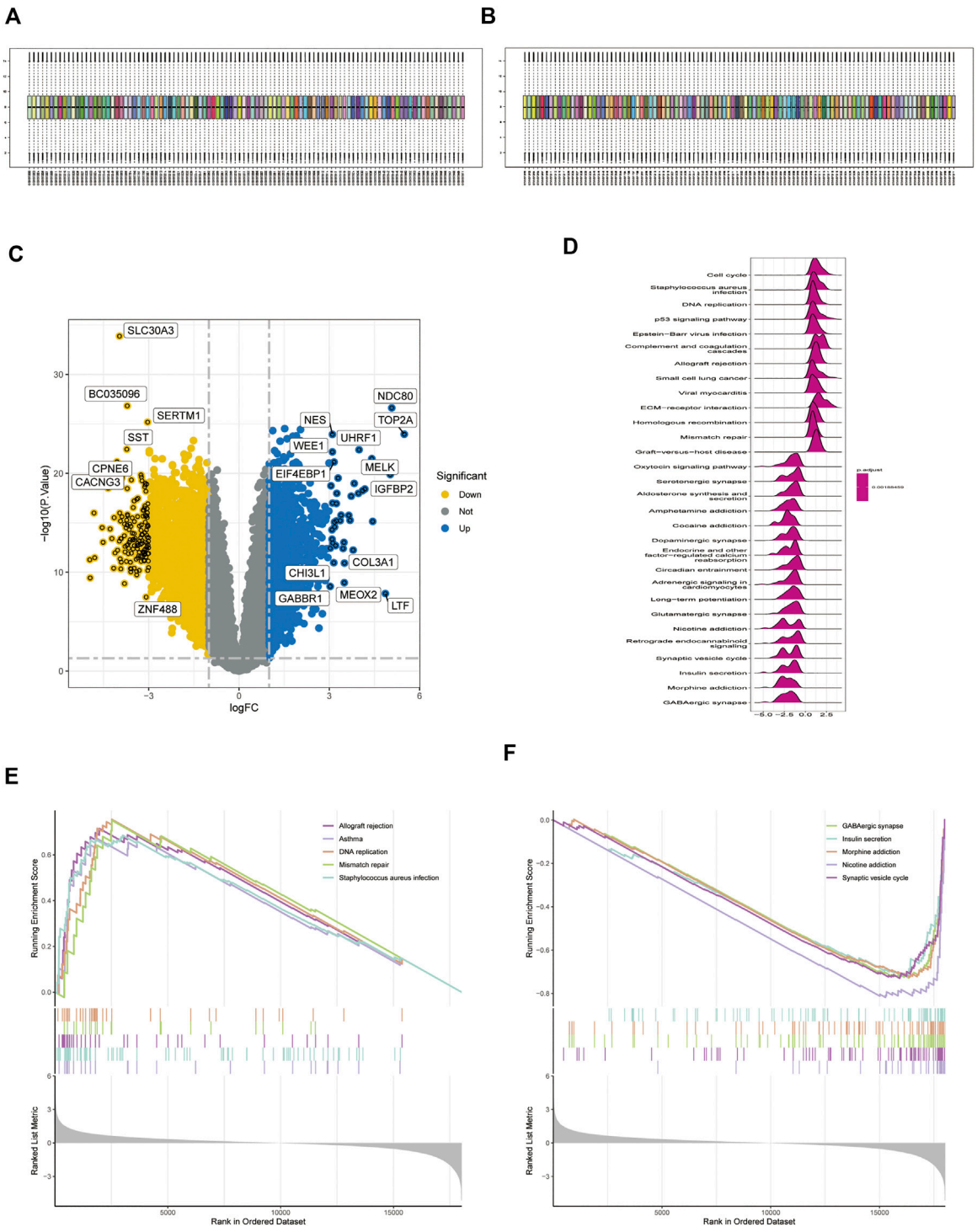


FIGURE 3
Expression of differentially expressed genes (DEGs) in the GSE4290 dataset. **(A,B)** Datasets were compared before and after normalization. **(C)** Volcano plot of DEGs in the GSE4290 dataset. **(D)** Ridge plots with normalized enrichment scores show the pathways where DEGs are most enriched in gene set enrichment analysis (GSEA). **(E,F)** GSEA analysis based on KEGG analysis.

peptides, and endopeptidase activity. The GO cellular component category was mainly enriched in collagen-containing extracellular matrix, vesicle lumen, and cytoplasmic vesicle lumen. The molecular function category was mainly enriched in protein serine/threonine/tyrosine kinase, endopeptidase, and protein serine/threonine kinase activities (Figure 2B). DO upregulation was mainly enriched in musculoskeletal system cancer, connective tissue cancer, non-small cell lung carcinoma, bone cancer, female reproductive organ cancer, and breast carcinoma (Figures 2C,D). KEGG analysis revealed enrichment mainly in the pathways of neurodegeneration, multiple diseases, Alzheimer's disease, PI3K-Akt signaling pathway, lipid and atherosclerosis, and Epstein-Barr virus infection (Figures 2E,F).

3.2 Target genes in glioblastoma

Using normalization between arrays based on the GSE28424 dataset (Figures 3A,B), 3,135 genes were screened for differential expression between glioblastoma samples and normal tissue. Among the genes, 1,345 were upregulated and 1790 were downregulated (Supplementary Table S2); the top 20 genes are shown in a volcano plot (Figure 3C) and ridge plot (Figure 3D) drawn using R language for the glioblastoma group. Pathway enrichment was evaluated using GSEA pathway between the glioblastoma and control groups. The results showed that allograft rejection, asthma, DNA replication, mismatch repair, and *Staphylococcus aureus* infection activation were enriched in glioblastoma (Figure 3E). GABAergic synapses, insulin secretion, morphine addiction, nicotine addiction, and synaptic vesicle cycle were significantly inhibited (Figure 3F), suggesting that immune dysfunction plays an essential role in glioma development.

3.3 WGCNA

GSE4290 microarray data and clinical information were downloaded and pre-processed to obtain a final expression matrix of 100 samples corresponding to 23,323 genes. The 5,000 genes with the highest average expression were selected to create a gene co-expression module. The dataset was processed for outlier detection, which showed no significant outliers. Next, we directly analyzed the gene clustering module against the clinical grouping phenotypes. The soft threshold power was to 1–30, with $R^2 > 0.9$. The soft threshold power and mean connectivity were close to zero, indicating that the network was scale-free. Therefore, a soft threshold of 9 was chosen (Figure 4A). The topological overlap matrix and correlation matrix between genes were also computed. The topological overlap matrix was used to build a hierarchical clustering tree between genes, and merging of similar modules produced eight modules. The turquoise module showed the strongest correlation with glioblastoma ($r = 0.78$, P 0.01), as shown in Figures 4B,C. The scatter plot revealed a strong correlation between GS and MM within the turquoise module (correlation = 0.93, $p < 0.01$) (Figure 4D). Thus, the turquoise module may be a pivotal module linked to glioblastoma.

3.4 Functional enrichment analysis of genes within modules

Genes in the turquoise module were compared with differential genes to identify disease-related genes in GO and KEGG analyses (Figure 5A). According to the GO enrichment results, the enriched pathways were mainly involved in modulation of chemical synaptic transmission, regulation of transsynaptic signaling, synapse organization, presynaptic membrane, and glutamatergic synapses. According to KEGG enrichment analysis, the enriched pathways were mainly involved in GABAergic synapses, glutamatergic synapses, MAPK signaling pathways, and morphine addiction (Figures 5B–F).

To construct the PPI network, 155 disease-related genes and molecular drug targets were imported into the STRING online database (version 11.0) (Figure 6A). Aberrant proteins were removed, resulting in a 154-protein interaction network. Cytoscape's plugin code was used to identify 17 essential subpopulation genes (score = 13) (Figure 6B). Key cluster genes were upregulated for proteoglycans in cancer, bladder cancer, PI3K-Akt signaling pathway, AGE-RAGE signaling pathway in diabetic complications, HIF-1 signaling pathway, Kaposi sarcoma-associated herpesvirus infection, endocrine resistance, small cell lung cancer, pancreatic cancer, and human cytomegalovirus infection (Figures 6C,D). In addition to acting on cellular metabolic pathways, paclitaxel may be useful for diagnosing glioblastoma.

3.5 Differential expression of critical genes in tumor tissue and controls and prognostic analysis

As shown in the box plots, individual essential sub-cluster genes were significantly different between the disease and control groups (Figure 7A), and the ROC curves showed that all 17 essential sub-cluster genes had excellent robustness for glioblastoma (area under the ROC curve >0.6) (Figure 7B). The immune heat map showed that the significant regulatory targets of the critical cluster genes were mainly in the immune pathways of Macrophages_M0, Macrophages_M2, Mast_cells_activated, and T_cells_follicular_helper (Figures 7C,D), indicating that these genes are involved in regulating glioma immune function, which is consistent with previous studies (Wang et al., 2022).

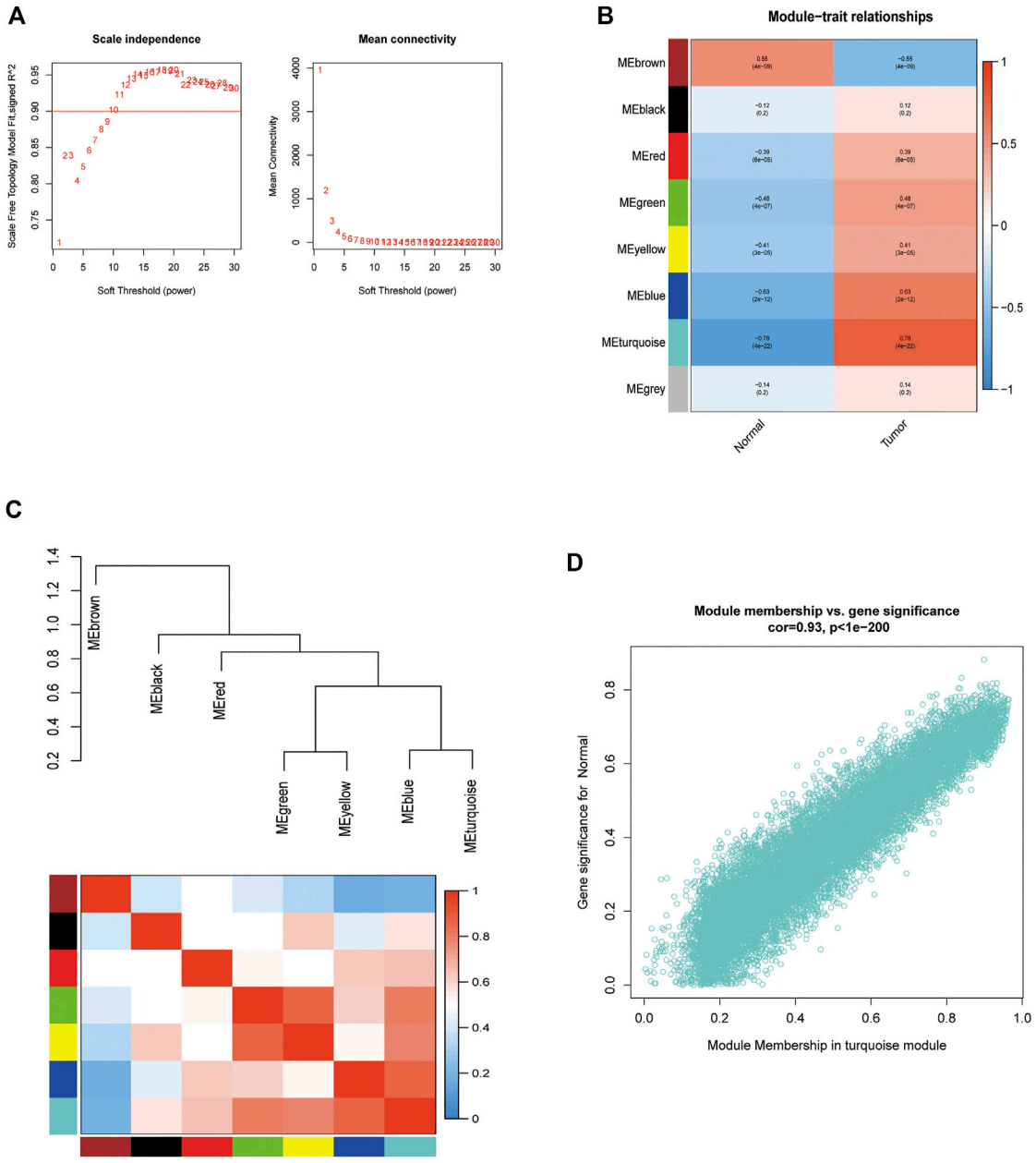


FIGURE 4
Enrichment levels in genomic weighted gene co-expression network analysis (WGCNA). **(A)**. Selection of soft thresholds. **(B,C)** Correlation of module eigengenes with glioblastoma. **(D)** Correlation of turquoise eigengenes with glioblastoma.

3.6 Single-cell assay analysis

Analysis of biopsy specimens from four patients with active glioblastoma showed a strong positive correlation between the measured gene expression and number of genes detected in the cells, both in normal and diseased tissues. In contrast, gene expression detected in the cells was not correlated with the percentage of mitochondria. Therefore, cells with >2,500 and <200 genes detected per cell and cells with a >5% mitochondrial percentage were filtered out to ensure the quality of the analyzed cells. Quality control and screening of single-cell sequencing of samples from patients with glioblastoma are shown in (Figures 8A,B). Principal component analysis plots were downscaled for cluster analysis (Figures 8C,D); the cluster tree was scaled to a resolution of 1.5 (Figure 8E), and the principal component

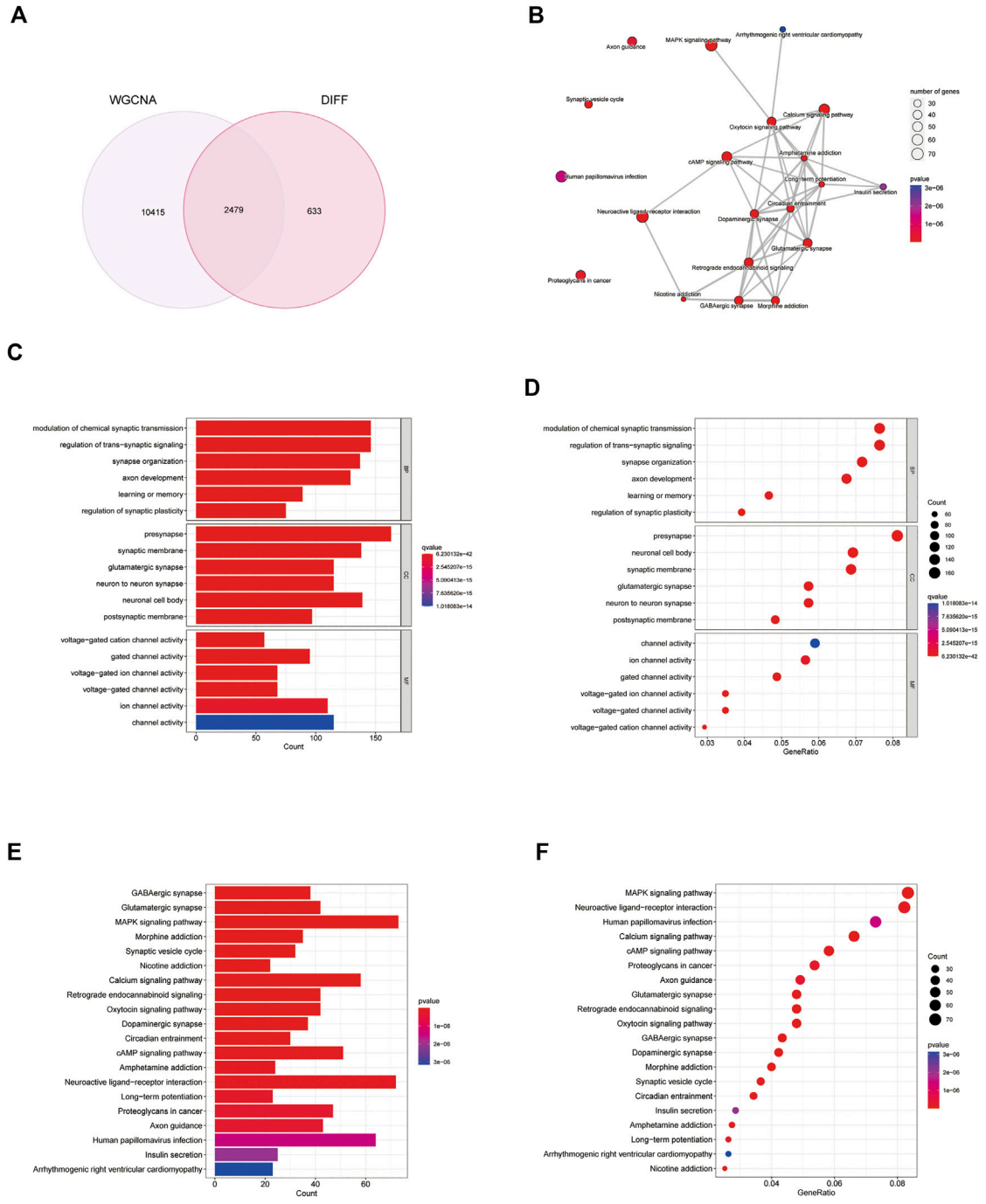


FIGURE 5
Analysis of differential genes and weighted gene co-expression network analysis (WGCNA) genes. **(A)** Venn diagram of differential genes and WGCNA genes. **(B–F)** Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of disease genes.

value was 16 (Figure 8F). The heatmap shows each gene type (Figure 9A). Unified manifold approximation and projection showed 24 cell clusters (Figure 9B), with different categories of cells labeled with different colors. Relevant genes were retrieved using the Cellmaker database and intersected with the gene corresponding to each unified manifold approximation and projection cluster (Figure 9C). The cellular distribution of drug target AUCell functional scores showed that paclitaxel acts mainly on neuronal cells (Figure 9D).

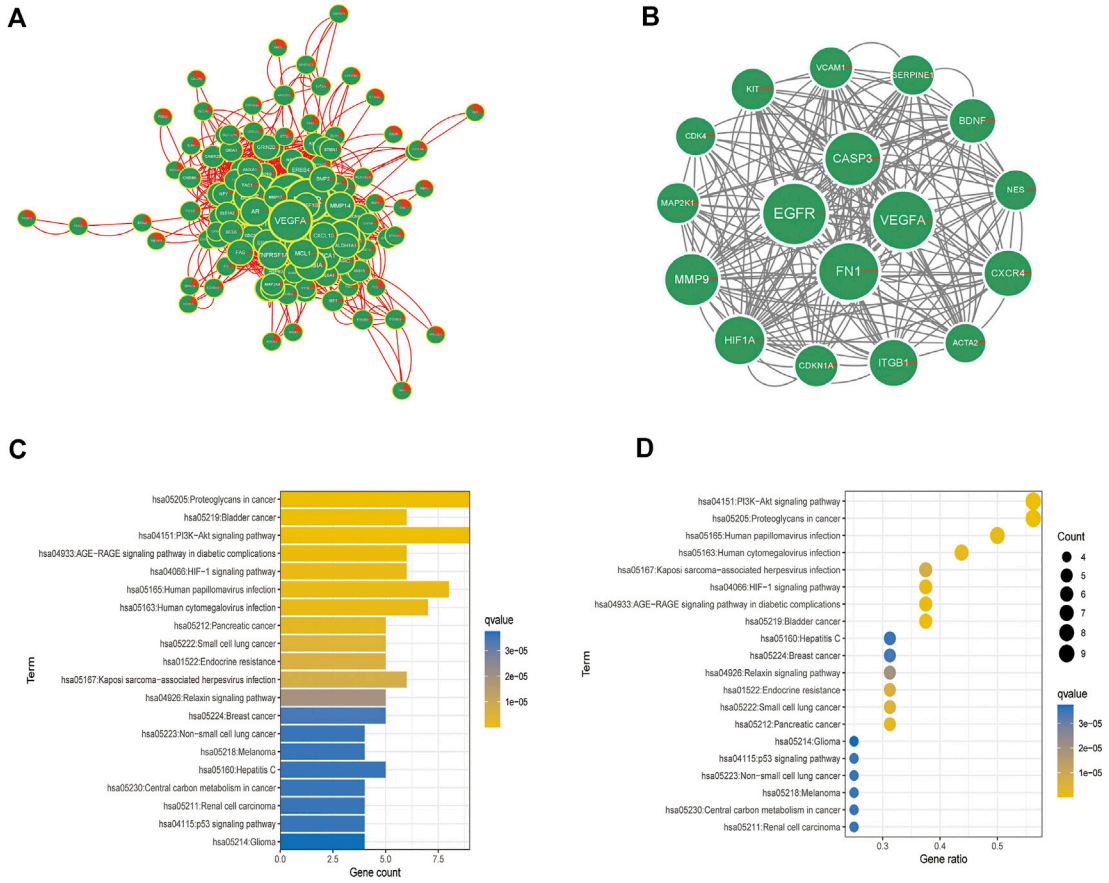


FIGURE 6
Analysis of crucial cluster genes. **(A,B)** Screening for essential subcluster genes. **(C,D)** Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of essential subcluster genes (top 10).

4 Discussion

Glioblastoma is a highly malignant primary malignant tumor; and WHO grade III and IV malignant glioblastoma is a common type of high-grade glioblastoma (Mitobe et al., 2022; Yilmaz et al., 2022), with no apparent boundary between the tumor tissue and surrounding tissue. Therefore, the efficacy of surgery alone in treating malignant glioblastoma is poor and results in a median survival of only approximately 10 months (Liu et al., 2017b). Paclitaxel is a traditional anti-tumor drug effective against ovarian cancer, colorectal cancer, breast cancer, and glioblastoma (Dorsey et al., 2009). In recent years, more and more studies have shown the therapeutic effect of paclitaxel on glioma, and it has been verified *in vivo* and *in vitro* experiments (Xie et al., 2006; Jiang et al., 2011b). Substantial progress has been made in the *in vitro* treatment of glioma with continuous delivery of paclitaxel through biodegradable materials (Xie and Wang, 2006). In addition, it has been shown that the treatment of glioblastoma with tumor-targeted gene vectors and brain-targeted micelles and paclitaxel co-delivery has achieved good efficacy in mouse experiments (Zhan et al., 2012). In this experiment, we used single-cell sequencing data for in-depth mining, and advances in this next-generation sequencing approach have enabled genomic analysis of single cells, which is beneficial to reveal heterogeneous tumors and has an important role in the treatment of cancer (Navin and Hicks, 2011; Zhang et al., 2016). In addition, single-cell sequencing is likely to improve several aspects of pharmacology, including precise targeting of drugs, cellular receptors, and deeper mechanisms of action (Lee et al., 2014; Heath et al., 2016).

We identified 1,010 target genes related to paclitaxel using the SwissTargetPrediction, CTD, BindingDB, and TargetNet databases. GO analysis showed enrichment in the regulation of peptidase activity, response to peptide, and regulation of endopeptidase activity in the biological process category. In the cell component category, enrichment was observed in collagen-containing extracellular matrix, vesicle lumen, and cytoplasmic vesicle lumen. For molecular function, we

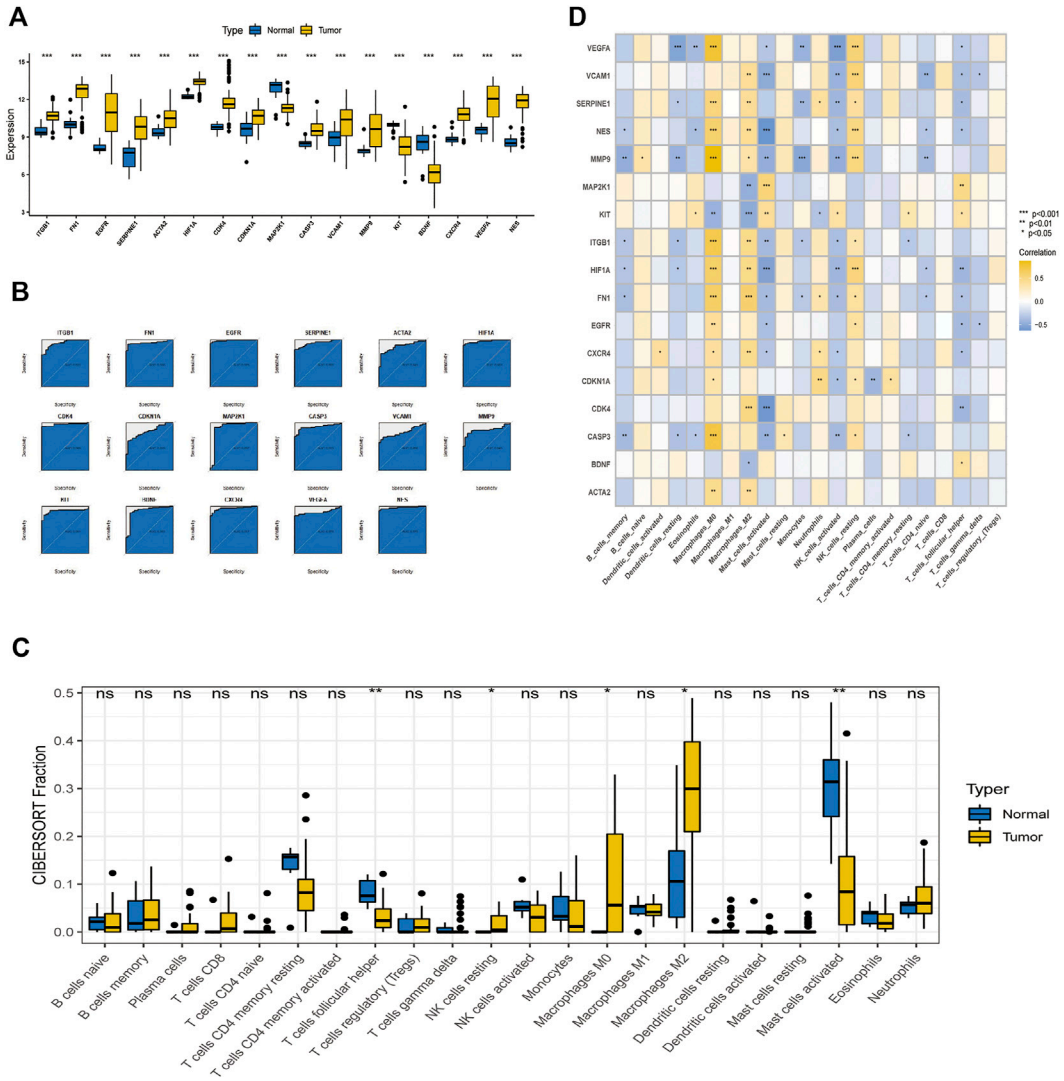


FIGURE 7
Relationship between crucial cluster genes and immune infiltration and receiver operating characteristic (ROC) prediction. **(A)** Differential expression of crucial subgroup genes in tumor and control tissues. **(B)** ROC curves of essential subcluster genes predicting disease onset. **(C)** Differential expression of immune function between tumor and control tissues. **(D)** Relationship between crucial cluster genes and immune infiltration.

observed enrichment in protein serine/threonine/tyrosine kinase activity, endopeptidase activity, and protein serine/threonine kinase activity (Figure 2B). DO enrichment analyses were mainly enriched in musculoskeletal system cancer, connective tissue cancer, non-small cell lung carcinoma, bone cancer, female reproductive organ cancer, and breast cancer; pathways related to neurodegeneration-multiple diseases, Alzheimer disease, PI3K-Akt signaling pathway, lipid and atherosclerosis, Epstein-Barr virus infection were enriched according to KEGG analysis. These results suggest that paclitaxel can treat a variety of tumor cells by regulating the body's peptidase activity and other signaling pathways. The 3,135 genes differentially expressed between glioblastoma samples and normal tissues, including 1,345 upregulated genes and 1790 downregulated genes, were enriched in allograft rejection, asthma, DNA replication, mismatch repair, and glioblastoma. Replication, mismatch repair, and *S. aureus* infection pathway were activated. GABAergic synapse, insulin secretion, morphine addiction, nicotine addiction, and synaptic vesicle cycle pathways were significantly inhibited, suggesting that immune function was overactivated in the tumor tissues (Perelroizen et al., 2022).

We also identified 2,479 key genes involved in disease progression. Fifty-three key subgroups of genes were enriched in proteoglycans in cancer, bladder cancer, and PI3K-Akt

10.3389/fphar.2022.1076958

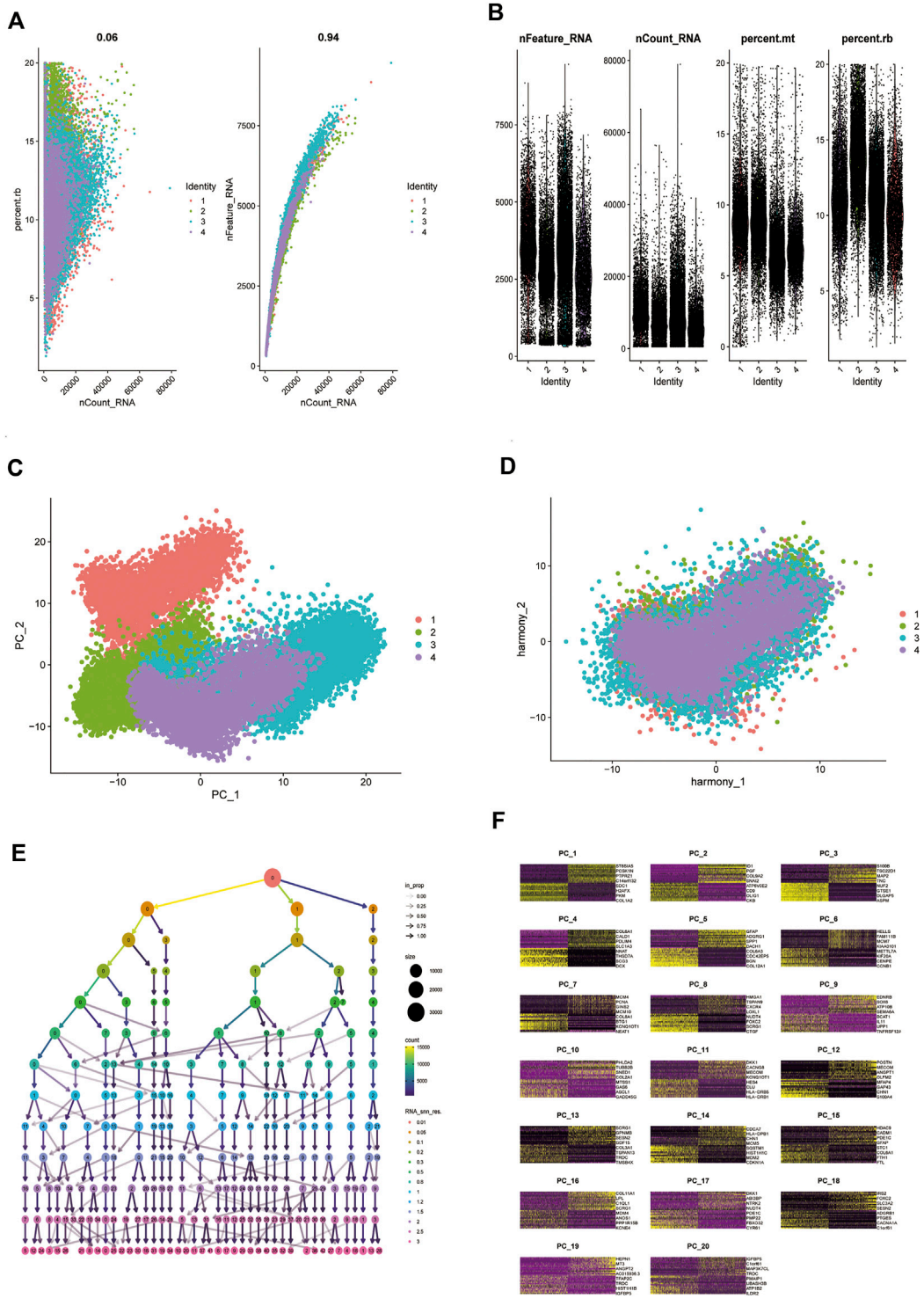


FIGURE 8
Comparison of single-cell analysis before and after normalization. **(A,B)** Quality control analysis of single-cell data sets. **(C,D)** Plots of principal component analysis (PCA) before and after standardization. **(E,F)** Resolution with principal components (PCs) to be confirmed.

signaling pathways. Cybersport analysis was used to calculate the number of genes in 22 immune cells. The results showed that the drug-disease critical cluster of genes was primarily targeted by one of the major immune pathways and other pathways to exert a therapeutic effect. Single-cell data analysis showed that the main target of paclitaxel was neuronal cells,

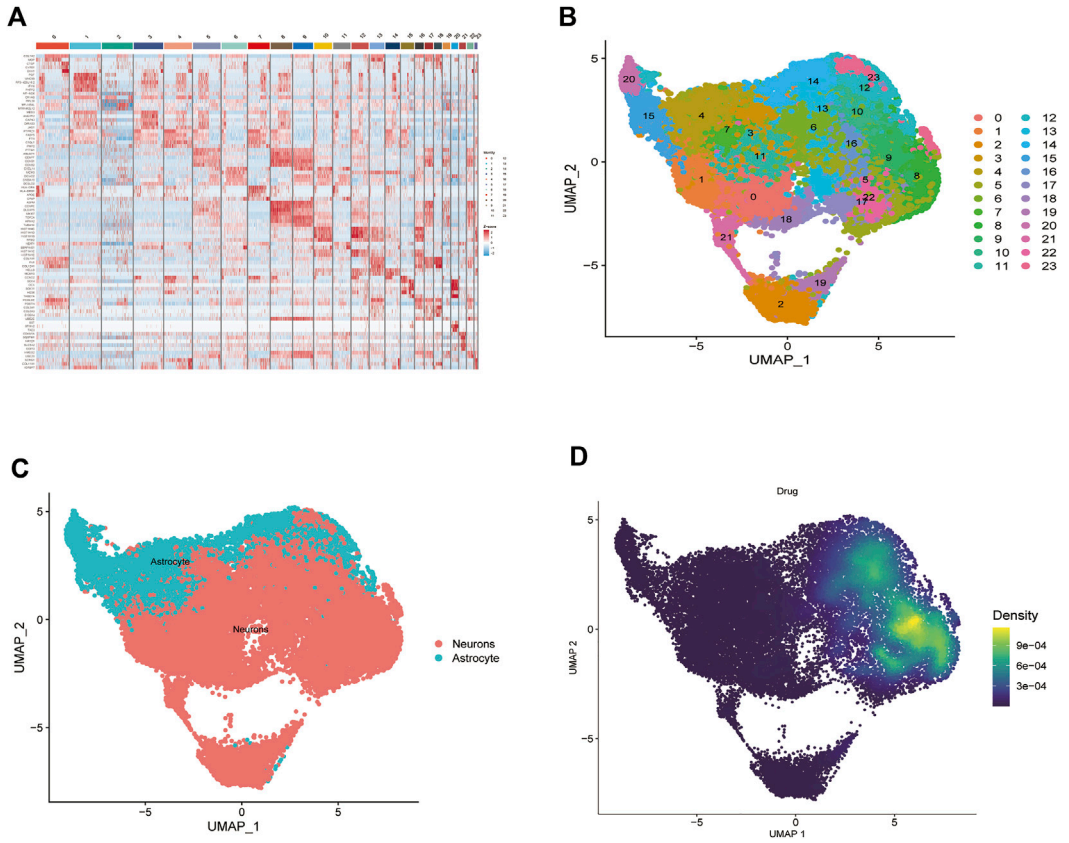


FIGURE 9
Paclitaxel pathways of action. **(A)** Heat map of each gene table level. **(B)** Unified manifold approximation and projection clustering into 24 clusters. **(C)** iTalk analysis identified two clusters. **(D)** Paclitaxel drug pathways of action.

which is consistent with previous results (Duhamel et al., 2018).

Our results suggest that paclitaxel improves the prognosis of glioblastoma by acting on neuronal cells and modulating immunity (Xue et al., 2017). We also identified *ITGB1*, *FN1*, *EGFR*, *SERPINE1*, *ACTA2*, *HIF1A*, *CDK4*, *CDKN1A*, *MAP2K1*, *CASP3*, *VCAM1*, *MMP9*, *KIT*, *BDNF*, *CXCR4*, *VEGFA*, and *NES* as essential cluster genes, which agrees with the results of previous related studies. Li et al. suggested that GLIPR1 enhances the proliferation, migration, and invasion of glioblastoma and may be involved in activation of the TIMP1-CD63-ITGB1-AKT signaling pathway, which is a potential target for the clinical prevention or management of glioblastoma (Jiang et al., 2022). Wu et al. (2021) found that the regulation of ITGB1 expression promotes progression, suggesting an essential role for ITGB1 in glioblastoma. Andersson et al. (2010) demonstrated that regulation of the EGFR pathway is involved in glioblastoma progression and that specific genotypes of the EGFR gene may be associated with glioblastoma risk. According to Ohtaki et al. (2017), ACTC1 is as an independent prognostic and aggressive marker of gliomas. In addition, Wu et al. (2022) demonstrated that CXCR4 promotes the proliferation of GICs through the KLF5/BCL2L12-dependent pathway. These essential cluster genes may have good predictive efficacy for glioblastoma and are essential for glioblastoma development. Further studies are needed to identify the molecular mechanisms involved in the immune response to glioblastoma.

Our experiments still have some limitations, lack prospective cohort and *in vitro* experiments, and the specific molecular mechanism of paclitaxel affecting glioma remains unclear, but our advantage lies in the clear description of paclitaxel target cells based on single-cell data, which makes an important contribution to the study of the specific molecular mechanism in the next step.

5 Conclusion

We examined the interactions and molecular mechanisms of paclitaxel in glioblastoma. Paclitaxel and glioblastoma synergistically

affected differentially regulated genes. We used modern network medicinal theories to investigate the molecular biological mechanisms of paclitaxel in glioblastoma, which may help guide clinical practice. In future studies, we will validate these results in pharmacological and molecular biology experiments.

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

JL and FX downloaded the dataset, analyzed the data, and was a major contributor in writing the manuscript. QL run the entire project. All authors read and approved the final manuscript.

Funding

This work was funded by the Medical Health Science and Technology Research Project of Zhejiang Province of China (2023KY888) and the Wenzhou Science and Technology Project under Grants (Y2020061).

Acknowledgments

We thank Liu (Nucleobase translocation of bioinformatics) for sharing their bioinformatics experience generously.

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/fphar.2022.1076958/full#supplementary-material

References

Andersson, U., Schwartzbaum, J., Wiklund, F., Sjöström, S., Liu, Y., Tsavachidis, S., et al. (2010). A comprehensive study of the association between the EGFR and ERBB2 genes and glioma risk. *Acta Oncol.* 49 (6), 767–775. doi:10.3109/0284186X.2010.480980

Bastiancich, C., Bozzato, E., Luyten, U., Danhier, F., Bastiat, G., and Préat, V. (2019). Drug combination using an injectable nanomedicine hydrogel for glioblastoma treatment. *Int. J. Pharm.* 559, 220–227. doi:10.1016/j.ijpharm.2019.01.042

Bush, N. A., Chang, S. M., and Berger, M. S. (2017). Current and future strategies for treatment of glioma. *Neurosurg. Rev.* 40 (1), 1–14. doi:10.1007/s10143-016-0709-8

Chen, R., Smith-Cohn, M., Cohen, A. L., and Colman, H. (2017). Glioma subclassifications and their clinical significance. *Neurotherapeutics* 14 (2), 284–297. doi:10.1007/s13311-017-0519-x

Dorsey, J. F., Mintz, A., Tian, X., Dowling, M. L., Plastaras, J. P., Dicker, D. T., et al. (2009). Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) and paclitaxel have cooperative *in vivo* effects against glioblastoma multiforme cells. *Mol. Cancer Ther.* 8 (12), 3285–3295. doi:10.1158/1535-7163.MCT-09-0415

Duhamel, M., Rose, M., Rodet, F., Murgoci, A. N., Zografidou, L., Régnier-Vigouroux, A., et al. (2018). Paclitaxel treatment and proprotein convertase 1/3 (PC1/3) knockdown in Macrophages is a promising antiglioma strategy as revealed by proteomics and cytotoxicity studies. *Mol. Cell. Proteomics* 17 (6), 1126–1143. doi:10.1074/mcp.RA117.000443

Fellner, S., Bauer, B., Miller, D. S., Schaffrik, M., Fankhänel, M., Spruß, T., et al. (2002). Transport of paclitaxel (Taxol) across the blood-brain barrier *in vitro* and *in vivo*. *J. Clin. Investig.* 110 (9), 1309–1318. doi:10.1172/JCI15451

Heath, J. R., Ribas, A., and Mischel, P. S. (2016). Single-cell analysis tools for drug discovery and development. *Nat. Rev. Drug Discov.* 15 (3), 204–216. doi:10.1038/nrd.2015.16

Hoang, T., Dahlberg, S. E., Schiller, J. H., Mehta, M. P., Fitzgerald, T. J., Belinsky, S. A., et al. (2012). Randomized phase III study of thoracic radiation in combination with paclitaxel and carboplatin with or without thalidomide in patients with stage III non-small-cell lung cancer: The ECOG 3598 study. *J. Clin. Oncol.* 30 (6), 616–622. doi:10.1200/JCO.2011.36.9116

Ito, K., and Murphy, D. (2013). Application of ggplot2 to pharmacometric graphics. *CPT. Pharmacometrics Syst. Pharmacol.* 2 (10), e79. doi:10.1038/psp.2013.56

Jiang, T., Tang, G. F., Lin, Y., Peng, X. X., Zhang, X., Zhai, X. W., et al. (2011). Prevalence estimates for primary brain tumors in China: A multi-center cross-sectional study. *Chin. Med. J.* 124 (17), 2578–2583.

Jiang, X., Xin, H., Sha, X., Gu, J., Jiang, Y., Law, K., et al. (2011). PEGylated poly (trimethylene carbonate) nanoparticles loaded with paclitaxel for the treatment of advanced glioma: *In vitro* and *in vivo* evaluation. *Int. J. Pharm.* 420 (2), 385–394. doi:10.1016/j.ijpharm.2011.08.052

Jiang, Z., Li, J., Feng, W., Sun, Y., and Bu, J. (2022). A ferroptosis-related lncRNA model to enhance the predicted value of cervical cancer. *J. Oncol.* 2022, 6080049. doi:10.1155/2022/6080049

Johnson, J. H., Jr., and Phillips, P. C. (1996). Malignant gliomas in children. *Cancer Investig.* 14 (6), 609–621. doi:10.3109/07357909609076905

Kawiak, A., Domachowska, A., and Lojkowska, E. (2019). Plumbagin increases paclitaxel-induced cell death and overcomes paclitaxel resistance in breast cancer cells through ERK-mediated apoptosis induction. *J. Nat. Prod.* 82 (4), 878–885. doi:10.1021/acs.jnatprod.8b00964

Kong, Q. W., Yang, J., Li, D., Ding, Y. W., Hu, Y. J., Xue, X. C., et al. (2023). Tongguanteng injection reverses paclitaxel resistance via upregulation of TAB1 expression in ovarian cancer *in vitro* and *in vivo*. *J. Ethnopharmacol.* 300, 115728. doi:10.1016/j.jep.2022.115728

Lee, M-C. W., Lopez-Diaz, F. J., Khan, S. Y., Tariq, M. A., Dayn, Y., Vaske, C. J., et al. (2014). Single-cell analyses of transcriptional heterogeneity during drug tolerance transition in cancer cells by RNA sequencing. *Proc. Natl. Acad. Sci. U. S. A.* 111 (44), E4726–E4735. doi:10.1073/pnas.1404656111

Li, Y., Zheng, X., Gong, M., and Zhang, J. (2016). Delivery of a peptide-drug conjugate targeting the blood brain barrier improved the efficacy of paclitaxel against glioma. *Oncotarget* 7 (48), 79401–79407. doi:10.18632/oncotarget.12708

Liu, Y., Yan, W., Zhang, W., Chen, L., You, G., Bao, Z., et al. (2012). MiR-218 reverses high invasiveness of glioblastoma cells by targeting the oncogenic transcription factor LEF1. *Oncol. Rep.* 28 (3), 1013–1021. doi:10.3892/or.2012.1902

Liu, H., Wang, J., Sheng, L., Zhang, Y., Tang, N., Li, Y., et al. (2017). Paclitaxel promotes cell apoptosis in uterine leiomyomas. *Pharmacology* 100 (5-6), 246–252. doi:10.1159/000479161

Liu, H., Liu, B., Hou, X., Pang, B., Guo, P., Jiang, W., et al. (2017). Overexpression of NIMA-related kinase 2 is associated with poor prognoses in malignant glioma. *J. Neurooncol.* 132 (3), 409–417. doi:10.1007/s11060-017-2401-4

Manhas, D., Mir, K. B., Tripathi, N., Bharti, S., Dhiman, S., Wazir, P., et al. (2022). Rottlerin promotes anti-metastatic events by ameliorating pharmacological parameters of paclitaxel: An *in-vivo* investigation in the orthotopic mouse model of breast cancer. *Chem. Biol. Interact.* 366, 110109. doi:10.1016/j.cbi.2022.110109

Mat Zin, A. A., and Zulkarnain, S. (2019). Diagnostic accuracy of cytology smear and frozen section in glioma. *Asian pac. J. Cancer Prev.* 20 (2), 321–325. doi:10.31557/APJCP.2019.20.2.321

Mitobe, Y., Nakagawa-Saito, Y., Togashi, K., Suzuki, S., Sugai, A., Matsuda, K. I., et al. (2022). CEP-1347 targets MDM4 protein expression to activate p53 and inhibit the growth of glioma cells. *Anticancer Res.* 42 (10), 4727–4733. doi:10.21873/anticanres.15977

Nance, E., Zhang, C., Shih, T-Y., Xu, Q., Schuster, B. S., and Hanes, J. (2014). Brain-penetrating nanoparticles improve paclitaxel efficacy in malignant glioma following local administration. *ACS Nano* 8 (10), 10655–10664. doi:10.1021/nn504210g

Navin, N., and Hicks, J. (2011). Future medical applications of single-cell sequencing in cancer. *Genome Med.* 3 (5), 31–12. doi:10.1186/gm247

Ohtaki, S., Wanibuchi, M., Kataoka-Sasaki, Y., Sasaki, M., Oka, S., Noshiro, S., et al. (2017). ACTC1 as an invasion and prognosis marker in glioma. *J. Neurosurg.* 126 (2), 467–475. doi:10.3171/2016.1.JNS152075

Perelroizen, R., Philosof, B., Budick-Harmelin, N., Chernobylsky, T., Ron, A., Katzir, R., et al. (2022). Astrocyte immunometabolic regulation of the tumour microenvironment drives glioblastoma pathogenicity. *Brain.* 145 (9), 3288–3307. doi:10.1093/brain/awac222

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinforma.* 12, 77. doi:10.1186/1471-2105-12-77

Song, T., Yan, L., Cai, K., Zhao, T., and Xu, M. (2018). Downregulation of long noncoding RNA PVT1 attenuates paclitaxel resistance in glioma cells. *Cancer Biomark.* 23 (3), 447–453. doi:10.3233/CBM-181573

Subramanian, A., Kuehn, H., Gould, J., Tamayo, P., and Mesirov, J. P. (2007). GSEA-P: A desktop application for gene set enrichment analysis. *Bioinformatics* 23 (23), 3251–3253. doi:10.1093/bioinformatics/btm369

Sun, X., Pang, Z., Ye, H., Qiu, B., Guo, L., Li, J., et al. (2012). Co-delivery of pEGFP-hTRAIL and paclitaxel to brain glioma mediated by an angiopep-conjugated liposome. *Biomaterials* 33 (3), 916–924. doi:10.1016/j.biomaterials.2011.10.035

Talibi, S. S., Talibi, S. S., Aweid, B., and Aweid, O. (2014). Prospective therapies for high-grade glial tumours: A literature review. *Ann. Med. Surg.* 3 (3), 55–59. doi:10.1016/j.amsu.2014.04.003

Tu, L., Zhang, W., Ni, L., Xu, Z., Yang, K., Gou, H., et al. (2022). Study of sox combined with intraperitoneal high-dose paclitaxel in gastric cancer with synchronous peritoneal metastasis: A phase II single-arm clinical trial. *Cancer Med.* doi:10.1002/cam4.5277

Wang, X., Chen, J. X., Zhou, Q., Liu, Y. H., Mao, Q., You, C., et al. (2016). Statistical report of central nervous system tumors histologically diagnosed in the sichuan Province of China from 2008 to 2013: A west China glioma center report. *Ann. Surg. Oncol.* 23 (5), 946–953. doi:10.1245/s10434-016-5410-1

Wang, X., Ye, L., He, W., Teng, C., Sun, S., Lu, H., et al. (2022). *In situ* targeting nanoparticles-hydrogel hybrid system for combined chemo-immunotherapy of glioma. *J. Control. Release* 345, 786–797. doi:10.1016/j.jconrel.2022.03.050

Wu, D., Sun, J., Wang, H., and Ma, C. (2021). LncRNA SOCS2-AS1 promotes the progression of glioma via regulating ITGB1 expression. *Neurosci. Lett.* 765, 136248. doi:10.1016/j.neulet.2021.136248

Wu, Y., Hu, Y., Tang, L., Yin, S., Lv, L., and Zhou, P. (2022). Targeting CXCR4 to suppress glioma-initiating cells and chemoresistance in glioma. *Cell Biol. Int.* 46 (9), 1519–1529. doi:10.1002/cbin.11836

Xia, T., Ji, Y., Lu, Y. N., Xie, H. J., You, Y. W., and You, B. (2022). Autophagy promotes recurrence of nasopharyngeal carcinoma via inducing the formation of dormant polyploid giant cancer cells. *Zhonghua Er Bi Yan Hou Tou Jing Wai Ke Za Zhi* 57 (9), 1102–1109. doi:10.3760/cma.j.cn115330-20220119-00034

Xie, J., and Wang, C-H. (2006). Electrospun micro-and nanofibers for sustained delivery of paclitaxel to treat C6 glioma *in vitro*. *Pharm. Res.* 23 (8), 1817–1826. doi:10.1007/s11095-006-9036-z

Xie, J., Marijnissen, J. C., and Wang, C-H. (2006). Microparticles developed by electrohydrodynamic atomization for the local delivery of anticancer drug to treat C6 glioma *in vitro*. *Biomaterials* 27 (17), 3321–3332. doi:10.1016/j.biomaterials.2006.01.034

Xue, J., Zhao, Z., Zhang, L., Xue, L., Shen, S., Wen, Y., et al. (2017). Neutrophil-mediated anticancer drug delivery for suppression of postoperative malignant glioma recurrence. *Nat. Nanotechnol.* 12 (7), 692–700. doi:10.1038/nnano.2017.54

Yang, P., Wang, Y., Peng, X., You, G., Zhang, W., Yan, W., et al. (2013). Management and survival rates in patients with glioma in China (2004-2010): A retrospective study from a single-institution. *J. Neurooncol.* 113 (2), 259–266. doi:10.1007/s11060-013-1103-9

Yasunaga, M., Yahata, H., Okugawa, K., Shimokawa, M., Maeda, Y., Hori, E., et al. (2022). Prognostic impact of adding bevacizumab to carboplatin and paclitaxel for recurrent, persistent, or metastatic cervical cancer. *Taiwan. J. Obstet. Gynecol.* 61 (5), 818–822. doi:10.1016/j.tjog.2022.06.005

Yilmaz, E., Emengen, A., Ceylan, E. C., Cabuk, B., Anik, I., and Ceylan, S. (2022). Endoscopic transnasal surgery in optic pathway gliomas located in the chiasma-hypothalamic region: Case series of ten patients in a single-center experience and endoscopic literature review. *Childs Nerv. Syst.* 38, 2071–2082. doi:10.1007/s00381-022-05665-7

Zhan, C., Wei, X., Qian, J., Feng, L., Zhu, J., and Lu, W. (2012). Co-delivery of TRAIL gene enhances the anti-glioblastoma effect of paclitaxel *in vitro* and *in vivo*. *J. Control. Release* 160 (3), 630–636. doi:10.1016/j.jconrel.2012.02.022

Zhang, P., Hu, L., Yin, Q., Feng, L., and Li, Y. (2012). Transferrin-modified c [RGDfK]-paclitaxel loaded hybrid micelle for sequential blood-brain barrier penetration and glioma targeting therapy. *Mol. Pharm.* 9 (6), 1590–1598. doi:10.1021/mp200600t

Zhang, X., Marjani, S. L., Hu, Z., Weissman, S. M., Pan, X., and Wu, S. (2016). Single-cell sequencing for precise cancer research: Progress and prospects. *Cancer Res.* 76 (6), 1305–1312. doi:10.1158/0008-5472.CAN-15-1907

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 19 December 2022
DOI 10.3389/fphar.2022.1081244

OPEN ACCESS

EDITED BY
Zhi-Qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Ranran Sun,
Zhengzhou University, China
Wei Li,
Shenzhen Longhua District Central Hospital, China

*CORRESPONDENCE
Changfa Qu,
2106@hrbmu.edu.cn

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 27 October 2022
ACCEPTED 07 December 2022
PUBLISHED 19 December 2022

CITATION
Liu M, Xiao Q, Yu X, Zhao Y and Qu C (2022), Characterization of lung adenocarcinoma based on immunophenotyping and constructing an immune scoring model to predict prognosis.
Front. Pharmacol. 13:1081244.
doi: 10.3389/fphar.2022.1081244

Characterization of lung adenocarcinoma based on immunophenotyping and constructing an immune scoring model to predict prognosis

Mengfeng Liu[1], Qifan Xiao[1], Xiran Yu[1], Yujie Zhao[2] and Changfa Qu[1]*

[1]Department of Thoracic Surgery, Harbin Medical University Cancer Hospital, Harbin Medical Sciences University, Harbin, China, [2]Regional Marketing Department, YuceBio Technology Co., Shenzhen, China

Background: Lung cancer poses great threat to human health, and lung adenocarcinoma (LUAD) is the main subtype. Immunotherapy has become first line therapy for LUAD. However, the pathogenic mechanism of LUAD is still unclear.

Methods: We scored immune-related pathways in LUAD patients using single sample gene set enrichment analysis (ssGSEA) algorithm, and further identified distinct immune-related subtypes through consistent clustering analysis. Next, immune signatures, Kaplan-Meier survival analysis, copy number variation (CNV) analysis, gene methylation analysis, mutational analysis were used to reveal differences between subtypes. pRRophetic method was used to predict the response to chemotherapeutic drugs (half maximal inhibitory concentration). Then, weighted gene co-expression network analysis (WGCNA) was performed to screen hub genes. Significantly, we built an immune score (IMscore) model to predict prognosis of LUAD.

Results: Consensus clustering analysis identified three LUAD subtypes, namely immune-Enrich subtype (Immune-E), stromal-Enrich subtype (Stromal-E) and immune-Deprived subtype (Immune-D). Stromal-E subtype had a better prognosis, as shown by Kaplan-Meier survival analysis. Higher tumor purity and lower immune cell scores were found in the Immune-D subtype. CNV analysis showed that homologous recombination deficiency was lower in Stromal-E and higher in Immune-D. Likewise, mutational analysis found that the Stromal-E subtype had a lower mutation frequency in TP53 mutations. Difference in gene methylation (ZEB2, TWIST1, CDH2, CDH1 and CLDN1) among three subtypes was also observed. Moreover, Immune-E was more sensitive to traditional chemotherapy drugs Cisplatin, Sunitinib, Crizotinib, Dasatinib, Bortezomib, and Midostaurin in both the TCGA and GSE cohorts. Furthermore, a 6-gene signature was constructed to predicting prognosis, which performed better than TIDE score. The performance of IMscore model was successfully validated in three independent datasets and pan-cancer.

KEYWORDS
LUAD, immune, immunophenotyping, immune scoring model, prognosis

1 Introduction

Lung cancer still serves as the most common malignancy and the foremost cause of cancer death in the world (Thandra et al., 2021). Lung adenocarcinoma (LUAD) as a main subtype of lung cancer belong to the larger group of non-small cell lung cancer, and accounts for about 40% of all lung cancer cases Moreover, it is strongly associated with smoking (Pintarelli et al., 2019; Spella and Stathopoulos, 2021). Although treatment modalities such as surgery, chemoradiotherapy, targeted therapy, and immunotherapy have been widely used for the treatment of lung cancer, but the prognosis remains dismal (Ladanyi and Pao, 2008; Bronte et al., 2010; Denisenko et al., 2018), with a 5-year survival rate as low as less than 10% (Hirsch et al., 2017).

Cancer evolution is influenced by complex interactions between tumor cells and host immune responses within the tumor microenvironment (Taube et al., 2018). Different immune cell populations are actively involved in tumor immune microenvironment (TIME), however, their relationship is currently unclear in LUAD therapy (Zhang et al., 2020). Immune cells are both positive and negative regulators of cancer progression (Zamarron and Chen, 2011). For example, B-cells exert antitumor functions by enhancing T-cell immunity, stimulating the production of interferon-γ and helping natural killer (NK) cells against tumors (Sorrentino et al., 2011). However, B-cells also suppress immune responses and promote angiogenesis (Schwartz et al., 2016). Considering the importance of TIME in cancers, we attempted to classify LUAD based on immune pathways and compared their characteristics.

Disease prediction models have been widely used to evaluate patient survival and other prognostic indicators in LUAD (Jiang et al., 2020; Jiang et al., 2022). Recently, a novel model related to lactate metabolism for predicting overall survival and immune signature in LUAD was reported and different immune signatures were built (Jiang et al., 2022). Moreover, an increasing number of LUAD prognostic biomarkers have been discovered by analyzing expression profiles from public databases and related clinical information (Jiang et al., 2020).

Due to advances in genetic technology, various molecular subtypes and gene expression studies have been published and popularized for the identification of prognostic biomarkers (Li and Wang, 2021). However, the prognostic manifestations of established biomarkers are controversial and limited. In the current study, we identified and classified three immune subtypes and a six-gene combination capable of predicting survival in patients with LUAD.

2 Materials and methods

2.1 Data collection

To better understand the pathogenesis of LUAD, RNA-Seq data of LUAD with clinical survival and characteristic information were downloaded from TCGA database by using TCGA GDC API. Fragments per kilobase of transcript per million fragments mapped (FPKM) were converted into transcripts per million (TPM), and then Log2 was transformed for subsequent analysis. In addition, we downloaded the GSE37745, GSE50081, GSE30219, GSE31210 microarray datasets with survival times from the GEO database (Yamauchi et al., 2012; Botling et al., 2013; Rousseaux et al., 2013; Der et al., 2014). For GEO, MINiML formatted family file(s) were downloaded and samples was preprocessed using RMA implemented in affy package. Normalized data values were transformed in log2 space and used in subsequent analysis.

2.2 Data quality control

To ensure the accuracy of downstream analysis, the RNA-Seq data of TCGA-LUAD without clinical follow-up information or survival time were removed. Next, the data without status were filtered. Then, only the genes with more than one expression in more than 50% of the samples were retained.

For GSE data, firstly, normal tissues, samples without clinical follow-up information, samples without OS data and samples without status were filtered. Then, we converted the probes to symbols according to the annotation file.

2.3 Batch effect processing

The *removeBatchEffect* function of the limma (R package) was used to remove batch effects between different datasets (Ritchie et al., 2015). Principal component analysis (PCA) was used to observe the batch effect (Supplementary Figure S1).

2.4 LUAD classification based on pathway score

In order to explore the molecular typing of LUAD, single sample gene set enrichment analysis (ssGSEA) analysis was used to calculate the score related to immune pathway (Supplementary Table S1) (Gao et al., 2021). Next, ConsensusclusterPlus (R Bioconductor/R package) software

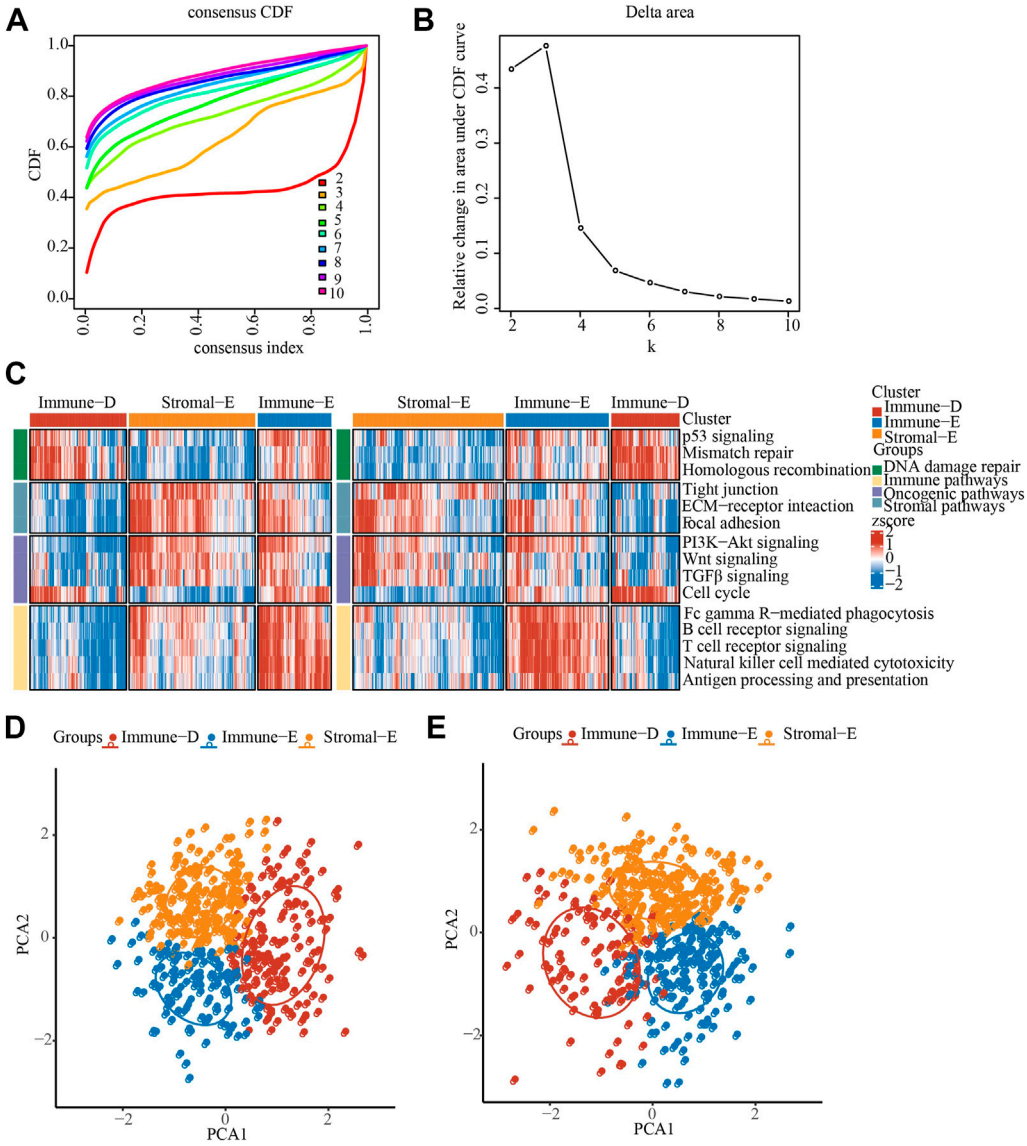


FIGURE 1
Identification of LUAD immunophenotyping. **(A)**, The cumulative distribution function (CDF) of different consensus index. **(B)**, The CDF Delta area curve. **(C)**, The heat map showing the immune pathway score in TCGA and GSE cohorts. **(D)**, PCA analysis in TCGA cohorts, different color means different subtypes. **(E)**, PCA analysis in GSE cohorts, different color means different subtypes.

was used for consensus clustering analysis with "pam" arithmetic and "pearson" distance, and the input was a sample matrix pool of immune pathway scores (Wilkerson and Hayes, 2010). Then, we determined the optimal number of clusters according to the cumulative distribution function (CDF).

2.5 Gene mutation analysis

The *mutect2* software was applied to perform gene mutation analysis (Benjamin et al., 2019). The tumor promoting genes were obtained from previous study. The fisher's test was used to screen genes with significantly high frequency mutations in each subtype, with a threshold of p-value < 0.05. Moreover, the *maftools* software (R package) was used to calculate the tumor mutational burden (TMB) score (Mayakonda et al., 2018).

2.6 Copy number variation (CNV) analysis

GISTIC2.0 was used to analyze the change of CNV (Mermel et al., 2011). If the ratio was greater than 0.2, it was considered as

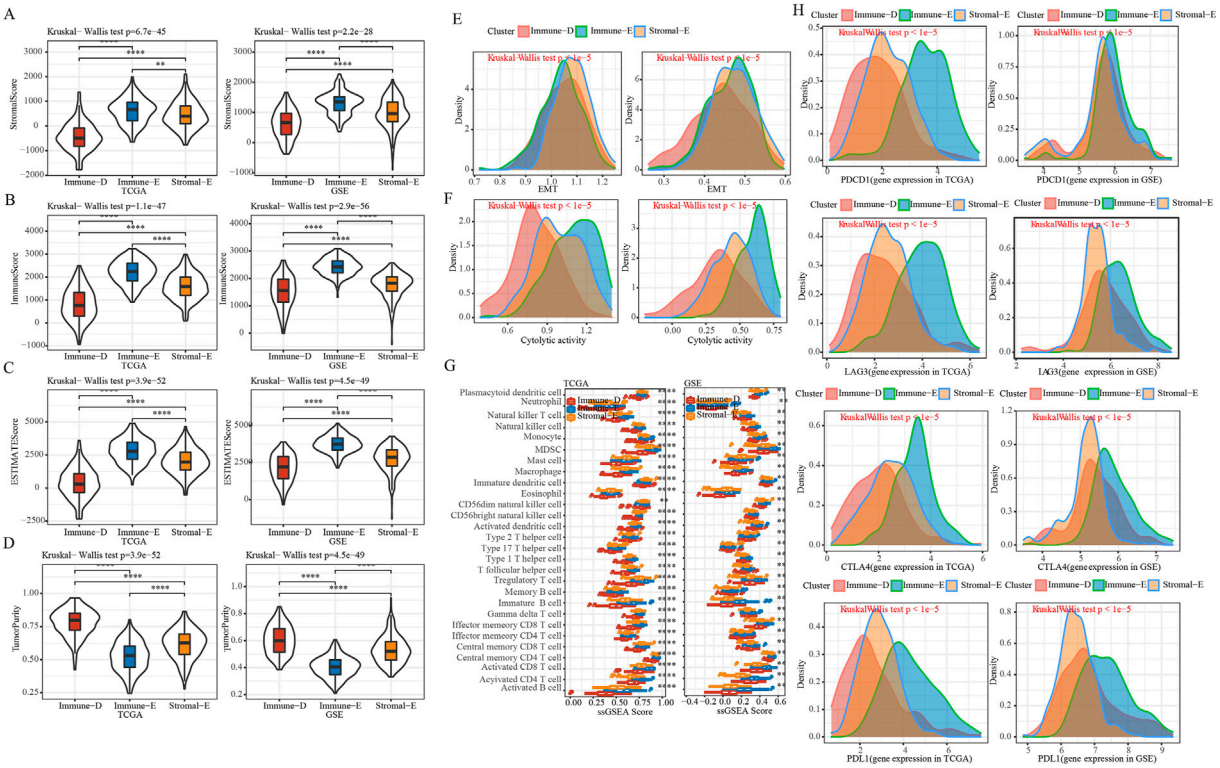


FIGURE 2
Immune features of different subtypes. **(A)**, Comparison of distribution of StromalScore in subtypes. **(B)**, Comparison of distribution of ImmuneScore in subtypes. **(C)**, Comparison of distribution of ESTIMATEScore in subtypes. **(D)**, Comparison of distribution of TumorPurity in subtypes. **(E)**, EMT score Comparison of distribution among subtypes. **(F)**, Comparison of distribution of Cytolytic activity scores among subtypes. **(G)**, Comparison of distribution of immune cell scores among subtypes. **(H)**, Density distribution map of gene PD1, CTLA4, LAG3 and PD-L1 in different subtypes. * means *p*-value < 0.05, ** means *p*-value < 0.01, *** means *p*-value < 0.001, **** means *p*-value < 0.0001.

Gain, if the ratio was less than −0.2, it was considered as Loss, and the rest was considered as Diploid.

2.7 Methylation analysis

The 450K methylation data of LUAD were used to perform methylation analysis of EMT-promoting genes (Wang and Zhou, 2013), and missing values were imputed using the KNN algorithm of the *impute* software (R package) (Hastie et al., 2011).

2.8 Treatment plan sensitivity analysis

We used the tumor immune dysfunction and exclusion (TIDE) (https://tide.dfci.harvard.edu/) algorithm to predict response of immunotherapy (Jiang et al., 2018). The higher the TIDE prediction score, the higher the likelihood of immune escape, and the lower the likelihood that the patient would benefit from immunotherapy. Moreover, IC50 (half maximal inhibitory concentration) analysis was used to determine the sensitivity of different subtypes to chemotherapy drugs using pRRophetic method.

2.9 Weighted gene co-expression network analysis (WGCNA)

WGCNA was performed by WGCNA (R package) to filter hub genes in the module related to different LUAD subtypes (Langfelder and Horvath, 2008). Hub genes refer to genes that play key roles in modules and are generally closely related to shape. The hub genes were further subjected to KEGG pathway analysis by using clusterProfiler (R Bioconductor/R package). *p*-value < 0.05 was considered significance.

2.10 IMscore predicted by prognostic model

IMScore model was constructed using univariate Cox regression analysis and LASSO analysis. Finally, six prognostic

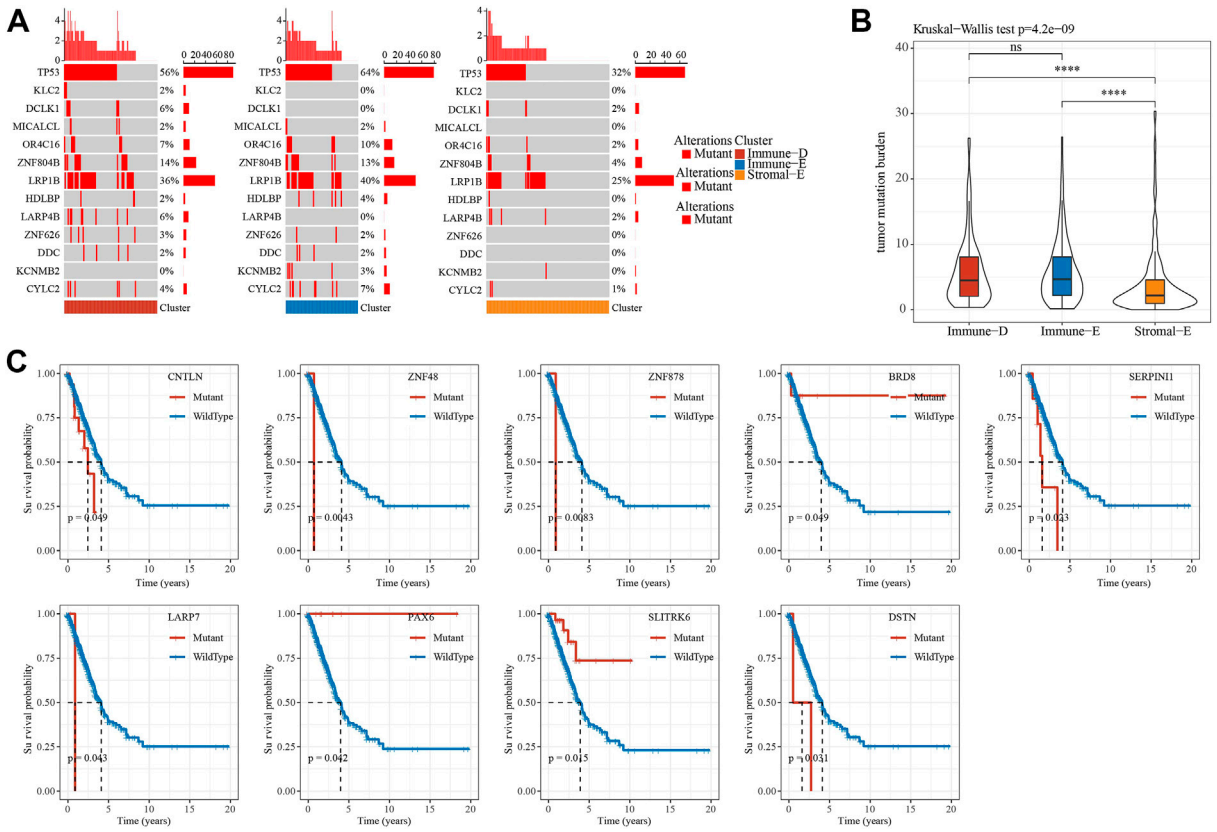


FIGURE 3
Mutation features of different subtypes. **(A)**, Waterfall plot of tumor mutated genes in different subtypes. **(B)**, Differential analysis of TMB distribution in different subtypes. **(C)**. Kaplan-Meier survival analysis of mutation and wild-type tumor driver genes. Ns means no significance, * means *p*-value < 0.05, ** means *p*-value < 0.01, *** means *p*-value < 0.001, **** means *p*-value < 0.0001.

genes were obtained, including MARCKS, CDK2, SFN, SSBP1, MRE11 and FZD7. IMScore = ΣExp(i)*β(i), where i refers to the immune-related prognostic genes, Exp refers to the expression levels of genes, and β refers to the LASSO coefficients.

3 Results

3.1 LUAD classification based on pathway scores

The pathway scores of different samples in the TCGA and GSE cohorts were calculated by ssGSEA algorithm, and then consensus clustering analysis was performed on the pathway scores. The optimal number of clusters was determined according to the CDF, and the CDF Delta area curve was observed. When it was selected as 3, it has a relatively stable clustering result, and finally we choose k = 3 to obtain three related subtypes, including immune-Enrich subtype (Immune-E), stromal-Enrich subtype (Stromal-E) and immune-Deprived subtype (Immune-D) (Figures 1A,B). A heatmap showing the trends in pathway scores for each sample demonstrated a good discrimination between the three different subtypes (Figure 1C). In addition, PCA analysis between different subtypes showed that in both datasets, there were distinct boundaries between different subtypes (Figures 1D,E).

3.2 Analysis of clinical characteristics of different subtypes

Survival analysis showed a better prognosis in the Stromal-E subtype and a worse prognosis for the Immune-D subtype, which was consistent in the TCGA cohort (left) and the GSE cohort (right) (Supplementary Figure S2A). Analysis of differences in the distribution of clinical features among subtypes in the TCGA dataset showed significant differences in age, sex as well as the distribution of T Stage among subtypes (Supplementary Figure S2B). Comparative analysis of the molecular subtypes with the six previously identified immunophenotypes (Thorsson et al., 2018)

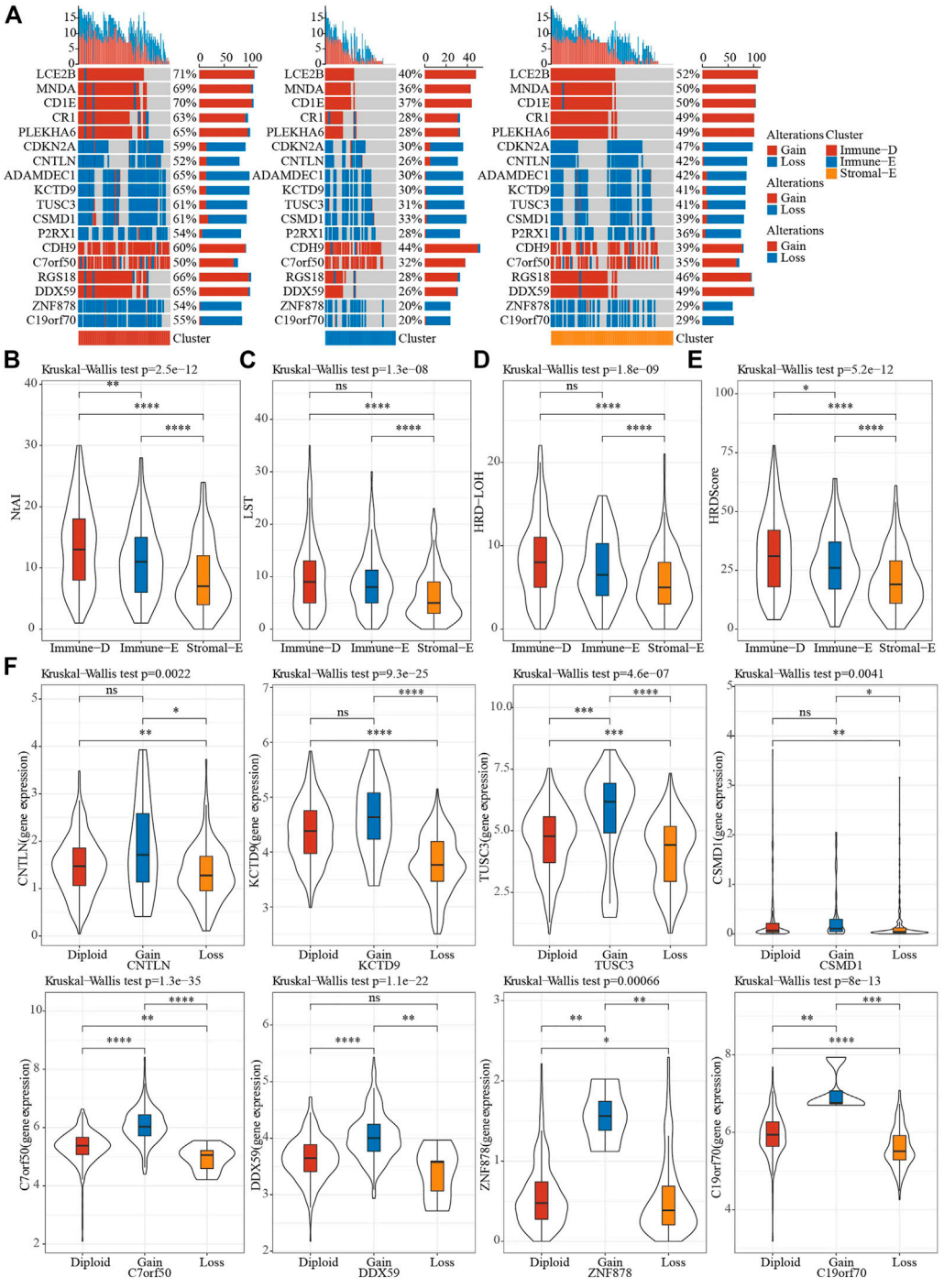


FIGURE 4
CNV features of different subtypes. **(A)**, CNV distribution of tumor driver genes. **(B–E)**, Differences in NtAI, LST, LOH and HRD score of different subtypes, respectively. **(F)**, Differences in gene expression by tumor driver gene CNV groupings. Ns means no significance, * means p-value < 0.05, ** means p-value < 0.01, *** means p-value < 0.001, **** means p-value < 0.0001.

showed that the previously published immunophenotypes were significantly different from our current study, for example, the Immune-E subtype had the highest proportion of the C2 subtype and C3 subtype had the best prognosis and also contributed the largest proportion in Stromal-E subtype. These results supported the reliability of our immunosubtyping (Supplementary Figure S2C)

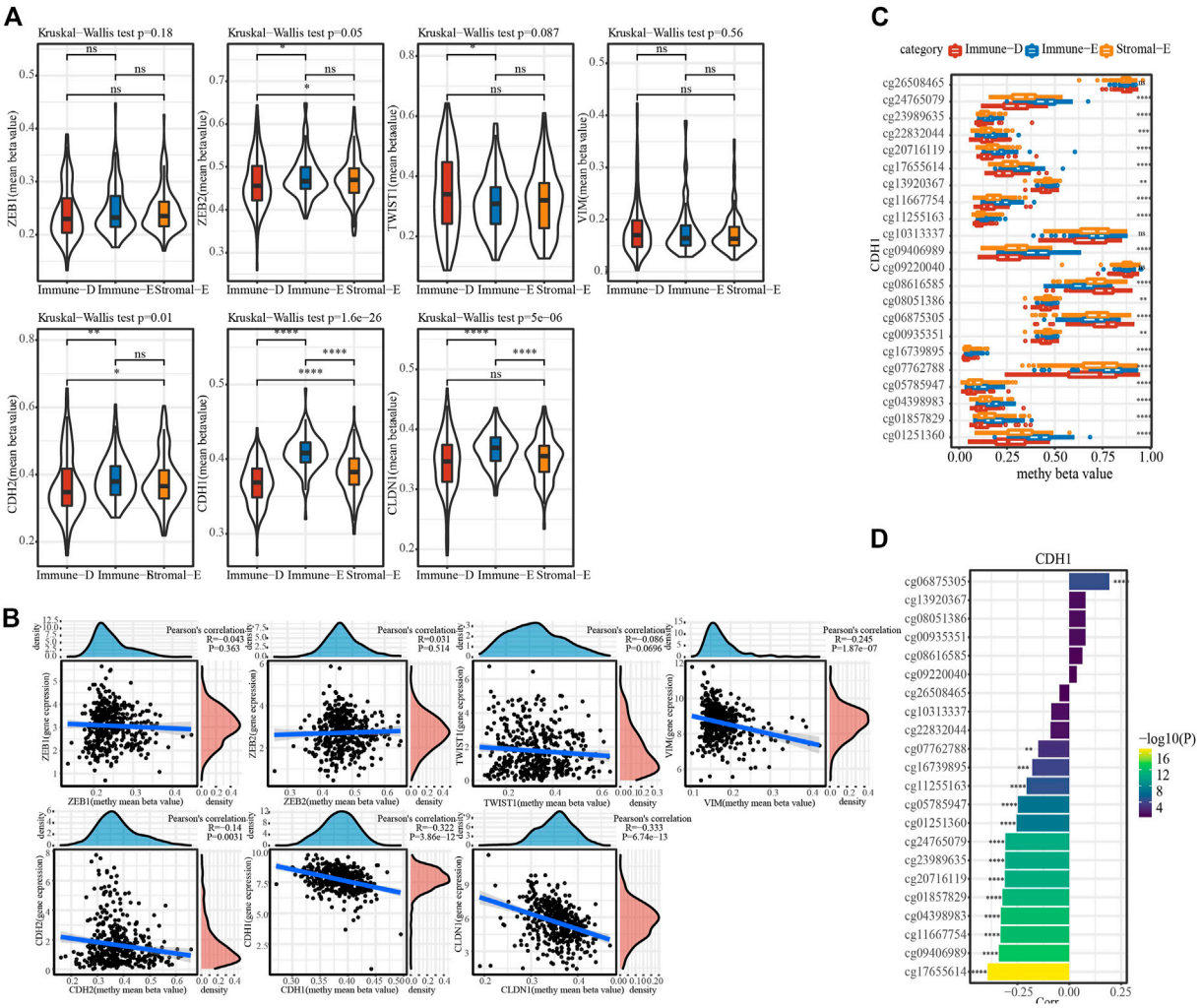


FIGURE 5
Methylation features of different subtypes. **(A)**, The distribution of methylation value of EMT-promoting gene in subtypes. **(B)**, Correlation analysis of methylation value and expression value of EMT-promoting gene. **(C)**, The beta value of cg probe site of gene CDH1 in subtype Differences in distribution among subtypes. **(D)**. Correlation of beta values of cg probe sites for CDH1 with CDH1 gene expression. Ns means no significance, * means *p*-value < 0.05, ** means *p*-value < 0.01, *** means *p*-value < 0.001, **** means *p*-value < 0.0001.

3.3 Immune characteristics among different LUAD subtypes

In order to further understand the immune characteristics among the various subtypes of LUAD, we first used the *Estimate* software to evaluate the immune scores and tumor purity score of the TCGA and GSE cohorts. The results demonstrated that the immune score, stromal score, and ESTIMATE score were the highest in Immune-E subtype both in the TCGA and GSE cohorts (Figures 2A–C). The opposite was true for the tumor purity score (Figure 2D). Epithelial-mesenchymal Transition (EMT) is closely related to tumor metastasis and recurrence (Wang and Zhou, 2013). Therefore, we applied the ssGSEA algorithm to evaluate the difference in EMT scores between different subtypes. The results showed that in the TCGA cohort, the Stromal-E subtype had the highest EMT score, and in the GSE cohort Immune-E subtype was the highest (Figure 2E). Cytolytic activity is associated with immunotherapy, and here ssGSEA was used to assess differences in Cytolytic activity score (Rooney et al., 2015). We found that both in TCGA and GSE cohorts, the score was highest in Immune-D subtype and the lowest in Immune-E subtype (Figure 2F). To further understand the status of different immune cells, we assessed immune cell infiltration between different subtypes and found the highest scores in Immune-E, such as

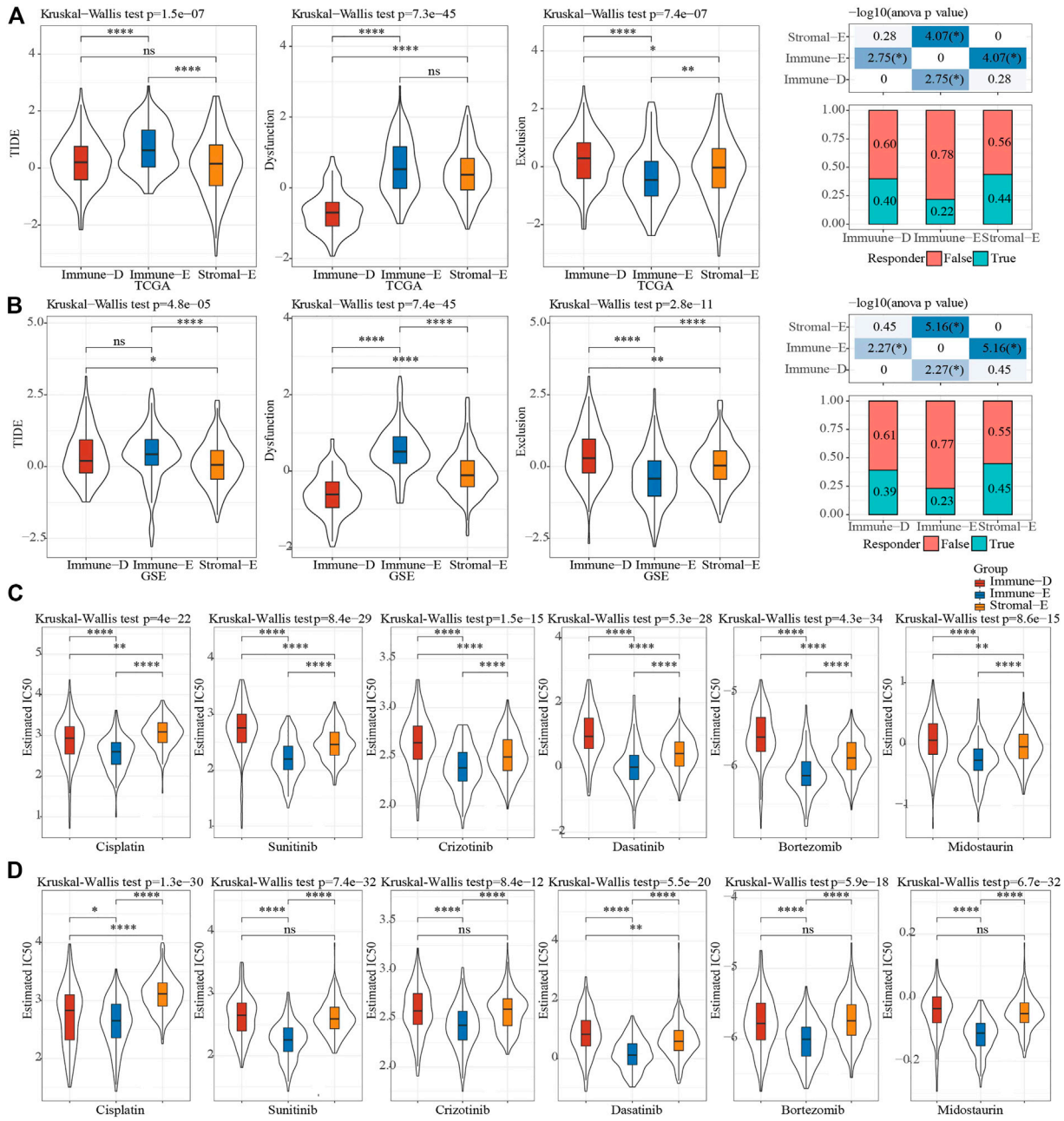


FIGURE 6
Sensitivity analysis of treatment options. **(A)**, Difference analysis of TIDE score of different immune subtypes in TCGA cohort. **(B)**, Difference analysis of TIDE score of different immune subtypes of GSE cohort. **(C)**, Difference analysis of drug IC50 of in TCGA cohort. **(D)**, Difference analysis of drug IC50 of in GSE cohort. Ns means no significance, * means *p*-value < 0.05, ** means *p*-value < 0.01, *** means *p*-value < 0.001, **** means *p*-value < 0.0001.

Plasmacytoid dendritic cell, Regulatory T cell and Gamma delta T cell. This was also the same in the TCGA and GSE cohorts (Figure 2G). Moreover, we determine the expression level of PD1, PD-L1, CTLA4 and LAG3, and observed that the expression level was higher in Immune-E subtype (Figure 2H).

3.4 Mutational analysis of tumor driver genes in different subtypes

A total of 172 tumor driver genes were obtained from previous study (Gao et al., 2013). Mutation analysis of tumor

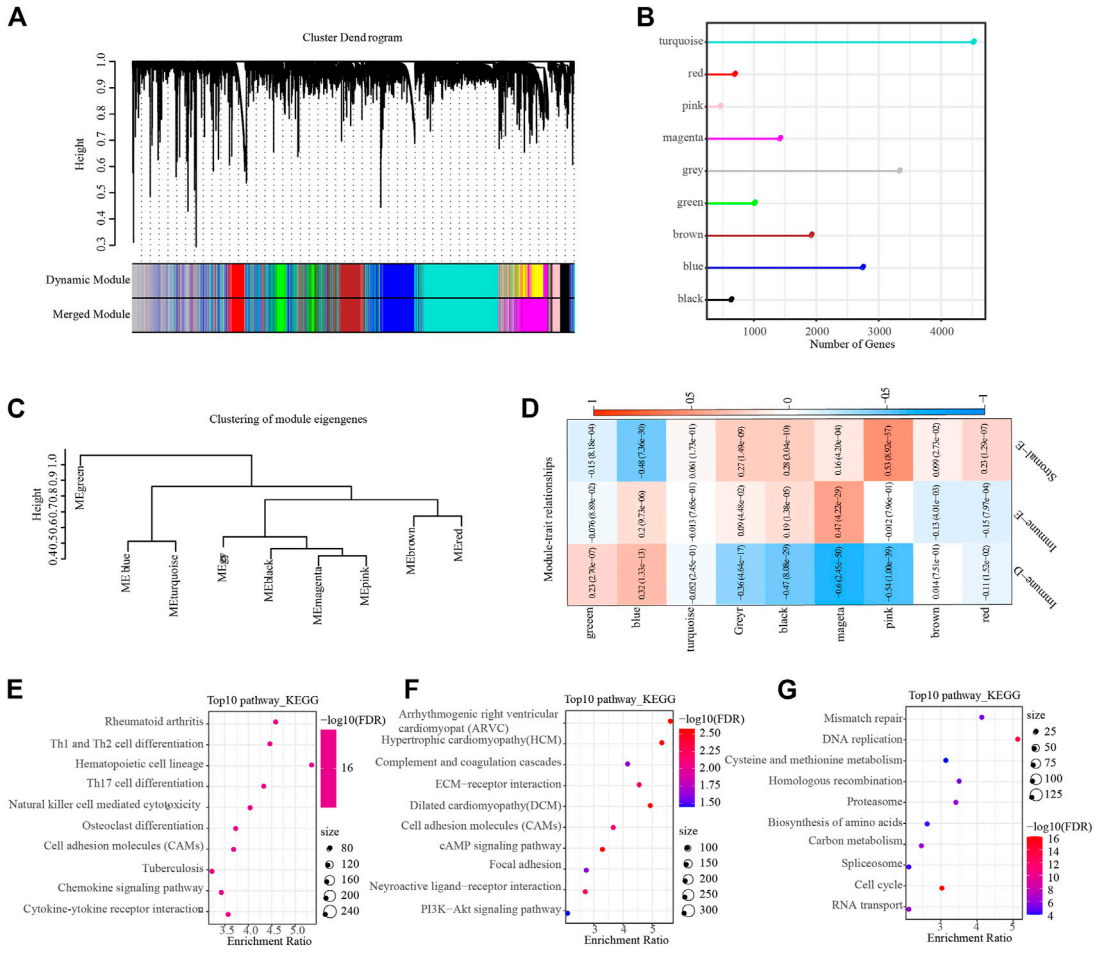


FIGURE 7
WGCNA identifies functional modules associated with different subtypes. **(A)**, Dendrogram of all genes clustered based on a dissimilarity measure. Different colors on the bottom panel represent different modules. **(B)**, Lollipop plot showing the number of genes in different modules. **(C)**, Clustering tree showing correlation between different modules. **(D)**, Heat map showing module correlation with different subtypes. **(E–G)**, The top 10 KEGG pathway in magenta (Immune-E), pink (Stromal-E) and blue (Immune-D) module, respectively.

driver genes in different subtypes of TCGA dataset found that 13 genes had different mutations in different groups and TP53 had the highest mutation frequency in the Immune-E subtype (Figure 3A). The results of TMB analysis in different subtypes found that TMB in Stroma-E subtype was significantly lower than that in Immune-E and Immune-D subtypes. There was no difference in TMB distribution between Immune-E and Immune-D subtypes in TCGA dataset (Figure 3B). In addition, survival curve (KM) analysis of mutations in driver genes and wild-type samples in TCGA dataset demonstrated that mutations in nine genes were significantly different from wild-type, including CNTLN, ZNF48, ZNF878, BRDS, SERPINI1, LARP7, SLITRK6 and DSTN (Figure 3C). Those analysis indicated that three subtypes may predict mutation status.

3.5 CNV analysis of tumor driver genes in different subtypes

A total of 159 of 172 tumor driver genes had CNV data. To understand CNV in tumor driver genes, GISTIC2 was used to perform CNV analysis. The results showed that the amplification and deletion of 159 driver genes were significant in different subtypes in TCGA dataset, and top 18 genes were visualized. Notably, immune-D had the largest amounts of CNVs (Figure 4A). Homologous recombination deficiency (HRD) is associated with a poorer cancer prognosis (Knijnenburg et al., 2018). Loss of heterozygosity (LOH), LST (large-scale state transitions), NtAI (number of telomeric allelic imbalances) score and HRD score were selected to assess the HRD status in different subtypes, and we found that the above scores were the

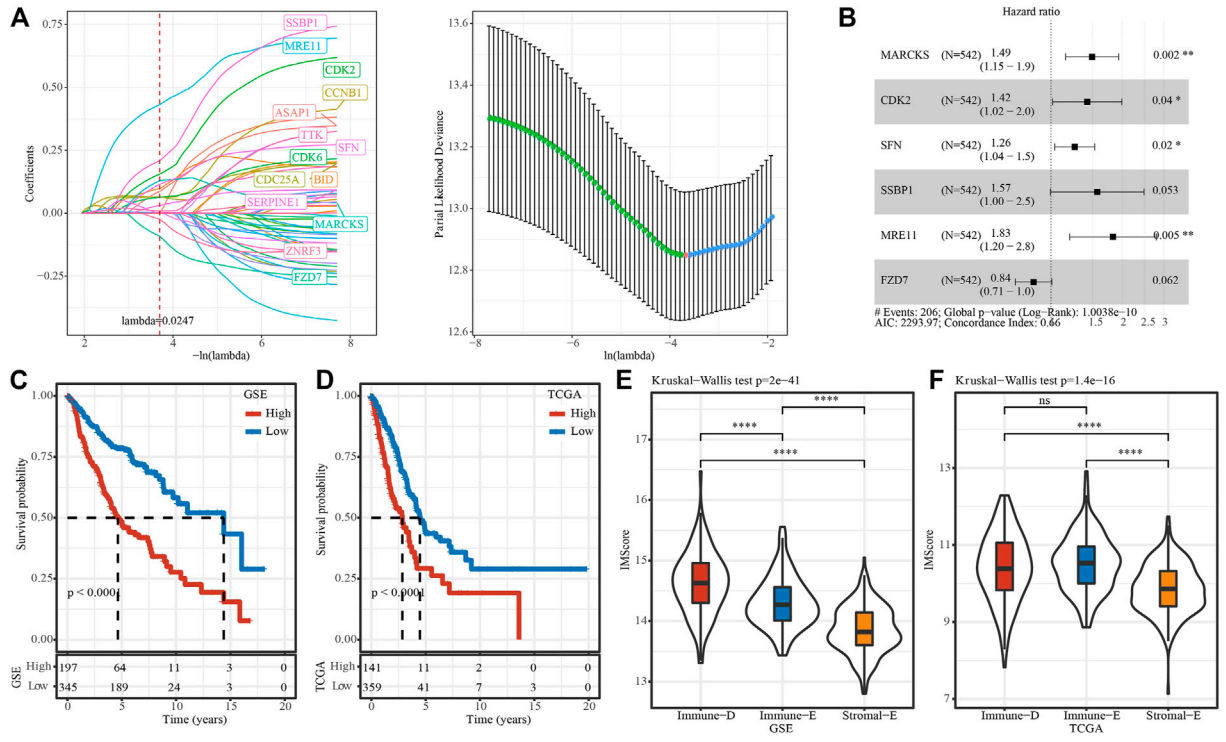


FIGURE 8
IMscore model to predict prognosis. **(A)**, Distribution of IMscore for each sample in the TCGA cohort. **(B)**, Distribution of LASSO coefficients for six genes. C and D, Kaplan-Meier survival analysis in different IMscore group in TCGA **(C)** and GSE **(D)** cohort. E and F, Differences in the distribution of IMscore among different subtypes (**(E)**, TCGA cohort; **(F)**, GSE cohort). Ns means no significance, * means *p*-value < 0.05, ** means *p*-value < 0.01, *** means *p*-value < 0.001, **** means *p*-value < 0.0001.

lowest in Stromal-E and highest in Immune-D (Figures 4B–E). The expression level of tumor driver genes in the gene CNV grouping showed that the expression of the genes corresponding to the grouping with Gain was higher, while that corresponding to the grouping with Loss was lower (Figure 4F).

3.6 Methylation analysis of genes in different subtypes

The methylation status of genes plays an important role in gene expression. DNA methylation analysis of seven EMT-promoting genes showed that the methylation status of ZEB1 and VIM did not differ among subtypes in TCGA dataset, while the methylation status of ZEB2, TWIST1, CDH2, CDH1 and CLDN1 were significantly different in each subtype (Figure 5A). The correlation analysis between gene expression value and methylation status showed that the gene expression values of VIM, CDH2, CDH1, and CLDN1 were significantly negatively correlated with methylation status (Figure 5B). The distribution of cg sites of CDH1 in different subtypes in TCGA dataset showed that some sites had significant differences in different subtype groups, and that the highest value in Immune-E was in a hypermethylated state (Figure 5C). From the correlation analysis of CDH1 cg loci and CDH1 gene expression, most of the cg loci and CDH1 gene expression showed a significant negative correlation (Figure 5D).

3.7 Immunotherapy and drug sensitivity of different subtypes

The TIDE (https://tide.dfci.harvard.edu/) algorithm was used to assess the potential clinical effects of immunotherapy on different molecular subtypes. In the TCGA cohort, the TIDE score of Immune-E was significantly higher than that of Stromal-E and Immune-D. The immunotherapy effect analysis indicated that the proportion of responses in Immune-E was only 22%, which was much lower than that of Stromal-E and Immune-D (Figure 6A). In GSE dataset, TIDE score was also higher in Immune-E. (Figure 6B). In addition, we also analyzed the response of different subtypes to the traditional chemotherapy drugs Cisplatin, Sunitinib, Crizotinib, Dasatinib, Bortezomib, Midostaurin. The results suggested that all the six drugs were

more sensitive to the Immune-E subtype in both the TCGA and GSE cohorts (Figures 6C, D).

3.8 WGCNA to identify the key gene of subtypes

Hierarchical cluster analysis of cohort samples showed no discrete samples in the TCGA cohorts (Supplementary Figure S3A). In order to ensure that the gene network we constructed conformed to the scale-free distribution, we set the β value to 9 at the time $R^2 > 0.85$ (Supplementary Figures S3A,S3C). After clustering the modules and merging the closer modules into a new module, we acquired a total of nine modules (Figure 7A). The lollipop graph showed the number of genes in each module. It can be observed that the turquoise color module had the largest number of genes, which was more than 4,000 genes (Figure 7B). The heat map of correlation analysis of each subtype and module showed that the Immune-D subtype had the highest positive correlation with the blue module, the Immune-E subtype had the highest positive correlation with the magenta module, and the Stromal-E subtype had the highest positive correlation with the pink module (Figures 7C,D). Further, we performed KEGG pathway analysis on the genes in the above three modules, and the results demonstrated that the genes in the magenta module were closely related to Natural killer cell mediated cytotoxicity, the genes in the pink module were closely related to ECM-receptor interaction, and the genes in the blue module were closely related to Genes were closely related to Mismatch repair, etc. (Figures 7E–G).

3.9 IMscore prognostic model construction

In GSE, six prognostic pathways (Fc gamma R-mediated phagocytosis mediated phagocytosis, p53 signaling, Mismatch repair, Homologous recombination, Wnt signaling, Cell cycle) from 15 pathways were identified by univariate COX analysis. Then, Pearson analysis on the genes in six pathways and six pathways score was used to select the top20 genes. Through univariate COX analysis and LASSO regression analysis, we obtained six genes form 20 genes as related genes affecting prognosis (Figures 8A, B), and the model score was developed based on the following formula: IMScore = 0.400*MARCKS + 0.353*CDK2 + 0.232*SFN + 0.450*SSBP1 + 0.604*MRE11 − 0.169*FZD7.

The survminer package was used to find the best cutoss of IMScore and divide the GSE and TCGA data set samples into high IMscore group (high group) and low IMscore group (low group). In the GSE cohort, survival analysis of high- and low-groups found that the prognosis of high groups was significantly worse than that of low groups. High group also had worse survival outcome in TCGA datset (Figures 8C, D). In addition, we compared the distribution differences of IMScore among subtypes in different datasets and found that: IMScore in Stromal-E had the lowest in GSE and TCGA, while Immune-D had the highest IMscore in GSE, Immune-E had highest IMscore in TCGA dataset (Figures 8E, F).

3.10 Prediction efficiency of IMscore prognostic model

The survival curve analysis of IMScore model in different cancer types showed that except ESCA, our IMScore had significant differences in high and low IMScore in all cancer types, and that the prognosis of high IMScore was significantly worse than that of low IMScore (Supplementary Figure S4). Furthermore, IMvigor210, GSE91061 and GSE135222 data were used to examine the efficiency of the IMscore model in immunotherapy. To evaluate the efficiency of IMscore, TIDE analysis was used as a control. For IMvigor210 cohort, survival analysis showed significant differences in survival among different groups (IMscore: p-value < 0.0001; TIDE, p-value = 0.012) (Supplementary Figures S5A, S5B). For GSE91061 cohort, survival analysis showed significant differences in survival among different IMscore groups (p-value < 0.023) (Supplementary Figures S5C). TIDE group was not statistically significant (p-value = 0.067) (Supplementary Figures S5D). For GSE135222 cohort, survival analysis showed no significant differences in IMscore or TIDE groups (IMscore: p-value = 0.055; TIDE, p-value = 0.051) (Supplementary Figures S5E, S5F). Notably, ROC curve analysis demonstrated that in the above three cohorts, the predictive efficiency of IMscore was higher than that of TIDE.

4 Conclusion

LUAD is one of the most common malignant tumors, with high metastasis rate and strong invasiveness. Its low 5-year survival rate seriously threatens the life and health of human beings (Charloux et al., 1997; Tan et al., 2016). Significantly, early-stage LUAD has been reported to be associated with a higher risk of postoperative recurrence and death (Bittner et al., 2014). The immune system has been shown to play a critical role and even determine different stages of cancer development and progression (Shurin, 2018). Hence, an accurate classification of LUAD patients according to immune characteristics and the identification of LUAD biomarkers have positive significance for the selection of LUAD treatment methods. A previous study reported six subtypes in LUAD, including Wound Healing, IFN-γ Dominant, Inflammatory, Lymphocyte Depleted, Immunologically Quiet, and TGF-β Dominant (Thorsson et al., 2018). However, in this research, three molecular

subtypes classified by immunological features were obtained and defined as Immune-E, Stromal-E and Immune-D.

Kaplan-Meier survival analysis was able to determine survival differences between different subtypes (Goel et al., 2010). We found that Stromal-E subtype had a better prognosis and Immune-D had a worse prognosis. To gain a deeper understanding of the differences in survival between the different subtypes, we used *Estimate* software for immune calculating index scores and tumor purity scores (Yoshihara et al., 2013). The results showed that higher tumor purity and lower immune cell scores were in the Immune-D subtype, which accounted for the poorer prognosis of the subtype.

Recently, immunotherapy has also been developed as a new treatment for LUAD (Saito et al., 2018). Moreover, antibodies against PD-1 and PD-L1 have been reported to be effective in the treatment of various malignancies (Antonia et al., 2017; Bie et al., 2021). Although the biology of the TIME driving these responses was not fully understood, it is critical for the design of immunotherapeutic strategies. We found that the expression level of PD1, PD-L1, CTLA4 and LAG3 was higher in Immune-E subtype. However, as we mentioned above, the survival analysis showed that the Stromal-E subtype had a better prognosis, which suggested the complexity of tumorigenesis mechanisms.

TMB is an emerging tumor biomarker and it is associated with response to PD-1/PD-L1 targeted therapies in lung cancer (Spigel et al., 2016). We found that TMB in the Stromal-E subtype was significantly lower than in the Immune-E and Immune-D subtypes. TP53, known as the guardian of the genome, is one of the most well-known tumor suppressor genes (Surget et al., 2014). Interestingly, the mutation analysis suggested that Stromal-E subtype had a lower TP53 mutation rate of only 32%, while the other two subtypes both exceeded 50%, which explained a better prognosis of Stromal-E subtype. Moreover, high levels of TIDE scores suggested that Immune-E was more likely to occur immune escape, suggesting that the Immune-E subtype had limited benefit from immunotherapy (Jiang et al., 2018). Interestingly, although the Immune-E subtype had little limited benefit from immunotherapy, IC50 analysis of chemotherapeutic agents showed that this subtype was more sensitive to chemotherapeutic agents, including Cisplatin, Sunitinib, Crizotinib, Dasatinib, Bortezomib and Midostaurin.

The above analysis showed a high heterogeneity among different immune subtypes. In order to better understand the differences between different subtypes, we used WGCNA analysis to identify hub genes among each subtype. Extensive analyses showed that WGCNA was an effective method for identifying phenotype-genotype linkages and biomarkers and therapeutic targets (Niemira et al., 2019; Ma et al., 2021; Zhang et al., 2022). Nine modules were obtained, and pink module was highly associated with Stromal-E subtype. Further function enrichment analysis showed that the hub genes in pink module were involved in ECM-receptor interaction pathway. Tumor progression depends not only on cell-autonomous changes in tumor cells, but also on the changes within the microenvironment (Götte and Kovalszky, 2018). An important feature of the dysregulated lung cancer microenvironment is the altered extracellular matrix (ECM), which can promote tumor angiogenesis, allow tumor cell immune escape, etc. (Mahale et al., 2016). In addition, magenta module was positively correlated with Immune-E subtype and participated in natural killer cell mediated cytotoxicity.

Notably, we developed an IMscore model containing six immune-related genes (MARCKS, CDK2, SFN, SSBP1, MRE11 and FZD7) to predict LUAD prognosis. CCNA2-CDK2 complex have been reported to inhibit LUAD progression (Li et al., 2021). CDK2 also was a biomarker for other cancers and next-generation CDK2 inhibitors play an increasingly pivotal role in the treatment of cancer (Wadler, 2001; Chohan et al., 2015; Tadesse et al., 2020). SFN gene encodes a protein participating in regulating epithelial-mesenchymal interaction (Asdaghi et al., 2012). Aya Shiba-Ishii reported that SFN promoted early progression of LUAD by activating cell proliferation (Shiba-Ishii, 2021). SSBP1, MRE11 and FZD7 have been considered as potential treatment sites of LUAD (Sun et al., 2017; Wang et al., 2017; Zeng et al., 2020). Kaplan-Meier survival analysis showed the high IMScore group had a significantly lower prognosis than the low IMScore group. Then, we detected the IMScore in three immune-related subtypes. As expected, the lowest score of IMScore in Stromal-E subtype was obtained. To explore the efficiency of the IMscore model, we used the expression profile data of the remaining 32 cancer types in the TCGA database for validation, and found that IMScore was significantly different in all cancer types except ESCA.

5 Conclusion

In this study, we used immune-related signaling pathways to classify LUAD and obtained three different subtypes, which had great differences in survival, gene mutation, CNV, gene methylation, etc. We further constructed an IMscore model to predict the prognosis among different subtypes of LUAD, and observed that the IMscore model had a higher efficiency. In addition, data in different cancers further confirmed the validity of the IMscore model.

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

All authors contributed to this present work: ML designed the study, QX acquired the data, XY drafted the manuscript, YZ and CQ revised the manuscript. All authors read and approved the manuscript.

Conflict of interest

YZ was employed by YuceBio Technology Co.,

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2022.1081244/full#supplementary-material

References

Antonia, S. J., Villegas, A., Daniel, D., Vicente, D., Murakami, S., Hui, R., et al. (2017). Durvalumab after chemoradiotherapy in stage III non–small-cell lung cancer. *N. Engl. J. Med.* 377 (20), 1919–1929. doi:10.1056/NEJMoa1709937

Asdaghi, N., Kilani, R. T., Hosseini-Tabatabaei, A., Odemuyiwa, S. O., Hackett, T-L., Knight, D. A., et al. (2012). Extracellular 14-3-3 from human lung epithelial cells enhances MMP-1 expression. *Mol. Cell. Biochem.* 360 (1), 261–270. doi:10.1007/s11010-011-1065-1

Benjamin, D., Sato, T., Cibulskis, K., Getz, G., Stewart, C., and Lichtenstein, L. (2019). Calling somatic SNVs and indels with Mutect2. *BioRxiv* 2019, 861054.

Bie, F., Tian, H., Sun, N., Zang, R., Zhang, M., Song, P., et al. (2021). Comprehensive analysis of PD-L1 expression, tumor-infiltrating lymphocytes, and tumor microenvironment in LUAD: Differences between asians and caucasians. *Clin. Epigenetics* 13 (1), 229–315. doi:10.1186/s13148-021-01221-3

Bittner, N., Ostoros, G., and Géczi, L. (2014). New treatment options for lung adenocarcinoma-in view of molecular background. *Pathol. Oncol. Res.* 20 (1), 11–25. doi:10.1007/s12253-013-9719-9

Botling, J., Edlund, K., Lohr, M., Hellwig, B., Holmberg, L., Lambe, M., et al. (2013). Biomarker discovery in non-small cell lung cancer: Integrating gene expression profiling, meta-analysis, and tissue microarray validation. *Clin. Cancer Res.* 19 (1), 194–204. doi:10.1158/1078-0432.CCR-12-1139

Bronte, G., Rizzo, S., La Paglia, L., Adamo, V., Siragusa, S., Ficorella, C., et al. (2010). Driver mutations and differential sensitivity to targeted therapies: A new approach to the treatment of lung adenocarcinoma. *Cancer Treat. Rev.* 36, S21–S29. doi:10.1016/S0305-7372(10)70016-5

Charloux, A., Quoix, E., Wolkove, N., Small, D., Pauli, G., and Kreisman, H. (1997). The increasing incidence of lung adenocarcinoma: Reality or artefact? A review of the epidemiology of lung adenocarcinoma. *Int. J. Epidemiol.* 26 (1), 14–23. doi:10.1093/ije/26.1.14

Chohan, T. A., Qian, H., Pan, Y., and Chen, J-Z. (2015). Cyclin-dependent kinase-2 as a target for cancer therapy: Progress in the development of CDK2 inhibitors as anti-cancer agents. *Curr. Med. Chem.* 22 (2), 237–263. doi:10.2174/0929867321666141106113633

Denisenko, T. V., Budkevich, I. N., and Zhivotovsky, B. (2018). Cell death-based treatment of lung adenocarcinoma. *Cell. Death Dis.* 9 (2), 117–214. doi:10.1038/s41419-017-0063-y

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

Gao, H., Liang, J., Duan, J., Chen, L., Li, H., Zhen, T., et al. (2021). A prognosis marker SLC2A3 correlates with EMT and immune signature in colorectal cancer. *Front. Oncol.* 11, 638099. doi:10.3389/fonc.2021.638099

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

Goel, M. K., Khanna, P., and Kishore, J. (2010). Understanding survival analysis: Kaplan-Meier estimate. *Int. J. Ayurveda Res.* 1 (4), 274–278. doi:10.4103/0974-7788.76794

Götte, M., and Kovalszky, I. (2018). Extracellular matrix functions in lung cancer. *Matrix Biol.* 73, 105–121. doi:10.1016/j.matbio.2018.02.018

Hastie, T., Tibshirani, R., Narasimhan, B., Chu, G., and Narasimhan, M. B. (2011). *Package 'impute'*. biocViews Bioinformatics.

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 (10066), 299–311. doi:10.1016/S0140-6736(16)30958-8

Jiang, H., Xu, S., and Chen, C. (2020). A ten-gene signature-based risk assessment model predicts the prognosis of lung adenocarcinoma. *BMC cancer* 20 (1), 782–811. doi:10.1186/s12885-020-07235-z

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, Z., Luo, Y., Zhang, L., Li, H., Pan, C., Yang, H., et al. (2022). A novel risk score model of lactate metabolism for predicting over survival and immune signature in lung adenocarcinoma. *Cancers* 14 (15), 3727. doi:10.3390/cancers14153727

Knijnenburg, T. A., Wang, L., Zimmermann, M. T., Chambwe, N., Gao, G. F., Cherniack, A. D., et al. (2018). Genomic and molecular landscape of DNA damage repair deficiency across the Cancer Genome Atlas. *Cell. Rep.* 23 (1), 239–254. e236. doi:10.1016/j.celrep.2018.03.076

Ladanyi, M., and Pao, W. (2008). Lung adenocarcinoma: Guiding EGFR-targeted therapy and beyond. *Mod. Pathol.* 21 (2), S16–S22. doi:10.1038/modpathol.3801018

Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9 (1), 559–613. doi:10.1186/1471-2105-9-559

Li, L., and Wang, X. (2021). Identification of gastric cancer subtypes based on pathway clustering. *NPJ Precis. Oncol.* 5 (1), 46. doi:10.1038/s41698-021-00186-z

Li, Z., Zhang, Y., Zhou, Y., Wang, F., Yin, C., Ding, L., et al. (2021). Tanshinone IIA suppresses the progression of lung adenocarcinoma through regulating CCNA2-CDK2 complex and AURKA/PLK1 pathway. *Sci. Rep.* 11 (1), 23681–23712. doi:10.1038/s41598-021-03166-2

Ma, Q., Geng, K., Xiao, P., and Zeng, L. (2021). Identification and prognostic value exploration of radiotherapy sensitivity-associated genes in non-small-cell lung Cancer. *Biomed. Res. Int.* 2021, 5963868. doi:10.1155/2021/5963868

Mahale, J., Smagurauskaite, G., Brown, K., Thomas, A., and Howells, L. M. (2016). The role of stromal fibroblasts in lung carcinogenesis: A target for chemoprevention? *Int. J. Cancer* 138 (1), 30–44. doi:10.1002/ijc.29447

Mayakonda, A., Lin, D-C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118

Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhim, R., and Getz, G. (2011). GISTIC2. 0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* 12 (4), R41–R14. doi:10.1186/gb-2011-12-4-r41

Niemira, M., Collin, F., Szalkowska, A., Bielska, A., Chwialkowska, K., Reszec, J., et al. (2019). Molecular signature of subtypes of non-small-cell lung cancer by large-scale transcriptional profiling: Identification of key modules and genes by weighted gene co-expression network analysis (WGCNA). *Cancers* 12 (1), 37. doi:10.3390/cancers12010037

Pintarelli, G., Noci, S., Maspero, D., Pettinicchio, A., Dugo, M., De Cecco, L., et al. (2019). Cigarette smoke alters the transcriptome of non-involved lung tissue in lung adenocarcinoma patients. *Sci. Rep.* 9 (1), 13039–13110. doi:10.1038/s41598-019-49648-2

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell.* 160 (1-2), 48–61. doi:10.1016/j.cell.2014.12.033

Rousseaux, S., Debernardi, A., Jacquiau, B., Vitte, A-L., Vesin, A., Nagy-Mignotte, H., et al. (2013). Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci. Transl. Med.* 5 (186), 186ra66. 186ra166-186ra166. doi:10.1126/scitranslmed.3005723

Saito, M., Suzuki, H., Kono, K., Takenoshita, S., and Kohno, T. (2018). Treatment of lung adenocarcinoma by molecular-targeted therapy and immunotherapy. *Surg. Today* 48 (1), 1–8. doi:10.1007/s00595-017-1497-7

Schwartz, M., Zhang, Y., and Rosenblatt, J. D. (2016). B cell regulation of the anti-tumor response and role in carcinogenesis. *J. Immunother. Cancer* 4 (1), 40–15. doi:10.1186/s40425-016-0145-x

Shiba-Ishii, A. (2021). Significance of stratifin in early progression of lung adenocarcinoma and its potential therapeutic relevance. *Pathol. Int.* 71 (10), 655–665. doi:10.1111/pin.13147

Shurin, M. R. (2018). Immunological targets for cancer therapy: New recognition. *Immunotargets Ther.* 7, 83–85. doi:10.2147/ITT.S191821

Sorrentino, R., Morello, S., Forte, G., Montinaro, A., Vita, G. D., Luciano, A., et al. (2011). B cells contribute to the antitumor activity of CpG-oligodeoxynucleotide in a mouse model of metastatic lung carcinoma. *Am. J. Respir. Crit. Care Med.* 183 (10), 1369–1379. doi:10.1164/rccm.201010-1738OC

Spella, M., and Stathopoulos, G. T. (2021). Immune resistance in lung adenocarcinoma. *Cancers* 13 (3), 384. doi:10.3390/cancers13030384

Spigel, D. R., Schrock, A. B., Fabrizio, D., Frampton, G. M., Sun, J., He, J., et al. (2016). *Total mutation burden (TMB) in lung cancer (LC) and relationship with response to PD-1/PD-L1 targeted therapies*. Virginia, United States: American Society of Clinical Oncology.

Sun, Y., Xu, T., Cao, Y., and Ding, X. (2017). Antitumor effect of miR-27b-3p on lung cancer cells via targeting Fzd7. *Eur. Rev. Med. Pharmacol. Sci.* 21 (18), 4113–4123.

Surget, S., Khoury, M. P., and Bourdon, J-C. (2014). Uncovering the role of p53 splice variants in human malignancy: A clinical perspective. *Onco. Targets. Ther.* 7, 57–68. doi:10.2147/OTT.S53876

Tadesse, S., Anshabo, A. T., Portman, N., Lim, E., Tilley, W., Caldon, C. E., et al. (2020). Targeting CDK2 in cancer: Challenges and opportunities for therapy. *Drug Discov. Today* 25 (2), 406–413. doi:10.1016/j.drudis.2019.12.001

Tan, Q., Cui, J., Huang, J., Ding, Z., Lin, H., Niu, X., et al. (2016). Genomic alteration during metastasis of lung adenocarcinoma. *Cell. Physiol. biochem.* 38 (2), 469–486. doi:10.1159/000438644

Taube, J. M., Galon, J., Sholl, L. M., Rodig, S. J., Cottrell, T. R., Giraldo, N. A., et al. (2018). Implications of the tumor immune microenvironment for staging and therapeutics. *Mod. Pathol.* 31 (2), 214–234. doi:10.1038/modpathol.2017.156

Thandra, K. C., Barsouk, A., Saginala, K., Aluru, J. S., and Barsouk, A. (2021). Epidemiology of lung cancer. *Contemp. Oncol.* 25 (1), 45–52. doi:10.5114/wo.2021.103829

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Yang, T-H. O., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830. e814. doi:10.1016/j.immuni.2018.03.023

Wadler, S. (2001). Perspectives for cancer therapies with cdk2 inhibitors. *Drug resist. updat.* 4 (6), 347–367. doi:10.1054/drup.2001.0224

Wang, Y., Hu, L., Zhang, X., Zhao, H., Xu, H., Wei, Y., et al. (2017). Downregulation of mitochondrial single stranded DNA binding protein (SSBP1) induces mitochondrial dysfunction and increases the radiosensitivity in non-small cell lung cancer cells. *J. Cancer* 8 (8), 1400–1409. doi:10.7150/jca.18170

Wang, Y., and Zhou, B. P. (2013). Epithelial-mesenchymal transition—A hallmark of breast cancer metastasis. *Cancer Hallm.* 1 (1), 38–49. doi:10.1166/ch.2013.1004

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

Yamauchi, M., Yamaguchi, R., Nakata, A., Kohno, T., Nagasaki, M., Shimamura, T., et al. (2012). Epidermal growth factor receptor tyrosine kinase defines critical prognostic genes of stage I lung adenocarcinoma. *PLoS One* 7, e43923. doi:10.1371/journal.pone.0043923

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 (1), 2612–2711. doi:10.1038/ncomms3612

Zamarron, B. F., and Chen, W. (2011). Dual roles of immune cells and their factors in cancer development and progression. *Int. J. Biol. Sci.* 7 (5), 651–658. doi:10.7150/ijbs.7.651

Zeng, Y., Jie, X., Wu, B., Wu, G., Liu, L., and Xu, S. (2020). IQGAP3 interacts with Rad17 to recruit the Mre11-Rad50-Nbs1 complex and contributes to radioresistance in lung cancer. *Cancer Lett.* 493, 254–265. doi:10.1016/j.canlet.2020.08.042

Zhang, F-L., Zhu, W-M., He, T-R., Zhao, Y-T., Ge, W., Tan, J-H., et al. (2022). Comparative transcriptomic analysis reveals that TPX2 and AURXA are involved in porcine PCV2 infection. *Gene* 834, 146649. doi:10.1016/j.gene.2022.146649

Zhang, L., Chen, J., Cheng, T., Yang, H., Li, H., and Pan, C. (2020). Identification of the key genes and characterizations of tumor immune microenvironment in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC). *J. Cancer* 11 (17), 4965–4979. doi:10.7150/jca.42531

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 06 January 2023
DOI 10.3389/fphar.2022.1095965

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Yongjun Guan,
Renmin Hospital of Wuhan University, China
Weicheng Lu,
Sun Yat-sen University Cancer Center (SYSUCC), China
Bole Tian,
Sichuan University, China

*CORRESPONDENCE
Jinbao Zhang,
jinbaozhang1988@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 11 November 2022
ACCEPTED 14 December 2022
PUBLISHED 06 January 2023

CITATION
Wang C, Yang L, Xu S, Guo H, Guan H, Wang Q, Jiang X, Fei M and Zhang J (2023), Systematic analysis of the role and significance of target genes of active ingredients of traditional Chinese medicine injections in the progression and immune microenvironment of hepatocellular carcinoma.
Front. Pharmacol. 13:1095965.
doi: 10.3389/fphar.2022.1095965

Systematic analysis of the role and significance of target genes of active ingredients of traditional Chinese medicine injections in the progression and immune microenvironment of hepatocellular carcinoma

Chao Wang[1†], Lili Yang[1†], Shaoheng Xu[1†], Hui Guo[1], Hewen Guan[2], Qiannan Wang[3], Xueyan Jiang[1], Mingyang Fei[4] and Jinbao Zhang[1*]

[1]Department of General Surgery, First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China, [2]Department of Dermatology, First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China, [3]Department of Plastic Surgery, First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China, [4]Department of Neurosurgery, First Affiliated Hospital of Dalian Medical University, Dalian, Liaoning, China

Background: Traditional Chinese medicine in China is an important adjuvant therapy for the treatment of hepatocellular carcinoma (HCC) and traditional Chinese medicines injections have a wide range of clinical applications. The purpose of this study was to identify the active ingredients and related genes of traditional Chinese medicine injections that can treat hepatocellular carcinoma.

Methods: Effective small molecule components were extracted from 14 types of traditional Chinese medicines from 8 injections and the main gene targets were identified. The 968 patients with HCC were classified based on the target gene set, and the characteristics of patients with different subtypes were analyzed. Patients with two subtypes of HCC were compared with normal tissues and cirrhosis to identify important gene targets related to traditional Chinese medicines in HCC progression.

Results: In this study, 138 important genes associated with traditional Chinese medicines were identified and two HCC subtypes were identified. By analyzing the differences between the two subtypes, 25 related genes were associated with HCC subtypes. Through clinical and pharmacological analysis, this study identified quercetin as an important traditional Chinese medicines small molecule and secreted phosphoprotein 1 (SPP1) as an important oncogene in HCC.

Conclusion: Traditional Chinese medicines injection is an important adjuvant treatment modality for HCC. SPP1 is an important oncogene in HCC.

KEYWORDS

hepatocellular carcinoma, traditional Chinese medicine, immune microenvironment, injection, adjuvant therapy

Introduction

Liver cancer is the sixth leading cause of death in the world, accounting for about 8.3% of cancer mortality (Sung et al., 2021). Liver cancer is a common chronic liver disease. Hepatitis viruses HBV and HCV are common causes of liver cancer. Liver cirrhosis caused by virus can develop into liver cancer under the action of various incentives. Liver cancer caused by hepatitis virus is also a common type of liver cancer in the Chinese population (Shi et al., 2005; Schietroma et al., 2018). At the same time, excessive intake of aflatoxin, alcoholic liver disease and non-alcoholic fatty liver disease are also common causes of liver cancer (Sagnelli et al., 2020; Llovet et al., 2021).Most of these diseases can cause liver fibrosis and chronic inflammation, thereby promoting the occurrence and development of liver cancer (Dhar et al., 2020). Hepatocellular carcinoma (HCC) is a liver tumor originating from hepatocytes and is the most common type of liver cancer. 90% of liver cancers are hepatocellular carcinoma (Gao et al., 2020). Traditional treatments for HCC include surgery, chemotherapy, radiation therapy, targeted therapy, etc., (Wang et al., 2019). In recent years, with the emergence of immunotherapy, many immunotherapy drugs, such as PD-L1/PD-1 monoclonal antibody, have been widely used in the treatment of liver cancer. At present, immunotherapy combined with traditional targeted therapy has become an important treatment method for HCC (Makarova-Rusher et al., 2015; De Martin et al., 2018; Xu et al., 2018; Faivre et al., 2020).

Traditional Chinese medicine is a traditional Chinese medicine treatment method. Many studies have shown that traditional Chinese medicine has a greater role in a variety of tumor microenvironments. Many active ingredients of traditional Chinese medicine can promote the generation of tumor immune cells and the occurrence of anti-tumor responses (Zhai et al., 2019; Wang et al., 2020a). Traditional Chinese medicine has a variety of anti-tumor effects in the treatment of HCC and can play a greater role in other therapeutic methods (Wei et al., 2022). However, traditional Chinese medicines are often administered orally or externally due to the complex composition of traditional Chinese medicines. In this way, the effect of some traditional Chinese medicines is not obvious and it is difficult to study the main active components of traditional Chinese medicines. In recent years, with the discovery of traditional Chinese medicine ingredients and the improvement of refining technology, more traditional Chinese medicine injections have appeared in clinical treatment. In this study, 8 kinds of commonly used and commercialized traditional Chinese medicine injections were collected, and these injections were all related to hepatocellular carcinoma. Ideas and results.

Method and material

Classification of traditional Chinese medicine and acquisition of active ingredients

In this study, the commonly used traditional Chinese medicine injections were screened, and 8 traditional Chinese medicine injections related to HCC were obtained. Eight kinds of traditional Chinese medicine injections contain a total of 14 kinds of traditional Chinese medicines. According to the medicinal properties and pharmacological effects of traditional Chinese medicines, we divided the 14 kinds of traditional Chinese medicines into three categories. The first category is heat-clearing traditional Chinese medicine (HCM), including marsdenia tenacissima, sarcandra glabra, semen coicis, venenum bufonis and mylabris. The second category is tonic traditional Chinese medicine (TCM), including ginseng, astragalus mongholicus, red ginseng and acanthopanax root. The third category of pain-relieving traditional Chinese medicine (PRCM), including caulis sinomenii, celandine, schefflera kwangsiensis, aconiti radix and radix aconiti agrestis.

We obtained the main components of relevant Chinese medicines from 4 Chinese medicine databases including SymMap, TCMID, TCMSP, and TCM-ID. The pharmacokinetics and pharmacokinetics of these components were obtained from the TCMSP database. We screened the main components with DL$\geqq$0.1 and TPSA$\leq$140 as the active ingredients of various traditional Chinese medicines. Among the three types of traditional Chinese medicines, 50 active ingredients were screened by HCM; 153 active ingredients were screened by TCM; 110 active ingredients were screened by PRCM.

Acquisition of related genes in active ingredients of traditional Chinese medicine

In this study, the target of 303 main components of three types of traditional Chinese medicines was determined through the four traditional Chinese medicine databases of SymMap, TCMID, TCMSP and TCM-ID. Then, we obtained the HCC-related target sets through the DisGNET database and the GeneCard database, and obtained the intersections with the three types of traditional Chinese medicines targets respectively. The three target gene sets were intersected and 138 target genes were found.

RNA-seq data of HCC patients and correlation analysis

In this study, RNA-seq data were collected from three online databases, TCGA(TCGA-LIHC), ICGC(ICGC-LIRI) and GEO (GSE14520,GSE116174,GSE54236), for 968 cases of patients with HCC. RNA-seq data from another 340 patients with cirrhosis were also collected for comparison in GEO (GSE15654,GSE84044). A total of 479 cases of RNA-seq data from normal liver tissues were obtained from three online databases, TCGA(TCGA-LIHC), GTEx and GEO (GSE14520,GSE54236). All sequencing data were transformed into FPKM data and normalized accordingly. Spearman test was used for all gene expression correlation analysis in this study. In this study, the expression of 120 target genes in HCC was demonstrated.

Consistent clustering and variance analysis

Based on the expression of 120 target genes, 968 patients were divided into Class I and Class II by consistent clustering method. Then, the differential genes between the two groups were detected by DeSeq2 differential gene analysis. Differential gene screening conditions were $p < 0.05$ and log2FC < -1 or log2FC > 1. WGCNA analysis was used to further identify important sets of differential genes between the two HCC subtypes.

Gene and pathway enrichment analysis

GSEA enrichment analysis was performed on the differential genes of the two types of samples, and gene functions were annotated by using the gene sets of GO and KEGG. At the same time, the basic enrichment analysis was performed on the target gene sets of three types of traditional Chinese medicines by using the gene sets of GO and KEGG. GSVA pathway analysis was used to analyze the differences in pathway enrichment between the two HCC subtypes.

Immune infiltration analysis and survival analysis

The Cibersort immune infiltration analysis was performed on the gene sets of the two types of samples, and the differences in the infiltration degrees of 22 immune cells in the two types of samples were obtained. 576 patients with HCC had clinical survival data. KM analysis was used to analyze the effect of HCC subtypes on survival of patients. Cox analysis was used to identify important genes associated with traditional Chinese medicines target genes that are associated with survival of patients with HCC.

Genetic multi-omics analysis and drug analysis

The GEPIA database is used to observe gene expression in a variety of cancers. The ULCAN database is used to analyze gene methylation. cBioPortal database is used to analyze gene mutations. HPA data is used to analyze gene expression and cellular localization at the protein level. Oncopredict package and easier package are used to analyze different subtypes of HCC patients' sensitivity to common small molecule drug treatments and immunotherapy. A molecular docking approach was used to validate the interactions between important small molecule components of traditional Chinese medicines and gene target proteins.

Statistical analysis

The software used in this study includes R studio (R4.2.1), Cytoscape 3.7.2, Prism 9, Autodock4, Pymol and SPSS 26. Statistical results at $p < 0.05$ were considered statistically significant.

Result

Classification and target screening of traditional Chinese medicines related to HCC

In this study, we included 8 clinically used traditional Chinese medicine injections and collected small molecules and effective targets of 14 classes of Chinese medicines (Supplementary Table S1). A detailed analysis flow chart has been shown in Supplementary Figure S1. We divided the 14 classes of traditional Chinese medicine into three major categories based on the pharmacology and clinical efficacy of Chinese medicine. The first group of traditional Chinese medicine is HCM, which has the effect of clearing heat and detoxifying the body, and can inhibit the development of inflammation and cancer. The number of HCC-related targets of HCM is 243 (Figure 1A). The second group of traditional Chinese medicine is TCM, which has the effect of benefiting the immune function of the body and promoting the body's anti-disease function. The number of HCC-related targets of TCM is 548 (Figure 1A). The third group of traditional Chinese medicine is PRCM, which has pain-relieving effect and is often used in the treatment of cancer pain in advanced cancer. The number of HCC-related targets of PRCM is 297 (Figure 1A). Gene

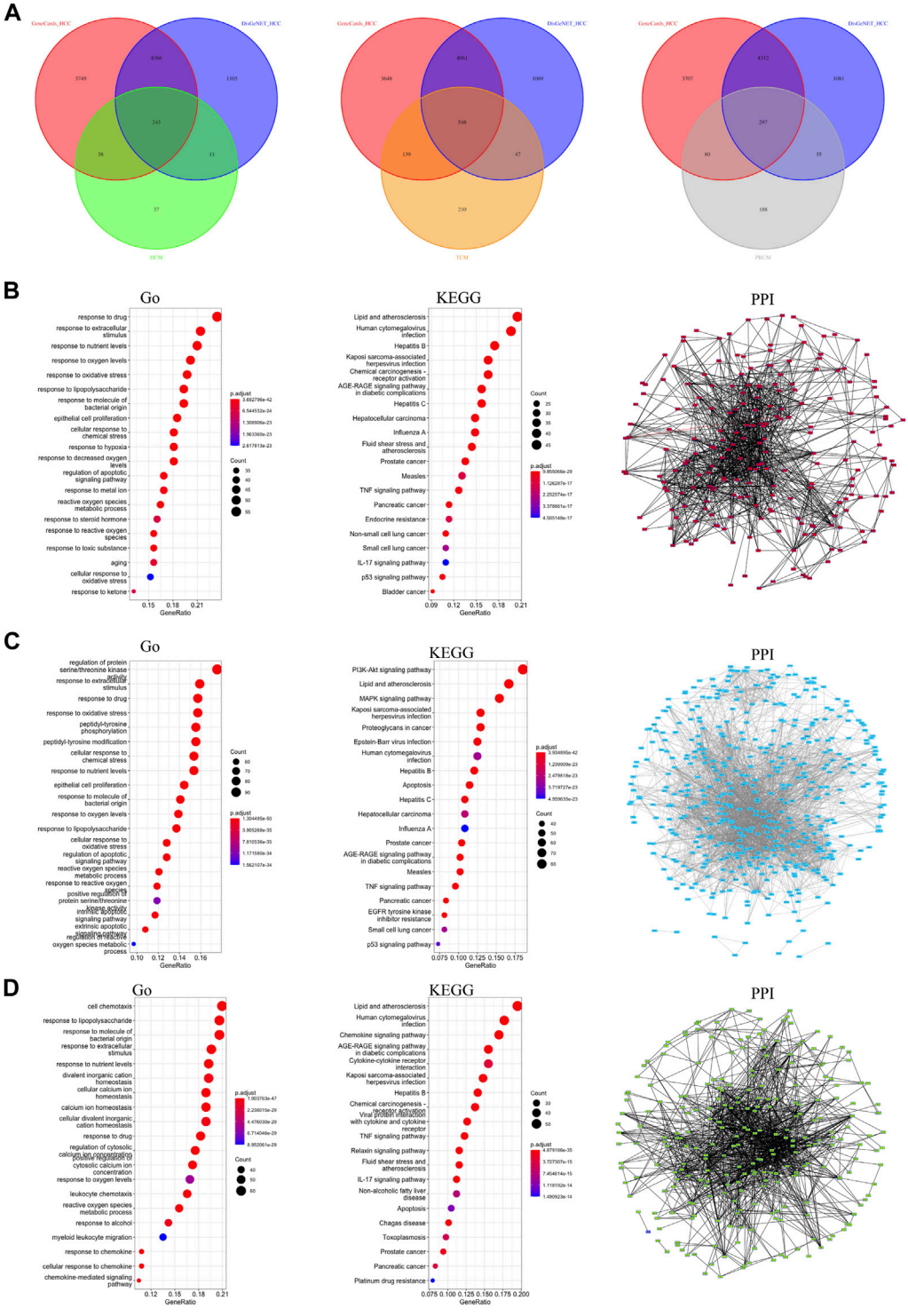


FIGURE 1
Classification and target screening of traditional Chinese medicines. **(A)** The number of HCC-related targets of three types. **(B)** Gene function analysis of HCM gene targets. **(C)** Gene function analysis of TCM gene targets. **(D)** Gene function analysis of PRCM gene targets.

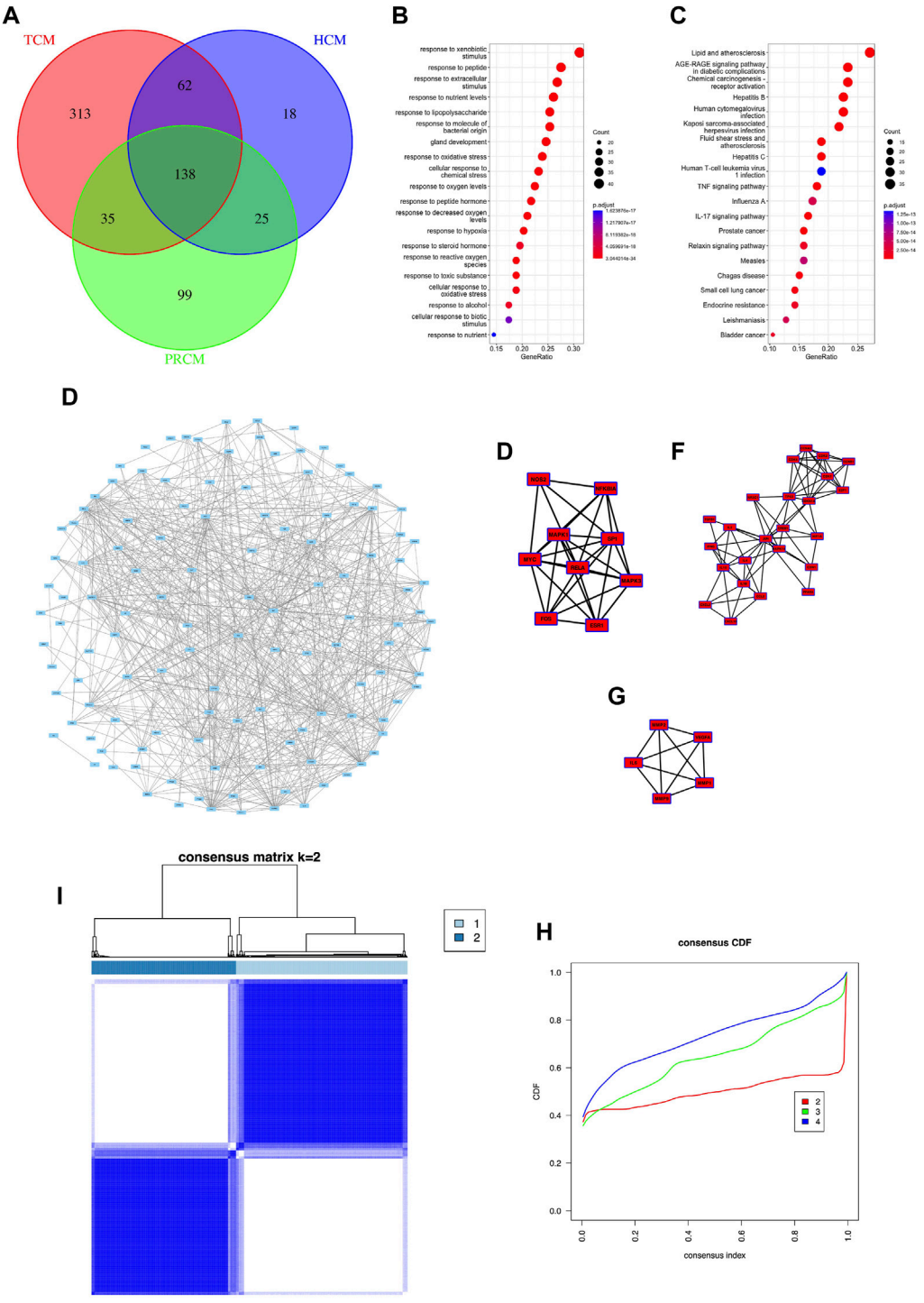


FIGURE 2
Common targets and classification subtypes. **(A)** The number of common targets. **(B)** Gene enrichment Go analysis of common targets. **(C)** Gene enrichment KEGG analysis of common targets. **(D)** PPI analysis of common targets. **(E)** Core networks 1 of PPI analysis. **(F)** Core networks 2 of PPI analysis.**(G)** Core networks 3 of PPI analysis. **(G)** Consensus matrix of consistent clustering. **(H)** Consensus CDF of consistent clustering.

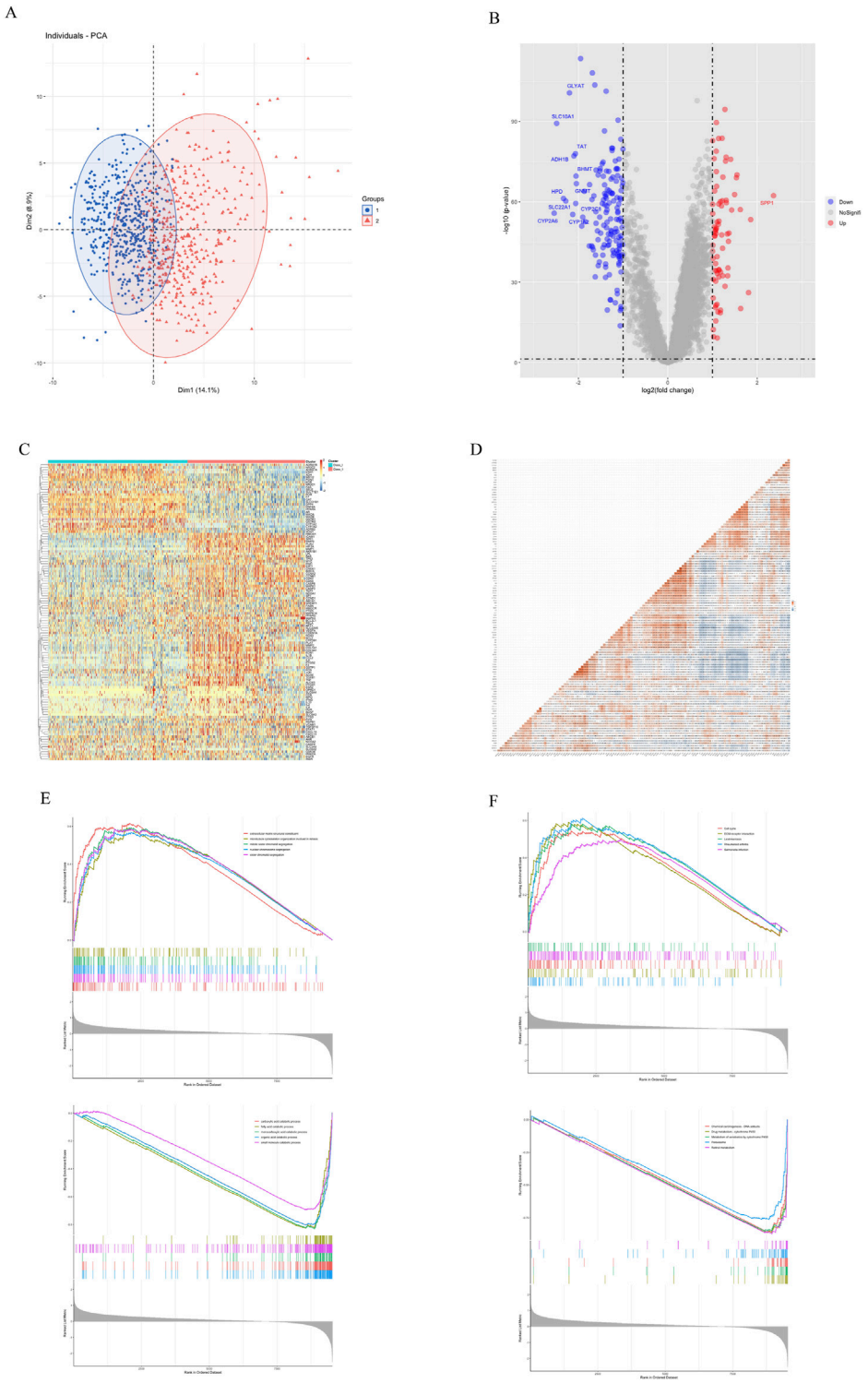


FIGURE 3
Gene differential analysis of two subtypes. **(A)** PCA analysis. **(B)** Volcano map of differential analysis. **(C)** Expression of 120 traditional Chinese medicine targets. **(D)** Correlation of 120 traditional Chinese medicine targets. **(E)** GSEA Go analysis. **(F)** GSEA KEGG analysis.

enrichment analysis showed that the targets of HCM were related to drug response (Figure 1B); the targets of TCM were related to extracellular stimulation (Figure 1C); and the targets of PRCM were related to cell chemotaxis (Figure 1D). All three types of TCM were associated with hepatitis virus (Figures 1B–D). PPI analysis showed that the target sets of all three types of traditional Chinese medicine had high correlation (Figures 1B–D).

Common targets of traditional Chinese medicine and classification subtypes of HCC

In this study, 138 gene targets were obtained from the target sets of three types of traditional Chinese medicine (Figure 2A). Based on 138 genes, we collected related traditional Chinese medicine and small molecule components (Supplementary Table S2).Gene enrichment analysis showed that most of the 138 targets were associated with cellular stress response, hepatitis virus (Figures 2B,C). PPI analysis showed a large correlation between targets (Figure 2D) and three core networks were obtained with clustering indices of 7.5 (Figure 2E), 6.87 (Figure 2F), and 5 (Figure 2G), respectively. Most of the genes in all three core networks were associated with cellular stress response and cellular inflammation. In this study, sequencing data from 968 HCC cases were used to find the expression of 120 targets (The details are shown in Supplementary Table S3). Therefore, based on the expression of 120 targets, consistent clustering was performed and patients with HCC were classified into two subtypes, Class I and Class II (Figures 2H,I).

Target expression and differential analysis of different subtypes of HCC

PCA analysis showed that based on the expression of traditional Chinese medicine targets, 968 patients could be clearly classified into two categories, with 525 cases in Class I and 443 cases in Class II (Figure 3A). Differential analysis showed that there were 76 genes with higher expression and 151 genes with lower expression in Class II compared with Class I. Among them, SPP1 was the most significantly differential gene with a LogFC of 2.36 (Figure 3B). We observed the expression and correlation of 120 traditional Chinese medicine targets in HCC (Figures 3C,D). VEGFA, an important target for HCC, was generally highly expressed in Class II (Figure 3C), so we hypothesized that Class II was more malignant than Class I. GSEA analysis showed that the gene set based on GO (Figure 3E) and KEGG (Figure 3F), Class II exhibits enhanced cell cycle and cell division processes, as well as reduced metabolic processes (especially P450 enzyme metabolism). The *p*-values of GSEA in this study were all less than 0.05, and the top5 items with higher NES were screened for presentation (Figures 3E,F; Supplementary table S4).This indicates that Class II is more malignant than Class I but less drug metabolic than Class I (Figure 3F).

Subtype characteristics and key genes of different subtypes of HCC

WGCNA analysis was performed on HCC sequencing data (Figures 4A–C). The soft threshold value obtained from the screening was 5 (Figure 4A). The analysis divided 9651 genes into 13 gene sets (Figure 4B). And the HCC subtype was found to be most correlated with the MEbrown gene set ($r = 0.7$, $p < 0.05$, Figure 4C). GSVA analysis was performed using three pathway gene sets, KEGG (Figure 4D), Reactome (Figure 4E) and Wiki (Figure 4F). The results of the study showed that there was a reduction in liver function-related metabolism in Class II compared to Class I, especially for P450-related drug metabolism. We also analyzed the metabolism of P450-related drugs in cirrhotic versus normal tissues by GSVA analysis (Figure 4G). The degree of Class II drug metabolism was found to be lower than cirrhosis versus normal tissue, but the degree of Class I drug metabolism was higher than cirrhosis and lower than normal tissue. This proximally confirms that Class II is less metabolizable and thus more malignant than Class I. The MEbrown gene set was intersected with the traditional Chinese medicine target set to obtain 25 key genes (Figure 4H). Among the 25 genes, Class II had low expression of 23 genes and high expression of 2 genes compared with Class I. SPP1 and MMP1 were more highly expressed in Class II and higher than in cirrhotic versus normal tissues (Figure 4I), and negatively correlated with the expression of other key genes (Figure 4J). Gene enrichment analysis revealed that these genes were associated with extracellular stimulation (Figure 4K) and with chemical carcinogenesis receptor activation (Figure 4L).

Immune microenvironment characteristics and survival analysis of different subtypes of HCC

Previous results demonstrated that Class II has a higher degree of malignancy and greater damage to liver function, as well as a lower metabolism. In this study, we first performed Estimate analysis of sequencing data from HCC patients (Figure 5A). We found that Class II had a higher Estimate score and had more immune cell components with stromal cell components. However, Class I had higher tumor purity than Class II. Cibersort immune infiltration analysis showed significant differences between Class I and Class II in terms of

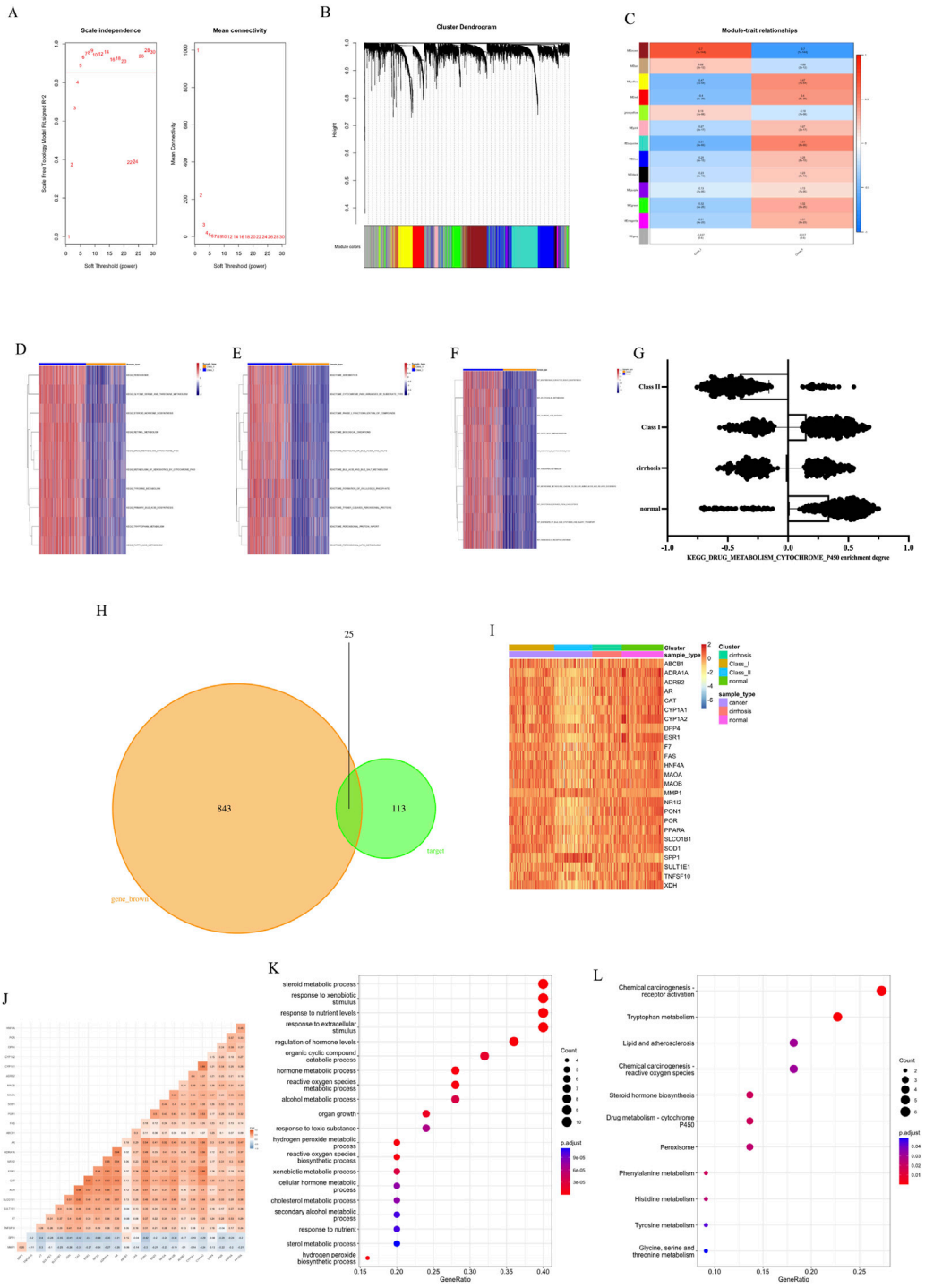


FIGURE 4
Subtype characteristics and key genes. **(A)** Soft threshold of WGCNA. **(B)** Cluster Dendrogram of WGCNA. **(C)** Module-trait relationships of WGCNA. **(D)** GSVA analysis by KEGG. **(E)** GSVA analysis by Reactome. **(F)** GSVA analysis by Wiki. **(G)** Metabolism of P450-related drugs by KEGG with GSVA analysis. **(H)** Intersection of MEbrown gene set and target genes. **(I)** Expression of 25 key genes. **(J)** Correlation of 25 key genes. **(K)** Gene enrichment Go analysis of key genes. **(L)** Gene enrichment KEGG analysis of key genes.

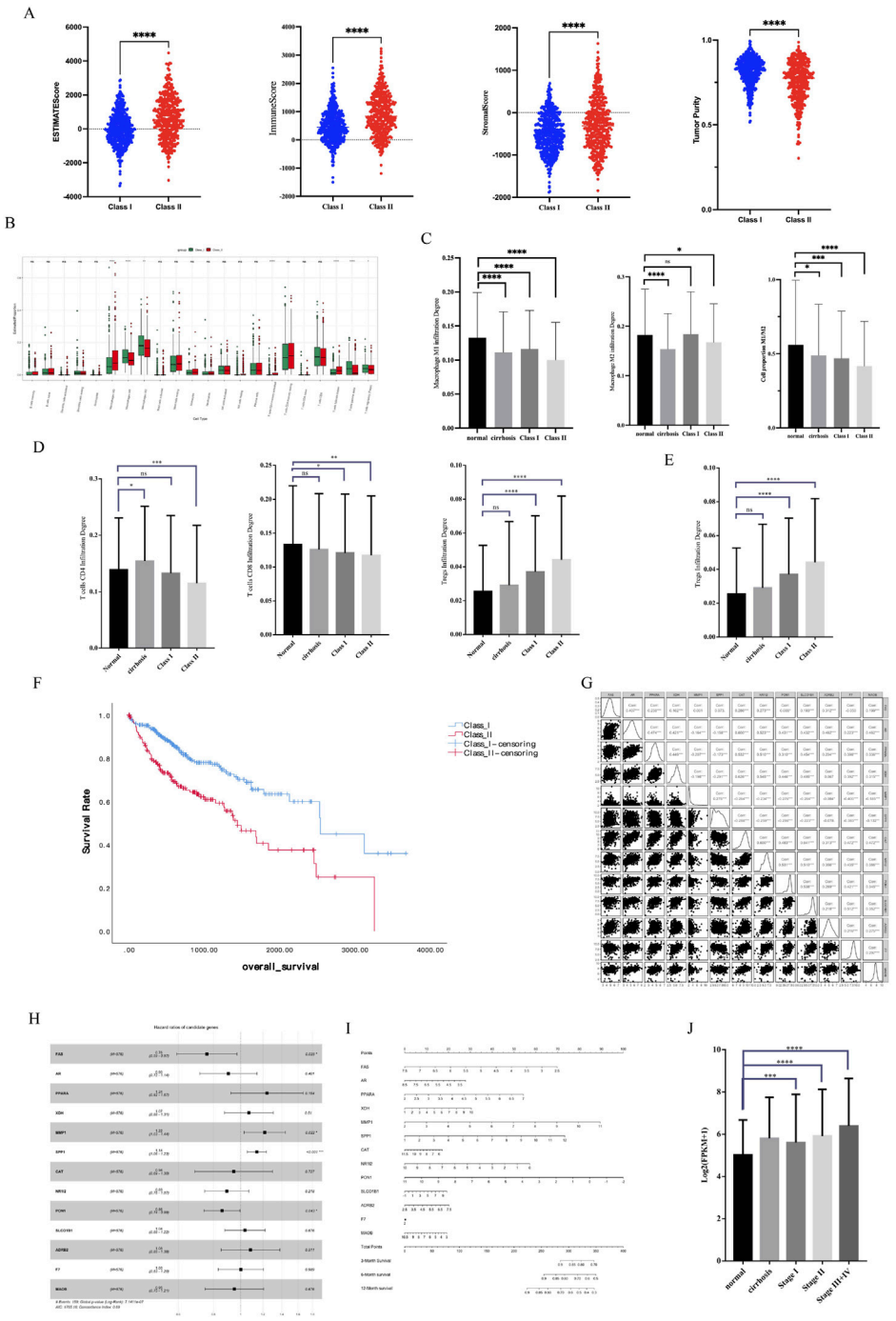


FIGURE 5
Immune microenvironment characteristics and survival analysis. **(A)** Estimate analysis of HCC. **(B)** Cibersort immune infiltration analysis. **(C)** Macrophage infiltration analysis. **(D)** CD4-positive and CD8-positive T-cell infiltration analysis. **(E)** Treg infiltration analysis. **(F)** KM analysis of different subtypes of HCC **(G)** Correlation of 13 key genes related to survival. **(H)** HR Forest Map of 13 key genes. **(I)** Nomogram of 13 key genes. **(J)** SPP1 expression in different stages of HCC.

macrophage and T-cell infiltration (Figure 5B). In terms of macrophage infiltration (Figure 5C), there was a decrease in the M1 type ratio and an increase in the M2 type ratio in Class II, resulting in a lower M1/M2 in Class II. In terms of T-cell, there was a decrease in CD4-positive T-cell, a decrease in CD8-positive T-cell, and an increase in the CD4/CD8 ratio in

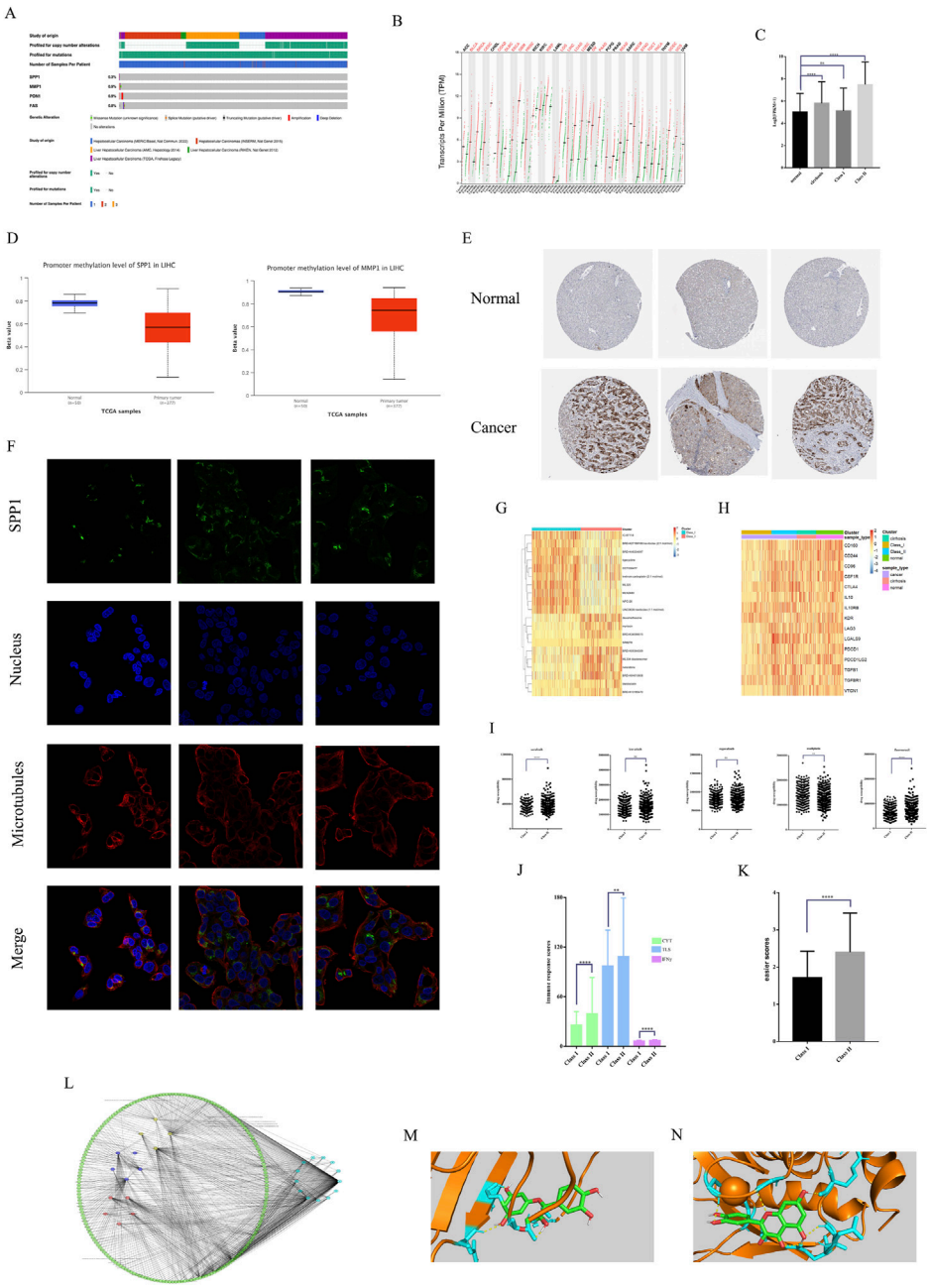


FIGURE 6
Multi-omics analysis and sensitive drug analysis. **(A)**Mutation analysis of key genes. **(B)** Pan-cancer expression of SPP1. **(C)** SPP1expression of different subtypes of HCC. **(D)** Gene methylation of SPP1 and MMP1 in HCC. **(E)** IHC analysis of SPP1 expression. **(F)** IF analysis of SPP1 subcellular localization. **(G)** The top 10 small molecule drugs sensitive in different subtypes. **(H)** ICIs expression in different subtypes. **(I)** Sensitivity analysis of commonly used HCC therapeutic drugs. **(J)** Three different immune response scores for different subtypes. **(K)** Easier scores for different subtypes. **(L)** Drug-component-target network for 13 key genes. **(M)** Molecular docking of SPP1 with quercetin. **(N)** Molecular docking of MMP1 with quercetin.

Class II (Figure 5D). Of note is the presence of a higher infiltration of Treg cells in Class II (Figure 5E). In summary Class II has more extracellular matrix and suppresses anti-tumor immune responses in the microenvironment.

We performed a survival analysis of 576 HCC patients with clinical data. KM analysis showed that Class II had a worse prognosis and shorter survival time than Class I (Figure 5F). We constructed Cox survival regression models using 25 key gene

expressions and screened 13 genes that were associated with survival time (Figure 5G). Among the 13 genes, four genes, SPP1, MMP1, PON1, and FAS, were significantly associated with survival (Figure 5H) and could predict the probability of patient survival (Figure 5I). In particular, it is important to note that SPP1, a key gene with significant differences, had high expression in advanced tumors (Figure 5J), suggesting that SPP1 may be an important oncogene in HCC.

Multi-omics analysis of key genes and sensitive drug analysis

We perform further analysis of important key genes. Mutation analysis showed that SPP1, MMP1, PON1, and FAS all had low mutation rates in HCC, which demonstrated that all four genes were stably expressed genes (Figure 6A). Pan-cancer analysis showed that SPP1 was in high expression in most of the 32 tumors (Figure 6B). And the expression of SPP1 in Class II was significantly higher than that in Class I with normal tissues. And there was high expression of SPP1 in cirrhosis, which proximately suggested that SPP1 was associated with malignant phenotype (Figure 6C). Methylation analysis showed low methylation of both SPP1 and MMP1 in HCC (Figure 6D). IHC analysis showed high expression of SPP1 in hepatic tissues of HCC, but almost no expression in the mesenchyme (Figure 6E). IF analysis showed that HepG2 SPP1 protein was widely present in the cytoplasm and had a large overlap with microtubule regions overlap with the microtubule region, suggesting that SPP1 may have a greater relationship with HCC metastasis (Figure 6F).

We also analyzed the sensitivity of small molecule drugs with immunotherapy in HCC patients of two subtypes. The results showed that Class I had greater sensitivity to carboplatin (Figure 6G) versus oxaliplatin (Figure 6I), while Class II had greater sensitivity to sorafenib, fluorouracil (Figure 6I). We show the top 10 small molecule drugs sensitive to both subtypes (Figure 6G). Class II showed higher sensitivity to lenvatinib and regorafenib than class I, but there was no significant difference (Figure 6I). Class II had higher expression of immunosuppressive checkpoints than class I vs. normal tissue (Figure 6H). And easier package immunotherapy sensitivity analysis showed that Class II had high immune response scores (Figure 6J) and easier scores (Figure 6K). Therefore Class II and higher immunotherapy sensitivity than Class I. This suggests that Class II may be prioritized for targeted therapy and immunotherapy, and that herbal therapy may enhance the efficacy of both treatments. We screened 13 key genes related to survival for effective small molecules of traditional Chinese medicine (Figure 6L). Molecular docking validation showed that both SPP1 (Figure 6M) and MMP1 (Figure 6N) could interact with quercetin through hydrogen bonds with binding energies of −7.58 kj/mol and −7.22 kj/mol, respectively. This suggests that quercetin may be an important adjunctive therapeutic agent for HCC.

Discussion

As the main type of liver cancer, HCC has a variety of treatment methods. At present, surgery is still the best treatment for HCC. But not all patients have the possibility of surgery, which requires other treatments to make patients have the opportunity for surgery. Chemotherapy and radiotherapy are traditional surgical adjuvant treatment methods, but the prognosis is poor and there are many adverse reactions. In recent years, with the development of genomics and tumor immunology, targeted therapy and immunotherapy have gradually replaced traditional therapy for adjuvant therapy and the treatment of advanced HCC patients. However, due to the single target of targeted therapy and immunotherapy, a large part of patients have no target, resulting in poor drug response and poor prognosis.

Traditional Chinese medicine is a national treasure of China and the main treatment method of traditional Chinese medicine. There are a variety of traditional Chinese medicines and prescriptions for the treatment of HCC. Relevant studies have shown that traditional Chinese medicine can affect the development, occurrence and spread of tumors and lead to imbalance within the tumor (Xi and Minuk, 2018). Traditional Chinese medicine can inhibit the invasion and metastasis of HCC, and is related to the EMT process of HCC (Jia et al., 2021; Li et al., 2021). At the same time, traditional Chinese medicine can promote tumor angiogenesis and the growth of tumor stem cells in HCC (Li, 2016; Wang et al., 2020b; Tang et al., 2020). The pharmaceutical preparations of traditional Chinese medicines are mostly extracts of traditional Chinese medicines, which have a variety of active ingredients and belong to different types of compounds. Among them, related studies have shown that alkaloids related to traditional Chinese medicine have anti-HCC effects and have certain development prospects (Liu et al., 2019a; Liu et al., 2019b). Most of the HCC-related components have immunomodulatory effects, which can promote the body's immune response and facilitate the immunotherapy of HCC (Jia and Wang, 2020). This is consistent with the conclusion of this study. Traditional Chinese medicine injections commonly used in clinic are also used in combination with targeted therapy and immunotherapy.

Although there are many targets for the action of traditional Chinese medicine, most of them are still used as adjuvant therapy in clinic due to the complex composition and the slow mode of action of the drug. One study showed that taking it for more than 3 months in a row increased the 3-year survival rate of patients with HCC (Hou et al., 2022). Another study also came to similar conclusions that taking traditional Chinese medicine for more than 6 months increased the survival time of patients, and this

was associated with multiple pathway enrichment (Zhang et al., 2022). However, traditional Chinese medicine still needs to be used in combination with a variety of therapies. Several studies have shown that the combined treatment of traditional Chinese medicine can improve the survival rate and quality of life of HCC patients compared with western medicine alone (Liao et al., 2015; Yang et al., 2017; Liu et al., 2019c; Liao et al., 2020).

However, traditional Chinese medicine cannot be used as the main treatment for HCC. This may be due to the fact that most of the traditional Chinese medicine preparations are oral and external medicines, which have a long action time and need to be taken for a long time. Compared with chemotherapy drugs, targeted drugs and immune preparations, traditional Chinese medicine lacks good preparations for intravenous infusion. Although the commonly used traditional Chinese medicine injection has a certain anti-cancer effect, it is prone to many adverse reactions due to its complex composition. This study explores the active ingredients of commonly used liver cancer injections in clinic, aiming to discover the main targets of traditional Chinese medicine in the treatment of HCC, which is conducive to the development of more comprehensive HCC-related traditional Chinese medicine injections, and provides ideas and preliminary ideas for the further development of traditional Chinese medicine anticancer drugs Research results.

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

All authors are solely responsible for the content and writing of the manuscript. The study's design, data collection and analysis, article preparation, and manuscript revision all benefited greatly from the efforts of all authors.

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/fphar.2022.1095965/full#supplementary-material

References

De Martin, E., Michot, J. M., Papouin, B., Champiat, S., Mateus, C., Lambotte, O., et al. (2018). Characterization of liver injury induced by cancer immunotherapy using immune checkpoint inhibitors. *J. Hepatol.* 68 (6), 1181–1190. doi:10.1016/j.jhep.2018.01.033

Dhar, D., Baglieri, J., Kisseleva, T., and Brenner, D. A. (2020). Mechanisms of liver fibrosis and its role in liver cancer. *Exp. Biol. Med. (Maywood)* 245 (2), 96–108. doi:10.1177/1535370219898141

Faivre, S., Rimassa, L., and Finn, R. S. (2020). Molecular therapies for HCC: Looking outside the box. *J. Hepatol.* 72 (2), 342–352. doi:10.1016/j.jhep.2019.09.010

Gao, Y. X., Yang, T. W., Yin, J. M., Yang, P. X., Kou, B. X., Chai, M. Y., et al. (2020). Progress and prospects of biomarkers in primary liver cancer (Review). *Int. J. Oncol.* 57 (1), 54–66. doi:10.3892/ijo.2020.5035

Hou, J., Shi, K., Liu, Y., Chen, J., Ran, C., and Wang, X. (2022). Effects of traditional Chinese medicine adjuvant therapy on the survival of patients with primary liver cancer. *Evid. Based Complement. Altern. Med.* 2022, 9810036. doi:10.1155/2022/9810036

Jia, W., Liang, S., Cheng, B., and Ling, C. (2021). The role of cancer-associated fibroblasts in hepatocellular carcinoma and the value of traditional Chinese medicine treatment. *Front. Oncol.* 11, 763519. doi:10.3389/fonc.2021.763519

Jia, W., and Wang, L. (2020). Using traditional Chinese medicine to treat hepatocellular carcinoma by targeting tumor immunity. *Evid. Based Complement. Altern. Med.* 2020, 9843486. doi:10.1155/2020/9843486

Li, H. M. (2016). Microcirculation of liver cancer, microenvironment of liver regeneration, and the strategy of Chinese medicine. *Chin. J. Integr. Med.* 22 (3), 163–167. doi:10.1007/s11655-016-2460-y

Li, J. J., Liang, Q., and Sun, G. C. (2021). Traditional Chinese medicine for prevention and treatment of hepatocellular carcinoma: A focus on epithelial-mesenchymal transition. *J. Integr. Med.* 19 (6), 469–477. doi:10.1016/j.joim.2021.08.004

Liao, Y. H., Lin, C. C., Lai, H. C., Chiang, J. H., Lin, J. G., and Li, T. C. (2015). Adjunctive traditional Chinese medicine therapy improves survival of liver cancer patients. *Liver Int.* 35 (12), 2595–2602. doi:10.1111/liv.12847

Liao, Y. P., Kung, P. T., Wang, Y. H., Chu, Y. R., Kao, S. T., and Tsai, W. C. (2020). Effects and relative factors of adjunctive Chinese medicine therapy on survival of hepatocellular carcinoma patients: A retrospective cohort study in taiwan. *Integr. Cancer Ther.* 19, 1534735420915275. doi:10.1177/1534735420915275

Liu, C., Yang, S., Wang, K., Bao, X., Liu, Y., Zhou, S., et al. (2019). Alkaloids from traditional Chinese medicine against hepatocellular carcinoma. *Biomed. Pharmacother.* 120, 109543. doi:10.1016/j.biopha.2019.109543

Liu, F., Lou, G., Zhang, T., Chen, S., Xu, J., Xu, L., et al. (2019). Anti-metastasis traditional Chinese medicine monomer screening system based on perinucleolar compartment analysis in hepatocellular carcinoma cells. *Am. J. Transl. Res.* 11 (6), 3555–3566.

Liu, X., Li, M., Wang, X., Dang, Z., Yu, L., Wang, X., et al. (2019). Effects of adjuvant traditional Chinese medicine therapy on long-term survival in patients with hepatocellular carcinoma. *Phytomedicine* 62, 152930. doi:10.1016/j.phymed.2019.152930

Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., et al. (2021). Hepatocellular carcinoma. *Nat. Rev. Dis. Prim.* 7 (1), 6. doi:10.1038/s41572-020-00240-3

Makarova-Rusher, O. V., Medina-Echeverz, J., Duffy, A. G., and Greten, T. F. (2015). The yin and yang of evasion and immune activation in HCC. *J. Hepatol.* 62 (6), 1420–1429. doi:10.1016/j.jhep.2015.02.038

Sagnelli, E., Macera, M., Russo, A., Coppola, N., and Sagnelli, C. (2020). Epidemiological and etiological variations in hepatocellular carcinoma. *Infection* 48 (1), 7–17. doi:10.1007/s15010-019-01345-y

Schietroma, I., Scheri, G. C., Pinacchio, C., Statzu, M., Petruzziello, A., and Vullo, V. (2018). Hepatitis C virus and hepatocellular carcinoma: Pathogenetic mechanisms and impact of direct-acting antivirals. *Open Virol. J.* 12, 16–25. doi:10.2174/1874357901812010016

Shi, J., Zhu, L., Liu, S., and Xie, W. F. (2005). A meta-analysis of case-control studies on the combined effect of Hepatitis B and C virus infections in causing hepatocellular carcinoma in China. *Br. J. Cancer* 92 (3), 607–612. doi:10.1038/sj.bjc.6602333

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

Tang, K. Y., Du, S. L., Wang, Q. L., Zhang, Y. F., and Song, H. Y. (2020). Traditional Chinese medicine targeting cancer stem cells as an alternative treatment for hepatocellular carcinoma. *J. Integr. Med.* 18 (3), 196–202. doi:10.1016/j.joim.2020.02.002Epub2020 Feb 7

Wang, M., Ye, Q., Mao, D., and Li, H. (2020). Research progress in liver-regenerating microenvironment and DNA methylation in hepatocellular carcinoma: The role of traditional Chinese medicine. *Med. Sci. Monit.* 26, e920310. doi:10.12659/MSM.920310

Wang, S., Liu, Y., Feng, Y., Zhang, J., Swinnen, J., Li, Y., et al. (2019). A review on curability of cancers: More efforts for novel therapeutic options are needed. *Cancers (Basel)* 11 (11), 1782. doi:10.3390/cancers11111782

Wang, Y., Zhang, Q., Chen, Y., Liang, C. L., Liu, H., Qiu, F., et al. (2020). Antitumor effects of immunity-enhancing traditional Chinese medicine. *Biomed. Pharmacother.* 121, 109570. doi:10.1016/j.biopha.2019.109570

Wei, L., Wang, Z., Jing, N., Lu, Y., Yang, J., Xiao, H., et al. (2022). Frontier progress of the combination of modern medicine and traditional Chinese medicine in the treatment of hepatocellular carcinoma. *Chin. Med.* 17 (1), 90. doi:10.1186/s13020-022-00645-0

Xi, S. Y., and Minuk, G. Y. (2018). Role of traditional Chinese medicine in the management of patients with hepatocellular carcinoma. *World J. Hepatol.* 10 (11), 799–806. doi:10.4254/wjh.v10.i11.799

Xu, F., Jin, T., Zhu, Y., and Dai, C. (2018). Immune checkpoint therapy in liver cancer. *J. Exp. Clin. Cancer Res.* 37 (1), 110. doi:10.1186/s13046-018-0777-4

Yang, Z., Liao, X., Lu, Y., Xu, Q., Tang, B., Chen, X., et al. (2017). Add-on therapy with traditional Chinese medicine improves outcomes and reduces adverse events in hepatocellular carcinoma: A meta-analysis of randomized controlled trials. *Evid. Based Complement. Altern. Med.* 2017, 3428253. doi:10.1155/2017/3428253

Zhai, B., Zhang, N., Han, X., Li, Q., Zhang, M., Chen, X., et al. (2019). Molecular targets of β-elemene, a herbal extract used in traditional Chinese medicine, and its potential role in cancer therapy: A review. *Biomed. Pharmacother.* 114, 108812. doi:10.1016/j.biopha.2019.108812

Zhang, Z., Li, J. W., Zeng, P. H., Gao, W. H., and Tian, X. F. (2022). Data mining and systems pharmacology to elucidate effectiveness and mechanisms of Chinese medicine in treating primary liver cancer. *Chin. J. Integr. Med.* 28 (7), 636–643. doi:10.1007/s11655-021-3449-8

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 10 January 2023
DOI 10.3389/fphar.2022.1091018

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Cong Chen,
Nantong University, China
Fangchao Zhao,
Second Hospital of Hebei Medical University, China

*CORRESPONDENCE
Li-mei Zhao,
Lmzhao19@163.com

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 06 November 2022
ACCEPTED 19 December 2022
PUBLISHED 10 January 2023

CITATION
Zhou H-m and Zhao L-m (2023), Wnt signaling pathway-derived score for predicting therapeutic resistance and tumor microenvironment in lung adenocarcinoma.
Front. Pharmacol. 13:1091018.
doi: 10.3389/fphar.2022.1091018

Wnt signaling pathway-derived score for predicting therapeutic resistance and tumor microenvironment in lung adenocarcinoma

Hao-min Zhou[1] and Li-mei Zhao[2]*

[1]Department of Intensive Care Unit, Shengjing Hospital of China Medical University, Shenyang, Liaoning, China, [2]Department of Pharmacy, Shengjing Hospital of China Medical University, Shenyang, Liaoning, China

Background: Lung adenocarcinoma (LUAD) is the most common subtype of lung cancer. Due to tumor heterogeneity, understanding the pathological mechanism of tumor progression helps to improve the diagnosis process and clinical treatment strategies of LUAD patients.

Methods: The transcriptome pattern, mutant expression and complete clinical information were obtained from the cancer genome atlas (TCGA) database and microarray data from gene expression omnibus (GEO) database. Firstly, we used single sample Gene Set Enrichment Analysis (ssGSEA) to estimate the activation of Wnt signaling pathway in each sample. Consensus clustering algorithm was used to classify LUAD samples into different subgroups according to the transcription patterns of 152 Wnt signaling pathway related genes. Then, ESTIMATE, ssGSEA and Gene Set Variation Analysis (GSVA) algorithms were used to assess the biological pathways and immunocytes infiltration between different subtypes. LASSO-COX algorithm was conducted to construct prognostic model. Kaplan-Meier and multivariate Cox analysis were performed to evaluate the predictive performance of risk model. Gene features were further confirmed using external datasets. Finally, we conducted vitro assay for validating hub gene (LEF1).

Results: Based on the transcription patterns of 152 Wnt signaling pathway related genes, four different subtypes of LUAD patients were screened out by consensus clustering algorithm. Subsequently, it was found that patients with cluster A and B had massive immunocytes infiltration, and the survival rate of patients with cluster B was better than that of other subgroups. According to the coefficients in the LASSO- Cox model and the transcriptome patterns of these 18 genes, the risk score was constructed for each sample. The degree of malignancy of LUAD patients with high-risk subgroup was remarkable higher than that of patients with low-risk subgroup ($p < 0.001$). Subsequently, five top prognostic genes (AXIN1, CTNNB1, LEF1, FZD2, FZD4.) were screened, and their expression values were different between cancer and normal tissues. FZD2 and LEF1 were negatively related to ImmunoScore, and AXIN1 was negatively related to ImmunoScore. The significant correlation between LUAD patient risk score and overall survival (OS) was verified in external datasets. In the A549 cell line, knockdown of LEF1 can reduce the invasive and proliferation ability of LUAD cells.

Conclusion: A innovative 18 genes predictive feature based on transcriptome pattern was found in patients with lung adenocarcinoma. These investigations further promote the insight of the prognosis of lung adenocarcinoma and may contribute to disease management at risk stratification.

KEYWORDS

lung adenocarcinoma, tumor microenvironment, immunotherapy targets, unsupervised clustering, molecular target

Introduction

Lung adenocarcinoma (LUAD) is the main cause of cancer-related deaths worldwide, accounting for about 40% of lung cancer patients (Nguyen et al., 2022). Over the past decade, treatment for epidermal growth factor receptors and anaplastic lymphoma kinases has benefited only a small proportion of LUAD patients (Wang et al., 2019; Wang et al., 2020). Clinically defined LUAD molecular subtypes urgently need precise treatment.

Previous studies have integrated different data types such as transcriptome, genome, and metabonomics to characterize the molecular mechanism of lung cancer and predict the survival status of cancer patients. Zou et al. (2020) quantified the infiltration of immunocytes in 32 types of cancer and observed considerable heterogeneity in the prognostic correlation of these immunocytes in different types of cancer. An immune cell feature score model with good prognosis was constructed for LUAD, which can further deepen our understanding of the LUAD and have an impact on immunotherapy (Zuo et al., 2020). On the basis of co-occurrence of KEAP1 mutation, Marinelli et al. (2020) identified four genes that may be related to the reduction of immunotherapy effect (KEAP1, PBRM1, SMARCA4, and STK11). This study suggested that co-existing changes in a limited set of genes were the characteristics of LUAD patients who are no response to immunotherapy and high TMB. The immune cold microenvironment may explain the clinical process of the disease. Similar to other cancers (Troiano et al., 2020), LUAD also exhibits peculiar molecular and clinical behavior compared to lung squamous cell carcinoma.

Recently, studies have been conducted to generate genetic signatures that predict prognostic risk in patients with lung adenocarcinoma. Zhang et al. (2020b) collected seven cohorts' of 1300 patients with LUAD and constructed the first tumor Necrosis Factor family-based model to predict the clinical outcome and immune status of LUAD patients (Zhang et al., 2020a). Zhao et al. (2020) obtained RNA sequencing data of LUAD from the cancer genome atlas (TCGA) database, as well as microarray data from the gene expression omnibus (GEO) database, and found a new type of 19 prognostic characteristics based on transcriptome pattern in LUAD patients. Compared with patients with high-risk scores, the mortality risk of patients with low-risk scores was reduced by 81%. The above investigation identified different genetic characteristics for prognostic risk prediction by using different algorism and presented different genomics profile.

Wnt signaling pathway is a classic tumor activation pathway related to tumor progression, which regulates cell growth, differentiation and migration. (Yang et al., 2017; Pan et al., 2020; Li et al., 2021a). There are three Wnt signaling pathways that have been described so far: classical β-catenin dependent pathway, non-classical Wnt/calcium pathway and non-classical planar cell polarity pathway (Zhao et al., 2019; Abplanalp et al., 2021). Wnt signal abnormalities are associated with some tumor disease, and the most significant ones are lung cancer, breast cancer, bladder cancer, clear cell renal cell cancer and prostate cancer (Xu et al., 2022; Xue et al., 2022; Zhao et al., 2022). Most investigation on Wnt signaling pathway in lung adenocarcinoma only focus on Wnt pathway as a downstream signaling pathway to regulate the proliferation and differentiation of lung adenocarcinoma cells (Tammela et al., 2017; Wu et al., 2021).

This investigation did not demonstrate the effect of Wnt signaling pathway as a whole regulatory profile on the results of lung adenocarcinoma. Therefore, current studies have identified the correlation between Wnt signaling pathway and clinicopathologic parameters of cancer by using transcriptome pattern downloaded from TCGA website, and identified the impact of Wnt pathway-related genes on lung adenocarcinoma results. Moreover, a predictive model was established based on transcriptome pattern, and its applicability and predictive performance in lung adenocarcinoma were evaluated.

Materials and methods

Dataset collecting and processing

The transcriptome pattern, mutant expression and complete clinical information were obtained from the TCGA website and microarray data from GEO website. Patients without complete clinical data were excluded from assessment. A total of two datasets (GSE68465 and GSE72094) were downloaded, and the combat method of "sva" R package was employed to remove the batch effect. The merged TCGA and GEO cohort were named as meta-cohort.

Consensus clustering of lung adenocarcinoma

We first collected 152 Wnt signaling pathway related genes from previous articles (Sun et al., 2021). According to the transcriptome pattern of 152 Wnt signaling pathway-related genes, the optimal k-means clustering ("kmeans" function in R tool) was employed to assign different distribution information to each patient, and each patient was classified into our subgroups. The "ConsensusClusterPlus" R package was employed for cluster analysis, and 1000 cycles were calculated to ensure stability and reliability. Kaplan-Meier algorism was performed to assess the total survival (OS) rate between different subgroups.

Identification of differential pathways

The GSVA algorithm was conducted to assess the differences in biological pathways between subgroups. GSVA algorism mainly evaluated gene set enrichment based on microarray nuclear transcriptome level. The principle of GSVA algorithm is to calculates the expression matrix of gene set among samples, so as to evaluate whether the mechanism pathway is enriched among different samples (Hänzelmann et al., 2013). c2. cp. kegg. v7.0. symbols as a reference gene set, FDR <0.05 as a screening threshold.

Comparison of immunocytes infiltration

In order to investigate the component of immunocytes infiltration between the four subtypes, we employed the "ESTIMATE" technological method to estimate immune score and stromal score for further prediction of immunoreaction and tumor microenvironment. Then we conducted the ssGSEA technological method to assess the enrichment level of immunocytes infiltration and immune-related pathways based on meta-cohort. Mann-Whitney U test was performed to compare the differences among the four subgroups.

LASSO-COX regression and prognostic signature verification

In order to improve the predictive performance and feasibility of the Wnt signal pathway related model, Lasso-Cox model was conducted to analyze the correlation between the clinical characteristics of Wnt signaling pathway and the risk score. Using "glmnet" R package, the best Wnt signaling pathway genes were screened and the prognosis model was constructed. Use the following formula to generate risk score:

Risk score = expression*$gene_i$ × coefficient*gene $_i$

Importantly, we used coefficient of each gene in multivariate cox regression. The according to the median risk, the patients were divided into high-risk group and low-risk group. Target genes in the model include risk genes and protective genes. HR > 1 was considered as risk factor, whereas it was a protective factor. Then, Kaplan-Meier survival model and ROC curve model were draw to assess the predictive performance of the model. We input the risk signature genes into the STRING database and employed MCC method to identify the key Top 5 molecular in the protein-protein interaction network.

Vitro assay

The shRNAs targeting LEF1 were designated by Biomics Biotechnologies Co. Ltd. (Nantong, China). The expression plasmids, pU6H1-GFP-shLEF1-1, -2, -3, and controls were constructed. Sequences of targeting LEF1 from references as follows (Xiao et al., 2021): shRNA1: GCGATTTAGCTGACA TCAA, shRNA2: AGATGTCAACTCCAAACAA, shRNA3: GTT GCTGAGTGTACTCTAA, shRNA-NC: TTCTCCGAACGTGTC ACGT. Additional experimental details are presented in our previous study (Feng et al., 2022a; Feng et al., 2022b; Cheng et al., 2022).

Results

Wnt signaling score based on ssGSEA in the multicenter study

The Wnt signaling pathway absolute enrichment score were calculated in each dataset (Figure 1A). The range of Wnt scores was approximated in most of the datasets except for GSE14814 dataset. In addition, cox regression analysis demonstrated that Wnt scores had prognostic value in the GSE31210, GSE30129, GSE29016, GSE31210, GSE72094 and GSE68465 datasets (Figure 1B). Interestingly, the mutational landscape (Figure 1C) as well as the activation pathways (Figure 1D) were significantly different in different Wnt groupings.

Relationship between Wnt signaling pathway related genes and phenotypic characteristics of lung adenocarcinoma

In order to explore whether Wnt signaling pathway related genes play a regulatory role in lung adenocarcinoma. We first explored the Wnt signaling pathway related genes in cancer tissues and normal tissues based on TCGA dataset. Heatmap showed the expression level of Wnt signaling pathway related genes (Figure 2A). We found that there were significant differences in molecular expression between normal tissues and tumor tissues. Based on the expression pattern of Wnt signaling pathway related genes, we classified the lung adenocarcinoma patients into different subgroups. Using the similarity of Wnt-related gene expression, we selected the value of k = 4 (Figures 2B,C). The lung adenocarcinoma patients from meta-cohort dataset were divided into four subgroups (Cluster A, Cluster B, Cluster C and Cluster D). The four subgroups contained 376 samples, 451 samples, 229 samples and 259 samples, respectively. As shown in Figure 2D, K-M survival curve analysis showed that cluster B had the best clinical outcomes, and cluster C had the shortest survival time ($p < 0.001$). We compared the expression patterns of Wnt signaling pathway-related genes in these subgroups (Figure 2E).

Differences of TME infiltration among four subtypes of Wnt signaling pathway

Subsequently, we explored the tumor immune microenvironment among different subtypes to attempt to explain that Wnt signaling pathway affects the clinical outcome of patients by regulating the tumor immune microenvironment. We analyzed the information of immune cell infiltration and found that activated innate immune cells in cluster A and B were abundant, including activated dendritic cells, CD56dim natural killer cells, macrophages, activated B cells, and activated CD4 and CD8 T cells, which had significant survival advantages (Figure 3A). Although some tumor tissues have a large number of immune cells, these immune cells cannot penetrate the tumor and are forced to remain in the surrounding stromal tissues. Therefore, stromal activation in tumor microenvironment is considered to be immunosuppressive (333). Therefore, the clinical outcome of Cluster A is not ideal. In this investigation, the "ESTIMATE" algorism also showed that the immune scores of cluster A and B were higher than those of cluster C and D (Figure 3B). We also noticed that although the types of immunocytes infiltration were consistent in different subtypes, the proportion of immunocytes in different subgroups was different. This indicates that tumor regulatory pathways such as Wnt, RAS and other signaling pathways do not alter the types of immune infiltrating cells, but they may change their proportions. Figure 3C verified the above results. Subsequently, we continued to detect the expression levels

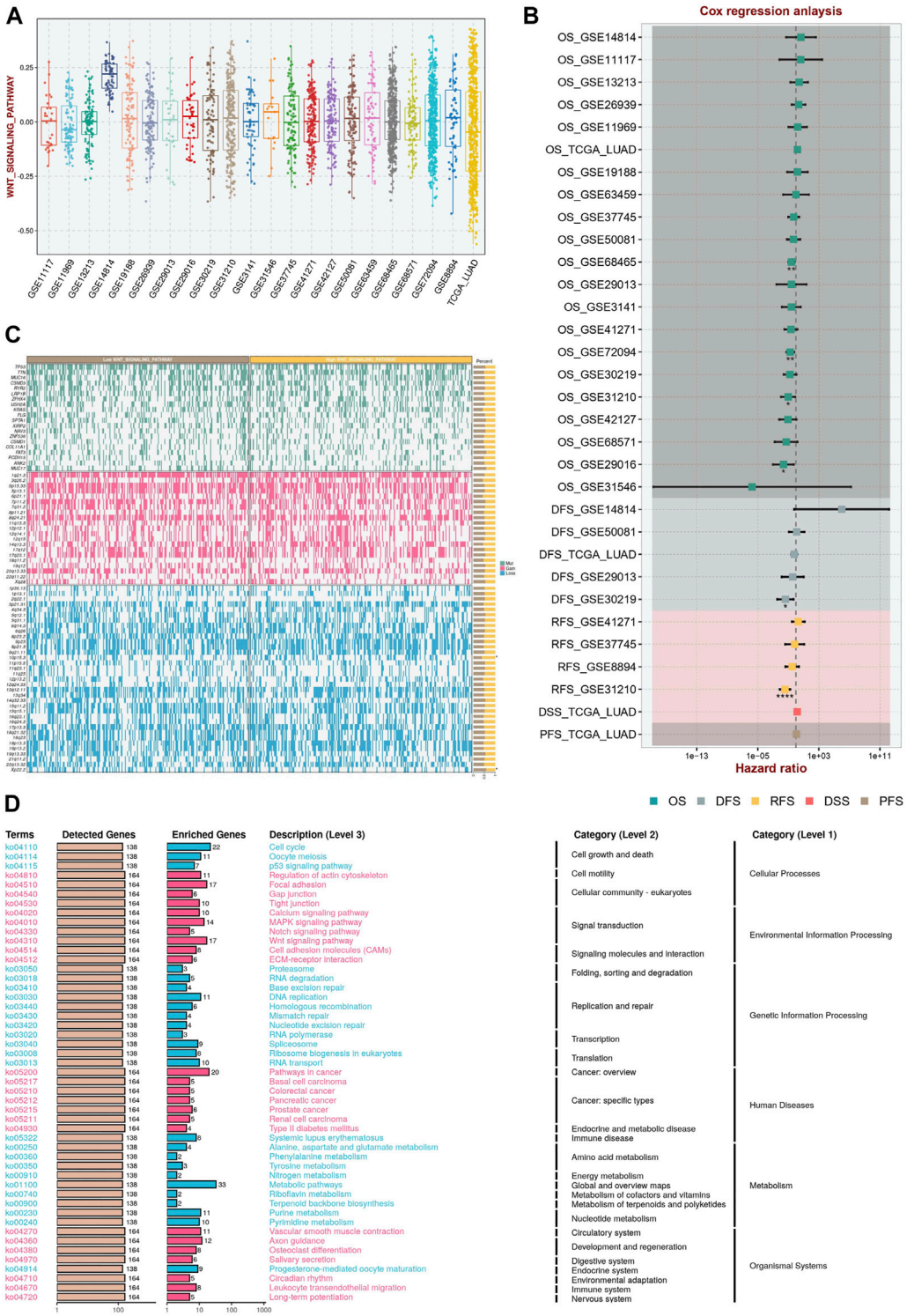


FIGURE 1
Wnt Signaling Score Based on ssGSEA in the Multicenter Study. **(A)** The Wnt signaling pathway absolute enrichment score based ssGSEA in different datasets. **(B)** Cox regression analysis in multicenter study demonstrated that Wnt scores had prognostic value. **(C)** The mutational landscape in different Wnt groupings. **(D)** The activation pathways in different Wnt groupings.

of immunomodulators among different subtypes. The Mann-Whitney U test was conducted to compare differences between the four subgroups. Immunomodulators such as CD27, CD86, CTLA4 and HLA family were highly expressed in cluster A and B. Immunomodulators induce immune response and regulate immune response. This further confirmed that cluster A and B have higher immune cell infiltration levels than cluster C and D (Figures 3D,E).

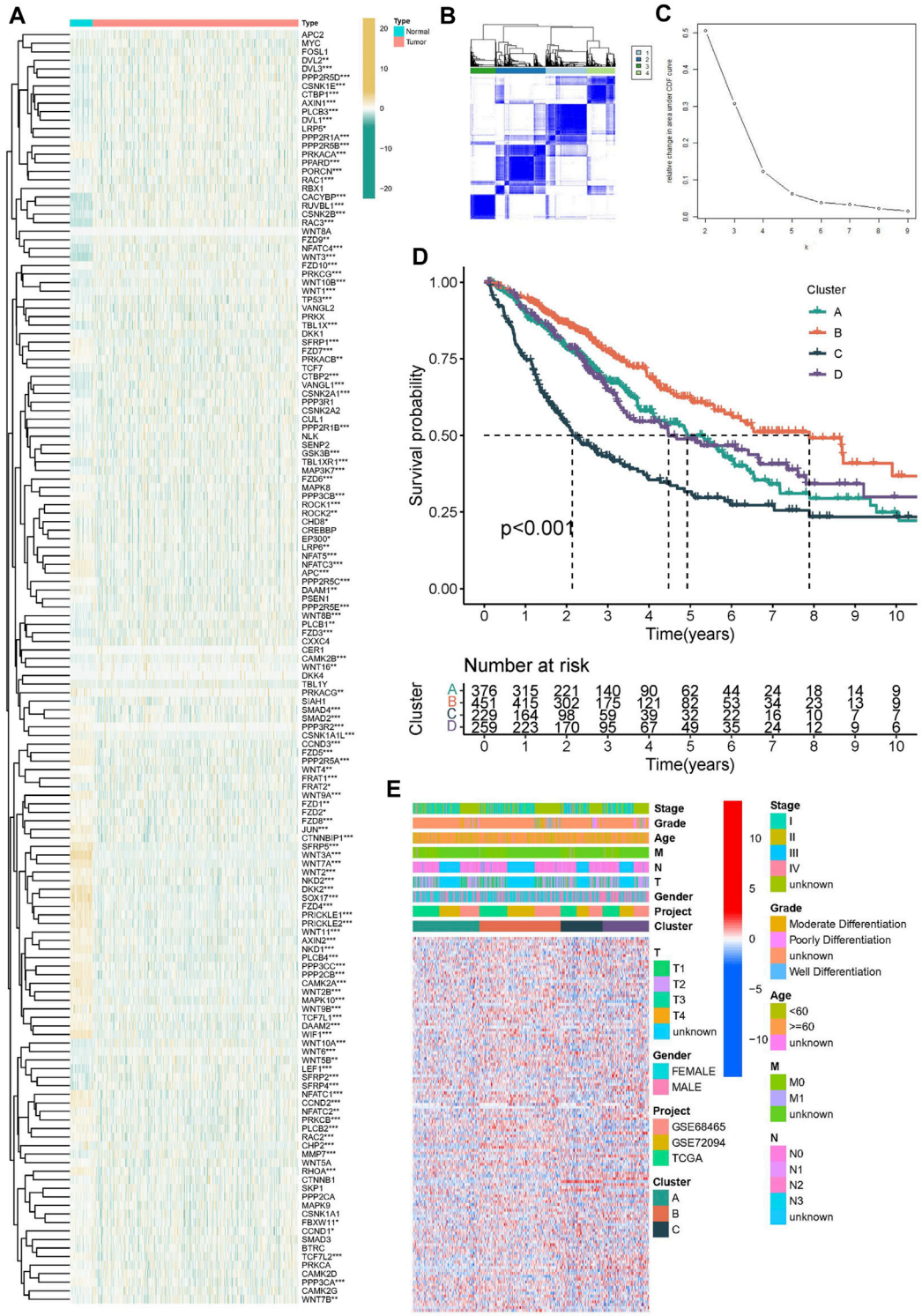


FIGURE 2
Construction of molecular subtypes in lung adenocarcinoma. **(A)** The heatmap was utilized to present the transcriptome pattern of 152 wtn signal pathway related genes in distinct samples. **(B,C)** Unsupervised clustering was conducted to divided samples into four subgroups performed according to the 152 wtn signal pathway related genes patterns. **(D)** Kaplan–Meier survival model for lung adenocarcinoma, $p < 0.001$. **(E)** The heatmap was utilized to present the transcriptome pattern of 152 wtn signal pathway related genes in distinct subtypes. Red/blue represented high/low expression of genes. The comments on the right include TNM stage, clinical grade/stage, gender, age, project and cluster, respectively.

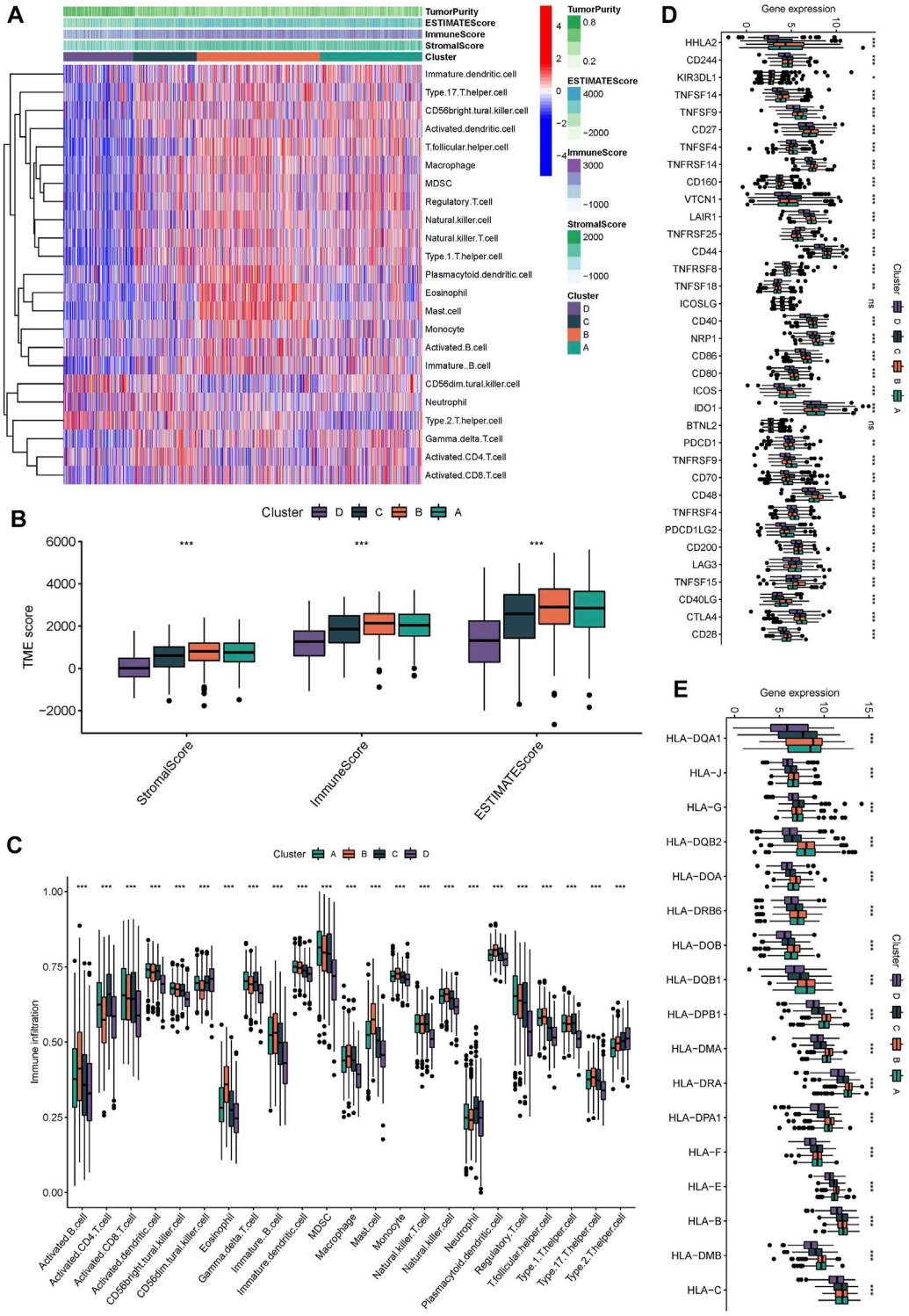


FIGURE 3
Immunological characteristic of different molecular subgroups. **(A)** the heatmap displayed the immunocytes infiltration in different molecular subgroups. **(B)** the stromalScore, immunoScore, ESTIMATEScore was calculated by "ESTIMATE" algorism. **(C)** the immunocytes infiltration in different molecular subgroups. **(D,E)** Expression levels of immunomodulators in different molecular subgroup, The median value: black lines in boxes, the outliers: black dots out boxes.

Identification of biological behavior patterns

Subsequently, we conducted GSVA enrichment analysis to identify the differences in biological behaviors among these four subgroups, so as to determine the pathways through which Wnt signaling pathway-related genes regulate malignant differentiation of tumor cells. Cluster A presented with Cell cycle, DNA replication, mismatch repair. Cluster B showed enrichment of metabolic pathways, including linoleic acid

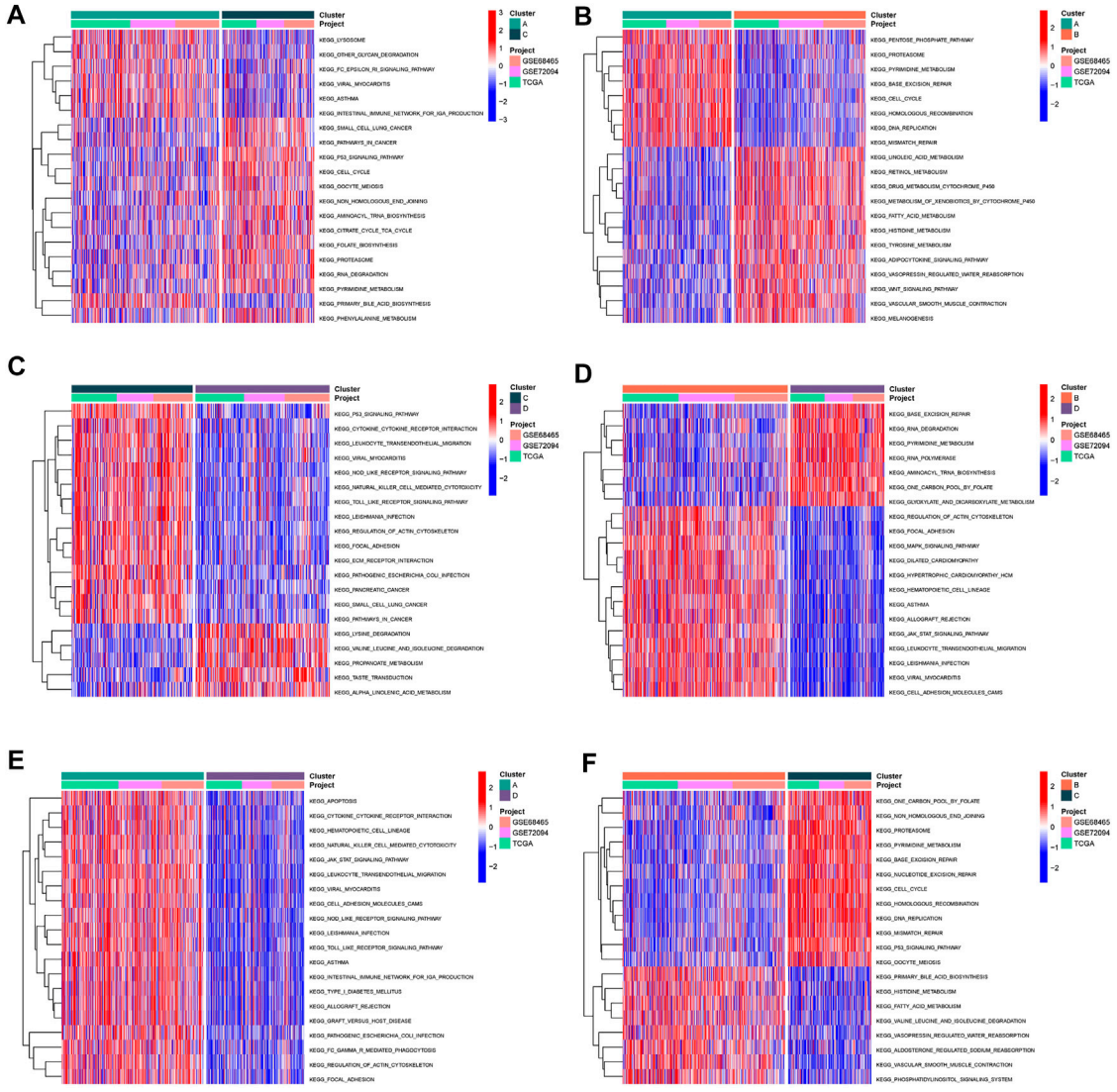


FIGURE 4
Pathway enrichment analysis **(A–F)** The heatmap displayed the biological processes among each cluster, plotted by GSVA algorithm.

metabolism, retinal metabolism, drug metabolism cytochrome P450 signaling pathways. Cluster C shows enrichment of carcinogenic activation pathways, such as small cell lung cancer, cell cycle, pathway in cancer, and P53 signature. Cluster D is enriched in RNA degradation, base excision repair and other signaling pathways (Figure 4).

Prognostic prediction model of Wnt pathway in lung adenocarcinoma

After discovering that Wnt signaling pathway was related to the clinical results of lung cancer, we tried to establish prognostic risk score based on Wnt signaling pathway. We collected 152 Wnt pathway related genes. After applying the LASSO algorithm to these 152 genes, we removed highly correlated genes and reduced the dimension of subsequent multivariate COX algorithm (Figure 5A). Subsequently, COX multivariate model was constructed to screen the final prognostic genes. A total of 18 genes were identified as independent prognostic genes, including FZD4, FZD7, LEF1, FZD9, CTNNBIP1, AXIN1, DKK4, CSNK2A1, TBL1Y, NFATC1, PLCB2, PLCB3, PRKCG, FOSL1, PSEN1, CTNNB1, PPARD, and FZD2. The LASSO-Cox regression coefficient was integrated with the corresponding gene expression values to establish a risk marker feature (Figure 5B). According to the median risk marker, we divided lung adenocarcinoma patients into high-risk group and low risk group. The heatmap showed the transcriptome pattern of prognostic genes in high and low risk patients. Red represents high expression, while green represents low expression (Figure 5C). The scatter plot displayed the risk score of patients, and the correlation between mortality and risk score. With the increase of risk value, mortality increases (Figures 5D,E). In addition, the K-M survival curve showed that the overall survival time of the low-risk group was longer than that of the high-risk group in TCGA cohort and GEO cohort (Figures 5F,G). In order to test the predictive performance of the prognostic model, we draw the ROC curve, and the area under the curve (AUC) was used to evaluate the 1, 3, and 5–year survival rates in TCGA cohort, 0.749, 0.756, and 0.743 respectively (Figure 6A). The multivariate Cox model proved that the prognostic model can independently predict the clinical outcomes of lung

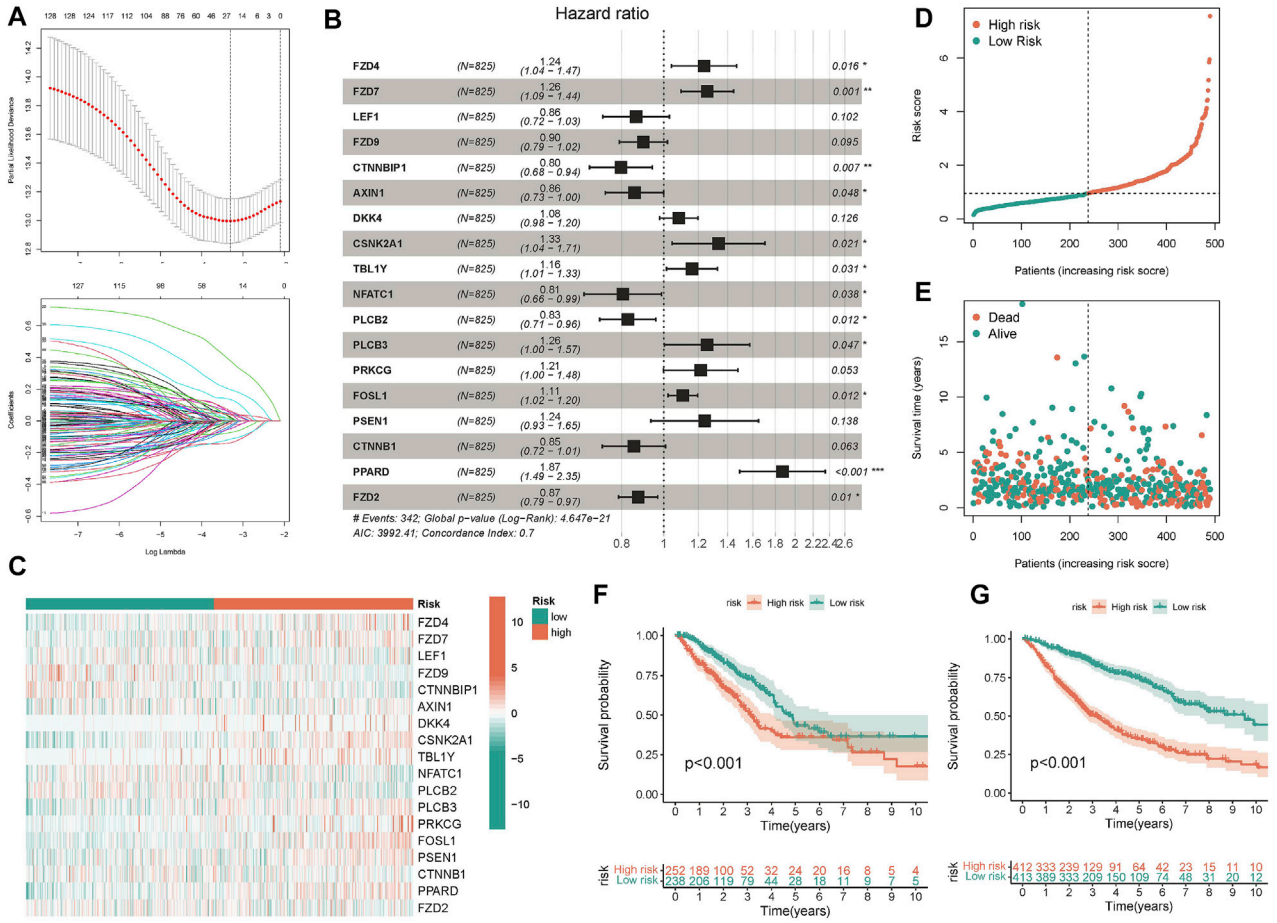


FIGURE 5
Construction and validation of predictive model. **(A)** 26 candidate genes were selected by LASSO regression **(B)** The forest map displayed the hazard ratios of the 18 target genes **(C)** The heatmap displayed the expression values of 18 target genes in different risk score groups. **(D,E)** The scatter plot depicts the distribution of patient risk values and mortality. **(F,G)** The K-M survival model was constructed to explore the predictive performance.

adenocarcinoma patients (Figures 6B,C). As an external validation, we performed ROC curve, univariate and multivariate Cox model analysis again. The area under the curve (AUC) was used to evaluate the 1, 3 and 5–year survival rates in GEO cohort, 0.694, 0.661, and 0.565 respectively (Figure 6D). The results of univariate and multivariate Cox analysis were consistent with the above (Figures 6E, F). Clinical pathological parameters, especially TNM staging, are closely related to the prognosis of patients. In order to improve the accuracy of the prediction model, we integrated M stage and risk score to construct a nomogram for improving the prediction performance (Figure 6G). The calibration curve of the established nomogram showed high accuracy between the actual observation value and the predicted value (Figures 6H, I).

Association of five genes with immune microenvironment in lung adenocarcinoma

In order to explore the mechanism of 18 prognostic genes, we constructed the interaction diagram of 18 prognostic genes (Figure 7A). MCC method was used to identify the key Top 5 proteins in the protein-protein interaction network, including AXIN1, CTNNB1, LEF1, FZD2, FZD4 (Figure 7B). Therefore, we speculate that the above five top genes may be the key molecules affecting the disease progression. We detected the expression level of these molecule in normal tissues and cancer tissues of lung adenocarcinoma patients in TCGA and GEO cohorts. We found that AXIN1, CTNNB1 and LEF1 were highly expressed in tumor tissues and FZD4 was highly expressed in normal tissues in TCGA cohort (Figure 7C). In GEO cohort, AXIN1 and LEF1 were highly expressed in tumor tissues, while FZD4 and FZD2 were highly expressed in normal tissues (Figure 7D). Subsequently, we detected the correlation between these five genes and immune scores. FZD2 and LEF1 were positively related to immune scores, and AXIN1 was negatively related to immune score (Figures 7E–G). Finally, we conducted vitro assay for validating the carcinogenic potency of LEF1 (Figures 7H–J). In the A549 cell line, knockdown of LEF1 can reduce the invasive and proliferation ability of LUAD cells.

Discussion

Lung adenocarcinoma is the most common form of lung cancer and the deadest cancer in the world. In order to characterize the

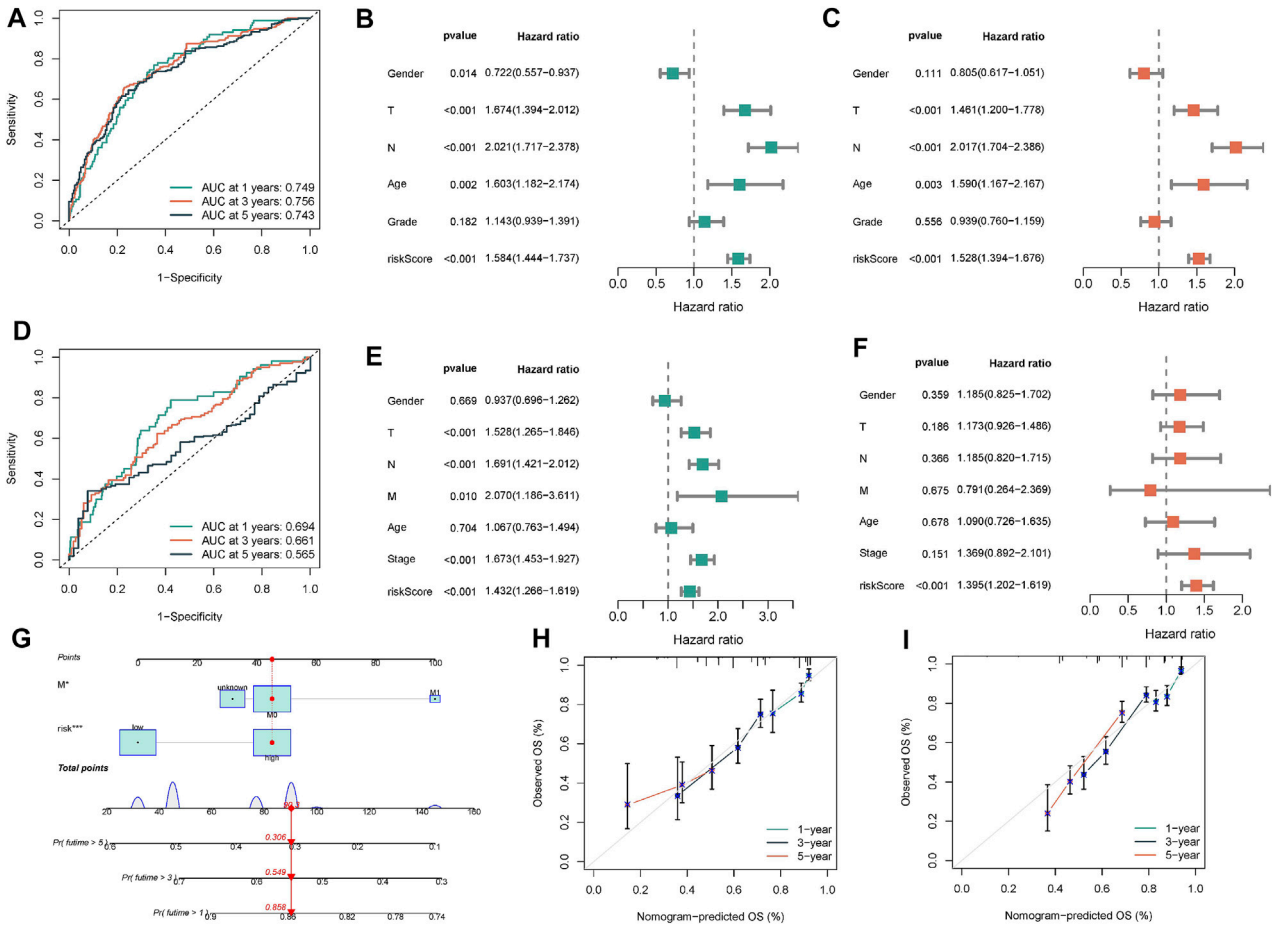


FIGURE 6
Construction and validation of nomogram. **(A)** The ROC curve was performed to verify the predictive performance of predictive model. **(B,C)** Univariate/multivariate Cox regression model was constructed to verify the independent predictive performance of each parameter in TCGA-BLCA cohort. **(D–F)** ROC curve and Univariate/multivariate Cox regression model were plotted to verify in GEO database. **(G)** Construction of a nomogram combining M stage and risk score. **(H,I)** Calibration plot showing that nomogram-predicted survival probabilities corresponded closely to the observed proportions.

genomic and transcriptome abnormalities of lung adenocarcinoma and identify the survival outcomes of patients, different data types such as transcriptome, genome, and metabonomics have been employed for comprehensive analysis (Jin et al., 2020; Sun et al., 2020). Recently, prognostic risk features have emerged to cluster lung adenocarcinoma patients. In our investigation, we used RNA-seq data which downloaded from the TCGA website to determine the correlation between Wnt signaling pathway and the clinical outcomes of lung adenocarcinoma. In addition, a prognostic model was established based on transcriptome pattern, and its applicability and value in lung adenocarcinoma were evaluated.

Abnormalities of Wnt signaling pathway is usually associated with cancer (Lei et al., 2021). The characteristics of lung adenocarcinoma include abnormal epigenetic regulation of Wnt pathway genes and inactivation of tumor suppressor genes (Hou et al., 2021). Tumor suppressor genes inhibit Wnt signaling pathway. However, the silencing of these tumor suppressor genes leads to the activation of Wnt signaling pathway, which is involved in the occurrence or progression of human malignant tumors (Líbalová et al., 2014). NKX2-1/ERK drived Wnt pathway to promote cell proliferation, shorten the value-added cycle, and increased the malignancy of lung adenocarcinoma. BRAF/MEK inhibitors drived NKX2-1 positive tumor cells into a stationary state, while NKX2-1 negative cells cannot exit the cell cycle after the same treatment. These data clarify the complex interrelationship between lineage specificity and carcinogenic signaling pathways, which may affect lineage-specific treatment strategies in regulating lung adenocarcinoma characteristics. Kerdidani et al. demonstrated that Wnt silenced chemokine genes in dendritic cells and induces adaptive immune resistance in lung adenocarcinoma. Moreover, we examined whether 18 target genes from classical and non-classical Wnt signaling pathways could accurately detect the risk of lung adenocarcinoma. The report of the Guidelines for Prognosis of Tumor Markers has recently been applied in many journals (Zhang et al., 2020b).

Based on the TCGA data, we constructed a lung adenocarcinoma prognosis model, and we found five key Wnt signaling pathway related genes, including AXIN1, CTNNB1, LEF1, FZD2, FZD4. These key genes were related to cancer and play an indispensable role in lung adenocarcinoma pathway. These markers were candidate genes for molecular targeting. Li et al. (2021b) confirmed that AXIN1 encoding the negative regulator of

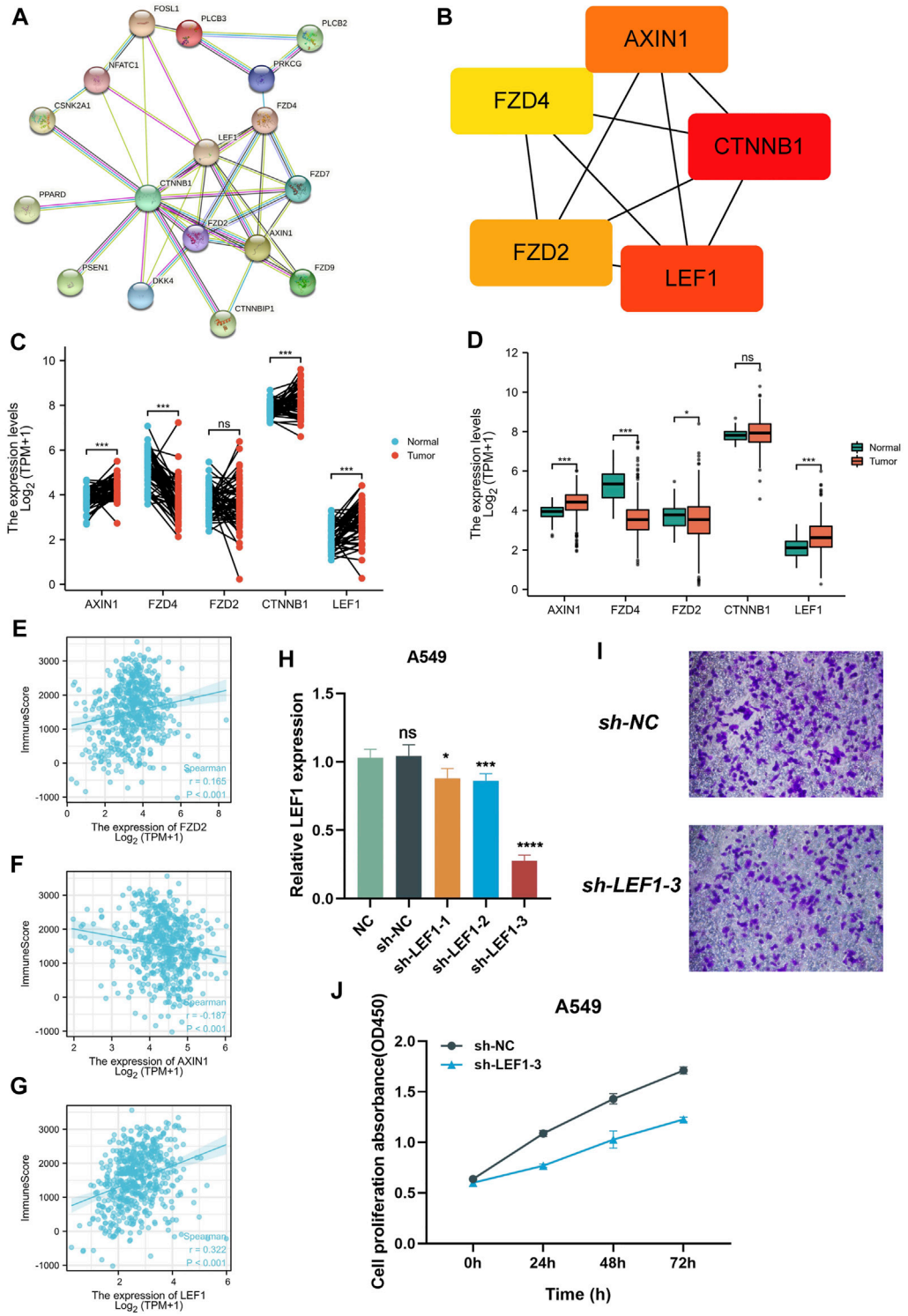


FIGURE 7
Collection of key prognostic genes. **(A)** Protein-protein interaction network diagram of 18 target genes **(B)** MCC method was performed to identify the key Top 5 molecular in the protein-protein Interaction network. **(C,D)** the expression level of 5 key prognostic genes in normal tissues and tumor tissues **(E–G)** the Correlation between 5 key prognostic genes and immune scores. **(H)** qPCR assay detect the expression of LEF1. **(I)** Transwell assay detect the invasion of the A549 cell line after knockdown LEF1. **(J)** CCK-8 assay detect the proliferation of the A549 cell line after knockdown LEF1.

Wnt/β-catenin signal was the direct target of YTHDF2. YTHDF2 promotes AXIN1 mRNA decay and subsequently activates Wnt/β-catenin signal. Knockout of AXIN1 fully rescued the inhibitory effect of YTHDF2 depletion on proliferation, colony formation and migration of lung cancer cells. These results suggest that YTHDF2 promotes the

development of LUAD by upregulating AXIN1/Wnt/β-catenin signaling, which may be a potential therapeutic target for LUAD. Zhou et al. 2022 collected 564 patients with lung adenocarcinoma, of which 30 (5.3%) carried CTNNB1 mutations. The study found that female patients and non-smokers had high CTNNB1 mutations, and the clinical outcomes of primary lung adenocarcinoma with CTNNB1 mutations was poor (Zhou et al., 2019). Nguyen et found that LEF1 and HOXB9 mediated Wnt/TCF signal transduction to promote lung adenocarcinoma metastasis, indicating that LEF1 may be a target for lung adenocarcinoma metastasis (Nguyen et al., 2009). In addition, Li et al. 2021a collected clinical information of LUSC patients from the Cancer Genome Atlas database and related methylation sequences from the University of California, Santa Cruz database to construct methylation subtypes and analyze clinical outcomes. The researchers constructed a predictive model based on the difference of DNA methylation level to classify the molecular subtypes of LUSC patients, and provided more personalized clinical treatment strategies according to different clinical subtypes. GNAS, FZD2 and FZD10 are the three core genes that may be related to the prognosis of LUSC patients (Li et al., 2021c).

First of all, our investigation was worthy of recognition, which obtained Wnt-related prognostic genes, and then constructed a Wnt-related scoring model. The function of Wnt is most common in embryogenesis and tumorigenesis, but previous literature has only described the carcinogenic effect of Wnt and rarely investigated its predictive ability. Therefore, our study opened up a new horizon for exploring the Wnt pathway. Our research has some limitations. First of all, our investigation was based on the public database for multi-omics analysis, which was a retrospective study with subjective bias. Future investigation needed to validate the diagnostic value of 18 candidate genes in fresh frozen biopsy or peripheral blood. Secondly, our research had a moderate sample size and needed a larger case-control study for LUAD patients.

Data availability statement

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

Author contributions

H-MZ designed and analyzed the research study. L-MZ wrote and revised the manuscript, and all authors have read and approved the manuscript.

Funding

This study is supported by outstanding scientific fund of Shengjing hospital (No. m0779).

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

Abplanalp, W. T., John, D., Cremer, S., Assmus, B., Dorsheimer, L., Hoffmann, J., et al. (2021). Single-cell RNA-sequencing reveals profound changes in circulating immune cells in patients with heart failure. *Cardiovasc. Res.* 117, 484–494. doi:10.1093/cvr/cvaa101

Cheng, X., Li, J., Feng, L., Feng, S., Wu, X., and Li, Y. (2022). The role of hypoxia-related genes in TACE-refractory hepatocellular carcinoma: Exploration of prognosis, immunological characteristics and drug resistance based on onco-multi-OMICS approach. *Front. Pharmacol.* 13, 1011033. doi:10.3389/fphar.2022.1011033

Feng, S., Xia, T., Ge, Y., Zhang, K., Ji, X., Luo, S., et al. (2022). Computed tomography imaging-based radiogenomics analysis reveals hypoxia patterns and immunological characteristics in ovarian cancer. *Front. Immunol.* 13, 868067. doi:10.3389/fimmu.2022.868067

Feng, S., Xu, Y., Dai, Z., Yin, H., Zhang, K., and Shen, Y. (2022). Integrative analysis from multicenter studies identifies a WGCNA-derived cancer-associated fibroblast signature for ovarian cancer. *Front. Immunol.* 13, 951582. doi:10.3389/fimmu.2022.951582

Hänzelmann, 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

Hou, M., Wu, N., and Yao, L. (2021). LncRNA CBR3-AS1 potentiates Wnt/β-catenin signaling to regulate lung adenocarcinoma cells proliferation, migration and invasion. *Cancer Cell Int.* 21, 36. doi:10.1186/s12935-020-01685-y

Jin, C. Y., Du, L., Nuerlan, A. H., Wang, X. L., Yang, Y. W., and Guo, R. (2020). High expression of RRM2 as an independent predictive factor of poor prognosis in patients with lung adenocarcinoma. *Aging* 13, 3518–3535. doi:10.18632/aging.202292

Lei, L., Wang, Y., Li, Z. H., Fei, L. R., Huang, W. J., Zheng, Y. W., et al. (2021). PHLDA3 promotes lung adenocarcinoma cell proliferation and invasion via activation of the Wnt signaling pathway. *Laboratory investigation; a J. Tech. methods pathology* 101, 1130–1141. doi:10.1038/s41374-021-00608-3

Li, H. J., Ke, F. Y., Lin, C. C., Lu, M. Y., Kuo, Y. H., Wang, Y. P., et al. (2021). ENO1 promotes lung cancer metastasis via HGFR and WNT signaling-driven epithelial-to-mesenchymal transition. *Cancer Res.* 81, 4094–4109. doi:10.1158/0008-5472.CAN-20-3543

Li, X. S., Nie, K. C., Zheng, Z. H., Zhou, R. S., Huang, Y. S., Ye, Z. J., et al. (2021). Molecular subtypes based on DNA methylation predict prognosis in lung squamous cell carcinoma. *BMC cancer* 21, 96. doi:10.1186/s12885-021-07807-7

Li, Y., Sheng, H., Ma, F., Wu, Q., Huang, J., Chen, Q., et al. (2021). RNA m(6)A reader YTHDF2 facilitates lung adenocarcinoma cell proliferation and metastasis by targeting the AXIN1/Wnt/β-catenin signaling. *Cell death Dis.* 12, 479. doi:10.1038/s41419-021-03763-z

Líbalová, H., Krčková, S., Uhlířová, K., Kléma, J., Ciganek, M., Rössner, P., Jr., et al. (2014). Analysis of gene expression changes in A549 cells induced by organic compounds from respirable air particles. *Mutat. Res.* 770, 94–105. doi:10.1016/j.mrfmmm.2014.10.002

Marinelli, D., Mazzotta, M., Scalera, S., Terrenato, I., Sperati, F., D'Ambrosio, L., et al. (2020). KEAP1-driven co-mutations in lung adenocarcinoma unresponsive to immunotherapy despite high tumor mutational burden. *Ann. Oncol. official J. Eur. Soc. Med. Oncol.* 31, 1746–1754. doi:10.1016/j.annonc.2020.08.2105

Nguyen, B., Fong, C., Luthra, A., Smith, S. A., DiNatale, R. G., Nandakumar, S., et al. (2022). Genomic characterization of metastatic patterns from prospective clinical sequencing of 25, 000 patients. *Cell* 185, 563–575. e11. doi:10.1016/j.cell.2022.01.003

Nguyen, D. X., Chiang, A. C., Zhang, X. H., Kim, J. Y., Kris, M. G., Ladanyi, M., et al. (2009). WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. *Cell* 138, 51–62. doi:10.1016/j.cell.2009.04.030

Pan, J., Fang, S., Tian, H., Zhou, C., Zhao, X., Tian, H., et al. (2020). lncRNA JPX/miR-33a-5p/Twist1 axis regulates tumorigenesis and metastasis of lung cancer by activating Wnt/β-catenin signaling. *Mol. cancer* 19, 9. doi:10.1186/s12943-020-1133-9

Sun, S., Guo, W., Wang, Z., Wang, X., Zhang, G., Zhang, H., et al. (2020). Development and validation of an immune-related prognostic signature in lung adenocarcinoma. *Cancer Med.* 9, 5960–5975. doi:10.1002/cam4.3240

Sun, S., Wang, Y., Wang, J., and Bi, J. (2021). Wnt pathway-related three-mRNA clinical outcome signature in bladder urothelial carcinoma: Computational biology and experimental analyses. *J. Transl. Med.* 19, 409. doi:10.1186/s12967-021-03061-4

Tammela, T., Sanchez-Rivera, F. J., Cetinbas, N. M., Wu, K., Joshi, N. S., Helenius, K., et al. (2017). A Wnt-producing niche drives proliferative potential and progression in lung adenocarcinoma. *Nature* 545, 355–359. doi:10.1038/nature22334

Troiano, G., Rubini, C., Togni, L., Caponio, V. C. A., Zhurakivska, K., Santarelli, A., et al. (2020). The immune phenotype of tongue squamous cell carcinoma predicts early relapse and poor prognosis. *Cancer Med.* 9, 8333–8344. doi:10.1002/cam4.3440

Wang, S., Shi, J., Ye, Z., Dong, D., Yu, D., Zhou, M., et al. (2019). Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur. Respir. J.* 53, 1800986. doi:10.1183/13993003.00986-2018

Wang, Y., Yang, N., Zhang, Y., Li, L., Han, R., Zhu, M., et al. (2020). Effective treatment of lung adenocarcinoma harboring EGFR-activating mutation, T790M, and cis-C797S triple mutations by brigatinib and cetuximab combination therapy. *J. Thorac. Oncol. official Publ. Int. Assoc. Study Lung Cancer* 15, 1369–1375. doi:10.1016/j.jtho.2020.04.014

Wu, X. T., Wang, Y. H., Cai, X. Y., Dong, Y., Cui, Q., Zhou, Y. N., et al. (2021). RNF115 promotes lung adenocarcinoma through Wnt/β-catenin pathway activation by mediating APC ubiquitination. *Cancer & metabolism* 9, 7. doi:10.1186/s40170-021-00243-y

Xiao, L., Zhang, C., Li, X., Jia, C., Chen, L., Yuan, Y., et al. (2021). LEF1 enhances the progression of colonic adenocarcinoma via remodeling the cell motility associated structures. *Int. J. Mol. Sci.* 22, 10870. doi:10.3390/ijms221910870

Xu, Q., Wang, C., Zhou, J. X., Xu, Z. M., Gao, J., Sui, P., et al. (2022). Loss of TET reprograms Wnt signaling through impaired demethylation to promote lung cancer development. *Proc. Natl. Acad. Sci. U. S. A.* 119, e2107599119. doi:10.1073/pnas.2107599119

Xue, C., Li, G., Zheng, Q., Gu, X., Bao, Z., Lu, J., et al. (2022). The functional roles of the circRNA/Wnt axis in cancer. *Mol. cancer* 21, 108. doi:10.1186/s12943-022-01582-0

Yang, S., Liu, Y., Li, M. Y., Ng, C. S. H., Yang, S. L., Wang, S., et al. (2017). FOXP3 promotes tumor growth and metastasis by activating Wnt/β-catenin signaling pathway and EMT in non-small cell lung cancer. *Mol. cancer* 16, 124. doi:10.1186/s12943-017-0700-1

Zhang, C., Zhang, G., Sun, N., Zhang, Z., Zhang, Z., Luo, Y., et al. (2020). Comprehensive molecular analyses of a TNF family-based signature with regard to prognosis, immune features, and biomarkers for immunotherapy in lung adenocarcinoma. *EBioMedicine* 59, 102959. doi:10.1016/j.ebiom.2020.102959

Zhang, H. J., Chang, W. J., Jia, C. Y., Qiao, L., Zhou, J., Chen, Q., et al. (2020). Destrin contributes to lung adenocarcinoma progression by activating wnt/β-catenin signaling pathway. *Mol. cancer Res. MCR* 18, 1789–1802. doi:10.1158/1541-7786.MCR-20-0187

Zhao, C. C., Jiao, Y., Zhang, Y. Y., Ning, J., Zhang, Y. R., Xu, J., et al. (2019). Lnc SMAD5-AS1 as ceRNA inhibit proliferation of diffuse large B cell lymphoma via Wnt/β-catenin pathway by sponging miR-135b-5p to elevate expression of APC. *Cell death Dis.* 10, 252. doi:10.1038/s41419-019-1479-3

Zhao, H., Ming, T., Tang, S., Ren, S., Yang, H., Liu, M., et al. (2022). Wnt signaling in colorectal cancer: Pathogenic role and therapeutic target. *Mol. cancer* 21, 144. doi:10.1186/s12943-022-01616-7

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 (Amsterdam, Neth.* 149, 90–96. doi:10.1016/j.lungcan.2020.09.014

Zhou, C., Li, W., Shao, J., Zhao, J., and Chen, C. (2019). Analysis of the clinicopathologic characteristics of lung adenocarcinoma with CTNNB1 mutation. *Front. Genet.* 10, 1367. doi:10.3389/fgene.2019.01367

Zuo, S., Wei, M., Wang, S., Dong, J., and Wei, J. (2020). Pan-cancer analysis of immune cell infiltration identifies a prognostic immune-cell characteristic score (ICCS) in lung adenocarcinoma. *Front. Immunol.* 11, 1218. doi:10.3389/fimmu.2020.01218

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 16 January 2023
DOI 10.3389/fphar.2023.1099927

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Xinglong Fan,
Qilu Hospital of Shandong University, China
Ji Lv,
Maternity and Child Health Hospital of Qinhuangdao, China
Joshua Ochieng,
University of Texas MD Anderson Cancer Center, United States

*CORRESPONDENCE
Anxin Gu,
guanxin@hrbmu.edu.cn
Mingyan E,
emingyan889@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 16 November 2022
ACCEPTED 03 January 2023
PUBLISHED 16 January 2023

CITATION
Jie Y, Wu J, An D, Li M, He H, Wang D, Gu A and E M (2023), Molecular characterization based on tumor microenvironment-related signatures for guiding immunotherapy and therapeutic resistance in lung adenocarcinoma. *Front. Pharmacol.* 14:1099927. doi: 10.3389/fphar.2023.1099927

Molecular characterization based on tumor microenvironment-related signatures for guiding immunotherapy and therapeutic resistance in lung adenocarcinoma

Yamin Jie[1†], Jianing Wu[2†], Dongxue An[1], Man Li[3], Hongjiang He[4], Duo Wang[5], Anxin Gu[2*] and Mingyan E[2*]

[1]Department of Radiation Oncology, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, China, [2]Department of Radiation Oncology, Harbin Medical University Cancer Hospital, Harbin, China, [3]Department of Endoscopy, Harbin Medical University Cancer Hospital, Harbin, China, [4]Department of Head and Neck Surgery, Harbin Medical University Cancer Hospital, Harbin, China, [5]Department of Neurology, The 2nd Affiliated Hospital of Harbin Medical University, Harbin, China

Background: Although the role of tumor microenvironment in lung adenocarcinoma (LUAD) has been explored in a number of studies, the value of TME-related signatures in immunotherapy has not been comprehensively characterized.

Materials and Methods: Consensus clustering was conducted to characterize TME-based molecular subtypes using transcription data of LUAD samples. The biological pathways and immune microenvironment were assessed by CIBERSORT, ESTIMATE, and gene set enrichment analysis. A TME-related risk model was established through the algorithms of least absolute shrinkage and selection operator (Lasso) and stepwise Akaike information criterion (stepAIC).

Results: Four TME-based molecular subtypes including C1, C2, C3, and C4 were identified, and they showed distinct overall survival, genomic characteristics, DNA methylation pattern, immune microenvironment, and biological pathways. C1 had the worst prognosis and high tumor proliferation rate. C3 and C4 had higher enrichment of anti-tumor signatures compared to C1 and C2. C4 had evidently low enrichment of epithelial–mesenchymal transition (EMT) signature and tumor proliferation rate. C3 was predicted to be more sensitive to immunotherapy compared with other subtypes. C1 is more sensitive to chemotherapy drugs, including Docetaxel, Vinorelbine and Cisplatin, while C3 is more sensitive to Paclitaxel. A five-gene risk model was constructed, which showed a favorable performance in three independent datasets. Low-risk group showed a longer overall survival, more infiltrated immune cells, and higher response to immunotherapy than high-risk group.

Conclusion: This study comprehensively characterized the molecular features of LUAD patients based on TME-related signatures, demonstrating the potential of TME-based signatures in exploring the mechanisms of LUAD development. The TME-related risk model was of clinical value to predict LUAD prognosis and guide immunotherapy.

KEYWORDS

lung adenocarcinoma, molecular subtypes, immunotherapy, cisplatin, paclitaxel, DNA methylation

Introduction

Lung cancer consists of the largest population of all cancers worldwide, where lung adenocarcinoma (LUAD), as the most common histological type, contributes to a proportion of approximately 40% in all lung cancer cases (Sung et al., 2021). The common risk factors are smoking, the exposure to environmental carcinogens, and genetic susceptibility (Gibelin and Couraud, 2016). A large number of lung cancer patients are diagnosed at a late stage, leading to a low 5-year survival rate no more than 20%. (Senosain and Massion, 2020). In the recent years, molecular profiling of lung cancer promotes the development and improvement of molecular-targeted therapy and immunotherapy (Network, 2014; Saito et al., 2018). A diversity of molecular biomarkers of LUAD have been discovered involving transcriptional alteration, genetic mutations, copy number variations, and epigenetics features (Devarakonda et al., 2015; Daugaard et al., 2016; Calvayrac et al., 2017; Hua et al., 2020).

Molecular biomarkers help to predict the prognosis of cancer patients or even are capable to assist decision-makings in clinical treatment. Advanced or metastatic LUAD patients can benefit little from traditional therapy, while the rising immunotherapy or other targeted therapy maybe can function on these patients. For example, immune checkpoint blockade is a hot therapeutic strategy, such as programmed cell death protein 1/programmed cell death ligand 1 (PD-1/PD-L1) inhibitors exhibiting a favorable performance in cancer immunotherapy (Borghaei et al., 2015; Herbst et al., 2016; Jain et al., 2018). However, resistance or immune escape to immunotherapy is a common issue resulting in its inefficiency and poor outcomes. The feature of tumor microenvironment (TME) is one of the critical factors contributing to different response to immunotherapy (Binnewies et al., 2018). For example, high expression of PD-1/PD-L1 is associated with high efficiency of anti-PD-1/PD-L1 therapy (Brody et al., 2017). Cytokines and chemokines released by immune cells, neoplastic or stromal cells can orchestrate and reconstruct the immune microenvironment, and thus lead to different anti-tumor responses (Nisar et al., 2021). Cytokines such as tumor necrosis factor (TNF)-α (Laha et al., 2021), interleukin (IL) family (Sato et al., 2011; Kitamura et al., 2017), and transforming growth factor (TGF)-β (Yao et al., 2010) play an important role in angiogenesis, immune evasion, resistance to immunotherapy, and epithelial-mesenchymal transition (EMT) process responsible for tumor progression and metastasis. Consequently, TME-related features largely determine the anti-tumor response and the activated response to immunotherapy.

Bagaev et al. collected a total of 29 knowledge-based functional gene expression signatures related to TME from previous studies, and grouped them into four classes including anti-tumor microenvironment (e.g., T cells), pro-tumor microenvironment (e.g., macrophages), angiogenesis fibrosis (e.g., angiogenesis), and malignant cell properties (e.g., EMT signature) (Bagaev et al., 2021). Based on these TME-related signatures, they identified four microenvironment subtypes and comprehensively elucidate the relation between TME and melanoma by using transcriptomic and genomic data. The four microenvironment subtypes were also conserved in other cancer types, and were correlated with the response to immunotherapy. The ASLC/ATS/ERS classification is a significant improvement in the classification criteria for lung adenocarcinoma, encompassing pathology, molecular biology, radiology, oncology and clinical practice to provide better clinical diagnosis (Yoshizawa et al., 2011; Gu et al., 2013). The current molecular classification still has limitations on the prognosis evaluation of lung adenocarcinoma patients, such as, the new classification content is too complex to apply. Inspiring by the above study, we sought to explore the TME in LUAD through analyzing these TME-related signatures, and identify effective prognostic genes for guiding immunotherapy or other therapy in LUAD patients.

Materials and Methods

Data collection and data preprocessing

TCGA-LUAD dataset (abbreviated as TCGA dataset in the following) containing the RNA sequencing (RNA-seq) data and clinical information of LUAD samples was obtained from The Cancer Genome Atlas (TCGA) database through Sangerbox platform (http://vip.sangerbox.com/) (Shen et al., 2022). mRNA expression was quantified with fragments per kilobase of exon per million reads mapped (FPKM), which converted into transcripts per million (TPM). GSE72094 (Schabath et al., 2016) and GSE50081 (Der et al., 2014) datasets containing microarray data were obtained from Gene Expression Omnibus (GEO) database. In TCGA dataset, LUAD samples with survival status and survival time were included. The average expression value was used when one gene had multiple ensemble IDs. For microarray data, probes were annotated by the annotation profile of corresponding chip platform. If a gene had multiple probes, the averaged value was used. After data preprocessing, 487 LUAD samples were remained in TCGA dataset, 442 and 127 LUAD samples were remained in GSE72094 and GSE50081 datasets respectively.

Identification of TME-based molecular subtypes

A total of 29 TME signatures were obtained from the previous research (Bagaev et al., 2021), including four groups of signatures, anti-tumor microenvironment (MHCI, MHCII, coactivation molecules, cytotoxic cells, T cells, T cells trafficking, B cells, M1 signature, NK cells, Th1 signature, and anti-tumor cytokines), pro-tumor microenvironment (Treg, Treg traffic, MDSC, MDSC traffic, neutrophil signature, granulocyte traffic, macrophages, Th2 signature, macrophages/DC traffic, and pro-tumor cytokines), angiogenesis fibrosis (angiogenesis, endothelium, cancer-associated fibroblasts (CAFs), matrix, matrix remodeling), and malignant properties (proliferation rate signature and EMT).

The enrichment score of 29 TME signatures was measured by single sample gene set enrichment analysis (ssGSEA) (Hänzelmann et al., 2013). ConsensusClusterPlus R package (Wilkerson and Hayes, 2010) was utilized to construct consensus matrix based on the ssGSEA score of TME signatures in TCGA dataset. 1 - Pearson correlation was selected as distance and KM algorithm was used for repeating 500 times of bootstraps with each bootstrap having 80% samples of TCGA dataset. The optimal cluster number (k) was selected according to the cumulative distribution function (CDF) curves and consensus matrix.

For the validation of the TME-based subtyping in the external datasets (GSE72094 and GSE50081), a support vector machine (SVM) model was used (Huang et al., 2018) (LUAD samples within TCGA dataset were randomly grouped into training set and testing set with a ratio of 7: 3).

Functional enrichment analysis

Gene sets of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways were obtained from Molecular Signature Database (MSigDB), and used for GSEA by "fgsea" algorithm (Liberzon et al., 2015). GSVA R package (Hänzelmann et al., 2013) was utilized to conduct ssGSEA on hallmark pathways obtained from MSigDB and 11 oncogenic pathways (EGFR, hypoxia, NFκB, PI3K, JAK-STAT, MAPK, TGF-β, Trail, VEGF, TNF-α, and P53) obtained from the previous research (Schubert et al., 2018).

DNA methylation analysis

In order to observe the methylation difference of different subtypes, we obtained the methylation data set of HumanMethylation450 from the TCGA database, extracted the methylation signals of each sample, and completed the missing values using the KNN method. Limma was used to analyze the methylation difference of each subtype (P.val< 0.05 and |logFC| >log2 (1.1)). In addition, we annotated the methylation sites to the gene promoter region to obtain genes regulated by methylation, The Biological Pathway to Obtain Methylation Disorder of Each Subtype by Function Enrichment Analysis.

Prediction of the response to immunotherapy and chemotherapeutic drugs

The gene signatures of T cell inflamed GEP (Ayers et al., 2017), Th1/IFN-γ (Danilova et al., 2019), and cytolytic activity (Rooney et al., 2015) were obtained from previous studies. SsGSEA was conducted on these gene signatures. The estimated sensitivity of different groups to chemotherapeutic drugs was evaluated by pRRophetic R package (Geeleher et al., 2014). TIDE algorithm (Jiang et al., 2018) was implemented to analyze immunosuppressive cells and T cell status for estimating immune escape to immunotherapy. Higher TIDE score represents higher immune escape. The immune infiltration and stromal infiltration were evaluated by ESTIMTAE analysis (Yoshihara et al., 2013). CIBERSORT algorithm (Chen et al., 2018) was performed to analyze the proportion of 22 immune-related cells. IMvigor210 dataset (Balar et al., 2017) (treated by anti-PD-L1 therapy) was included to assess the effectiveness of the risk model in predicting prognosis and response to immunotherapy. Drug sensitivity data was downloaded from Genomics of Drug Sensitivity in Cancer (GDSC) database (https://www.cancerrxgene.org/) (Yang et al., 2013).

Construction and validation of a TME-based risk model

Firstly, differential analysis was performed between different subtypes through limma R package (Ritchie et al., 2015), and differentially expressed genes (DEGs) were screened under criterions of |log2 (fold change, FC)| > 1 and false discovery rate (FDR) < .05. Functional analysis of DEGs including the enrichment analysis of Gene Ontology (GO) terms and KEGG pathways was carried out by ClusterProfiler R package (Yu et al., 2012). The DEGs significantly associated with prognosis ($p < 0.05$) was screened by univariate Cox regression analysis. Least absolute shrinkage and selection operator (Lasso) regression analysis (Friedman et al., 2010) and stepwise Akaike information criterion (stepAIC) (Zhang, 2016) were conducted to compress the number of prognostic genes. The formula of risk model was defined as: risk score = Σ(βi*expi), where i represents genes, β represents Lasso coefficients, and exp represents gene expression levels.

TCGA dataset was set as the training set. GSE72094 and GSE50081 datasets were set as the validation set. Each sample obtained a risk score and the risk score was transferred to z-score. The samples were stratified into high-risk (z-score > 0) and low-risk (z-score <0) groups. The effectiveness and efficiency of the risk model was validated by Kaplan-Meier survival analysis and receiver operation characteristic (ROC) curve analysis.

Statistical analysis

Statistical analysis in this study was conducted in R software (v4.2.0). Wilcoxon test was employed to detect the difference between two groups. The difference among multiple groups was examined by Kruskal–Wallis test. Log-rank test was conducted in survival analysis. We determined $p < 0.05$ as statistically significant.

Results

TME signatures were associated with LUAD prognosis

We compared the enrichment of 29 TME signatures in normal and LUAD samples using ssGSEA. 18 TME signatures were significantly different in normal and tumor samples (Figure 1A). Stromal-related signatures such as CAFs ($p < 0.01$), matrix ($p < 0.01$), and matrix remodeling ($p < 0.0001$) were more enriched in tumor samples compared to the normal. In addition, pro-tumor signatures such as regulatory T cells (Tregs) and malignant cell properties such as tumor proliferation rate were more accumulated in tumor samples ($p < 0.0001$). Correlation analysis on these 29 TME signatures revealed evidently positive correlations among them, suggesting close interactions among these signatures (Figure 1B). We assessed the relation between the signatures and clinical characteristics, and found that MHC II and Th2 signature were positively correlated with age (Figures 1C,D). Anti-tumor signatures such as T cells, B cells, coactivation molecules, and MHC II were negatively correlated with gender, T stage, N stage, M stage, and Stage. Notably, tumor proliferation rate was significantly upregulated in N1-N3 stages and Stage III+IV (Figure 1D). In the relation of 29 TME signatures to LUAD overall survival, we found that some of them were risk factors such as tumor proliferation rate, matrix remodeling, and EMT signature and some were protective factors such as B cells, Th2 signature, T cells, and MHC II ($p < 0.05$, Figure 1E), indicating a close relation between TME signatures and LUAD prognosis.

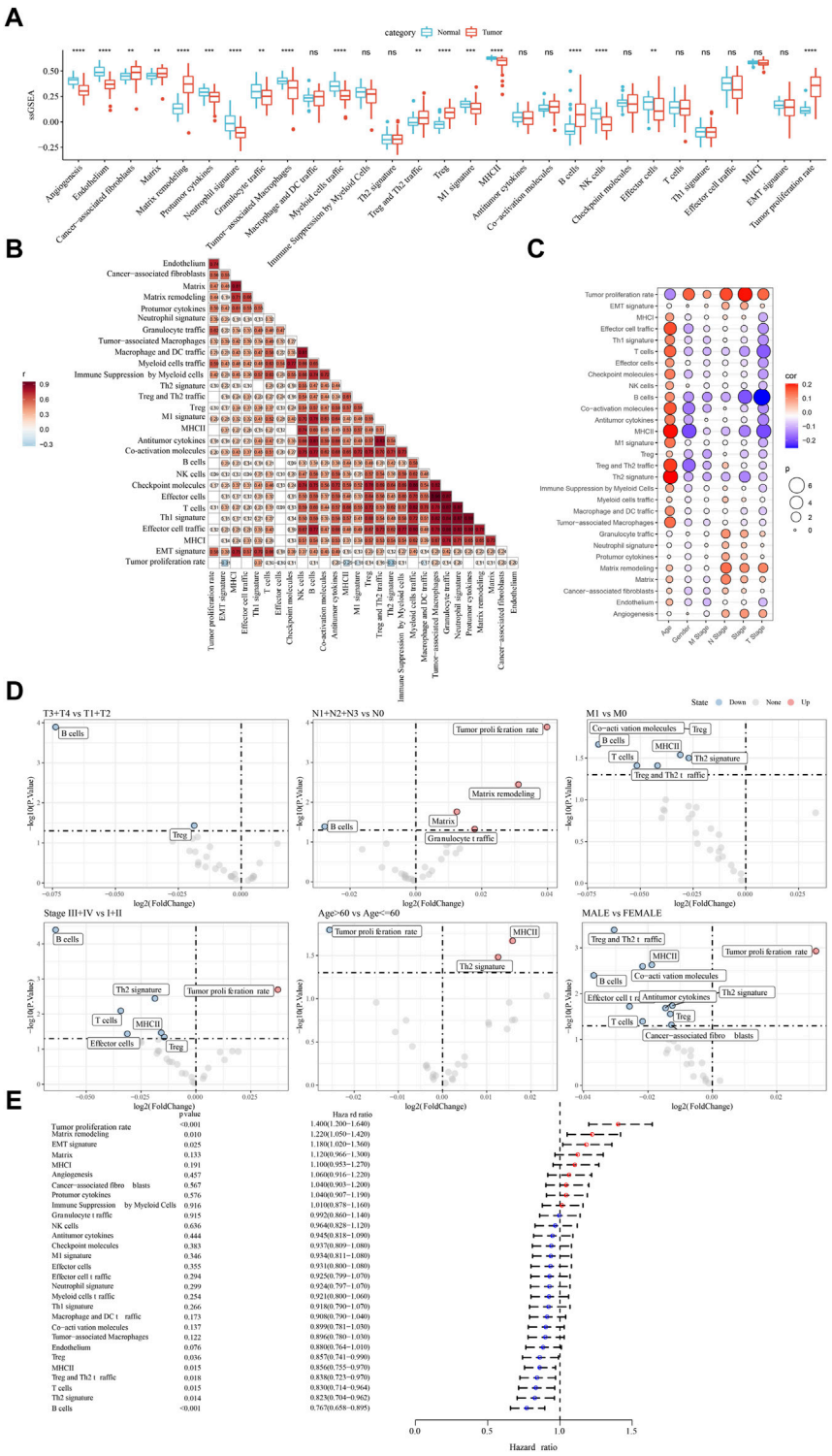


FIGURE 1
The relation between TME signatures and LUAD analyzed in TCGA dataset **(A)** The ssGSEA score of TME signatures in tumor and normal samples **(B)** Correlation analysis among TME signatures **(C)** Correlation between TME signatures and clinical characteristics **(D)** Fold change of the enrichment of TME signatures in different stages, ages, and genders **(E)** Hazard ratio of 29 TME signatures. ns, not significant. $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$.

Identification of TME-based molecular subtypes

Given that TME signatures were significantly related to LUAD prognosis, we attempted to identify molecular subtypes based on their enrichment scores. By using consensus clustering, we determined four molecular subtypes (C1, C2, C3, and C4) according to CDF and consensus matrix (Supplementary Figure S1). Four subtypes had distinct enrichment patterns of 29 TME signature as shown in the heatmap (Figure 2A). C3 and C4 subtypes had higher enrichment of

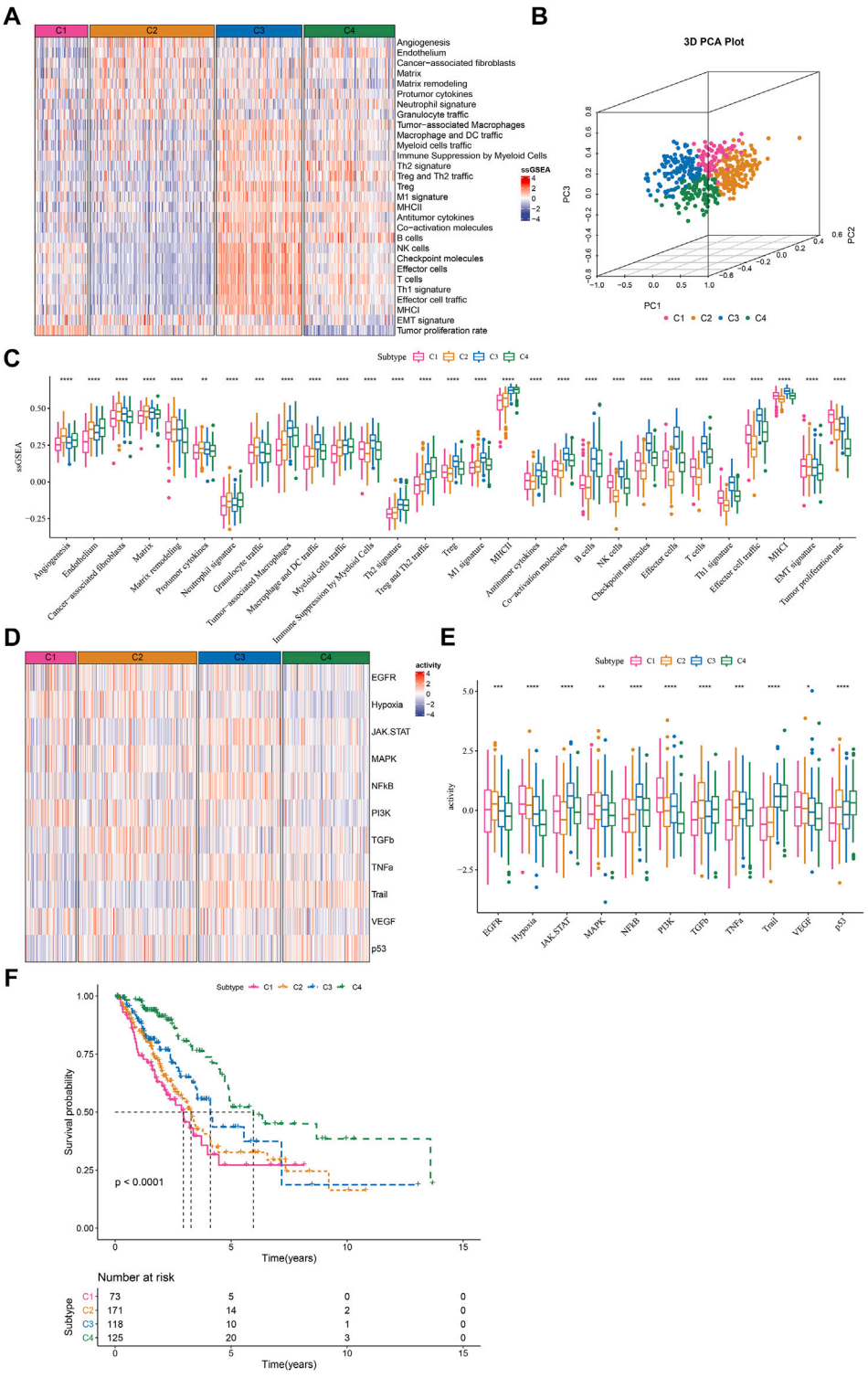


FIGURE 2
The immune and pathway difference of four TME-based molecular subtypes in TCGA dataset **(A)** The heatmap showed the enrichment of TME signatures in four subtypes **(B)** 3D PCA plot of four subtypes based on TME signatures **(C)** The ssGSEA score of TME signatures in four subtypes **(D)** The heatmap showed the enrichment of oncogenic pathways in four subtypes **(E)** The ssGSEA score of oncogenic pathways in four subtypes **(F)** Kaplan-Meier survival plot of four subtypes. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

anti-tumor signatures compared to C1 and C2. C4 subtype had evidently lower enrichment of EMT signature and tumor proliferation rate compared to C1, C2 and C3 subtypes. Principle component analysis (PCA) displayed the different distribution of four subtypes based on the TME signatures (Figure 2B). Significant differences were shown among four subtypes on the enrichment of

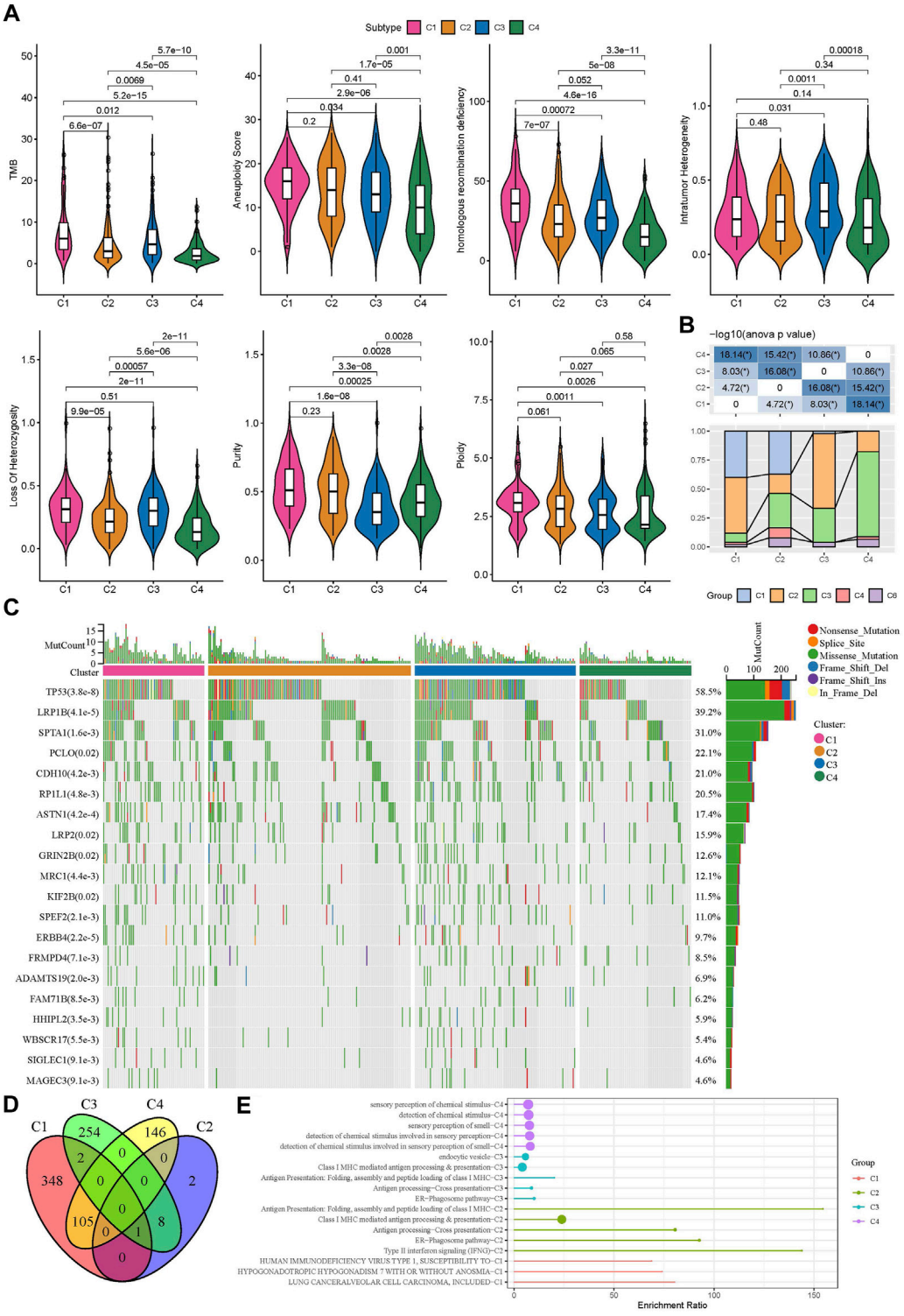


FIGURE 3
Gene mutations and genomic features of four subtypes in TCGA dataset **(A)** The score of TMB, aneuploidy, homologous recombination deficiency, intratumor heterogeneity, loss of heterozygosity, tumor purity, and ploidy in four subtypes **(B)** The distribution of previously reported immune subtypes (C1, C2, C3, C4, and C6) in TME-based subtypes **(C)** The top 20 mostly mutated genes in LUAD **(D)** Cross Venn Diagram of Different Methylation Sites of Each Subtype **(E)** Biological pathway of significant enrichment of methylation sites in different subtypes (top5).

all TME signatures (Figure 2C). In addition, we analyzed the oncogenic activity of four subtypes, and the result showed different activation of these oncogenic pathways (Figures 2D,E). PI3K and hypoxia were activated in C1, EGFR and TGF-β were activated in C2, JAK-STAT and NFκB were activated in C3, and P53 signaling was activated in C4. Moreover, survival analysis revealed that C1 subtype had the worst prognosis while C4 had the longest overall survival ($p < 0.0001$, Figure 2F). Different activation of these pathways may indicate

different TME-based molecular mechanisms of tumor progression. Furthermore, we evaluated the distribution of different clinical characteristics in four subtypes. The result exhibited an evident trend that the samples with advanced stages were more distributed in C1 (Supplementary Figure S2). Female patients and the patients with age >60 had a higher proportion in C4 compared with that in C1-C3 (Supplementary Figure S2). Not surprisingly, C1 had the largest number of the samples with dead status than other subtypes (Supplementary Figure S1).

Genomic landscape and DNA methylation of four TME-based subtypes

Genomic instability has been demonstrated to be associated with tumor development. We obtained a series of genomic characteristics of TCGA-LUAD data from a pan-cancer research (pan-cancer) (Thorsson et al., 2018). C1 subtype had relatively high scores of tumor mutation burden (TMB), aneuploidy, homologous recombination deficiency, loss of heterozygosity, purity, and ploidy, while C3 showed relatively high score of intratumor heterozygosity (Figure 3A). In the previous pan-cancer research, they identified six immune subtypes of LUAD (C1, C2, C3, C4, and C6 immune subtypes). We analyzed the distribution of previous immune subtypes in our TME-based subtypes (Figure 3B). C1 immune subtype (also known as wound healing) mostly distributed in C1 and C2 TME-based subtypes. C2 immune subtype (also known as IFN-γ dominant) mostly accumulated in C1 and C3 TME-based subtypes. C4 TME-based subtype had the highest proportion of C3 immune subtype (also known as inflammatory). The different distribution of previous immune subtypes in our TME-based subtypes also supported the distinct TME characteristics of four subtypes. In addition, we evaluated the gene mutations in four subtypes, and observed that TP53, LRP1B, and SPTA1 contributed high somatic mutation frequencies (Figure 3C). In addition, we analyzed the different DNA methylation sites of each subtype in the genome. Among them, C1 has the most differential methylation sites, C2 has only a small amount of DNA methylation differences, and the different methylation sites of each subtype overlap less (Figure 3D). Further functional analysis showed that the methylation site of C1 imbalance was mainly related to LUNG CANCERALVEOLAR CELL CARCINOMA, INCLUDED, the methylation site of C2 imbalance was mainly related to Type II interaction signaling (IFNG), the methylation site of C3 imbalance was mainly related to ER Phagosome pathway, and the methylation site of C4 imbalance was mainly related to detection of chemical stimulus involved in sensor performance of smart (Figure 3E), These results indicate that different molecular subtypes may have different apparent disorder patterns.

Four TME-based subtypes had differently activated pathways

To further understand the different molecular mechanism of tumor development in four subtypes, we analyzed the biological pathways using GSEA. Different pathways were enriched in four subtypes. In C1 subtype, cell cycle and DNA repair-related pathways were relatively activated such as mismatch repair, base excision repair, homologous recombination, and DNA replication (Figure 4A). In C2 subtype, EMT-related pathways were significantly enriched such as ECM receptor interaction, tight junction, TGF-β signaling pathway, and focal adhesion (Figure 4B). In C3 subtype, cell cycle and immune-related pathways were activated such as DNA replication, cell cycle, cytokine-cytokine receptor interaction, antigen processing and presentation, chemokine signaling pathway, and Toll-like receptor signaling pathway (Figure 4C). In C4 subtype, immune-related pathways were also evidently activated such as chemokine signaling pathway, antigen processing and presentation, cytokine-cytokine receptor interaction, and complement and coagulation cascades (Figure 4D).

Additionally, similar results were carried out in hallmark pathways (Figure 4E). Cell cycle-related pathways were much enriched in C1 subtype. C2 subtype showed activated EMT, angiogenesis, hypoxia, Notch signaling, and TGF-β signaling. Immune-related pathways were significantly enriched in both C3 and C4 subtypes. Besides, C4 subtype also had relatively high enrichment of metabolic pathways such as heme metabolism, fatty acid metabolism, adipogenesis, xenobiotic metabolism, and bile acid metabolism. The above results implied that these differently enriched pathways may result in different TME characteristics in four subtypes.

Different response of four TME-based subtypes to immunotherapy and chemotherapeutic drugs

TME characteristics can decide the outcomes of clinical treatment to some extent especially immunotherapy. We selected three immune-related signatures including T cell inflamed gene expression profiles (GEP), Th1/IFN-γ, and cytolytic activity from previous studies to evaluate the predicted response to immunotherapy. T cell inflamed GEP has been illustrated to reflect the response to immune checkpoint inhibitors (ICIs) (Ott et al., 2019). IFN-γ is an important cytokine in modulating immune response and anti-tumor activity (Danilova et al., 2019). Cytolytic activity reflects the cytotoxicity of activated T cells (Rooney et al., 2015). The above three signatures manifested differences in four subtypes, with that C3 subtype had the highest ssGSEA score of T cell inflamed GEP, IFN-γ, and cytolytic activity (Figures 5A–C). Immune checkpoints are also important in the response to ICIs. High expression of PD-1/PD-L1 indicates high response to ICIs. Analysis on immune checkpoints clarified that C3 subtype had the highest expression levels of PDCD1 (PD-1), CD274 (PD-L1), CTLA4, LAG3, PDCD1LG2, BTLA, HAVCR2, and TIGIT ($p < 0.0001$, Figure 5D), meaning that C3 subtype was predicted to have the highest sensitivity to immune checkpoint blockade treatment. Furthermore, we examined the estimated IC50 of four chemotherapeutic drugs including Docetaxel, Vinorelbine, Paclitaxel, and Cisplatin. C1 subtype had the lowest estimated IC50 of Docetaxel, Vinorelbine, and Cisplatin, and C3 subtype had the lowest IC50 of Paclitaxel ($p < 0.01$, Figure 5E). The results suggested that C1 may benefit much from the treatment of Docetaxel, Vinorelbine, and Cisplatin, and C3 may benefit much from Paclitaxel.

Validation the robustness of TME-based subtypes in two external datasets

LUAD samples in the TCGA dataset were randomly grouped into training cohort (n = 343) and test cohort (n = 144). TME score was

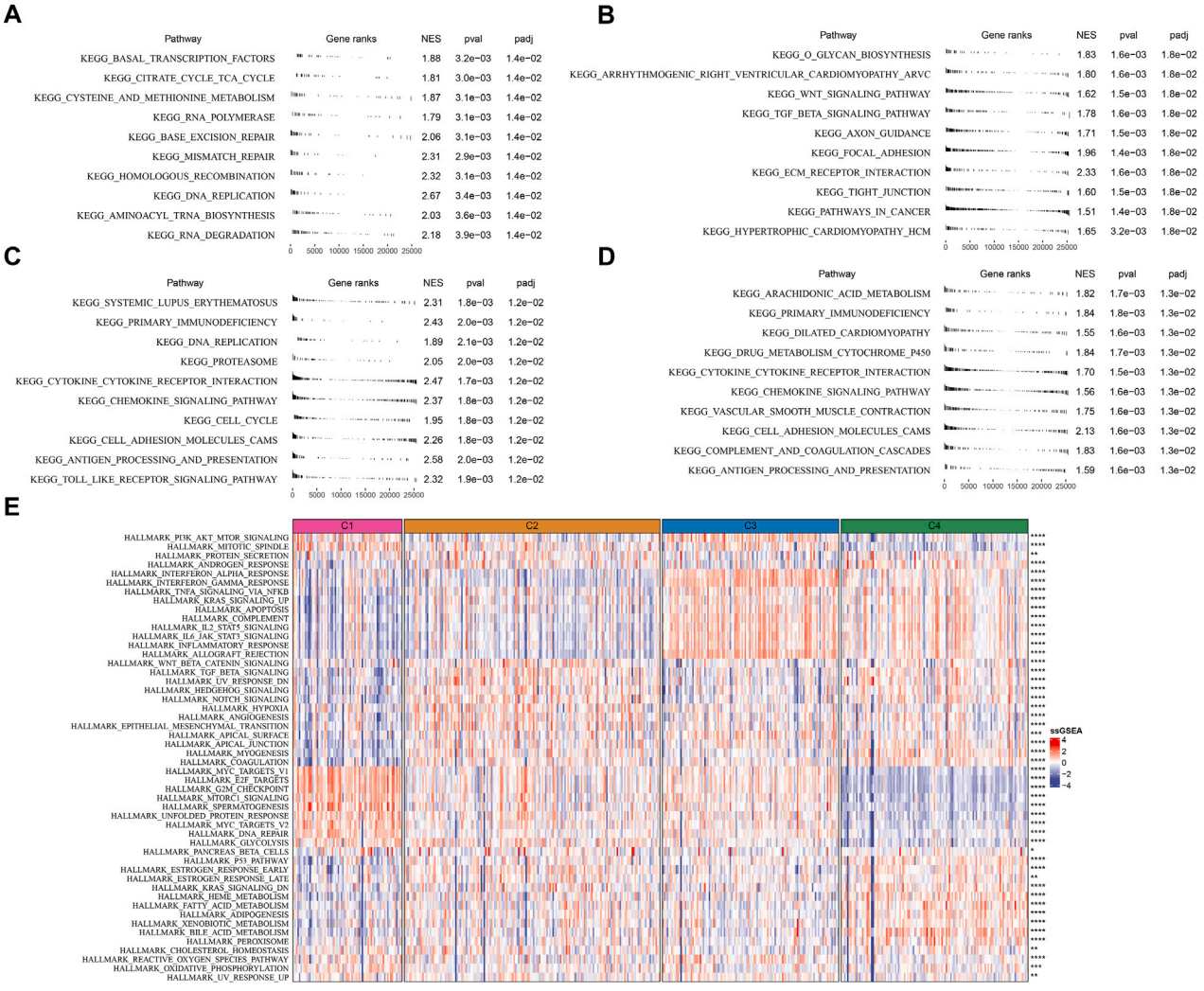


FIGURE 4
Analysis of KEGG and hallmark pathways in TCGA dataset **(A–D)** GSEA revealed the top 10 enriched KEGG pathways of C1 **(A)**, C2 **(B)**, C3 **(C)**, and C4 **(D, E)** The heatmap showed the enrichment of hallmark pathways. Comparison of the enrichment score among four subtypes was performed and the significance was shown in the right. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

inputted to SVM model for determining TME-based subtypes in the training cohort. The accuracy of the SVM model in the test cohort was 1. We then used the SVM model to examine the TME-based subtypes in the external datasets (GSE72094 and GSE50081). The external datasets showed the similar results on the TME score of four subtypes compared to the result in the TCGA dataset (Figures 6A–D). C1 subtype showed the worst prognosis and C4 subtype had the best prognosis in both two external datasets (Figures 6A,C). The enrichment patterns of 29 TME signatures in GSE72094 and GSE50081 datasets were similar to that in TCGA dataset (Figures 6B,D; Figure 2A). Tumor proliferation rate was highly enriched in C1 subtype. Anti-tumor and pro-tumor signatures were both more enriched in C3 compared with other subtypes. The ssGSEA scores of most TME signatures were different among four subtypes in two external datasets (Figures 6E,F). The validation of TME-based subtypes in the external datasets supported the reliability and robustness of the subtyping, and suggested the important role of these TME signatures in LUAD.

Development of a TME-based prognostic model

As four TME-based subtypes showed different TME scores, prognosis and activated pathways, we then identified the DEGs among four subtypes. A total of 353 DEGs (135 upregulated and 218 downregulated) were screened in C1 vs other, 91 DEGs (9 upregulated and 82 downregulated) were screened in C2 vs other, 171 DEGs (161 upregulated and 10 downregulated) were screened in C3 vs other, and 396 DEGs (223 upregulated and 173 downregulated) were screened in C4 vs other. Functional analysis on all upregulated DEGs unveiled different biological function of upregulated DEGs in different subtypes. Cell cycle-related pathways and processes were enriched in C1; stromal-related processes were enriched in C2; immune-related pathways and processes were enriched in C3 (Supplementary Figure S3).

We screened a total of 648 DEGs among four subtypes in TCGA dataset after removing the duplicate DEGs. Univariate Cox regression analysis detected a total of 164 risk genes and 254 protective genes within 648 DEGs ($p < 0.05$, Supplementary Table S1). Next, we applied

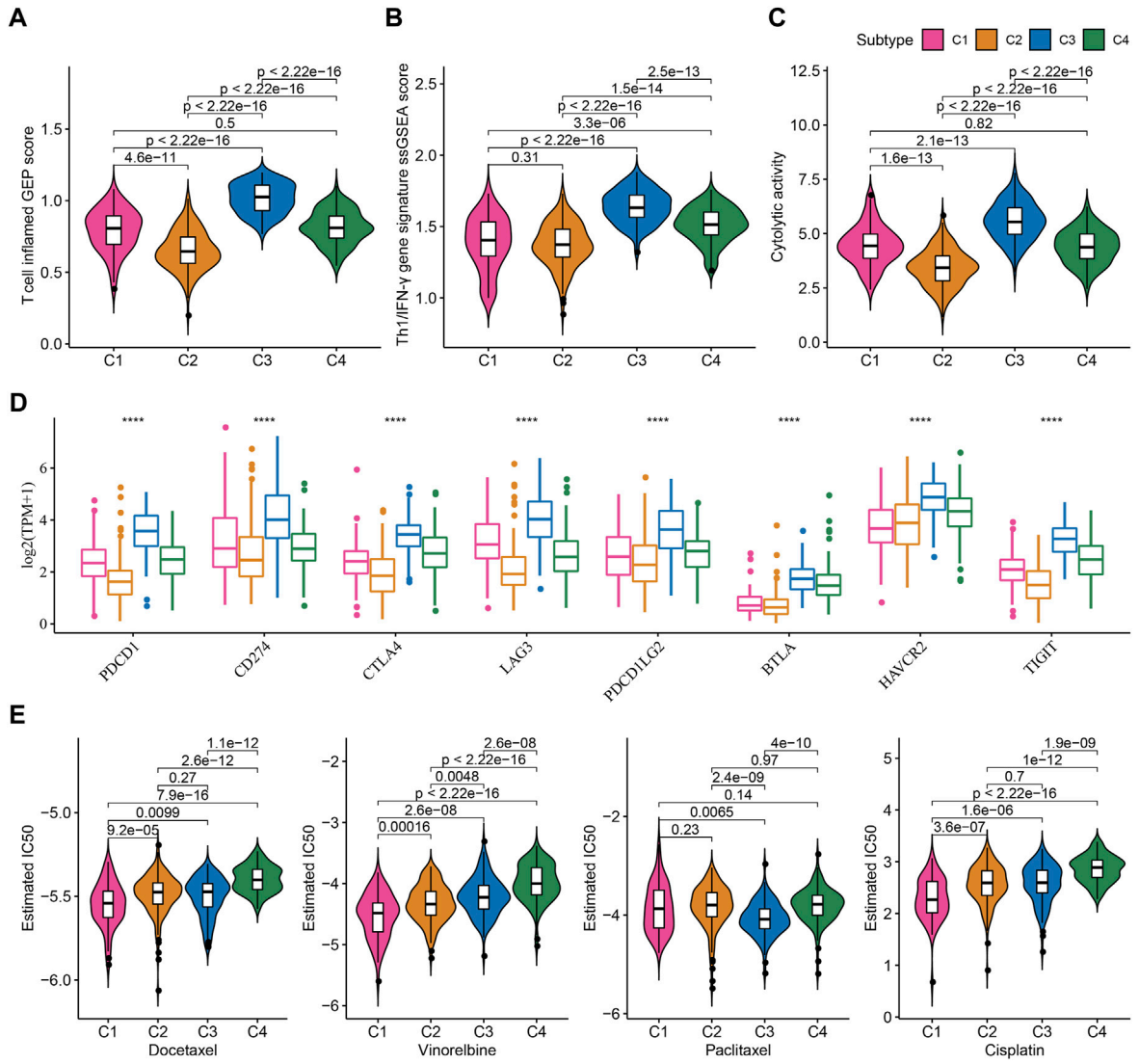


FIGURE 5
Prediction of the response to immunotherapy and chemotherapeutic drugs in TCGA dataset **(A–C)** The ssGSEA score of T cell inflamed GEP, Th1/IFN-γ, and cytolytic activity in four subtypes **(D)** The expression of immune checkpoint genes in four subtypes **(E)** The estimated IC50 of four chemotherapeutic drugs in four subtypes. ****$p < 0.0001$.

Lasso regression and stepAIC to dig out key prognostic genes from the above risk and protective genes. Lasso analysis compressed the coefficients to zero and remained nine prognostic genes when the lambda value = 0.0689 (Figures 7A,B). Subsequently, stepAIC was performed on the nine prognostic genes and further compress the number of genes. Consequently, five prognostic genes were remained, including PTTG1, MS4A1, ZNF750, RHOV, and KRT6A. The 5-gene prognostic model was determined as: Risk Score = 0.206*PTTG1—0.155*MS4A1—0.12*ZNF750 + 0.136*RHOV + 0.05*KRT6A.

Examination the performance of the 5-gene prognostic model

Each LUAD sample obtained a risk score, and was stratified into high-risk and low-risk groups referring z-score = 0 as a cut-off (Figure 7D). Two risk groups exhibited different enrichment pattern of survival status, with a higher density of dead samples in high-risk groups. Five prognostic genes showed distinct expression patterns in two risk groups, where ZNF570 and MS4A1 were relatively upregulated in low-risk group, while RHOV, PTTG1, and KRT6A were upregulated in high-risk group. There were 242 and 245 LUAD samples in high- and low-risk groups respectively, and two groups exhibited distinct prognosis ($p < 0.0001$, Figure 7E). ROC curve analysis revealed that the risk model had favorable AUC in predicting survival at 1, 3, and 5 years with the scores of 0.75, 0.72, and 0.67 respectively (Figure 7F). Furthermore, we validated the 5-gene risk model in two independent datasets (GSE72094 and GSE50081). The validation results were consistent with the TCGA dataset (Figures 7G–J), which demonstrated the effectiveness and reliability of the risk model.

The relation of risk score to different clinical characteristics was assessed and a trend showed that the risk score was higher in the advanced stages compared with early stages (Supplementary Figure S4). A significant difference was also observed in different ages and

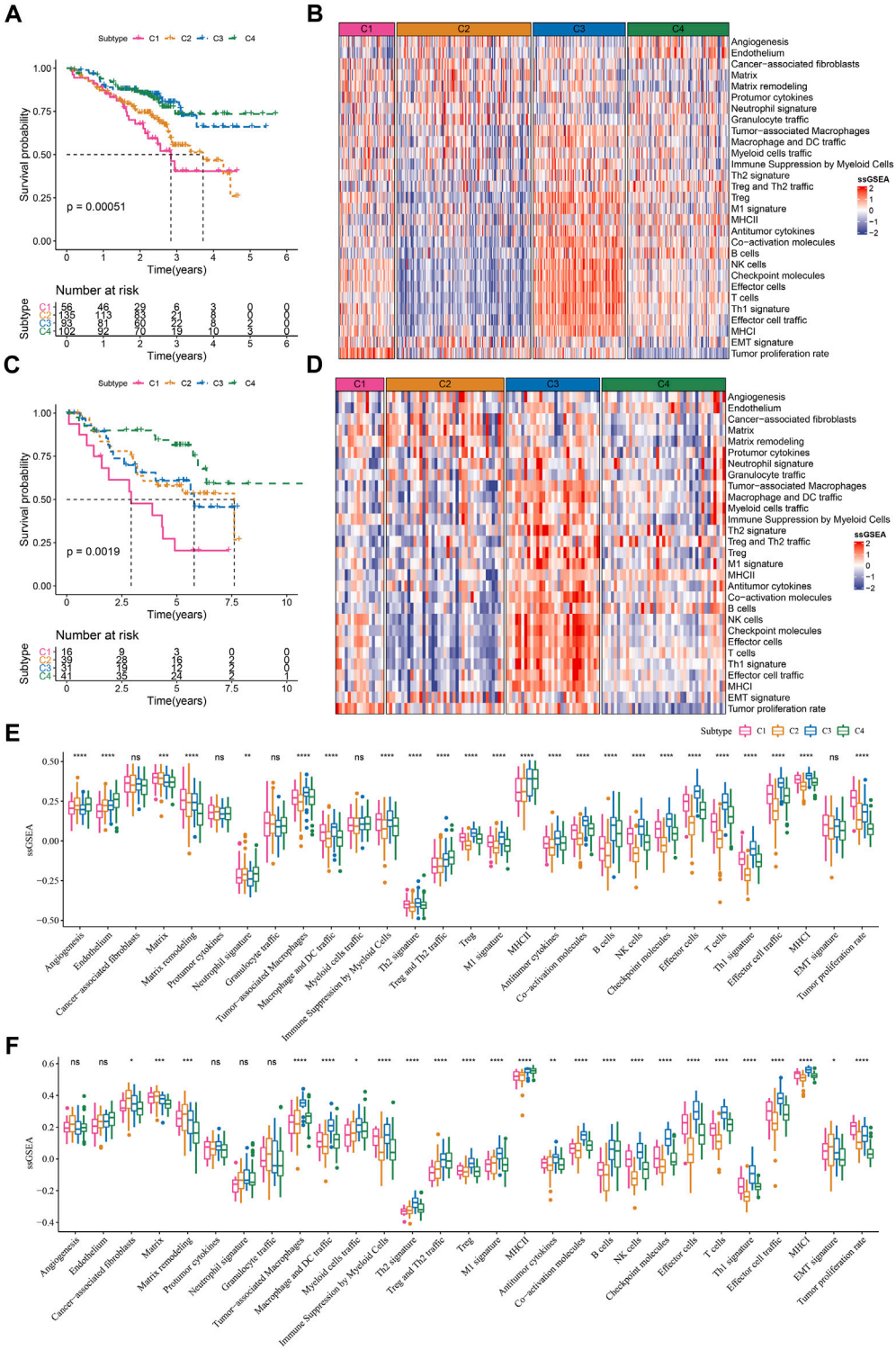


FIGURE 6
Validation of TME-based molecular subtypes in two external datasets **(A)** Survival plot of four subtypes in GSE72094 dataset **(B)** The heatmap showed the enrichment of TME signatures in GSE72094 dataset **(C)** Survival plot of four subtypes in GSE50081 dataset **(D)** The heatmap showed the enrichment of TME signatures in GSE50081 dataset **(E–F)** The ssGSEA score of 29 TME signatures in four subtypes in GSE72094 and GSE50081 datasets. ns, not significant. *p < .05, **p < 0.01, ***p < 0.001, ****p < 0.0001.

genders. Moreover, we compared the risk score in four TME-based subtypes, and the results showed that the prognosis of subtypes was consistent with their risk levels. C1 with the worst prognosis showed the highest risk score, which was consistent with the previous results (Supplementary Figure S4; Supplementary Figure S2). In different clinical characteristics, the risk model also showed a favorable performance in dividing samples into high-risk and low-risk groups (Supplementary Figure S4).

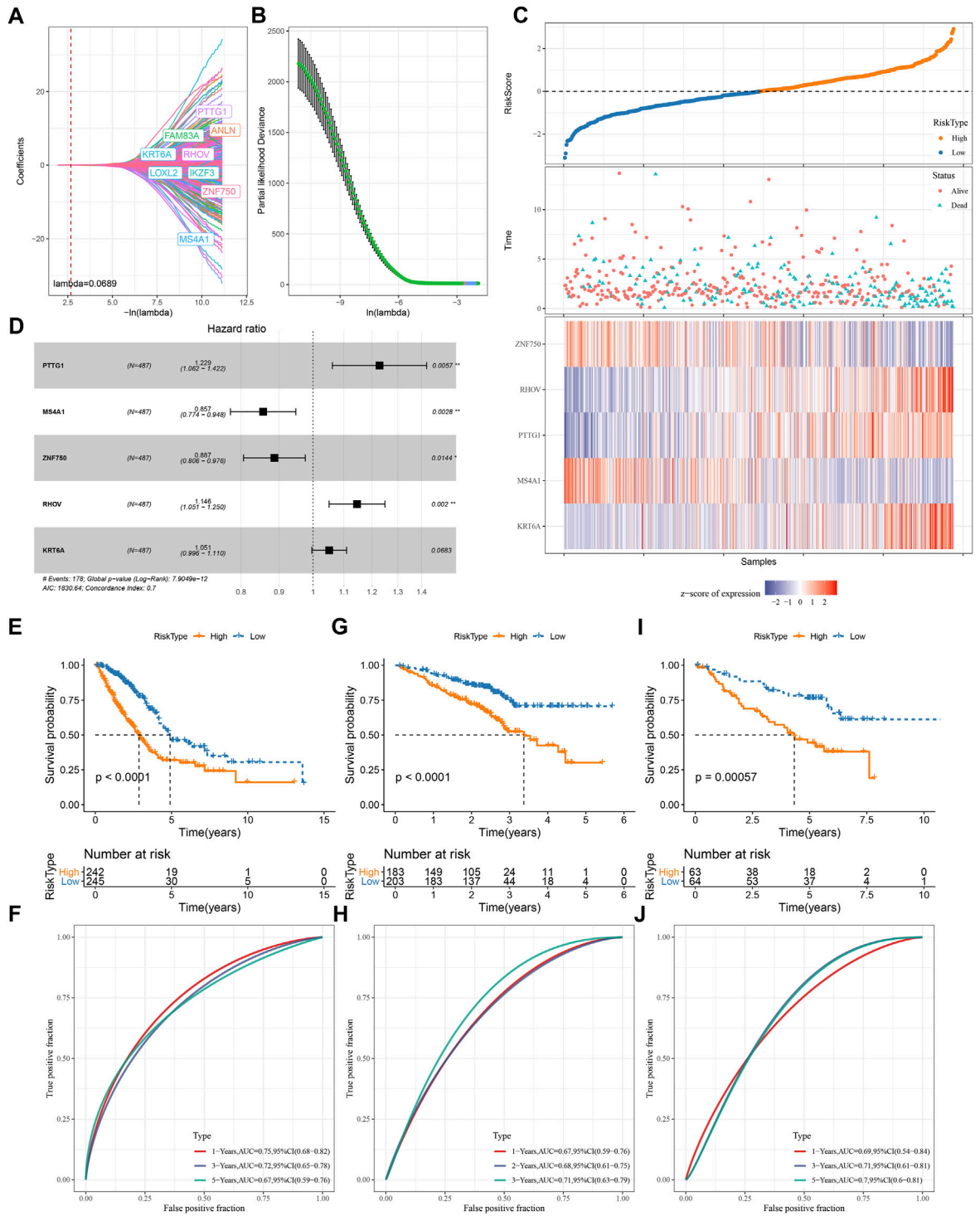


FIGURE 7
Construction and validation of a TME-based risk model **(A–B)** Lasso regression analysis of prognostic genes. The coefficients changed with increasing lambda value **(A)**. Partial likelihood deviance under different lambda values **(B)**. When lambda = 0.0689 (red dotted line in A and red dot in B), the model reached the optimal **(C)** The forest plot of the final five prognostic genes in the risk model **(D)** The risk score, survival status and expression of five prognostic genes of tumor samples in TCGA dataset **(E–I)** Survival plot of high-risk and low-risk groups in TCGA **(E)**, GSE72094 **(G)**, GSE50081 **(I)** datasets **(F–J)** ROC curve of the risk model in predicting survival in TCGA **(F)**, GSE72094 **(H)**, GSE50081 **(J)** datasets.

Biological pathways and immune characteristics of two risk groups

Next, we compared the difference of two risk groups in biological pathways and immune characteristics. GSEA on KEGG pathways revealed that high-risk group had relatively activated pathways of purine metabolism, pyrimidine metabolism, citrate cycle TCA cycle, oxidative phosphorylation, and DNA replication (Figure 8A). Consider that immunity is related to the tumor, thus we determine

10.3389/fphar.2023.1099927

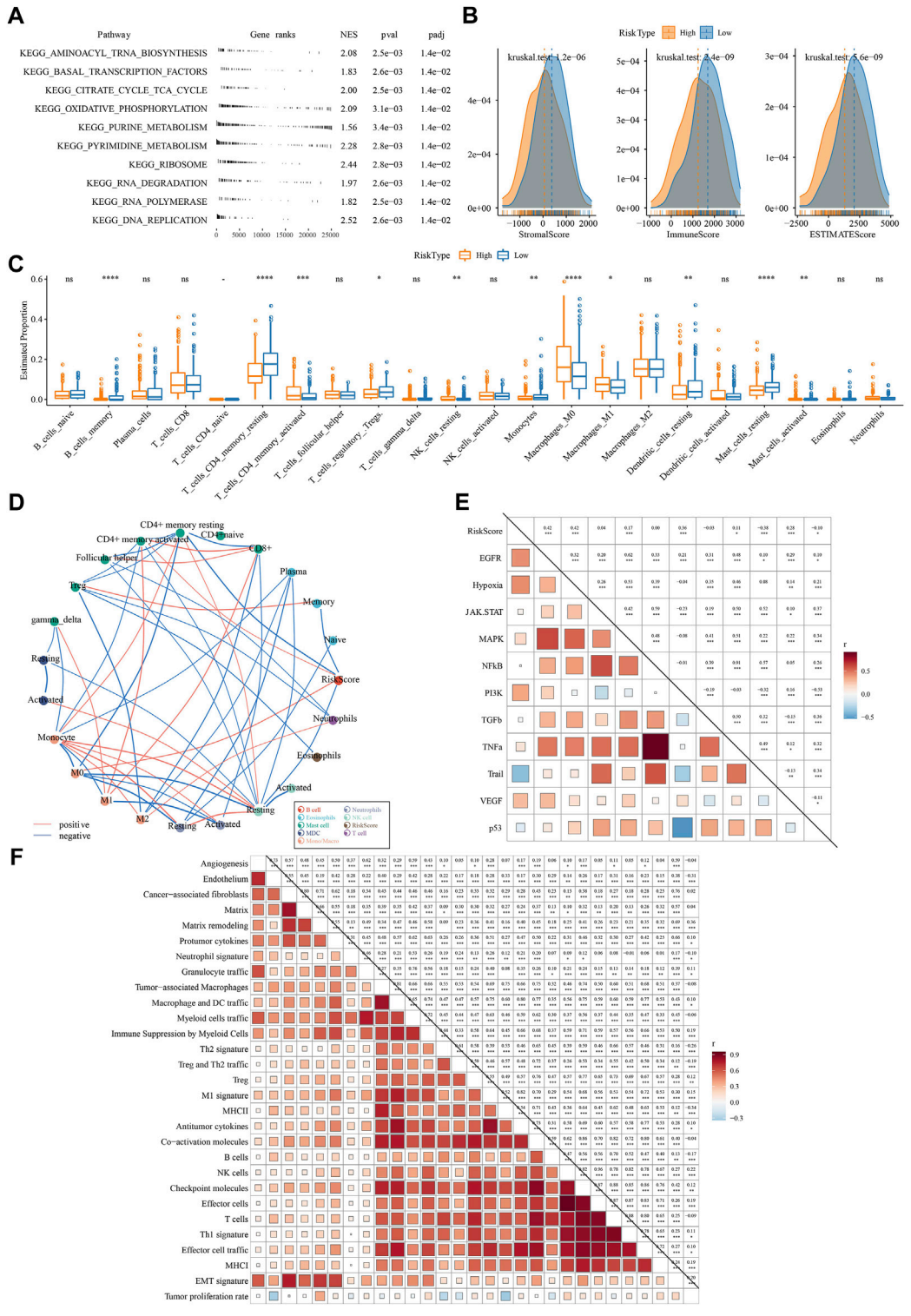


FIGURE 8
The difference of two risk groups on biological pathways and immune characteristics analyzed in TCGA dataset **(A)** GSEA result showed the significantly enriched KEGG pathways in high-risk group **(B)** The immune score, stromal score and ESTIMATE score calculated by ESTIMATE analysis **(C)** The estimated proportion of 22 immune-related cells analyzed by CIBERSORT **(D)** Correlation analysis among immune cells and risk score. Red and blue lines indicate the positive and negative correlations respectively. The thicker line indicates the stronger correlation **(E)** Correlation analysis among oncogenic pathways and risk score. Red and blue indicate positive and negative correlations respectively. The darker color indicates the stronger correlation. **(F)** the correlation analysis between risk score and 29 TME pathways.

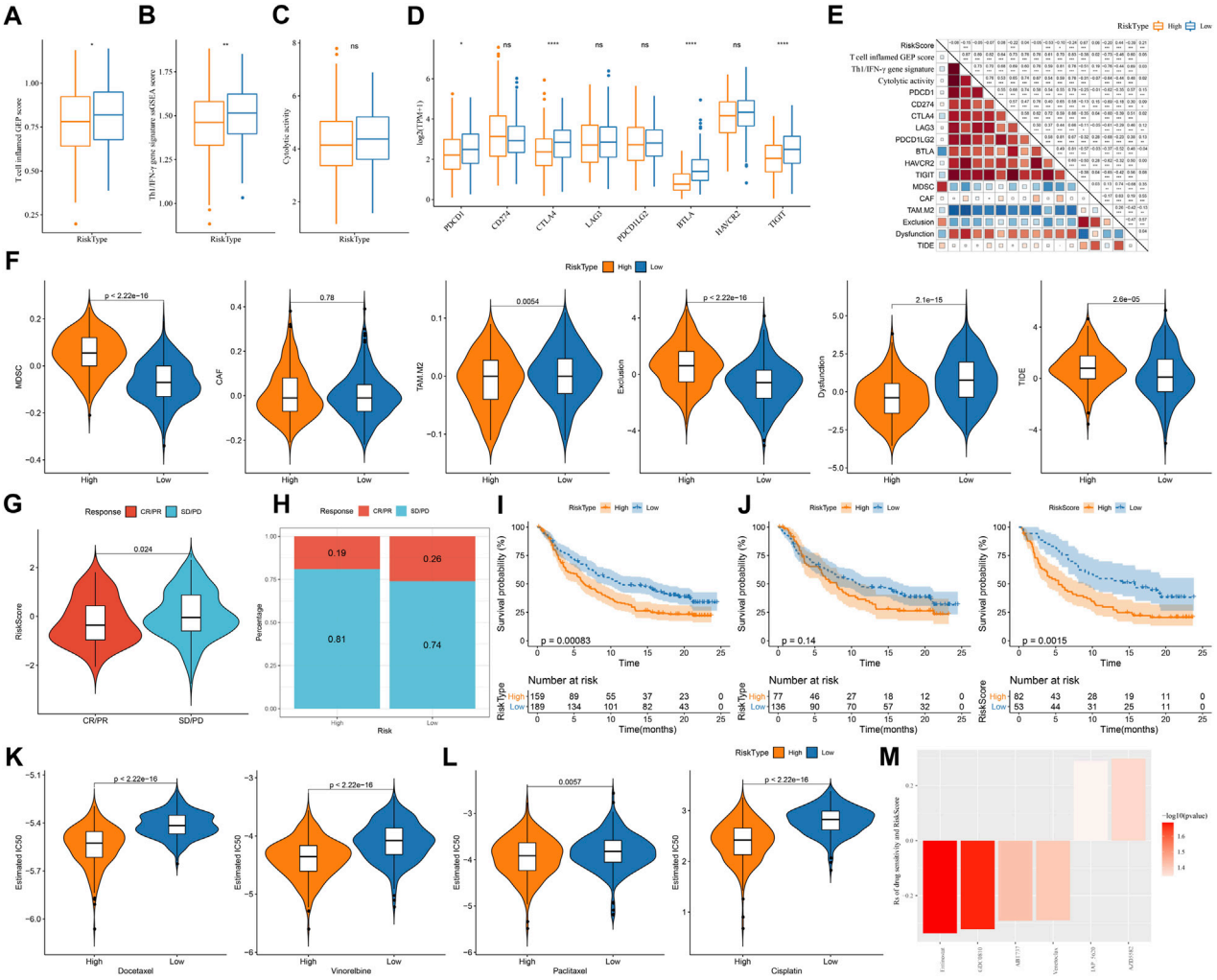


FIGURE 9
The predicted response to immunotherapy and chemotherapeutic drugs analyzed in TCGA and IMvigor210 datasets **(A–C)** The score of T cell inflamed GEP, Th1/IFN-γ signature, and cytolytic activity in two risk groups **(D)** The expression of immune checkpoint genes in two risk groups **(E)** TIDE analysis calculated the enrichment of immunosuppressive cells, T cell exclusion, T cell dysfunction, and TIDE score in two risk groups **(F)** Correlation of risk score with immunotherapy-related indicators **(G)** The risk score in CR/PR and SD/PD groups in IMvigor210 dataset **(H)** The proportion of CR/PR and SD/PD in two risk groups in IMvigor210 dataset **(I–K)** Survival plot of high-risk and low-risk groups in all stages **(I)**, early stage (I+II) **(J)**, and late stage (III+IV) **(K)** in IMvigor210 dataset **(L)** The estimated IC50 of four chemotherapeutic drugs in two risk groups **(M)** Drug sensitivity of potential drugs in the relation to the risk score. Rs < 0 indicates drug sensitivity and Rs > 0 indicates drug resistance.

immune characteristics. In terms of tumor microenvironment, two risk groups had distinguished infiltration levels that low-risk group had higher infiltration of both immune cells and stromal cells than high-risk group (Figure 8B). Of 22 immune cells, 11 immune cells were differentially distributed in two risk groups (Figure 8C). Low-risk group had higher infiltration of resting dendritic cells, memory B cells, resting memory CD4 T cells, and resting mast cells than high-risk group, while M0 macrophages and M1 macrophages were lower enriched in low-risk group. Supportively, risk score was significantly correlated with resting memory CD4 T cells and M0 macrophages (Figure 8D). In the relation of risk score with oncogenic pathways, EGFR, hypoxia, PI3K, and VEGF pathways were positively correlated with risk score (R = 0.42, 0.42, 0.36, and 0.28, respectively) (Figure 8E), suggesting that these pathways may be highly involved in the TME modulation and tumor progression. Moreover, risk score was strongly associated to 29 TME pathways (Figure 8F).

Different response of two risk groups to immunotherapy and chemotherapeutic drugs

Assessment on immunotherapy-related indicators unveiled that low-risk group had higher score of T cell inflamed GEP, Th1/IFN-γ, and cytolytic activity (Figures 9A–C), indicating a higher response of low-risk group to immunotherapy than high-risk group. Immune checkpoint analysis showed PD-1, CTLA-4, BTLA, and TIGIT were higher expressed in low-risk group than that in high-risk group (Figure 9D), suggesting that low-risk group was more responsive to immune checkpoint inhibitors. TIDE analysis predicted that high-risk

group was more prone to escape from immunotherapy, which may result from its high enrichment of myeloid-derived suppressor cells (MDSCs) and high T cell exclusion (Figure 9E). Correlation analysis of risk score with the above immunotherapy-related indicators showed that risk score was positively correlated with MDSC (R = .67), T cell exclusion (R = .44), and TIDE score (R = .21) but was negatively correlated with BTLA (R = -0.53) and T cell dysfunction (R = -0.39) (Figure 9F). Moreover, risk score was positively correlated to TMB (Supplementary Figure S5), and high group had enhanced TMB (Supplementary Figure S5).

Furthermore, we used an immunotherapy dataset (IMvigor210) to validate the reliability of the risk model in predicting immune response. Risk score was significantly higher in SD/PD group compared with that in CR/PR group (p = 0.024, Figure 9G), and low-risk group also had a higher proportion of CR/PR (p = 0.2325302, Figure 9H). The risk model was also effective to distinguish high-risk patients receiving immunotherapy in IMvigor210 dataset, especially in the patients with late stages (Figures 9I–K). In addition, we evaluated the ability of the risk model in predicting the response to chemotherapeutic drugs in TCGA dataset. High-risk group had significantly lower estimated IC50 of all four chemotherapeutic drugs (Figure 9L), implying that high-risk group was more sensitive to these four drugs than low-risk group. By utilizing the drug sensitivity data in GDSC database, we identified six drugs significantly correlating with risk score where four drugs (entinostat, GDC0810, ABT737, and venetoclax) may serve as therapeutic drugs for LUAD (Figure 9M).

Discussion

In the present study, we used 29 TME-related signatures as a basis to identify TME-based molecular subtypes for LUAD. Four TME-based subtypes were identified and their clinical and molecular features such as survival time, gene mutations, genomic characteristics, immune infiltration, and biological pathways were characterized. Four subtypes showed distinct clinical and molecular features, as well as different response to immunotherapy and chemotherapeutic drugs. By comparing the expression profiles between different subtypes, we identified DEGs and screened five key prognostic genes to construct a TME-related risk model for predicting LUAD prognosis.

Among 29 TME-related signatures, tumor proliferation rate, EMT signature, and matrix remodeling were shown to be positively correlated with poor prognosis. EMT has been widely known as a promotive process in inducing tumor cell invasion and metastasis through weakening cell-cell adhesion (Ye and Weinberg, 2015). The junctions of mesenchymal cells with extracellular matrix are loose, which enable tumor cells easily to migrate. Tumor proliferation rate had the highest HR (1.40) among these signatures, in accordant with the close relation between tumor proliferate rate and stage. Evidently positive correlations were observed among 29 TME-related signatures, suggesting a complicated regulation system of TME. Therefore, we used these TME-related signatures as a basis to perform molecular subtyping for LUAD patients.

We identified four TME-based molecular subtypes and each subtype showed different enrichment patterns of TME-related signatures. C1 subtype had the highest enrichment of tumor proliferation rate, which was considered as a pro-tumor phenotype. C2 subtype had the least infiltration of anti-tumor immune cells or molecules, and relatively high enrichment of angiogenesis (Voron et al., 2014), CAFs (Ziani et al., 2018), and pro-tumor cytokines, which was suggested as an immune-suppressed phenotype. C3 subtype had the highest enrichment of anti-tumor cells but the immunosuppressive cells or signatures such as tumor-associated macrophages (Pan et al., 2020), Treg (Tanaka and Sakaguchi, 2017), and checkpoint molecules were also highly enriched. C3 subtype was considered as an immune infiltrated phenotype. C4 subtype had the lowest enrichment of tumor proliferation rate and EMT signature. Therefore, we suggested C4 subtype as a tumor-silent phenotype. Survival analysis of four subtypes showed that C1 had the worst survival and C4 had the longest survival, which was consistent with their TME-related features.

Genomic instability is an important feature and is considered as a hallmark in cancers (Negrini et al., 2010). The mutation of oncogenes promotes DNA damage and genomic arrangements in cancer (Tubbs and Nussenzweig, 2017). High non-synonymous TMB was demonstrated to be associated with favorable prognosis in resected non-small cell lung cancer patients (Devarakonda et al., 2018). In our results, C1 had the highest score of aneuploidy, homologous recombination deficiency, loss of heterozygosity and ploidy, indicating high genomic instability thus contributing to poor prognosis of C1. Although high TMB was also shown in C1, the large number of genomic alterations covered the beneficial effect of TMB. The contribution of TME in genomic instability has been revealed in recent years, and hypoxia is a main factor causing DNA damage and genomic instability (Sonugür and Akbulut, 2019). In pathway analysis, we found that two hypoxia-related pathways, reactive oxygen species pathway and oxidative phosphorylation, were relatively activated in C1, which supported the above observation.

Biological pathway analysis revealed that four TME-based subtypes had different activated pathways that may lead to their different outcomes. In C1 subtype, cell cycle-related pathways such as E2F targets, G2M checkpoint, MYC targets, and DNA repair were strikingly enriched, while immune response-related pathways were relatively inhibited, which was consistent with high tumor proliferation rate of C1. The crosstalk among activated cell cycle pathways, genomic instability and oxidative stress promoted the tumor progression and thus led to unfavorable outcome in C1 subtype. Oncogenic pathways such as WNT, TGF-β, Notch, Hedgehog, angiogenesis, hypoxia, and EMT were more enriched in C2 subtype compared with other subtypes. Lines of evidence have verified the role of these oncogenic pathways in the regulation of TME and response to immunotherapy (Yang et al., 2010; Albini et al., 2018; Meurette and Mehlen, 2018; Patel et al., 2019; Gampala and Yang, 2021). Immune response pathways such as interferon response, IL2-STAT5 signaling, complement, IL6-JAK-STAT3 signaling, and inflammatory response were much enriched in C3 and C4 subtypes, which were responsible for their favorable prognosis. In addition to immune response pathways, metabolic pathways such as fatty acid metabolism, adipogenesis, and bile acid metabolism were also enriched in C4 subtype. The metabolic alterations have been demonstrated to shape TME components thereby influencing tumor progression and immunotherapy efficiency (Lyssiotis and Kimmelman, 2017; DeBerardinis, 2020).

Cancer patients with different TME may have different response to immunotherapy. We estimated the potential response of four subtypes to immunotherapy by using immunotherapy-related indicators (T cell inflamed GEP, Th1/IFN-γ, and cytolytic activity).

Four subtypes showed distinct enrichment of these indicators where C3 subtype was predicted to benefit most from immunotherapy. In addition, four subtypes also displayed differential expression of key immune checkpoints such as PD-1, CD274, CTLA-4, and LAG3. C3 subtype exhibited the highest expression of these checkpoints, suggesting C3 subtype was sensitive to ICIs. From the above results, we concluded that the TME-based subtyping was effective to provide a guidance for LUAD patients receiving immunotherapy.

Furthermore, we established a TME-related risk model containing five prognostic genes (PTTG1, MS4A1, ZNF750, RHOV, and KRT6A) for predicting LUAD survival. PTTG1 was found to promote lung cancer migration and invasion (Li et al., 2013), and knockdown of PTTG1 could enhance anti-tumor activity in LUAD (Chen et al., 2021). The expression of MS4A1 was shown to be positively correlated with the survival of colorectal carcinoma (Mudd et al., 2021), which was consistent with our result that MS4A1 was highly expressed in low-risk group. ZNF750 is a tumor suppressor in squamous cell carcinoma, which can suppress cell migration (Hazawa et al., 2017). In our study, ZNF750 expression level was downregulated in high-risk group, which supported its protective role in inhibiting tumor progression. RHOV was shown to facilitate tumor cell growth and metastasis in LUAD (Zhang et al., 2021). In our results, RHOV was evidently elevated in high-risk group. Overexpression of KRT6A was able to promote LUAD cell proliferation through EMT process (Yang et al., 2020), which may lead to poor prognosis in high-risk group. Previous studies have illustrated that the five prognostic genes are involved in tumor progression and migration in lung cancer or other cancer types, implying that our TME-related risk model was reliable to predict LUAD prognosis. ROC curve analysis showed a high AUC and validated the efficiency of the risk model. In addition, we evaluated the predictive value of the risk model in guiding immunotherapy and chemotherapy. Two risk groups showed differential immune responses to immunotherapy and differential IC50 to four chemotherapeutic drugs (docetaxel, vinorelbine, paclitaxel, and cisplatin), which illustrated that the risk model also had a potential in assisting the decision-makings in immunotherapy and chemotherapy.

Conclusion

In conclusion, our study revealed the molecular characteristics of LUAD patients based on TME-related signatures. The distinct biological pathways and TME features of four TME-based subtypes laid a foundation for the further exploration of the crosstalk among TME, genomic instability, and oncogenic pathways in LUAD. The TME-related risk model was efficient and reliable to predict LUAD prognosis and assist clinical treatment.

Data availability statement

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

Author contributions

DW and HH conducted statistical analyses of the data and prepared the draft manuscript YJ, DA, JW, and ML edited the manuscript AG and ME provide critical comments to the manuscript. All authors checked and proofread the final version of the manuscript.

Funding

The present study was supported by the National Natural Science Foundation of China (82072027); Heilongjiang Postdoctoral Fund (LBH-Z18120); Heilongjiang Applied Technology Research and Development Program Project (GA20C009); Haiyan Medical Research Foundation (JJZD 2021-03); Beijing Medical Award Foundation (YXT2-2021-0166-0183).

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2023.1099927/full#supplementary-material

References

Albini, A., Bruno, A., Noonan, D. M., and Mortara, L. (2018). Contribution to tumor angiogenesis from innate immune cells within the tumor microenvironment: Implications for immunotherapy. *Front. Immunol.* 9, 527. doi:10.3389/fimmu.2018.00527

Ayers, M., Lunceford, J., Nebozhyn, M., Murphy, E., Loboda, A., Kaufman, D. R., et al. (2017). IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. investigation* 127 (8), 2930–2940. doi:10.1172/JCI91190

Bagaev, A., Kotlov, N., Nomie, K., Svekolkin, V., Gafurov, A., Isaeva, O., et al. (2021). Conserved pan-cancer microenvironment subtypes predict response to immunotherapy. *Cancer Cell* 39 (6), 845–865.e7. e7. doi:10.1016/j.ccell.2021.04.014

Balar, A. V., Galsky, M. D., Rosenberg, J. E., Powles, T., Petrylak, D. P., Bellmunt, J., et al. (2017). Atezolizumab as first-line treatment in cisplatin-ineligible patients with locally advanced and metastatic urothelial carcinoma: A single-arm, multicentre, phase 2 trial. *Lancet (London, Engl.* 389 (10064), 67–76. doi:10.1016/S0140-6736(16)32455-2

Binnewies, M., Roberts, E. W., Kersten, K., Chan, V., Fearon, D. F., Merad, M., et al. (2018). Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* 24 (5), 541–550. doi:10.1038/s41591-018-0014-x

Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D. R., Steins, M., Ready, N. E., et al. (2015). Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N. Engl. J. Med.* 373 (17), 1627–1639. doi:10.1056/NEJMoa1507643

Brody, R., Zhang, Y., Ballas, M., Siddiqui, M. K., Gupta, P., Barker, C., et al. (2017). PD-L1 expression in advanced NSCLC: Insights into risk stratification and treatment selection from a systematic literature review. *Lung cancer (Amsterdam, Neth.* 112, 200–215. doi:10.1016/j.lungcan.2017.08.005

Calvayrac, O., Pradines, A., Pons, E., Mazières, J., and Guibert, N. (2017). Molecular biomarkers for lung adenocarcinoma. *Eur. Respir. J.* 49 (4), 1601734. doi:10.1183/13993003.01734-2016

Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol. Clift. NJ)* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Chen, Z., Cao, K., Hou, Y., Lu, F., Li, L., Wang, L., et al. (2021). PTTG1 knockdown enhances radiation-induced antitumour immunity in lung adenocarcinoma. *Life Sci.* 277, 119594. doi:10.1016/j.lfs.2021.119594

Danilova, L., Ho, W. J., Zhu, Q., Vithayathil, T., De Jesus-Acosta, A., Azad, N. S., et al. (2019). Programmed cell death ligand-1 (PD-L1) and CD8 expression profiling identify an immunologic subtype of pancreatic ductal adenocarcinomas with favorable survival. *Cancer Immunol. Res.* 7 (6), 886–895. doi:10.1158/2326-6066.CIR-18-0822

Daugaard, I., Dominguez, D., Kjeldsen, T. E., Kristensen, L. S., Hager, H., Wojdacz, T. K., et al. (2016). Identification and validation of candidate epigenetic biomarkers in lung adenocarcinoma. *Sci. Rep.* 6, 35807. doi:10.1038/srep35807

DeBerardinis, R. J. (2020). Tumor microenvironment, metabolism, and immunotherapy. *N. Engl. J. Med.* 382 (9), 869–871. doi:10.1056/NEJMcibr1914890

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

Devarakonda, S., Morgensztern, D., and Govindan, R. (2015). Genomic alterations in lung adenocarcinoma. *Lancet Oncol.* 16 (7), e342–e351. doi:10.1016/S1470-2045(15)00077-7

Devarakonda, S., Rotolo, F., Tsao, M. S., Lanc, I., Brambilla, E., Masood, A., et al. (2018). Tumor mutation burden as a biomarker in resected non-small-cell lung cancer. *J. Clin. Oncol.* 36 (30), 2995–3006. doi:10.1200/JCO.2018.78.1963

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33 (1), 1–22. doi:10.18637/jss.v033.i01

Gampala, S., and Yang, J. Y. (2021). Hedgehog pathway inhibitors against tumor microenvironment. *Cells* 10 (11), 3135. doi:10.3390/cells10113135

Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS one* 9 (9), e107468. doi:10.1371/journal.pone.0107468

Gibelin, C., and Couraud, S. (2016). Somatic alterations in lung cancer: Do environmental factors matter? *Lung cancer (Amsterdam, Neth.* 100, 45–52. doi:10.1016/j.lungcan.2016.07.015

Gu, J., Lu, C., Guo, J., Chen, L., Chu, Y., Ji, Y., et al. (2013). Prognostic significance of the IASLC/ATS/ERS classification in Chinese patients-A single institution retrospective study of 292 lung adenocarcinoma. *J. Surg. Oncol.* 107 (5), 474–480. doi:10.1002/jso.23259

Hänzelmann, 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

Hazawa, M., Lin, D. C., Handral, H., Xu, L., Chen, Y., Jiang, Y. Y., et al. (2017). ZNF750 is a lineage-specific tumour suppressor in squamous cell carcinoma. *Oncogene* 36 (16), 2243–2254. doi:10.1038/onc.2016.377

Herbst, R. S., Baas, P., Kim, D. W., Felip, E., Pérez-Gracia, J. L., Han, J. Y., et al. (2016). Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): A randomised controlled trial. *Lancet (London, Engl.* 387 (10027), 1540–1550. doi:10.1016/S0140-6736(15)01281-7

Hua, X., Zhao, W., Pesatori, A. C., Consonni, D., Caporaso, N. E., Zhang, T., et al. (2020). Genetic and epigenetic intratumor heterogeneity impacts prognosis of lung adenocarcinoma. *Nat. Commun.* 11 (1), 2459. doi:10.1038/s41467-020-16295-5

Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., and Xu, W. (2018). Applications of support vector machine (SVM) learning in cancer genomics. *Cancer genomics and proteomics* 15 (1), 41–51. doi:10.21873/cgp.20063

Jain, P., Jain, C., and Velcheti, V. (2018). Role of immune-checkpoint inhibitors in lung cancer. *Ther. Adv. Respir. Dis.* 12, 1753465817750075. doi:10.1177/1753465817750075

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

Kitamura, H., Ohno, Y., Toyoshima, Y., Ohtake, J., Homma, S., Kawamura, H., et al. (2017). Interleukin-6/STAT3 signaling as a promising target to improve the efficacy of cancer immunotherapy. *Cancer Sci.* 108 (10), 1947–1952. doi:10.1111/cas.13332

Laha, D., Grant, R., Mishra, P., and Nilubol, N. (2021). The role of tumor necrosis factor in manipulating the immunological response of tumor microenvironment. *Front. Immunol.* 12, 656908. doi:10.3389/fimmu.2021.656908

Li, H., Yin, C., Zhang, B., Sun, Y., Shi, L., Liu, N., et al. (2013). PTTG1 promotes migration and invasion of human non-small cell lung cancer cells and is modulated by miR-186. *Carcinogenesis* 34 (9), 2145–2155. doi:10.1093/carcin/bgt158

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004

Lyssiotis, C. A., and Kimmelman, A. C. (2017). Metabolic interactions in the tumor microenvironment. *Trends Cell Biol.* 27 (11), 863–875. doi:10.1016/j.tcb.2017.06.003

Meurette, O., and Mehlen, P. (2018). Notch signaling in the tumor microenvironment. *Cancer Cell* 34 (4), 536–548. doi:10.1016/j.ccell.2018.07.009

Mudd, T. W., Jr., Lu, C., Klement, J. D., and Liu, K. (2021). MS4A1 expression and function in T cells in the colorectal cancer tumor microenvironment. *Cell. Immunol.* 360, 104260. doi:10.1016/j.cellimm.2020.104260

Negrini, S., Gorgoulis, V. G., and Halazonetis, T. D. (2010). Genomic instability--an evolving hallmark of cancer. *Nat. Rev. Mol. Cell Biol.* 11 (3), 220–228. doi:10.1038/nrm2858

Network, C. G. A. R. (2014). Comprehensive molecular profiling of lung adenocarcinoma. *Nature* 511 (7511), 543–550. doi:10.1038/nature13385

Nisar, S., Yousuf, P., Masoodi, T., Wani, N. A., Hashem, S., Singh, M., et al. (2021). Chemokine-cytokine networks in the head and neck tumor microenvironment. *Int. J. Mol. Sci.* 22 (9), 4584. doi:10.3390/ijms22094584

Ott, P. A., Bang, Y. J., Piha-Paul, S. A., Razak, A. R. A., Bennouna, J., Soria, J. C., et al. (2019). T-Cell-Inflamed gene-expression profile, programmed death ligand 1 expression, and tumor mutational burden predict efficacy in patients treated with pembrolizumab across 20 cancers: KEYNOTE-028. *J. Clin. Oncol.* 37 (4), 318–327. doi:10.1200/JCO.2018.78.2276

Pan, Y., Yu, Y., Wang, X., and Zhang, T. (2020). Tumor-associated macrophages in tumor immunity. *Front. Immunol.* 11, 583084. doi:10.3389/fimmu.2020.583084

Patel, S., Alam, A., Pant, R., and Chattopadhyay, S. (2019). Wnt signaling and its significance within the tumor microenvironment: Novel therapeutic insights. *Front. Immunol.* 10, 2872. doi:10.3389/fimmu.2019.02872

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160 (1-2), 48–61. doi:10.1016/j.cell.2014.12.033

Saito, M., Suzuki, H., Kono, K., Takenoshita, S., and Kohno, T. (2018). Treatment of lung adenocarcinoma by molecular-targeted therapy and immunotherapy. *Surg. today* 48 (1), 1–8. doi:10.1007/s00595-017-1497-7

Sato, T., Terai, M., Tamura, Y., Alexeev, V., Mastrangelo, M. J., and Selvan, S. R. (2011). Interleukin 10 in the tumor microenvironment: A target for anticancer immunotherapy. *Immunol. Res.* 51 (2-3), 170–182. doi:10.1007/s12026-011-8262-6

Schabath, M. B., Welsh, E. A., Fulp, W. J., Chen, L., Teer, J. K., Thompson, Z. J., et al. (2016). Differential association of STK11 and TP53 with KRAS mutation-associated gene expression, proliferation and immune surveillance in lung adenocarcinoma. *Oncogene* 35 (24), 3209–3216. doi:10.1038/onc.2015.375

Schubert, M., Klinger, B., Klünemann, M., Sieber, A., Uhlitz, F., Sauer, S., et al. (2018). Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat. Commun.* 9 (1), 20. doi:10.1038/s41467-017-02391-6

Senosain, M. F., and Massion, P. P. (2020). Intratumor heterogeneity in early lung adenocarcinoma. *Front. Oncol.* 10, 349. doi:10.3389/fonc.2020.00349

Shen, W., Song, Z., Xiao, Z., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 3 (1), e36. doi:10.1002/imt2.36

Sonugür, F. G., and Akbulut, H. (2019). The role of tumor microenvironment in genomic instability of malignant tumors. *Front. Genet.* 10, 1063. doi:10.3389/fgene.2019.01063

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

Tanaka, A., and Sakaguchi, S. (2017). Regulatory T cells in cancer immunotherapy. *Cell Res.* 27 (1), 109–118. doi:10.1038/cr.2016.151

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830.e14. e14. doi:10.1016/j.immuni.2018.03.023

Tubbs, A., and Nussenzweig, A. (2017). Endogenous DNA damage as a source of genomic instability in cancer. *Cell* 168 (4), 644–656. doi:10.1016/j.cell.2017.01.002

Voron, T., Marcheteau, E., Pernot, S., Colussi, O., Tartour, E., Taieb, J., et al. (2014). Control of the immune response by pro-angiogenic factors. *Front. Oncol.* 4, 70. doi:10.3389/fonc.2014.00070

Wilkerson, M. D., and Hayes, D. N. (2010). ConsensusClusterPlus: A class discovery tool with confidence assessments and item tracking. *Bioinforma. Oxf. Engl.* 26 (12), 1572–1573. doi:10.1093/bioinformatics/btq170

Yang, B., Zhang, W., Zhang, M., Wang, X., Peng, S., and Zhang, R. (2020). KRT6A promotes EMT and cancer stem cell transformation in lung adenocarcinoma. *Technol. cancer Res. Treat.* 19, 1533033820921248. doi:10.1177/1533033820921248

Yang, L., Pang, Y., and Moses, H. L. (2010). TGF-Beta and immune cells: An important regulatory axis in the tumor microenvironment and progression. *Trends Immunol.* 31 (6), 220–227. doi:10.1016/j.it.2010.04.002

Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., et al. (2013). Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic acids Res.* 41, D955–D961. doi:10.1093/nar/gks1111

Yao, Z., Fenoglio, S., Gao, D. C., Camiolo, M., Stiles, B., Lindsted, T., et al. (2010). TGF-beta IL-6 axis mediates selective and adaptive mechanisms of resistance to molecular targeted therapy in lung cancer. *Proc. Natl. Acad. Sci. U. S. A.* 107 (35), 15535–15540. doi:10.1073/pnas.1009472107

Ye, X., and Weinberg, R. A. (2015). Epithelial-mesenchymal plasticity: A central regulator of cancer progression. *Trends Cell Biol.* 25 (11), 675–686. doi:10.1016/j.tcb.2015.07.012

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

Yoshizawa, A., Motoi, N., Riely, G. J., Sima, C. S., Gerald, W. L., Kris, M. G., et al. (2011). Impact of proposed IASLC/ATS/ERS classification of lung adenocarcinoma: Prognostic subgroups and implications for further revision of staging based on analysis of 514 stage I cases. *Mod. Pathol.* 24 (5), 653–664. doi:10.1038/modpathol.2010.232

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

Zhang, D., Jiang, Q., Ge, X., Shi, Y., Ye, T., Mi, Y., et al. (2021). RHOV promotes lung adenocarcinoma cell growth and metastasis through JNK/c-Jun pathway. *Int. J. Biol. Sci.* 17 (10), 2622–2632. doi:10.7150/ijbs.59939

Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. *Ann. Transl. Med.* 4 (7), 136. doi:10.21037/atm.2016.03.35

Ziani, L., Chouaib, S., and Thiery, J. (2018). Alteration of the antitumor immune response by cancer-associated fibroblasts. *Front. Immunol.* 9, 414. doi:10.3389/fimmu.2018.00414

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 06 February 2023
DOI 10.3389/fphar.2023.1126916

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Xue Zhan,
Chongqing Medical University, China
Aierpati Maimaiti,
First Affiliated Hospital of Xinjiang Medical University, China

*CORRESPONDENCE
Jiahao Ma,
ma1106013028@163.com
Jinjian Chen,
chenjinjian@wmu.edu.cn

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 18 December 2022
ACCEPTED 23 January 2023
PUBLISHED 06 February 2023

CITATION
Hu P, Ma J and Chen J (2023), A systematic and comprehensive analysis of T cell exhaustion related to therapy in lung adenocarcinoma tumor microenvironment. *Front. Pharmacol.* 14:1126916. doi: 10.3389/fphar.2023.1126916

A systematic and comprehensive analysis of T cell exhaustion related to therapy in lung adenocarcinoma tumor microenvironment

Peipei Hu[1], Jiahao Ma[1,2]* and Jinjian Chen[1]*

[1]Department of General Medicine, First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China, [2]Key Laboratory of Nano-carbon Modified Film Technology of Henan Province, Diagnostic Laboratory of Animal Diseases, School of Pharmacy, Xinxiang University, Xinxiang, China

Background: T cell exhaustion (TEX) is an important immune escape mechanism, and an in-depth understanding of it can help improve cancer immunotherapy. However, the prognostic role of TEX in malignant lung adenocarcinoma (LUAD) remains unclear.

Methods: Through TCGA and GEO datasets, we enrolled a total of 498 LUAD patients. The patients in TCGA-LUAD were unsupervised clustered into four clusters according to TEX signaling pathway. WGCNA analysis, survival random forest analysis and lasso regression analysis were used to select five differentially expressed genes among different clusters to construct a TEX risk model. The risk model was subsequently validated with GEO31210. By analyzing signaling pathways, immune cells and immune checkpoints using GSEA, GSVA and Cibersortx, the relationship between TEX risk score and these variables was evaluated. In addition, we further analyzed the expression of *CCL20* at the level of single-cell RNA-seq and verified it in cell experiments.

Results: According to TEX signaling pathway, people with better prognosis can be distinguished. The risk model constructed by *CD109, CCL20, DKK1, TNS4, and TRIM29* genes could further accurately identify the population with poor prognosis. Subsequently, it was found that dendritic cells, *CD44* and risk score were closely related. The final single-cell sequencing suggested that *CCL2O* is a potential therapeutic target of TEX, and the interaction between TEX and CD8 + T is closely related.

Conclusion: The classification of T cell depletion plays a crucial role in the clinical decision-making of lung adenocarcinoma and needs to be further deepened.

KEYWORDS

T exhausted, lung adenocarcinoma, single cell sequencing, biomarkers, tumor microenvironment

1 Introduction

The mortality rate of lung adenocarcinoma remains high throughout the world (Relli et al., 2019). LUAD is the most common form of primary lung cancer. Smoking-primary or secondary exposure, are the main causes (Hutchinson et al., 2019). The traditional treatment of LUAD includes surgical resection, chemotherapy, and radiotherapy. A number of new therapeutic approaches have also been discovered that can be used to treat LUAD, such as immunotherapy (Succony et al., 2021).

As a subset of T immunocytes, $CD8^+$ T lymphocytes are responsible for mediating the activity of the T immunocytes to chronic infections and cancers (Zhang and Bevan, 2011). Chronic infections and tumor antigens cause differentiated CD8 + T cells to exhaust (Speiser et al., 2014; McLane et al., 2019). The expression of cytokine suppression, decreased killing, and hypoproliferation of T cells are all symptoms of T cell exhaustion (TEX), which occurs as a result of these processes (Freeman et al., 2006; Zhang et al., 2022). The immune checkpoint inhibitors working mechanism is not depleting T cells in the immune microenvironment, and TEX is thought to be a pathway of resistance (Chow et al., 2022). In parallel, Immunotherapy to restore TEX responses has transformed the current clinical decision for cancer treatment (Hudson and Wieland, 2022). There has been evidence that inhibiting the PD-1 inhibitory receptor pathway can reactivate the TEX response and active the immune anti-tumor effect (McLane et al., 2019).

Although T lymphocytes in the body can attack tumors, the latter often present a highly reactive microenvironment that shuts down the killing capacity of T cells (Ma et al., 2019). The tumor microenvironment (TME) is a key factor in the escape of tumor cells from the immune system, and this environment plays a key role in cancer development (Gholami et al., 2017). In the TME, T cells are regulated by a complex immunosuppressive network consisting of cancer cells, inflammatory cells, stromal cells and cytokines (Jiang et al., 2015). Among these TME components, cancer cells, inflammatory cells, and suppressor cytokines have key roles in regulating T cell phenotype and function (Speiser et al., 2016). These components contribute to the eventual differentiation of T cells into "exhausted" T cells. Eventually, the majority of T cells in the TME differentiate into exhausted T cells that express high levels of suppressor receptors, produce fewer effector cytokines, and lose the ability to eliminate cancer.

In the initial characterization of exhausted T cells, the levels of transcription factors T cell factor (TCF1) and programed cell death protein (PD-1) expression were used to distinguish between the progenitor and terminally differentiated subtypes (Im et al., 2016; Siddiqui et al., 2019). As a result of progenitor exhaustion, T cells exhibit stem cell characteristics or memory characteristics, which enable them to self-renew and transform into terminally differentiated cells (Akbar and Henson, 2011; Utzschneider et al., 2016). Comparatively, the terminally differentiated branching subtype does not have a functional recovery potential and is limited in its expansion potential (Philip et al., 2017; Khan et al., 2019). In another study, TEX was divided into four stages based on Ly108 and CD69 expression (TEX^{prog1}: Ly108 + $CD69^+$; TEX^{prog2}: Ly108 + $CD69^-$; TEX^{int}: Ly108-$CD69^-$; TEX^{term}: Ly108-$CD69^+$) (Beltra et al., 2020). These studies have shown the TEX process is dynamic, with a phenotypic and functional continuum of intermediate states, indicating a developmental hierarchy (Zhang et al., 2022). Further researches showed that individual patients displayed different levels of T cell exhaustion (Kim et al., 2021) and the presence of T cell activation or exhaustion biomarkers such as sTIM-3, CD25 in patients is evidence of this, these markers are associated with a poor outcome (Berg et al., 2022). In a pan-cancer analysis, Zhang et al. obtained TEX-related genes through machine learning to classify tumors in different things TEX for clinical decision-making (Zhang et al., 2022).

In this study, we performed clustering analysis on the TCGA-LUAD data through TEX-related pathways, and further WGCNA and random survival forest and lasso regression analyses to construct TEX risk scores. Subsequently, the relationships between TEX risk scores and GSEA pathway enrichment analysis, GSVA pathway enrichment, and CIBERSORTX immune infiltrating cells were analyzed. 348 urothelial cancer patients which treated with atezolizumab (PD-L1) were collected to examined the effect of TEX risk score on immunotherapy effectiveness. Single-cell sequencing data and experiment were finally used to analyze potential therapeutic targets and cell communication in TEX.

2 Materials and methods

2.1 Data acquisition

The expression data, gene mutations, and clinical information were collected from the Cancer Genome Atlas (TCGA) website for 284 patients with LUAD (Tomczak et al., 2015). And 214 LUAD patients' information were collected through dataset GSE31210 in the GEO database. Single-cell sequencing data (GSE176021) of tumor-infiltrating T lymphocytes from six NSCLC patients were obtained from the GEO database. For data inclusion criteria, we selected patient samples with RNA transcriptome sequencing data and complete clinical data. For data from different datasets de-batching effects were performed and normalized, and we used fragment per kilobase transcript/fragment per million mapping (FPKM) expression values for further analysis. In FPKM, RNA-seq data were normalized to the length of each gene and the total number of aligned reads in the library (Trapnell et al., 2010). FPKM values were transformed using log2 (FPKM + 1). The flowchart of our investigation was displayed in Figure 1.

2.2 Unsupervised cluster analysis

The molecular signature database provided information on TEX signaling pathways and marker genes (Wherry, 2011; Liberzon et al., 2015). Similar to previous studies, we performed an unsupervised cluster analysis of LUAD patients using IFN-γ, TNF, and IL-2 signaling pathways to represent the TEX pathway (Zhang et al., 2022). The specific method is to use ssGSEA through the "GSVA" R package to estimate the activity score of each patient's TEX pathway (Hänzelmann et al., 2013). The percentage of patients at different stages in different clusters is also shown.

2.3 Comparison of overall survival between different clusters

To further explore the differences among different subgroups, we first drew Kaplan-Meier (K-M) survival curves for different clusters with the mark of 50% survival rate. Then the K-M survival curve of pairwise subgroups was drawn.

2.4 CIBERSORTX

To further explore the abundance of immunocytes in different classifications, we used CIBERSORTX algorithm to evaluate

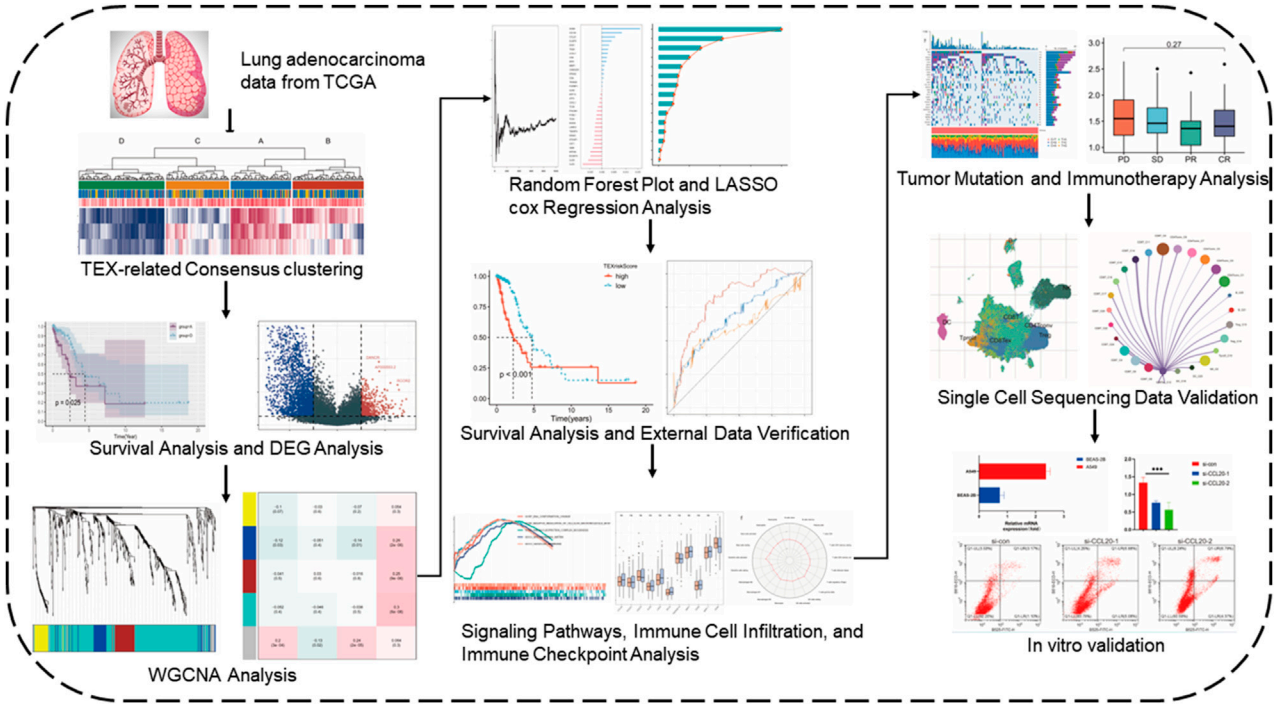


FIGURE 1
Flow chart of the study.

22 immunocytes in samples from different clusters (Steen et al., 2020). After cell infiltration of each sample was obtained, COX regression analysis was conducted to explore the prognostic value of various cells in each TEX cluster. Based on median immune cell content, LUAD was divided into diverse subgroups, and survival rates were compared between groups.

2.5 Weighted gene co-expression network analysis

The correlation patterns among genomics can be described using a systems biology method known as a weighted correlation network analysis (WGCNA) (Langfelder and Horvath, 2008). The R package repository had the package WGCNA 3.6.1 that was used for the WGCNA. WGCNA analysis was performed after deunion of the four cluster differential genes. The significance of each gene was taken into account when calculating the association between the gene expression profile and the TEX score, and the relationship between module eigengenes and gene expression profiles was taken into consideration when determining module membership. The soft threshold parameters were set at a power of 4 and a scale-free R2 of 0.9, in order to ensure the topology network was scale-free despite the number of nodes. The analysis consisted of retrieving an initial set of six modules, and the Grey modules that showed the strongest correlations were applied for further investigation.

2.6 RandomForest

Using the survival random forest of 1,000 trees by the R package randomSurvivalForest version 3.6.4, it was possible to validate the results and rank the importance of 7 genes obtained from Lasso regression using the R package randomSurvivalForest Version 3.6.4 (Taylor, 2011). The relative importance of gene > 0.2 is considered the ultimate hub gene.

2.7 Construction and validation of risk models

Gene expression tends to show significant collinearity between genes, which means it is necessary to use prognostic models if needed. LASSO regression and other methods reduce the number of variables to further reduce the redundancy of the model and increase the convenience of clinical use. Based on the following formula, we were able to calculate the risk score according to the following (Tibshirani, 1997):

$$Risk\ score = \sum_{i=1}^{n} Exp_i * Coef_i$$

Here, the TCGA-LUAD data set was employed as the training set to construct the risk model of LUAD patients based on survival random forest screening genes, took OS as the outcome event, and p value less than 0.05 as the limit of statistical significance. We then categorized the patients into diverse subgroups based on the formula generated by the risk model. Meanwhile, the K-M survival curve was drawn for the high and low risk group. Receiver operating characteristic (ROC) curves have a wide range of uses in identifying the diagnostic power of threshold changes. To further analyze the predictive power of prognostic models, we plotted the ROC. An analysis of multivariate cox regression was undertaken in

order to determine the independent prognostic significance of risk scores. To ensure the value of external generalization of the prognostic model, we used another LUAD dataset (GSE31210) for validation.

2.8 Biological function in relation to risk score

A follow-up analysis looked at genes with differential expression (DEGs) between high and low risk groups. It was established that the cut-off criteria for the study were |FC| > 2 and adj. p.val > 0.05. GSEA software was used to analyze three data sets of HALMARK, KEGG and GO for the biological functional differences among high-risk patients (Shi and Walker, 2007; Subramanian et al., 2007).

2.9 GSVA analysis between high-And low-risk groups

By using the GSVA analysis, it was possible to explore the differences between subgroups in signaling pathways for disease development (Hänzelmann et al., 2013). Moreover, a correlation analysis was conducted between the partial signal pathway score and the risk score.

2.10 The predictive significance of TEX risk model

A boxplot was made to illustrate the expression of 11 immune checkpoints between different subgroups. An analysis of 22 different types of immune cell infiltration was performed using the CIBERSORTX algorithm.

The detailed gene mutation statuses of the subgroups were displayed using the R package "maftools" to make comparisons between the two subgroups (Mayakonda et al., 2018). With the IMvigor210 package, we were able to determine gene expression and immunotherapeutic effectiveness in the IMvigor210 cohort (Mariathasan et al., 2018). IMvigor210 cohort is widely used to analyze the efficacy of immunotherapy.

2.11 Single cell sequencing analysis

It is a standard processing procedure that is used to do downstream processing on scRNA-seq data which is carried out using Seurat R software package, version 3.0.2, and a standard downstream processing package for this analysis (Stuart et al., 2019). In addition, genes detectable in fewer than 3 cells and genes detected in fewer than 200 cells were excluded, and the percentage of mitochondria detected was limited to no more than 20% of the total number of genes. Then, t Data was normalized using LogNormalize. A non-linear method used for reducing the dimensions of a sample is t-distributed stochastic neighborhood embedding (t-SNE) that is used for unsupervised clustering and unbiased visualization of cell populations on a two-dimensional map after principal component analysis (PCA) (Van Der Maaten and Hinton, 2008). A minimum fraction of 0.25 cell population fraction was used in both populations in order to identify marker genes in each cluster using the "FindAllMarkers" function. The filtering criterion was filter value of absolute log2 fold change (FC) ≥1. To visualize each marker gene's expression patterns within the cluster, the "DotPlot" function in Seurat was used. Then, the SingleR package (version 1.0.0) was utilized for annotating cell types based on marker-based information (Aran et al., 2019).

2.12 Cell culture

The A549 and BEAS-2B cell lines were obtained from Dr Liu. A549 cell lines were cultured in RPMI-1640 (Invitrogen) and BEAS-2B cells were cultured in DMEM medium, The medium was supplemented with 10% FBS (Gibco).

2.13 Molecular expression verification

The expression of *CCL20* in tumor and normal tissues of LUAD patients was compared through GEPIA2 online website, and we analyzed the overall survival rate of high expression group and low expression group (Tang et al., 2017).

The total RNA was extracted using the Trizol reagent. RT was performed with DNA-free total RNA in Revert Aid First Strand cDNA Synthesis Kit (Thermo). For PCR amplification, specific primers were used to amplify the transcribed cDNA. *CCL20* Forward: ATGTGC TGTACCAAGAGTTTGC; *CCL20* Reverse: CCAATTCCATTCCAG AAAAGCC.

Integrated DNA technologies (Coralville, IA, United States) provided us the synthetic siRNA and the scrambled negative control siRNA. This experiment consisted of transfecting cells with LipofectamineTM RNAiMAX (Thermo Fisher Scientific, Massachusetts, United States) in opti-MEM according to the procedure given by the manufacturer.

2.14 Flow cytometry

The manufacturer's instructions were followed when performing flow cytometry. Apoptosis was detected with the Annexin V-PE/7-ADD Apoptosis Detection Kit (Vazyme, A213-01). The B525 nm wavelength was selected for the Fluorescein (FITC) signal channel, and the B610 nm signal channel was selected for the ECDPE-TR (ECD) signal channel.

2.15 Statistical analysis

It was determined that two groups with normally distributed variables and those with variables that were not normally distributed were statistically significant using independent t-tests and Mann-Whitney U tests. In order to make a comparison between the two groups on the basis of differences between the groups, we conducted an analysis of variance (ANOVA) and a Kruskal–Wallis test (Hazra and Gogtay, 2016). We performed Spearman correlation and distance correlation analyses using the R package Hmisc 4.4.1. To analyze the correlation between

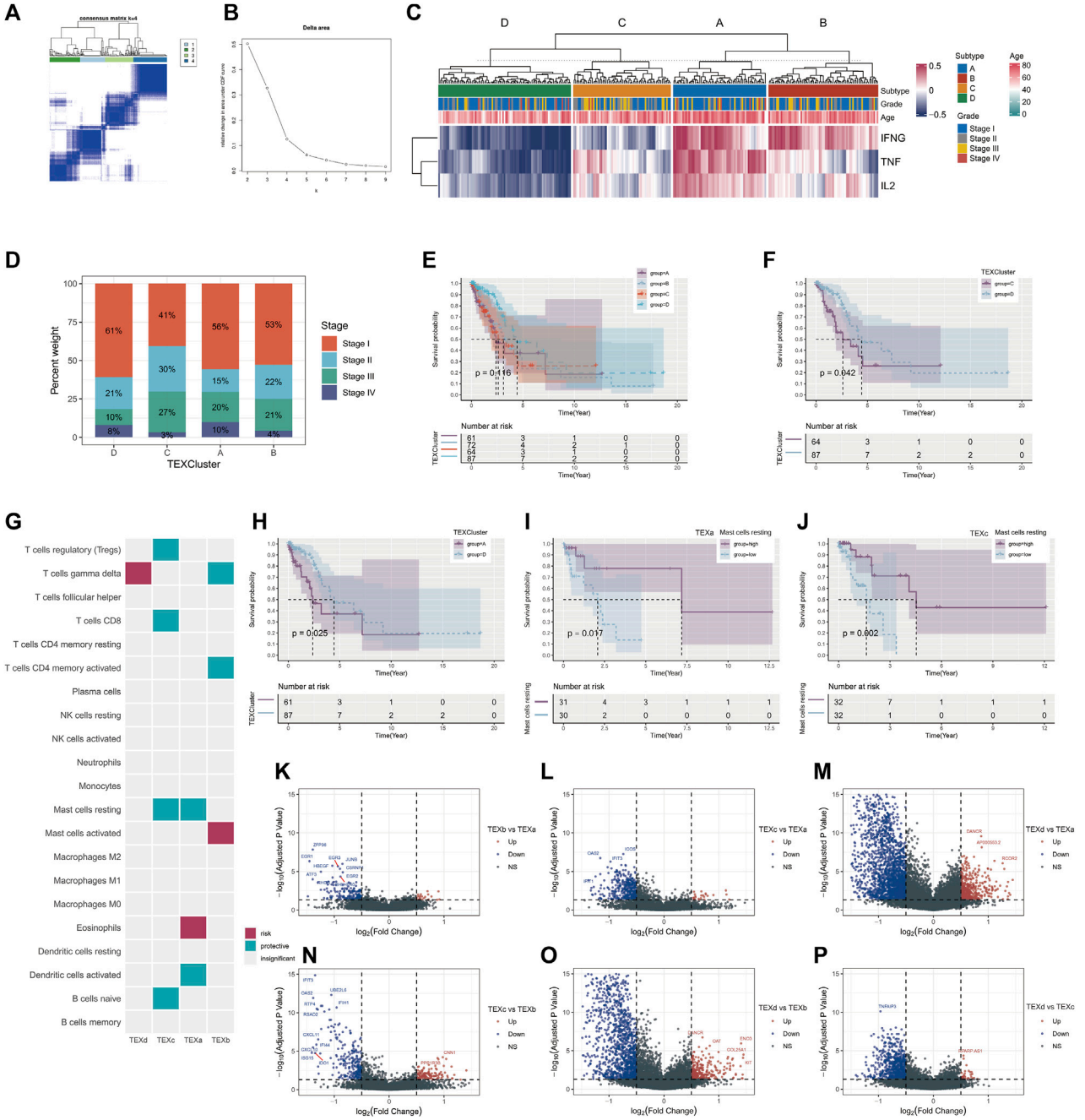


FIGURE 2
Unsupervised cluster analysis was performed on LUAD patients according to TEX signaling pathway. Unsupervised cluster analysis of patients with TCGA□LUAD (consensus matrix k = 4) **(A)**. Delta area of unsupervised consensus cluster analysis **(B)**. Heatmap of scores for four T cell exhaustion and IFNG/TNF/IL-2 signaling pathways **(C)**. Percentages of different clinical stages in TEX clusters **(D)**. Kaplan-Meier survival curves for the four TEX clusters **(E)**. Kaplan-Meier survival curves for TEX C and TEX D **(F)**. Kaplan-Meier survival curves for TEX A and TEX D **(H)**. The cibersortX algorithm in four TEX clusters was used to analyze the infiltration results of 22 immune cells **(G)**. K-M survival curves of patients with high and low abundance of mast cells resting in TEXa **(I)** and TEXc **(J)**. Volcano plot of differential genes between TEXa and TEXb **(K)**, TEXc and TEXa **(L)**, TEXd and TEXa **(M)**, TEXc and TEXb **(N)**, TEXd and TEXb **(O)**, TEXd and TEXc **(P)**.

the objects. Those objects whose coefficient was greater than 0.5 were considered highly correlated (Faul et al., 2007). For the purpose of identifying the prognostic factors, Cox regression analyses were conducted. A survival curve with all survivorship curves generated by the R package survminer was also used to determine the overall survival (OS) and TEX riskScore values before generating any survival curves with the R package survminer. As a means of plotting the heatmaps, the R package Complex Heatmap 2.4.3 was used. R package ggplot2 was used for visualizing data comparisons. There were two-sided statistical analyses conducted using R software, which was used for all statistical analyses.

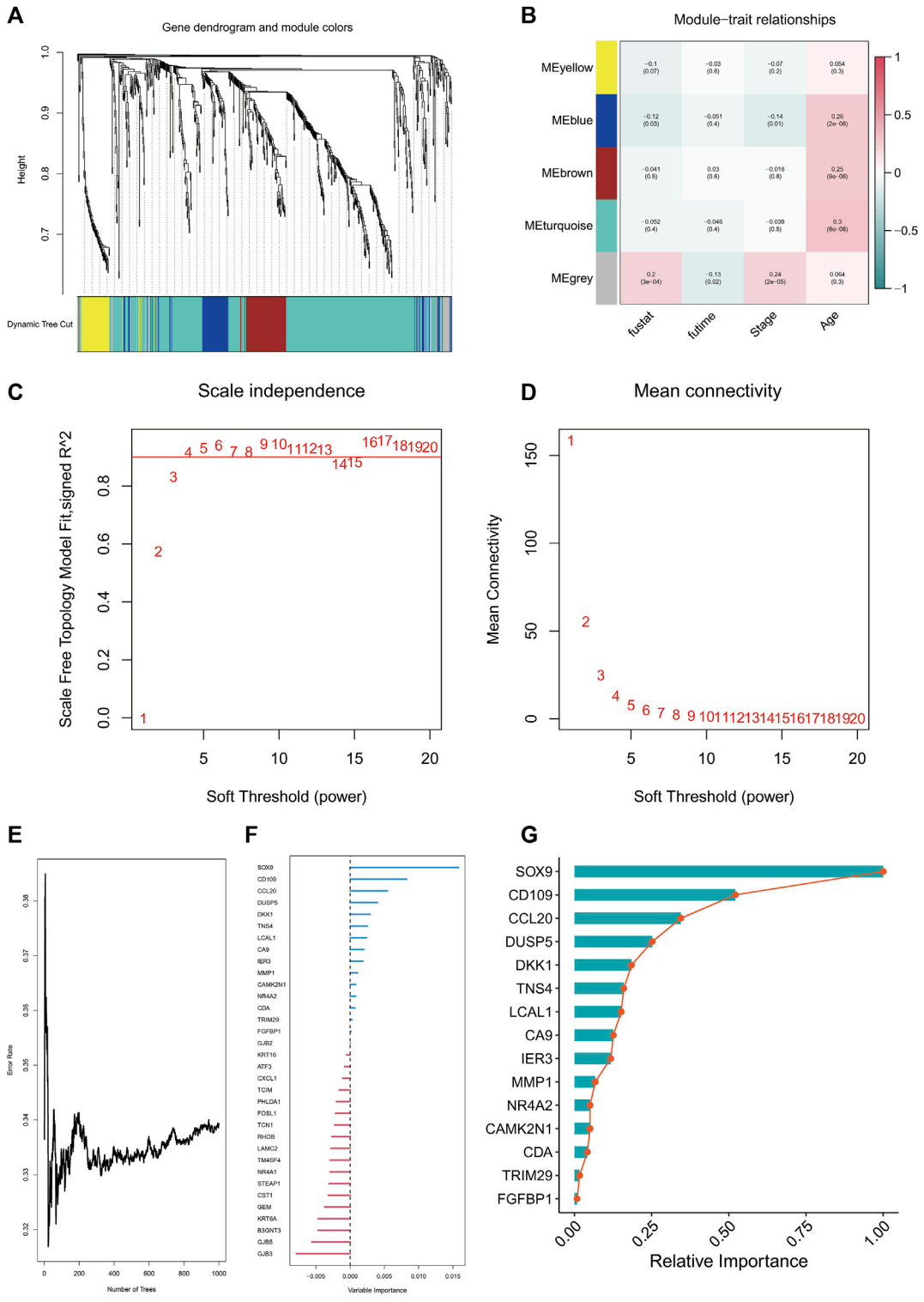


FIGURE 3
WGCNA analysis and random survival analysis. Clustering dendrogram of TCGA-LUAD **(A)**. Heatmap of correlation between WGCNA modules and clinical features **(B)**. Various soft thresholding powers are calculated according to their scale-free fit index **(C)**. Soft-threshold power mean connectivity analysis **(D)**. Plot of random survival forest based on number of trees and error rate **(E)**. Variable Importance ranking of genes in random survival forests **(F)**. Relative Importanc ranking of genes in random survival forests **(G)**.

3 Results

3.1 Unsupervised cluster analysis

The outcome of unsupervised cluster analysis were displayed in Figures 2A, B, where the best result was to classify the 284 LUAD patients into four TEX cluster (consensus matrix k = 4) according to the GSVA scores of the three IFNG/TNFA/IL-2 pathways. Figure 2C displayed the GSVA scores of the three pathways within the four clusters. Cluster B has the highest IFNG/TNF/IL-1 pathway score, and cluster D has the lowest score. The number of patients with stage1 in cluster b was larger than in cluster A, B and C. The number of patients with stage1 in cluster B was larger than that in cluster A, B and C (Figure 2D), suggesting that the overall survival of patients in cluster b may be better than that in other cluster populations. We plotted K-M survival curves for the overall survival of the four cluster populations, and there was no statistically significant difference (p = 0.116) in survival between the four clusters (Figure 2E). We plotted the K-M survival curves separately for cluster B and cluster A C D, and the results suggested that the overall survival rate of cluster B was higher than that of cluster A C D (Figures 2F-H; Supplementary Figure S1) (Cluster C vs. D, p = 0.042; cluster A vs. D, p = 0.025; cluster B vs. D, p = 0.073).

To further analyze the abundance of immune cells in different cluster, we applied the CIBERSORTX algorithm to evaluate the 22 immune cells in the samples in different clusters. After cell infiltration score was obtained for each sample, COX regression analysis was performed to explore the prognostic value of various cells in each TEX cluster. Mast cell resting was a protective factor in both TEXa and TEXc (Figures 2I, J).

For further analysis of the transcriptome differences between different cluster, we will contrast between different cluster differences in gene analysis, analysis of the standard is greater than or equal to | logFC | = 0.5, rectify the p value is less than 0.05, and mapped the volcano map is used to display the results of the analysis (Pearson correlation coefficient = 0.2, p value < 0.001) (Figures 2K–P).

3.2 WGCNA and survival random forest results

WGCNA analysis results suggested that the grey module was most relevant to survivals related information, and the grey module was selected for subsequent analysis (Figures 3A–D). There were 36 genes chosen as hub genes in the Grey module since they had absolute values of module membership [MM] that were greater than 0.5 and absolute values of gene significance [GS] that were greater than 0.5 within the module (Supplementary Table S1). Variable selection based on minimum depth values above the threshold (0.001) and importance values above the threshold (0.2) yielded seven tentative (*SOX9*, *CD109*, *CCL20*, *DUSP5*, *DKK1*, *TNS4*, *and LCAL1*) candidate prognostic markers for LUAD. This suggests that these seven genes are most relevant to the prognosis of LUAD (Figures 3E–G).

3.3 Development and validation of TEX risk model

A TEX risk model that includes five genes was constructed using lasso regression analysis. The formula for the risk score is as follows: risk socre = (0.2628**CD109* + 0.0464**CCL20* + 0.0163**DKK1* + 0.0359**TNS4* + 0.0348**TRIM29*). The TCGA-LUAD patients were divided into high- and low-risk groups based on their risk scores. The K-M survival curve between high and low risk groups suggested that the high-risk group had worse overall survival (p < 0.001) (Figure 4A). The AUC values of the TEX risk model were 0.823 in the first year, 0.688 in the third year, and 0.619 in the fifth year (Figure 4B). Multivariate COX analysis showed that TEX Score was an independent prognostic factor (p < 0.05, Hazard Ratio :1.625 [1.329–1.986]) (Figure 4C). In the validation set GSE31210, we also found that high-risk LUAD patients had worse OS (Figure 4D). The AUC values of TEX risk model in the validation set were 0.643 in the first year, 0.655 in the third year, and 0.700 in the fifth year. These results TEX risk model have good predictive power (Figure 4E).

3.4 TEX risk score and biological function GSEA analysis

Based on the HALLMARK, KEGG, and GO datasets, we performed an enrichment analysis of biological functions in high-risk patients using GSEA software. The results showed that the five HALLMARK pathways with the highest enrichment were bile acid metabolism, heme metabolism, MYC target v1, peroxisome, and protein secretion (Figure 4F). The five most enriched pathways in the GO database were DNA conformational changes, negative regulation of cellular macromolecular biosynthetic processes, ribonucleoprotein complex biogenesis, mitochondrial matrix, and vacuolar membrane (Figure 4G). The top five enriched KEGG pathways were insulin signaling pathway, melanoma, peroxisome, T cell receptor signaling pathway, and vascular smooth muscle contraction (Figure 4H). The results showed that high-risk patients were highly associated with many tumor proliferation and metabolism-related pathways, suggesting that targeted TEX affects the prognosis of LUAD patients mainly through tumor proliferation and metabolic pathways.

3.5 TEX risk score and GSVA analysis

We selected several gene sets for GSVA analysis based on the above GSEA results and found that TEX score was positively correlated with glycosaminoglycan degradation, linoleic acid metabolism, o glycan biosynthesis, leukocyte transendothelial migration, focal adhesion, ECM receptor interaction and p53 signaling pathway (Figures 5A, B). This suggests a potential pathway through which TEX exerts its effects.

3.6 Relationship between TEX risk score and immunity

Considering the great potential of TEX for immunotherapy, In both high and low risk groups, we plotted the expression levels of 11 immune checkpoints. In the high-risk group, CD44 expression was higher, which may be a therapeutic target in the future (Figure 5C). Based on 22 immune cell infiltrations, the high-risk group had a higher percentage of resting Dendritic cells and a lower percentage of activated Dendritic cells (Figure 5D). Further correlations of TEX

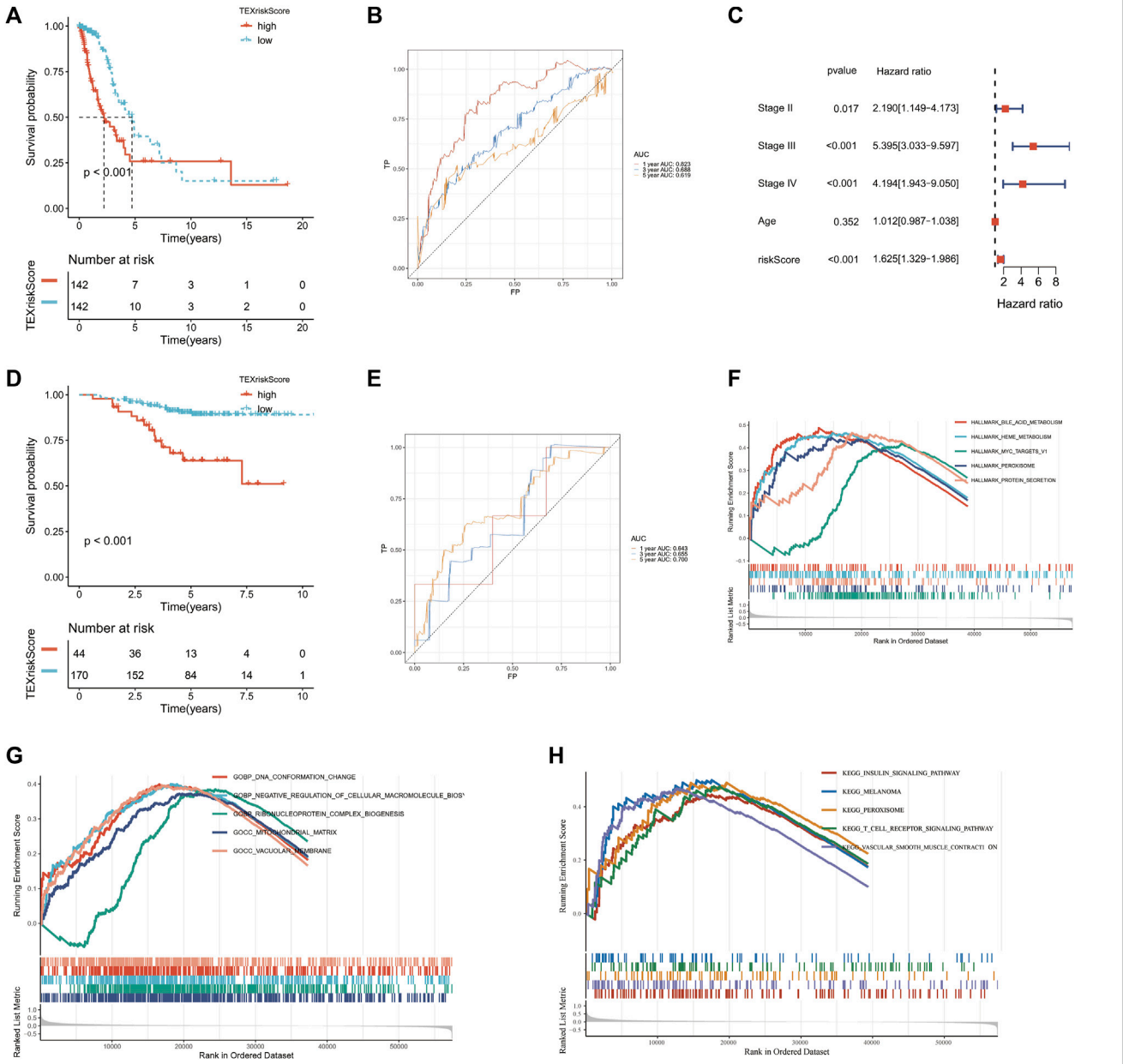


FIGURE 4
Construction and validation of TEX risk model. Kaplan-Meier (K–M) survival curves of patients in the high and low risk groups in TCGA-LUAD **(A)** and GSE31210 **(D)**. Multivariate Cox analysis in TCGA-LUAD cohort **(C)**. TEX risk model AUC values at year 1, 3, and 5 in TCGA-LUAD **(B)** and GSE32120 **(E)**. HALLMARK pathway enrichment analysis **(F)**, GO pathway enrichment analysis **(G)** and KEGG pathway enrichment analysis **(H)** of the high-risk group in TCGA-LUAD.

risk scores and 22 immunocytes are shown in Figure 5E. The radar chart further showed the contents of 22 immunocytes in the high-risk group (Figure 5F), suggesting that TEX may affect the prognosis of LUAD by regulating the state of Dendritic cells.

3.7 Relationship between TEX risk score and genetic mutations

We used a map to determine the landscape of gene mutations in high and low-risk subgroups of patients (Figures 6A, B). There was no statistically significant difference between the gene mutation frequencies between the groups that were analyzed, but it was noted that TP53 and TTN had the highest mutation frequencies.

3.8 Relationship between TEX risk score and immunotherapy

Based on the TCGA-LUAD data set and the IMvigor210 data set, subgraph analyses were conducted to evaluate immunotherapy and chemotherapy in high-risk and low-risk groups. The high-risk group demonstrated a lower percentage of responders to immunotherapy (Figure 7A), and TEXscore was lower in those

10.3389/fphar.2023.1126916

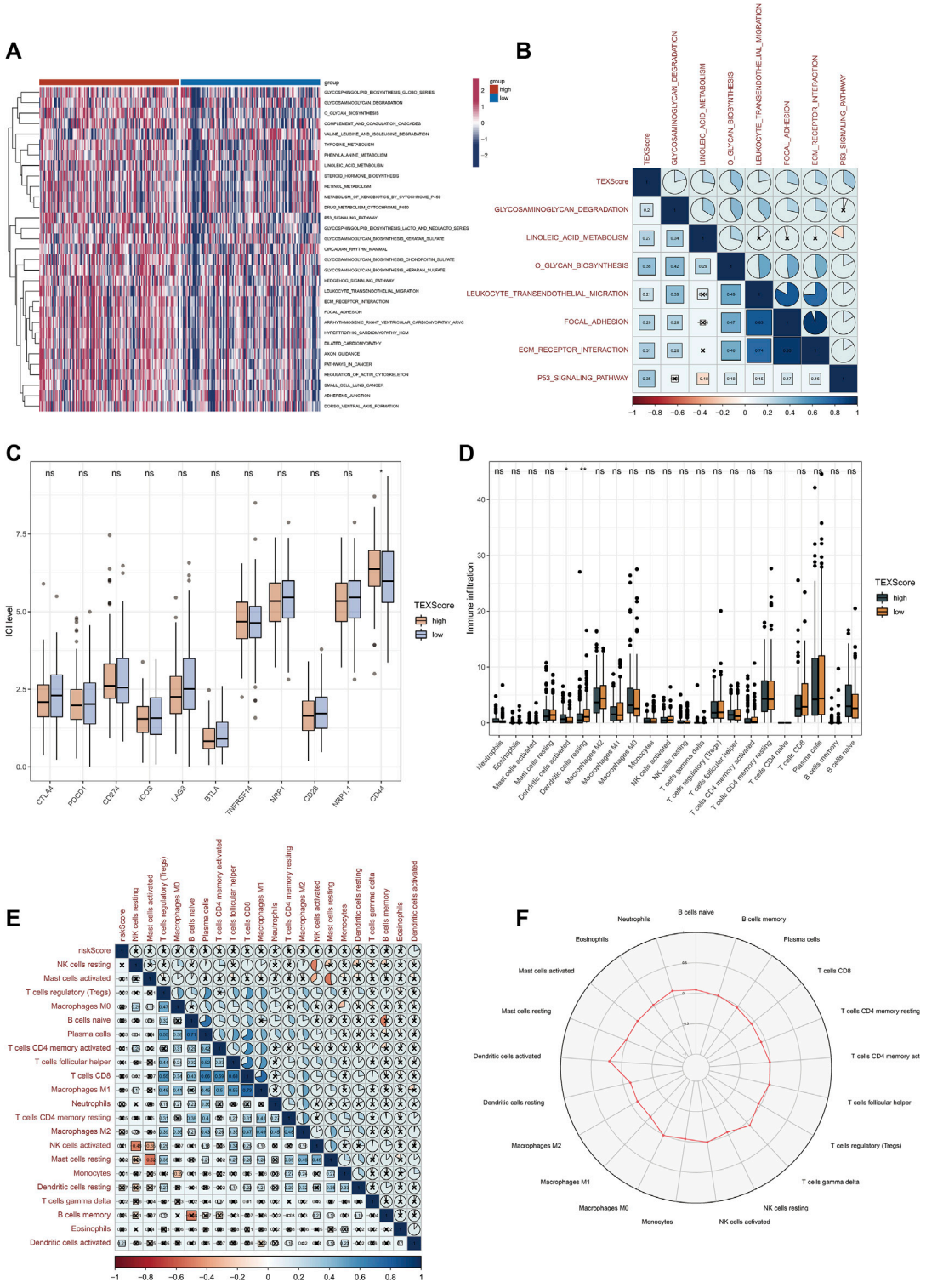


FIGURE 5
Relationship between TEX risk scores and immunity. GSVA analysis between high and low risk groups in TCGA-LUADA **(A)**. Correlation scores of TEX risk scores and signaling pathways **(B)**. Expression of 11 immune checkpoints in the high and low risk groups **(C)**. The abundance of 22 immune cells in cibersortx high and low risk groups was analyzed **(D)**. Association of TEX risk scores and 22 immune cells **(E)**. Radar plot of the abundance of 22 immune cells in the high-risk group **(F)**.

with low response (Figure 7B). These results suggest that patients with lower TEX risk scores are able to achieve a better immunotherapy response. Specific immunotherapy responses fall into four types: CR: complete response; PR, partial response; SD: stable disease; PD: progressive disease. There was no statistically significant difference in TEX risk scores among the four types of

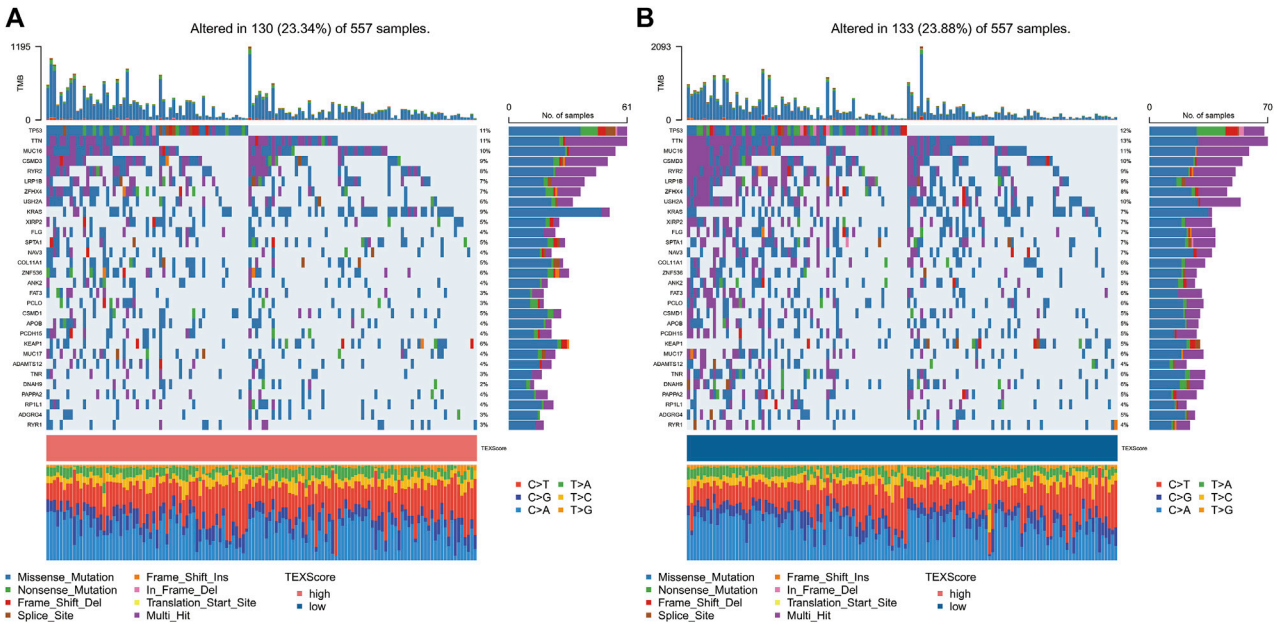


FIGURE 6
Gene mutation landscape. Gene mutation landscape in high **(A)** and low risk groups **(B)**.

response, suggesting that the specific immunotherapy response was not related to the risk score (Figure 7C).

It was found that the high-risk group had also a poorer overall survival rate than the low-risk group when they received immunotherapy, regardless of the median risk score (Figure 7D). These results suggest that the TEX risk score has a role in predicting the efficacy of immunotherapy and the prognosis of patients receiving immunotherapy.

3.9 Single-cell sequencing analysis revealed the therapeutic targets of TEX

To further search for the potential therapeutic target-cell interactions of TEX, a total of 8 cell subtypes were identified in the single-cell sequencing dataset of T cells (B, CD4 Tconv, CD8T cell, CD8Tex, DC, NK, T prolif, Treg) (Figure 8A). The key gene in the risk model, *CCL20*, was most highly expressed on TEX cells (Figures 8B,C), *CCL20*, a key gene in the risk model, was most highly expressed on TEX cells, suggesting that *CCL20* plays an important role in the TEX process in LUAD patients and is a potential therapeutic target. GSEA analysis showed that TEX cells were mainly enriched in cell adhesion (Figure 8D). The results of cell communication showed that TEX mainly interacted with CD8T cells (Figures 8E,F). These results provide new explanatory theories and therapeutic targets for TEX depletion in LUAD.

3.10 GEPIA2, real-time quantitative PCR, and flow cytometry validation

The GEPIA2 website contained 483LUAD patients and 347 normal lung tissues, and we found the expression level of *CCL20* was higher in tumors tissues. Subsequently, the LUAD patients were classified into diverse subgroups based on *CCL20* expression value (Figure 9A), and the results also showed that the high expression group had a shorter overall survival ($p = 0.022$) (Figure 9B). In addition, we detected *CCL20* mRNA values in both normal and tumor cell lines. The results displayed that the mRNA expression level of *CCL20* in A549 was more than twice that in BEAS-2B (Figure 9C). Subsequently, we knocked down *CCL20* in A549 cells by siRNA (Figure 9D), and the *CCL20* knockdown cells had more apoptosis than the control cells (Figure 9E).

4 Discussion

A growing body of evidence suggests that TEX is the result of delayed phenotypic differentiation as well as intermediate functional stages within T cells that follow a sustained state of hierarchy dysfunction. Like other forms of cellular differentiation, it is believed that TEX is the result of T cell hierarchical dysfunction over a prolonged period of time (Wherry, 2011; Jiang et al., 2015; Blank et al., 2019). By understanding CD8 + T cell dysregulation and exhaustion in the tumor microenvironment (TIME), we can overcome the TEX barrier and improve immune checkpoint blockade therapies in the clinic, regardless of whether the type of tumor is the same or different (Kurtulus et al., 2019). The dynamics and heterogeneity of TEX in the TIME are not well studied across LUAD.

In this study, we preferred unsupervised CLUSTER analysis of 284 LUAD patients based on the three most closely TEX signaling pathways (IFNG, TNF, and IL-2), and the patients were divided into four clusters. In cluster D, the number of LUAD patients with the lowest signal pathway score and the largest number of stage1 was the highest. The K-M curve showed that cluster D patients had a better prognosis. PD-1 overexpression leads to inhibitory signaling and induces TEX, leading to tumor immune escape (Zwergel et al., 2022), which suggested that we can use TEX related pathways for prognosis judgment and precise

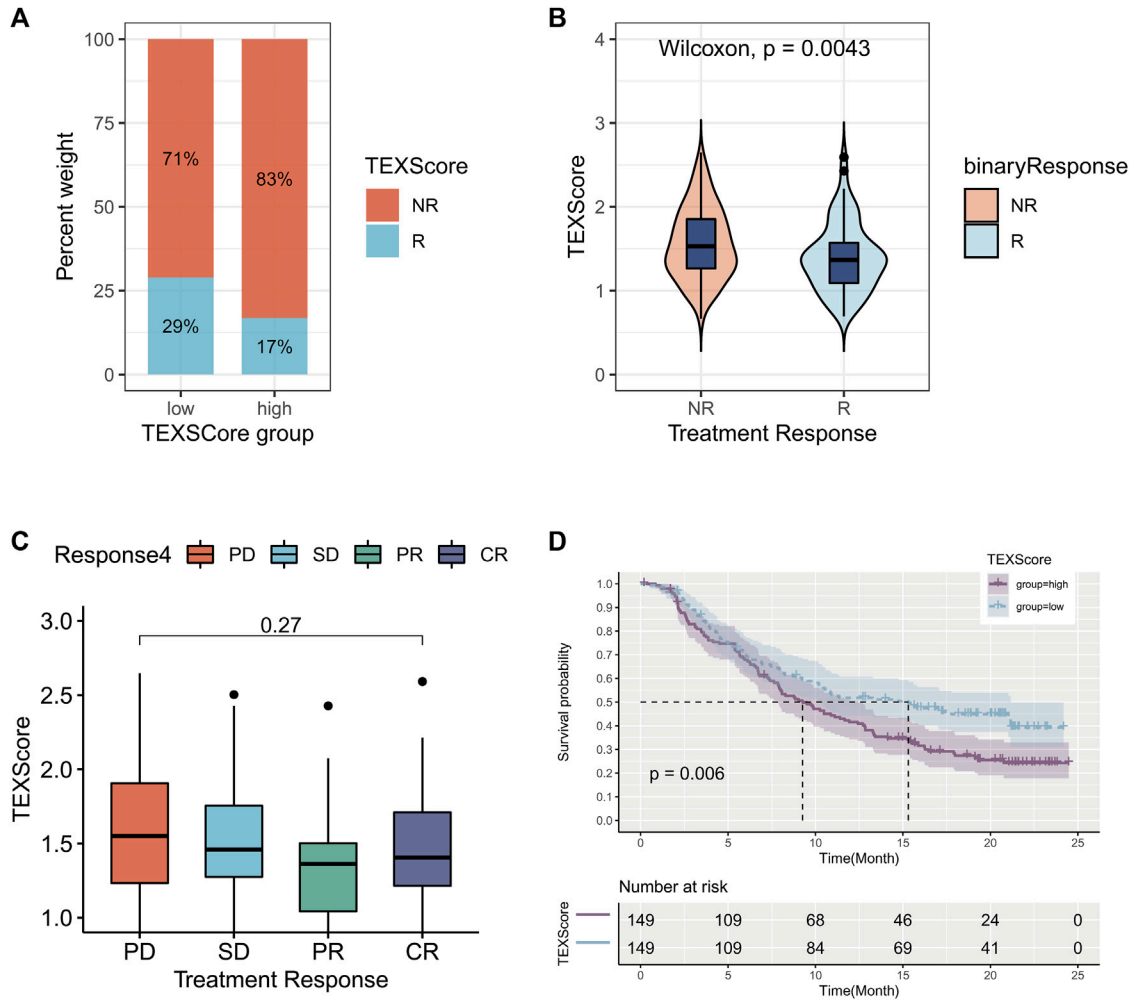


FIGURE 7
Relationship between TEX risk score and immunotherapy. Percentage weight with response (R) and no response (NR) in the high and low risk groups **(A)**. Wilcoxon test for TEX scores in response and no response populations **(B)**. Boxplots of TEX scores in the four treatment responses **(C)**. K-M survival curves for the high and low risk groups in the immunoresponsive population **(D)**. PR, partial response; CR:complete response; PD: progressive disease; SD: stable disease.

treatment of LUAD. Analysis of immune cell infiltration in four TEXclusters 22 by cibersortx revealed that higher abundance of mast cells resting in TEXa and TEXc was associated with worse prognosis. The existence of mast cells is associated with the prognosis of patients with lung adenocarcinoma, as exosomes derived from mast cells have been shown to promote the proliferation of lung adenocarcinoma cells (Xiao et al., 2014; Bao et al., 2020). However, the specific mechanism of TEX and mast cells needs to be further studied.

WGCNA analysis of the differentially expressed genes and survival random forest analysis obtained 7 key genes (*SOX9*, *CD109*, *CCL20*, *DUSP5*, *DKK1*, *TNS4*, *and LCAL1*). Then we selected 5 genes (*CD109*, *CCL20*, *DKK1*, *TNS4 and TRIM29*) by lasso regression algorism to build a TEX risk model. In the training set TCGA-LUAD and the validation set GSE, high-risk patients had worse overall survival. The AUC value and multivariate cox regression analysis of TEX risk model in training set and validation set showed that Tex risk model had good predictive value and clinical application value. Cluster of differentiation 109 (*CD109*) is a glycosylphosphatidylinositol-anchored protein (Lee et al., 2020). Further studies showed that *CD109* promoted lung adenocarcinoma invasion and metastasis *in vivo* through TGF-β signaling pathway (Chuang et al., 2017; Lee et al., 2020; Taki et al., 2020). However, there is no study on *CD109* and TEX.

Through the enrichment analysis of GSEA and GSVA, we found that TEX score was positively correlated with glycosaminoglycan degradation, linoleic acid metabolism, o glycan biosynthesis, leukocyte transendothelial migration, focal adhesion, ECM receptor interaction and P53 signaling pathway. Targeting P53 has been shown to restore CD8 + T cells depleted in hepatitis C virus infection. However, other pathways and TEX pathways are still worthy of further exploration in LUAD. Subsequent immune checkpoint analysis revealed that the high risk group had increased expression of *CD44*, a stemness marker of non-small cell lung cancer, and activation of *CD44* related pathways promoted squamous cell lung cancer resistance to FGFR1 inhibition (Elakad et al., 2022; Panda and Biswal, 2022). These results suggest that TEX may be involved in the stemness and other phenotypes of LUAD resulting in a poorer prognosis in high-risk patients.

10.3389/fphar.2023.1126916

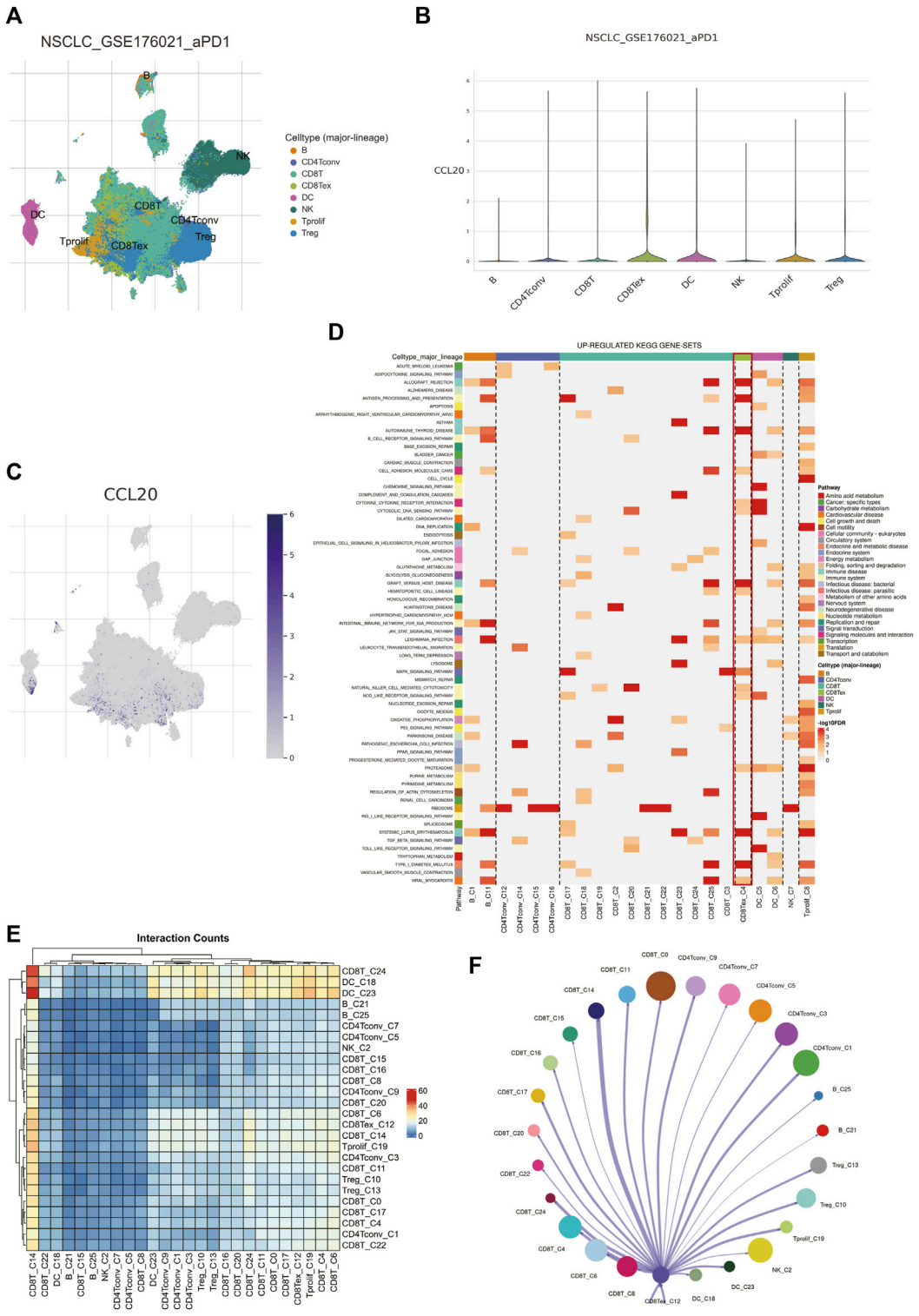


FIGURE 8
Analysis of TEX by single-cell sequencing. Major subtypes of cells **(A)**. The amount of CCL20 expression on different cells **(B,C)**. Up-regulated kegg pathways in different cell types **(D)**. Interaction conunts of different cell subtypes **(E)**. Diagram of the interaction network between TEX and other cell types, with the width of the network edge being the total number of ligand and receptor pairs **(F)**.

Furthermore, in the pathway enrichment analysis, we found that bile acid metabolism, peroxisome, and T cell receptor signaling pathways were significantly enriched. It has been shown that bile acids can regulate cell growth and proliferation and that alterations in bile acid levels in disease states are associated with liver injury/regeneration and tumorigenesis (Li and Apte, 2015). Peroxisomes can regulate various biological processes and play an important role in several diseases and conditions, and some studies suggest that they

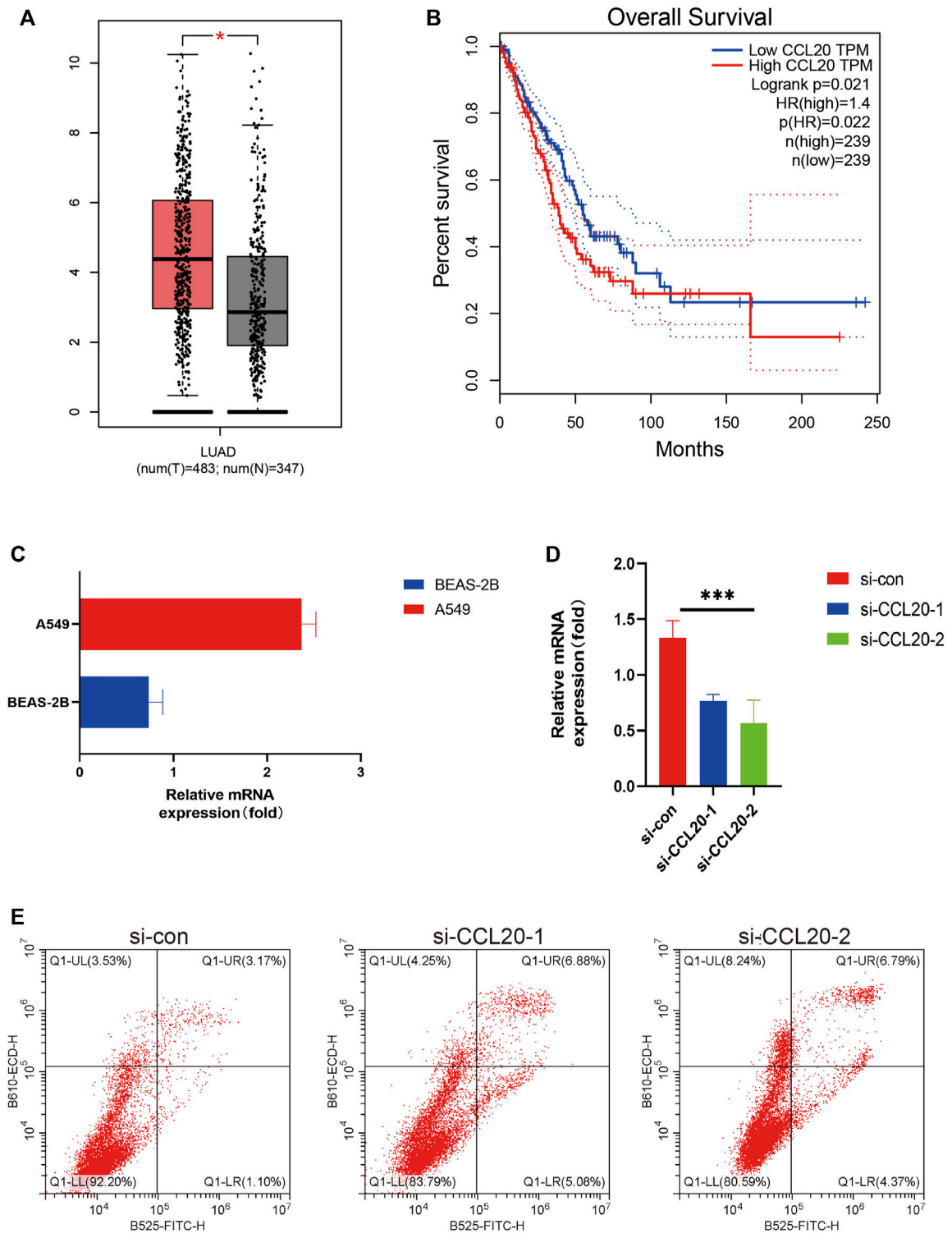


FIGURE 9
GEPIA2, Real-time quantitative PCR, and flow cytometry validation. *CCL20* expression in LUAD tissues and normal lung tissues in GEPIA **(A)**. K-M curves of overall survival of LUAD patients with high and low *CCL20* expression in GEPIA2 **(B)**. Relative mRNA expression of *CCL20* in A549 cell line and BEAS-2B cell line **(C)**. After knocking down *CCL20* in A549 cells by siRNA, the expression of *CCL20* gene in the three groups of cells was detected **(D)**. The number of apoptotic cells in the three groups was counted by flow cytometry **(E)**. * <0.05.

may also have an important role in the development and progression of cancer and may represent a new opportunity for cancer therapy (Peters et al., 2005; Youssef and Badr, 2011). In contrast, T cell receptor-based immunotherapy has been shown to be a promising approach for the treatment of various types of cancer. TCRs can recognize epitopes of proteins from any subcellular compartment,

including the membrane, cytoplasm and nucleus, and these advantages allow TCRs to detect a wide range of targets, such as neoantigens, cancer germline antigens and viral oncoproteins, and in the clinical setting TCR-based immunotherapy can mediate solid regression of malignant tumors, including immune checkpoint inhibitor-refractory cancers (Schmitt et al., 2009; Kirsch et al., 2015; Chandran and Klebanoff, 2019).

To accurately describe the mechanism of TEX at the single-cell level, we found that the major cells could be divided into eight types through single-cell sequencing data GSE of T cells. It has been found that C-C motif chemokine ligand 20 (*CCL20*) is involved in the occurrence and development of various types of cancer. We found that CCL20, a key gene in the TEX risk score, was highly expressed on TEX cells. In LUAD patients, high expression of *CCL20* is related to epithelial-mesenchymal transition (EMT), which is associated with poor prognosis. Patients responding to anti-PD-L1 therapy were significantly better when *CCL20* expression was low rather than high (Fan et al., 2022). Notably, TNF signaling is also a key pathway in TEX, suggesting that targeting *CCL20* in TEX may have potential clinical value. KEGG signaling pathway analysis identified multiple gene sets up-regulated in TEX, and three signaling pathways attracted our attention. The first is the antigen processing and presentation pathway. Previous findings suggested that patients with higher TEX risk scores had a higher proportion of dendritic cells. The single-cell analysis here further confirms the possible interaction between TEX and DC. Studies have shown that immune checkpoint therapy can restore the immune function of TEX, but it depends on the depleted precursor state of T cells. Dendritic cells provide an important niche for TPEX and prevent its excessive activation (Dähling et al., 2022). Cell communication shows that TEX mainly interacts with CD8T cells, CD8 + T cells differentiate and deplete to TEX, and TEX further acts on CD8 + T cells. This suggests that if we can stop this process, it may provide new ideas for immunotherapy. Subsequent GEPIA2 data analysis, RT-PCR and flow cytometry results similarly indicated *CCL20* as a prognostic indicator for LUAD.

Clinically, there are a number of available risk models based on multiple genes that can predict the prognosis of cancer patients. For example, 21 gene expression analysis (Oncotype DX, Genomic Health) is one of several commercially available gene expression assays that provide prognostic information in hormone receptor-positive breast cancer (Sparano et al., 2018). In clinical practice guidelines for breast cancer, the National Comprehensive Cancer Network (NCCN) strongly recommends 21-gene expression testing (Sparano et al., 2018). Our study now consists of 5 genes and represents a clinically convenient test. Moreover, our model is based on TEX-related genes, which means that our model also has unique potential for predicting immune function in patients.

However, our experiments still have some limitations. Our model performs well, but additional experiments are needed to further validate our model. In addition, although basic experiments were performed to validate one gene in the model, the specific mechanism by which it exerts its function still needs to be explored clearly.

Compared with other traditional models, our model still has great advantages. Our model has not only been validated using different datasets, but also an in-depth analysis based on single cell sequencing data, which will greatly affirm the reliability of our model. Our model can well predict the prognosis and immune control of LUAD patients and provide help for individual precision treatment.

5 Conclusion

We comprehensively described the prognostic significance, immunotherapy value and possible mechanism of TEX in LUAD patients for the first time. Nevertheless, the study has certain limitations. Firstly, we defined TEX only according to the scores of three TEX-related signaling pathways, which may simplify the definition of TEX. Secondly, we used public data to analyze the relationship between TEX and LUAD, and there is a lack of molecular biology experiments and *in vivo* results to further confirm our conclusion. In conclusion, our results provide a new insight into the role of TEX in LUAD.

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

PH performed data analysis and wrote the manuscript. JC contributed to the conception of the study and data analysis. JM contributed to data analysis. All authors have read and approved the final manuscript.

Acknowledgments

The authors thank the participants and staff of the First Affiliated Hospital of Wenzhou Medical University and Xinxiang College for their contributions.

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/fphar.2023.1126916/full#supplementary-material

SUPPLEMENTARY FIGURE S1
K-M survival curves for TEXB and TEXD.

SUPPLEMENTARY TABLE S1
Thirty-six genes in the Grey module in WGCNA.

References

Akbar, A. N., and Henson, S. M. (2011). Are senescence and exhaustion intertwined or unrelated processes that compromise immunity? *Nat. Rev. Immunol.* 11 (4), 289–295. doi:10.1038/nri2959

Aran, D., Looney, A. P., Liu, L., Wu, E., Fong, V., Hsu, A., et al. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nat. Immunol.* 20 (2), 163–172. doi:10.1038/s41590-018-0276-y

Bao, X., Shi, R., Zhao, T., and Wang, Y. (2020). Mast cell-based molecular subtypes and signature associated with clinical outcome in early-stage lung adenocarcinoma. *Mol. Oncol.* 14 (5), 917–932. doi:10.1002/1878-0261.12670

Beltra, J. C., Manne, S., Abdel-Hakeem, M. S., Kurachi, M., Giles, J. R., Chen, Z., et al. (2020). Developmental relationships of four exhausted CD8(+) T cell subsets reveals underlying transcriptional and epigenetic landscape control mechanisms. *Immunity* 52 (5), 825–841. doi:10.1016/j.immuni.2020.04.014

Berg, J., Halvorsen, A. R., Bengtson, M. B., Lindberg, M., Halvorsen, B., Aukrust, P., et al. (2022). Circulating T cell activation and exhaustion markers are associated with radiation pneumonitis and poor survival in non-small-cell lung cancer. *Front. Immunol.* 13, 875152. doi:10.3389/fimmu.2022.875152

Blank, C. U., Haining, W. N., Held, W., Hogan, P. G., Kallies, A., Lugli, E., et al. (2019). Defining 'T cell exhaustion'. *Nat. Rev. Immunol.* 19 (11), 665–674. doi:10.1038/s41577-019-0221-9

Chandran, S. S., and Klebanoff, C. A. (2019). T cell receptor-based cancer immunotherapy: Emerging efficacy and pathways of resistance. *Immunol. Rev.* 290 (1), 127–147. doi:10.1111/imr.12772

Chow, A., Perica, K., Klebanoff, C. A., and Wolchok, J. D. (2022). Clinical implications of T cell exhaustion for cancer immunotherapy. *Nat. Rev. Clin. Oncol.* 19 (12), 775–790. doi:10.1038/s41571-022-00689-z

Chuang, C. H., Greenside, P. G., Rogers, Z. N., Brady, J. J., Yang, D., Ma, R. K., et al. (2017). Molecular definition of a metastatic lung cancer state reveals a targetable CD109-Janus kinase-Stat axis. *Nat. Med.* 23 (3), 291–300. doi:10.1038/nm.4285

Dähling, S., Mansilla, A. M., Knöpper, K., Grafen, A., Utzschneider, D. T., Ugur, M., et al. (2022). Type 1 conventional dendritic cells maintain and guide the differentiation of precursors of exhausted T cells in distinct cellular niches. *Immunity* 55 (4), 656–670.e8. doi:10.1016/j.immuni.2022.03.006

Elakad, O., Häupl, B., Labitzky, V., Yao, S., Küffer, S., von Hammerstein-Equord, A., et al. (2022). Activation of CD44/PAK1/AKT signaling promotes resistance to FGFR1 inhibition in squamous-cell lung cancer. *NPJ Precis. Oncol.* 6 (1), 52. doi:10.1038/s41698-022-00296-2

Fan, T., Li, S., Xiao, C., Tian, H., Zheng, Y., Liu, Y., et al. (2022). CCL20 promotes lung adenocarcinoma progression by driving epithelial-mesenchymal transition. *Int. J. Biol. Sci.* 18 (11), 4275–4288. doi:10.7150/ijbs.73275

Faul, F., Erdfelder, E., Lang, A. G., and Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39 (2), 175–191. doi:10.3758/bf03193146

Freeman, G. J., Wherry, E. J., Ahmed, R., and Sharpe, A. H. (2006). Reinvigorating exhausted HIV-specific T cells via PD-1-PD-1 ligand blockade. *J. Exp. Med.* 203 (10), 2223–2227. doi:10.1084/jem.20061800

Gholami, M. D., Saeedi, Y., Heydari, S., Garssen, J., and Falak, R. (2017). Exhaustion of T lymphocytes in the tumor microenvironment: Significance and effective mechanisms. *Cell. Immunol.* 322, 1–14. doi:10.1016/j.cellimm.2017.10.002

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14 (1), 7–15. doi:10.1186/1471-2105-14-7

Hazra, A., and Gogtay, N. (2016). Biostatistics series module 3: Comparing groups: Numerical variables. *Indian J. Dermatol* 61 (3), 251–260. doi:10.4103/0019-5154.182416

Hudson, W. H., and Wieland, A. (2022). Technology meets TILs: Deciphering T cell function in the -omics era. *Cancer Cell* 41, 41–57. doi:10.1016/j.ccell.2022.09.011

Hutchinson, B. D., Shroff, G. S., Truong, M. T., and Ko, J. P. (2019). Spectrum of lung adenocarcinoma. *Semin. Ultrasound CT MR* 40 (3), 255–264. doi:10.1053/j.sult.2018.11.009

Im, S. J., Hashimoto, M., Gerner, M. Y., Lee, J., Kissick, H. T., Burger, M. C., et al. (2016). Defining CD8+ T cells that provide the proliferative burst after PD-1 therapy. *Nature* 537 (7620), 417–421. doi:10.1038/nature19330

Jiang, Y., Li, Y., and Zhu, B. (2015). T-cell exhaustion in the tumor microenvironment. *Cell Death Dis.* 6 (6), e1792. doi:10.1038/cddis.2015.162

Khan, O., Giles, J. R., McDonald, S., Manne, S., Ngiow, S. F., Patel, K. P., et al. (2019). TOX transcriptionally and epigenetically programs CD8(+) T cell exhaustion. *Nature* 571 (7764), 211–218. doi:10.1038/s41586-019-1325-x

Kim, C. G., Kim, G., Kim, K. H., Park, S., Shin, S., Yeo, D., et al. (2021). Distinct exhaustion features of T lymphocytes shape the tumor-immune microenvironment with therapeutic implication in patients with non-small-cell lung cancer. *J. Immunother. Cancer* 9 (12), e002780. doi:10.1136/jitc-2021-002780

Kirsch, I., Vignali, M., and Robins, H. (2015). T-cell receptor profiling in cancer. *Mol. Oncol.* 9 (10), 2063–2070. doi:10.1016/j.molonc.2015.09.003

Kurtulus, S., Madi, A., Escobar, G., Klapholz, M., Nyman, J., Christian, E., et al. (2019). Checkpoint blockade immunotherapy induces dynamic changes in PD-1–cd8+ tumor-infiltrating T cells. *Immunity* 50 (1), 181–194. doi:10.1016/j.immuni.2018.11.014

Langfelder, P., and Horvath, S. (2008). Wgcna: an R package for weighted correlation network analysis. *BMC Bioinforma.* 9 (1), 559. doi:10.1186/1471-2105-9-559

Lee, K. Y., Shueng, P. W., Chou, C. M., Lin, B. X., Lin, M. H., Kuo, D. Y., et al. (2020). Elevation of CD109 promotes metastasis and drug resistance in lung cancer via activation of EGFR-AKT-mTOR signaling. *Cancer Sci.* 111 (5), 1652–1662. doi:10.1111/cas.14373

Li, T., and Apte, U. (2015). Bile acid metabolism and signaling in cholestasis, inflammation, and cancer. *Adv. Pharmacol.* 74, 263–302. doi:10.1016/bs.apha.2015.04.003

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004

Ma, X., Bi, E., Lu, Y., Su, P., Huang, C., Liu, L., et al. (2019). Cholesterol induces CD8+ T cell exhaustion in the tumor microenvironment. *Cell metab.* 30 (1), 143–156. doi:10.1016/j.cmet.2019.04.002

Mariathasan, S., Turley, S. J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., et al. (2018). TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 554 (7693), 544–548. doi:10.1038/nature25501

Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118

McLane, L. M., Abdel-Hakeem, M. S., and Wherry, E. J. (2019). CD8 T cell exhaustion during chronic viral infection and cancer. *Annu. Rev. Immunol.* 37, 457–495. doi:10.1146/annurev-immunol-041015-055318

Panda, M., and Biswal, B. K. (2022). Evodiamine inhibits stemness and metastasis by altering the SOX9-β-catenin axis in non-small-cell lung cancer. *J. Cell Biochem.* 123 (9), 1454–1466. doi:10.1002/jcb.30304

Peters, J. M., Cheung, C., and Gonzalez, F. J. (2005). Peroxisome proliferator-activated receptor-α and liver cancer: Where do we stand? *J. Mol. Med.* 83 (10), 774–785. doi:10.1007/s00109-005-0678-9

Philip, M., Fairchild, L., Sun, L., Horste, E. L., Camara, S., Shakiba, M., et al. (2017). Chromatin states define tumour-specific T cell dysfunction and reprogramming. *Nature* 545 (7655), 452–456. doi:10.1038/nature22367

Relli, V., Trerotola, M., Guerra, E., and Alberti, S. (2019). Abandoning the notion of non-small cell lung cancer. *Trends Mol. Med.* 25 (7), 585–594. doi:10.1016/j.molmed.2019.04.012

Schmitt, T. M., Ragnarsson, G. B., and Greenberg, P. D. (2009). T cell receptor gene therapy for cancer. *Hum. gene Ther.* 20 (11), 1240–1248. doi:10.1089/hum.2009.146

Shi, J., and Walker, M. G. (2007). Gene set enrichment analysis (GSEA) for interpreting gene expression profiles. *Curr. Bioinforma.* 2 (2), 133–137. doi:10.2174/157489307780618231

Siddiqui, I., Schaeuble, K., Chennupati, V., Fuertes Marraco, S. A., Calderon-Copete, S., Pais Ferreira, D., et al. (2019). Intratumoral tcf1(+)pd-1(+)cd8(+) T cells with stem-like properties promote tumor control in response to vaccination and checkpoint blockade immunotherapy. *Immunity* 50 (1), 195–211. doi:10.1016/j.immuni.2018.12.021

Sparano, J. A., Gray, R. J., Makower, D. F., Pritchard, K. I., Albain, K. S., Hayes, D. F., et al. (2018). Adjuvant chemotherapy guided by a 21-gene expression assay in breast cancer. *N. Engl. J. Med.* 379 (2), 111–121. doi:10.1056/NEJMoa1804710

Speiser, D. E., Ho, P-C., and Verdeil, G. (2016). Regulatory circuits of T cell function in cancer. *Nat. Rev. Immunol.* 16 (10), 599–611. doi:10.1038/nri.2016.80

Speiser, D. E., Utzschneider, D. T., Oberle, S. G., Münz, C., Romero, P., and Zehn, D. (2014). T cell differentiation in chronic infection and cancer: Functional adaptation or exhaustion? *Nat. Rev. Immunol.* 14 (11), 768–774. doi:10.1038/nri3740

Steen, C. B., Liu, C. L., Alizadeh, A. A., and Newman, A. M. (2020). Profiling cell type abundance and expression in bulk tissues with CIBERSORTx. *Methods Mol. Biol.* 2117, 135–157. doi:10.1007/978-1-0716-0301-7_7

Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W. M., et al. (2019). Comprehensive integration of single-cell data. *Cell* 177 (7), 1888–1902. doi:10.1016/j.cell.2019.05.031

Subramanian, A., Kuehn, H., Gould, J., Tamayo, P., and Mesirov, J. P. (2007). GSEA-P: A desktop application for gene set enrichment analysis. *Bioinformatics* 23 (23), 3251–3253. doi:10.1093/bioinformatics/btm369

Succony, L., Rassl, D. M., Barker, A. P., McCaughan, F. M., and Rintoul, R. C. (2021). Adenocarcinoma spectrum lesions of the lung: Detection, pathology and treatment strategies. *Cancer Treat. Rev.* 99, 102237. doi:10.1016/j.ctrv.2021.102237

Taki, T., Shiraki, Y., Enomoto, A., Weng, L., Chen, C., Asai, N., et al. (2020). CD109 regulates *in vivo* tumor invasion in lung adenocarcinoma through TGF-β signaling. *Cancer Sci.* 111 (12), 4616–4628. doi:10.1111/cas.14673

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 (W1), W98–W102. doi:10.1093/nar/gkx247

Taylor, J. M. (2011). Random survival forests. *J. Thorac. Oncol.* 6 (12), 1974–1975. doi:10.1097/JTO.0b013e318233d835

Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Stat. Med.* 16 (4), 385–395. doi:10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3

Tomczak, K., Czerwińska, P., and Wiznerowicz, M. (2015). The cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp. Oncol. Pozn.* 19 (1), A68–A77. doi:10.5114/wo.2014.47136

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., Van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28 (5), 511–515. doi:10.1038/nbt.1621

Utzschneider, D. T., Charmoy, M., Chennupati, V., Pousse, L., Ferreira, D. P., Calderon-Copete, S., et al. (2016). T cell factor 1-expressing memory-like CD8(+) T cells sustain the immune response to chronic viral infections. *Immunity* 45 (2), 415–427. doi:10.1016/j.immuni.2016.07.021

Van Der Maaten, L., and Hinton, G. (2008). Visualizing high-dimensional data using t-sne. journal of machine learning research. *J. Mach. Learn Res.* 9 (26), 5.

Wherry, E. J. (2011). T cell exhaustion. *Nat. Immunol.* 12 (6), 492–499. doi:10.1038/ni.2035

Xiao, H., Lässer, C., Shelke, G. V., Wang, J., Rådinger, M., Lunavat, T. R., et al. (2014). Mast cell exosomes promote lung adenocarcinoma cell proliferation - role of KIT-stem cell factor signaling. *Cell Commun. Signal* 12, 64. doi:10.1186/s12964-014-0064-8

Youssef, J., and Badr, M. (2011). Peroxisome proliferator-activated receptors and cancer: Challenges and opportunities. *Br. J. Pharmacol.* 164 (1), 68–82. doi:10.1111/j.1476-5381.2011.01383.x

Zhang, N., and Bevan, M. J. (2011). CD8(+) T cells: Foot soldiers of the immune system. *Immunity* 35 (2), 161–168. doi:10.1016/j.immuni.2011.07.010

Zhang, Z., Chen, L., Chen, H., Zhao, J., Li, K., Sun, J., et al. (2022). Pan-cancer landscape of T-cell exhaustion heterogeneity within the tumor microenvironment revealed a progressive roadmap of hierarchical dysfunction associated with prognosis and therapeutic efficacy. *EBioMedicine* 83, 104207. doi:10.1016/j.ebiom.2022.104207

Zwergel, C., Fioravanti, R., and Mai, A. (2022). PD-L1 small-molecule modulators: A new hope in epigenetic-based multidrug cancer therapy? *Drug Discov. Today* 28 (2), 103435. doi:10.1016/j.drudis.2022.103435

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 28 February 2023
DOI 10.3389/fphar.2023.1121634

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Wei Li,
Shenzhen Longhua District Central Hospital, China
Zhe Wang,
Third Military Medical University, China

*CORRESPONDENCE
Fengyihuan Fu,
fufengyihuan@126.com
Yuqiang Nie,
eynieyuqiang@scut.edu.cn

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 12 December 2022
ACCEPTED 14 February 2023
PUBLISHED 28 February 2023

CITATION
Feng J, Fu F and Nie Y (2023), Comprehensive genomics analysis of aging related gene signature to predict the prognosis and drug resistance of colon adenocarcinoma.
Front. Pharmacol. 14:1121634.
doi: 10.3389/fphar.2023.1121634

Comprehensive genomics analysis of aging related gene signature to predict the prognosis and drug resistance of colon adenocarcinoma

Jubin Feng[1,2†], Fengyihuan Fu[3*†] and Yuqiang Nie[1,4*]

[1]The First Affiliated Hospital, Jinan University, Guangzhou, China, [2]Department of Gastroenterology, The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, [3]The Second Affiliated Hospital of Guangzhou Medical University, Guangzhou, China, [4]Department of Gastroenterology, School of Medicine, Guangzhou First People's Hospital, South China University of Technology, Guangzhou, China

Background: Colon adenocarcinoma (COAD) is a heterogeneous tumor and senescence is crucial in the occurrence of cancer. This study aimed to identify senescence-based subtypes and construct a prognostic signature to predict the prognosis and guide immunotherapy or chemotherapy decisions for COAD patients.

Methods: Based on the single-cell RNA sequencing (scRNA-seq) data of 13 samples from the Gene Expression Omnibus (GEO) database, we assessed cellular senescence characteristics. Transcriptome data, copy number variations (CNVs) and single nucleotide variations (SNVs) data were obtained from The Cancer Genome Atlas (TCGA) database. GSE39582 and GSE17537 were used for validation. Senescence subtypes were identified using unsupervised consensus clustering analysis, and a prognostic signature was developed using univariate Cox analysis and least absolute shrinkage and selection operator (LASSO). Response of risk groups to chemotherapy was predicted using the half-maximal inhibitory concentration (IC50) values. We further analyzed the relationship between risk gene expression and methylation level. The prediction performance was assessed by nomogram.

Results: Senescence-related pathways were highly enriched in malignant cells and bulk RNA-seq verified cellular senescence. Three senescence subtypes were identified, in which patients in clust3 had poorest prognosis and higher T stage, accompanied with higher tumor mutation burden (TMB) and mutations, activated inflammatory response, more immune cell infiltration, and higher immune escape tendency. A senescence-based signature using 11 genes (MFNG, GPRC5B, TNNT1, CCL22, NOXA1, PABPC1L, PCOLCE2, MID2, CPA3, HSPA1A, and CALB1) was established, and accurately predicted

Abbreviations: CRC, colorectal cancer; COAD, colon adenocarcinoma; scRNA-seq, single-cell RNA sequencing; TME, tumor microenvironment; TCGA, the cancer genome atlas; GEO, gene-expression omnibus; PCA, principal component analysis; ICI, immune checkpoint inhibition; TIDE, tumor immune dysfunction exclusion; OS, overall survival; KEGG, kyoto encyclopedia of genes and genomes; CDF, cumulative distribution function; ssGSEA, single-sample gene set enrichment analysis; ROC, receiver operating characteristic analysis; AUC, area under ROC curve; LASSO, least absolute shrinkage and selection operator; AIC, akaike information criterion; EMT, epithelial-to-mesenchymal transition; SNVs, single nucleotide variations; CNVs, copy number variations; TMB, tumor mutation burden; DEGs, differential expressed genes; TSNE, t-distributed stochastic neighbor embedding; PAM, partitioning around medoids; MDSC, myeloid-derived suppressor cells; NOX, NADPH oxidase; DCA, decision curve analysis.

a lower prognosis in high risk patients. Its robustness was validated by external cohort. Low risk patients were more sensitive to small molecule drugs including Erlotinib, Sunitinib, MG-132, CGP-082996, AZ628, Sorafenib, VX-680, and Z-LLNle-CHO. Risk score was an independent prognostic factor and nomogram confirmed its reliability. Four risk genes (CALB1, CPA3, NOXA1, and TNNT1) had significant positive correlation with their methylation level, while six genes (CCL22, GPRC5B, HSPA1A, MFNG, PABPC1L, and PCOLCE2) were negatively correlated with their methylation level.

Conclusion: This study provides novel understanding of heterogeneity in COAD from the perspective of senescence, and develops signatures for prognosis prediction in COAD.

KEYWORDS

colon adenocarcinoma, senescence, subtypes, prognosis, risk score, nomogram, chemotherapy drugs

Introduction

Colorectal cancer (CRC), is the most common diagnosed gastrointestinal malignant tumor, and ranks third in the morbidity and second in the mortality with an estimated 3.2 million new COAD cases in 2040 worldwide (Xi and Xu, 2021). The prevalence of CRC is 0.56 million in 2020, and will increase to 0.91 million in 2040 in China (Xi and Xu, 2021). Among these, colon adenocarcinoma (COAD) accounts for 90% of cases (Munro et al., 2018). Patients with Stage 1-2 have a 5-year survival rate of 82%–94%, while it reduces to 67% for patients with stage 3 and advanced metastatic or stage 4 have a dismal 5-year survival rate of only 11% (Sagaert et al., 2018). Various treatments such as radical surgery followed by adjuvant chemotherapies can be used for treatment of resectable COAD patients, and palliative chemo- or radiotherapy is optimal for unresectable COAD patients to prolong their life. It has been recognized that COAD is a malignancy with intertumor and intratumor heterogeneity, which contribute to difference of prognosis and therapy response (Punt et al., 2017). Hence, it is great of importance to stratify patients with COAD and develop novel markers to accurately predict prognosis and therapy response.

Over the past decades, high throughput sequencing technology has been widely used in various fields of biology and medicine, greatly promoting relevant research and clinical application (Lightbody et al., 2019). The traditional RNA sequencing technology (bulk RNA-seq) is applied to determine gene expression profiles, isoform expression, alternative splicing and single-nucleotide polymorphisms on basis of tissue samples, which contains various cell types (Kuksin et al., 2021). On the contrast, single-cell RNA sequencing (scRNA-seq), a novel technology can detect the gene expression patterns for each transcript within single cell and distinguish cell subtypes (Lähnemann et al., 2020). Recently, scRNA-seq has been employed widely used in different cell type of various species, especially in human and mouse, to assess biological variability (Papalexi and Satija, 2018).

Cellular senescence is a cell state of cell cycle arrest that can eliminate damaged cells and promote tissue remodeling. Cellular senescence is predominantly elicited in response to intrinsic and extrinsic stimulus, such as oncogene activation, stress, DNA damage, CDKN2A locus derepression, mitochondrial dysfunction (Hernandez-Segura et al., 2018). Unfortunately, compelling evidence has suggested that cellular senescence is implicated in pathological status, in which senescence-associated secretory phenotype (SASP) affects the clearance of senescent cells and further results the decline of tissue function (Muñoz-Espín and Serrano, 2014), and secret pro-inflammatory cytokines including interleukin (IL)-6 and IL-8, chemokines and growth factors, which contributes to tumorigenesis in aged organisms (Herranz and Gil, 2018). Cellular senescence has been studied in various cancer types and compelling evidences have revealed that cellular senescence is associated with cancer prognosis (Dai et al., 2022a; Domen et al., 2022a; Domen et al., 2022b). Development of senescence-related classification and characterization of senescence-based signature have attracted much attention in tumor research (Feng et al., 2022a; Hong et al., 2022). However, the mechanisms of cellular senescence in COAD, as well as the specific prognostic signatures are poorly understood. Therefore, this study identified senescence-based subtypes based on scRNA-seq and shed novel insights into potential roles of cellular senescence in COAD heterogeneity. We further constructed a prognostic risk model in The Cancer Genome Atlas (TCGA)-COAD, which offered a novel approach to predict clinical outcomes in patients with COAD.

Material and methods

Single-cell RNA sequencing (scRNA-seq) data collection and pre-processing

The scRNA-seq expression profiles of 13 samples (GSE161277) (Zheng et al., 2022) were downloaded from Gene-Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) database. To comprehensively understand the profile of cellular senescence-related genes in COAD patients, we filtered scRNA-seq data by setting each gene expressed in at least three cells, and each cell expressing at least 250 genes. The percentage of mitochondria and rRNA in each cell was calculated using the PercentageFeatureSet function ensuring 100 < genes < 6,000 and mitochondrial content <5% in each cell. Data of 13 samples were normalized using log-normalization method, and the FindVariableFeatures function was used to identify variable features based on variance stabilization transformation ("vst") and select highly variable genes.

Transcriptome data collection and pre-processing

The gene expression profiles and clinical information of COAD were obtained from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/) project, including 432 tumor samples

and 41 para-carcinoma tissue samples. The RNA-seq data standardization method was TPM normalization. To process TCGA-COAD data, samples lacking clinical follow-up information, survival time, and survival status were eliminated from further analysis, and all samples with survival time more than 0 days. Ensembl gene IDs were further transformed into gene symbol IDs. Then, the gene with multiple gene symbol IDs was normalized as median. We also downloaded the gene expression profiles of 573 COAD samples in GSE39582 (Marisa et al., 2013) and 55 COAD samples in GSE17537 (Xiao et al., 2022) from GEO database. Among these, clinical follow-up information, survival time, and survival status were excluded from this study. We converted ensembl gene IDs to gene symbol IDs. The probe related to several genes was removed, and the gene with multiple probes was expressed as median.

Masked copy number segment data of COAD were collected from TCGA and progressed by gistic2 software. Single nucleotide variations (SNVs) data of COAD that was derived using mutect2 software were obtained from TCGA cohort. Moreover, we obtained methylation data from TCGA. Methylation data was processed with following steps: 1) KNN function in "impute" R package was used to complete the NA value. 2) We converted beta value to M value. 3) We removed the cross-reactive CpG sites as previously reported (Chen et al., 2013). 4) We removed the unstable genomic methylation sites, that was, removed the CpG sites and single nucleotide sites on the sex chromosome. 5) Tumor samples (solid tumors) were retained in this study.

Collection of senescence-related pathways

Senescence-related pathways were retrieved in the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org/gsea/index. jsp).

Screening for cell subpopulations and marker genes

Subsequently, all genes were scaled through the ScaleData function, and principal components analysis (PCA) was conducted to reduce the dimensionality. The FindNeighbors and FindClusters functions were used to cluster cells (Resolution = 0.1). Further, we reduced the t-distributed stochastic neighbor embedding (TSNE) dimensionality using RunTSNE founction and then annotated the cell subpopulations with some classic markers of immune cells (Zheng et al., 2022). The FindAllMarkers function was employed to identify marker genes with logFC = 0.5 and Minpct = 0.5 under the statistical threshold of adjusted $p < 0.05$. "clusterProfiler" package (Yu et al., 2012) was implemented for Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis.

Cellular senescence characteristics in tumor microenvironment (TME) of single cell

The number of DNA copies was calculated by "copycat" package (Gao et al., 2021) under the threshold of at least 5 genes in each chromosome. We distinguished aneuploidy (malignant cells) and diploid (non-malignant cells) with at least 25 genes selected for each segment and KS. cut = 0.15. We downloaded the cellular senescence-related pathways from gene set enrichment analysis (GSEA, http://www.gsea-msigdb.org/gsea/index.jsp), and calculated single sample GSEA (ssGSEA) scores of aneuploidy and diploid through "GSVA" package (Hänzelmann et al., 2013). The distribution was compared using the wilcox. test, and $p < 0.05$ was considered statistically significant.

Verification of cellular senescence based on bulk RNA-seq data

Furthermore, we used bulk RNA-seq data to analyze abnormal cellular senescence in tumor and normal COAD samples. GSEA was applied to performed pathway enrichment analysis, and ssGSEA scores of cellular senescence-related pathways were calculated in tumor and normal COAD samples. The distribution was compared using the wilcox. test.

Identification of senescence subtypes

Based on the above analysis, genes in GOBP_REPLICATIVE_SENESCENCE, REACTOME_CELLULAR_SENESCENCE, REACTOME_DNA_DAMAGE_TELOMERE_STRESS_INDUCED_SENESCENCE, and KEGG_P53_SIGNALING_PATHWAY were selected for univariate Cox regression analysis using "survival" package (Therneau and Lumley, 2015) in R. Candidates with $p < 0.05$ were considered as prognosis-related genes. A consensus clustering analysis was performed to categorize the 432 TCGA-COAD samples based on the expression profiles of the 16 senescence-related genes using "Consensus ClusterPlus" package (Wilkerson et al., 2013) with "Partitioning Around Medoids" (PAM) algorithm (Kaufman and Rousseeuw, 1990) and Euclidean distancing, in procedures with 500 bootstraps containing 80% COAD patients. 2–10 clusters were tested. The cumulative distribution function (CDF) and consensus matrix were performed identify the optimal subtypes. Kaplan-Meier curves of identified subtypes were generated in TCGA cohort and GSE39582 cohort.

Analysis of clinicopathologic characteristics among senescence subtypes

We further compared the distributions of clinicopathologic characteristics (gender, T stage, N stage, M stage, Stage, age, and survival status) among three senescence subtypes in TCGA cohort using Chi square test. Besides, the distributions of subtypes in T stage and survival status (alive or dead) were also analyzed using Sankey diagram.

Differences in mutation characteristics among senescence subtypes

We integrated copy number variations (CNVs) of TCGA-COAD patients through gistic2 software with a confidence level of 0.9 and hg38 as the reference genome to analyze the differences of CNVs among the three subtypes. "maftools" package (Mayakonda et al., 2018) was employed to analyze SNVs data in TCGA cohort.

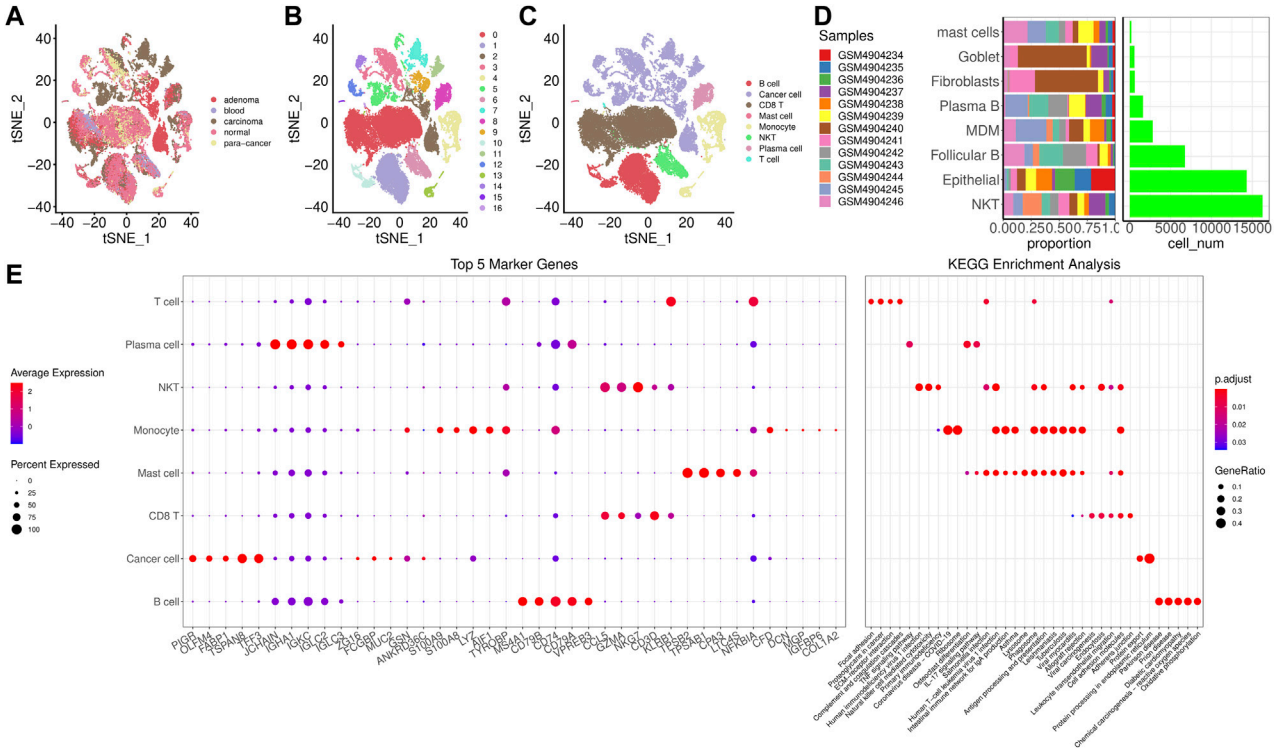


FIGURE 1
Screening for cell subpopulations and marker genes. **(A)**, TSNE diagrams of 5 samples. **(B)**, TSNE diagrams of 17 cell subpopulations. **(C)**, TSNE diagrams of 8 subpopulations after annotation. **(D)**, The number of cells of each subpopulation and its proportion in different samples. **(E)**, Top 5 maker genes in each subpopulation and the enrichment analysis.

Additionally, comparisons of TMB and the number of genetic mutations were carried out using wilcox. test among three subtypes.

Relationship between senescence subtypes and enriched pathway characteristics

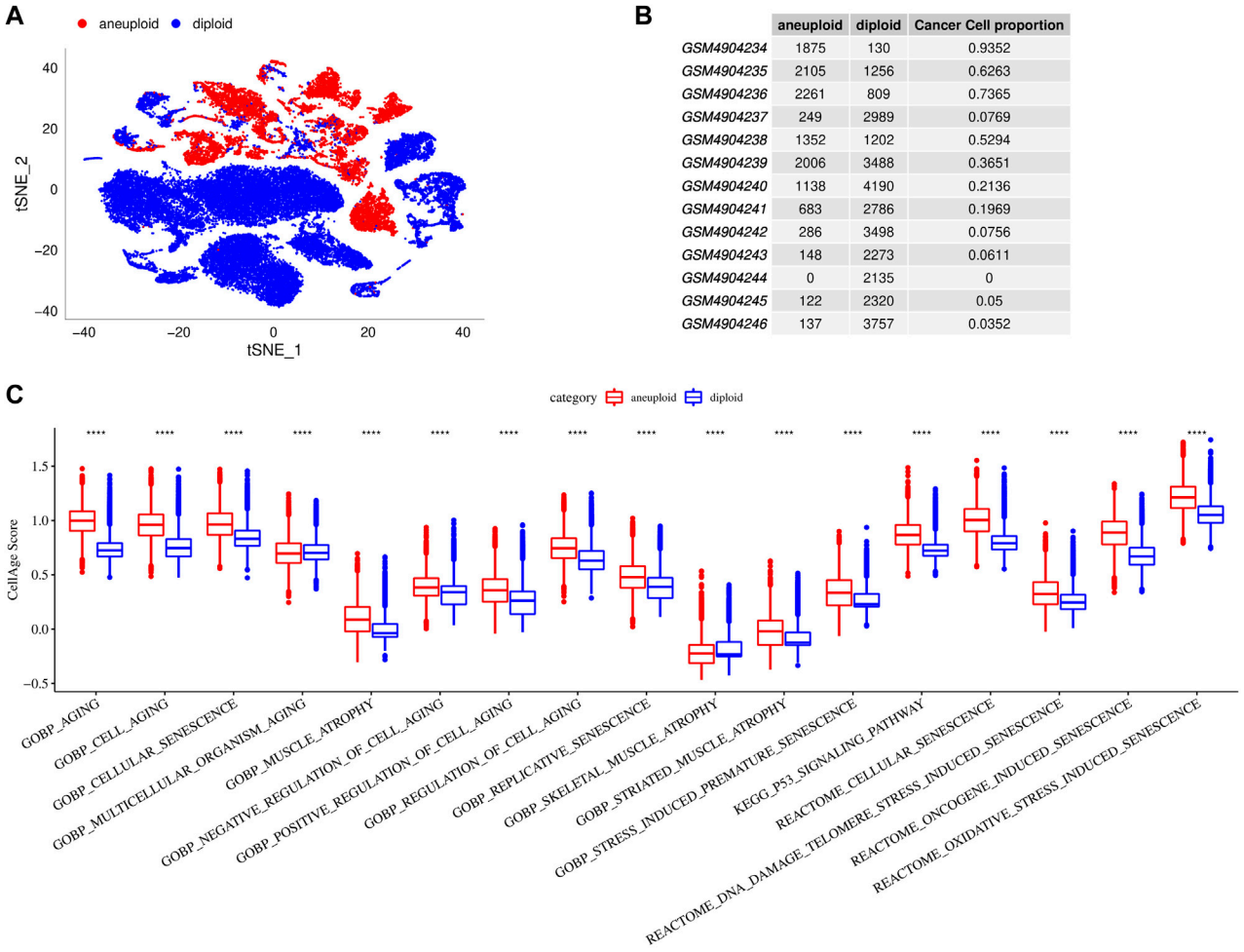
To evaluate the relationship between senescence subtypes and epithelial-to-mesenchymal transition (EMT), we calculated ssGSEA scores of EMT in each TCGA-COAD sample on basis of 200 genes of HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION in MSigDB (Yu et al., 2021). We calculated hypoxia score of genes of HALLMARK_HYPOXIA using ssGSEA method. Based on 24 genes as previously reported (Masiero et al., 2013), we scored angiogenesis by ssGSEA method. Differential analysis of these ssGSEA scores were performed using wilcox. test. Meanwhile, 10 tumor-related pathways were obtained (Sanchez-Vega et al., 2018) and the enrichment score was calculated by ssGSEA, followed by kruskal. test for comparisons.

Relationship between senescence subtypes and immune characteristics

We evaluated the immune cell infiltration in TCGA cohort by ESTIMATE algorithm, and calculated the score of 28 kinds of immune cells (Charoentong et al., 2017) by ssGSEA. Afterwards, we downloaded the genes related to inflammation through GSEA and calculated their ssGSEA scores. Comparisons were analyzed using kruskal. test. The tumor immune dysfunction and exclusion (TIDE) is a computational method that can determine the signatures of T cell dysfunction by using gene expression profiling in tumors interacts with the cytotoxic T lymphocytes infiltration level to affect patient survival and response to immunotherapy (Jiang et al., 2018). A high TIDE score indicates a low response rate to immune checkpoint inhibition (ICI) therapy. Thus, the TIDE algorithm (http://tide.dfci.harvard.edu/) was employed to predict the potential clinical effects of immunotherapy in subtypes.

Construction and validation of senescence-based risk model

To identify the differential expressed genes (DEGs), "limma" package (Ritchie et al., 2015) in R was applied to perform differential analysis when clust1 vs. non-clust1, clust2 vs. non-clust2 and clust3 vs. non-clust3. Under the threshold of $p < 0.05$ and $|\log2\ (\text{Fold Chage})| > \log2\ (1.5)$, 2,085 DEGs were identified and selected for univariate Cox regression analysis using coxph function embedded in "survival" package, and candidates with $p < 0.005$ were selected as genes that have greater impact on prognosis. To reduce the number of genes, the LASSO Cox regression was performed using "glmnet" package (Hastie et al., 2021) in R. Stepwise multivariate



	aneuploid	diploid	Cancer Cell proportion
GSM4904234	1875	130	0.9352
GSM4904235	2105	1256	0.6263
GSM4904236	2261	809	0.7365
GSM4904237	249	2989	0.0769
GSM4904238	1352	1202	0.5294
GSM4904239	2006	3488	0.3651
GSM4904240	1138	4190	0.2136
GSM4904241	683	2786	0.1969
GSM4904242	286	3498	0.0756
GSM4904243	148	2273	0.0611
GSM4904244	0	2135	0
GSM4904245	122	2320	0.05
GSM4904246	137	3757	0.0352

FIGURE 2
Cellular senescence characteristics in single cell TME. **(A)**, TSNE diagrams of malignant cells and non-malignant cells in single cell. **(B)**, Cell proportions of malignant cells and non-malignant cells in 13 samples. **(C)**, Comparisons of senescence-related pathway scores between malignant cells and non-malignant cells.

regression analysis with stepwise Akaike information criterion (AIC) was used to determine genes for risk model construction.

The risk score formula related to the prognostic signature was as follows: RiskScore = 0.417*MFNG + 0.424*GPRC5B + 0.137*TNNT1−0.389*CCL22 + 0.308*NOXA1 + 0.149*PABPC1L + 0.338*PCOLCE2 + 0.337*MID2−0.215*CPA3 + 0.261*HSPA1A + 0.161*CALB1. After calculating risk score in TCGA cohort, "timeROC" package (Blanche, 2015) was employed to carry out receiver operating characteristic (ROC) analysis with areas under the ROC curve (AUCs) for 1, 3, and 5 years. Finally, risk score was standardized as zscore, and TCGA-COAD samples were divided into high-risk group (zscore >0) and low-risk group (zscore <0). Kaplan-Meier curves were generated between high- and low-risk groups.

Associations of senescence-based risk score with clinicopathologic characteristics and biological characteristics.

To explore the relationship between RiskScore score and clinical characteristics of COAD patients, we analyzed the differences of risk score among clinicopathologic characteristics including gender, age, T stage, N stage, M stage, Stage, and clusters in TCGA-COAD cohort. Additionally, we performed correlation analysis between senescence-based risk score and biological characteristics (hypoxia, angiogenesis, and metastasis) with rcorr function in "Hmisc" package (Harrell and Harrell, 2019). Further, we used "GSVA" package to score pathways in KEGG, and performed correlation analysis between senescence-based risk score and pathways with $|cor| > 0.2$ and $p < 0.05$. We compared the scores of senescence-related pathways between high- and low risk groups. Wilcox. test was applied for comparisons.

Prediction of responsiveness to chemotherapy

To predict the responsiveness to traditional chemotherapy drugs, the half-maximal inhibitory concentration (IC50) values were evaluated using the "pRRophetic" package. Comparisons of IC50 values between high- and low-risk groups were performed using wilcox. tests.

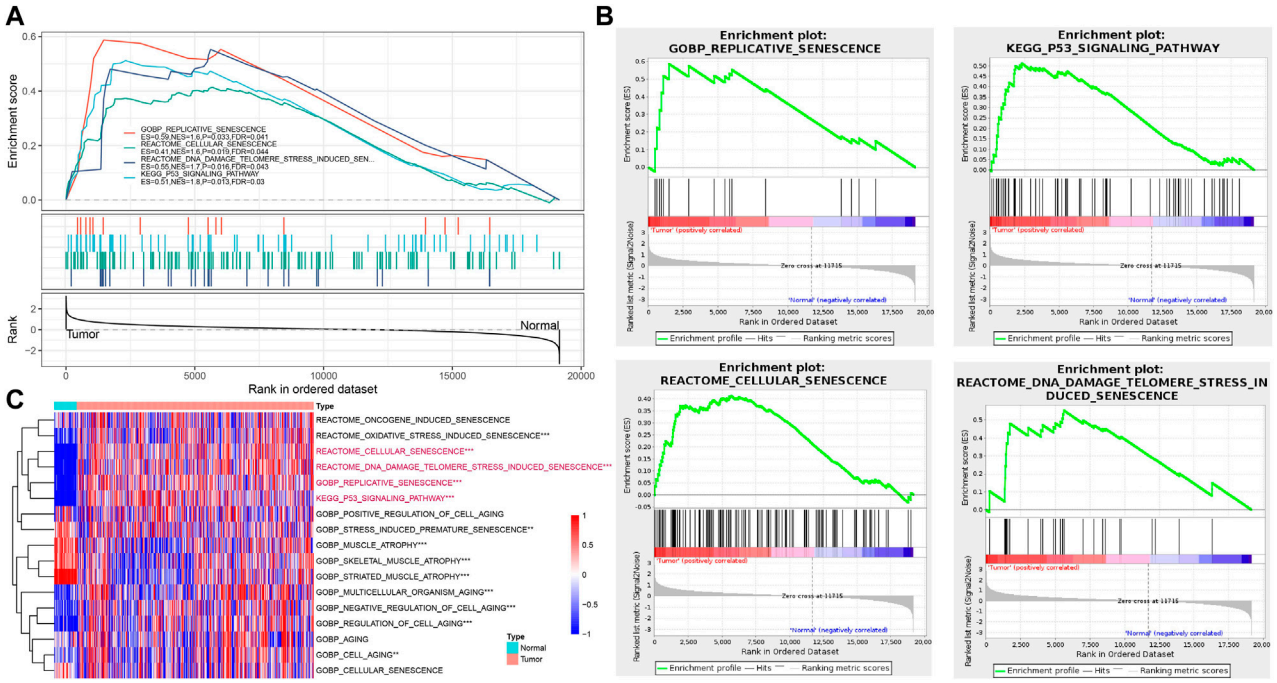


FIGURE 3
Verification of cellular senescence based on bulk RNA-seq data. **(A)**, The results of GSEA enrichment analysis in TCGA cohort. **(B)**, Four key pathways of GSEA analysis in TCGA cohort. **(C)**, Heatmap of ssGSEA scores of senescence-related pathways between tumor and para-cancer tissues in TCGA cohort. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

Relationship between risk gene expression and methylation

Based on the methylation data of the TCGA dataset, we constructed the methylation level of the CpG sites in the risk model and calculated the mean values of methylation level at different CpG sites of the same gene. The relationship between risk gene expression and the methylation level was analyzed using Pearson correlation analysis.

Construction of nomogram

Furthermore, the univariate and multivariate Cox regression analysis were utilized to determine whether senescence-based risk score is an independent predictor of prognosis. To predict the clinical outcomes of COAD patients, a nomogram based on risk score and clinicopathological characteristics was constructed with calibration curve. To evaluate the accuracy and reliability of this model, decision curve analysis (DCA) was established.

Statistical analysis

Data was processed and analyzed using (version 3.6.0, https://www.r-project.org/) and Seurat R package (Gribov et al., 2010) (version 3.6.3, https://satijalab.org/seurat/). Wilcox. test or kruskal. test was applied to determine the significant differences and $p < 0.05$ was considered statistically significant. Log-rank test was used to determine the statistically significant for Kaplan-Meier curves.

Results

Single cell RNA-seq analysis and marker gene recognition of COAD

Supplementary Figure S1A showed the cell number of 13 samples before and after filtering. As displayed in Supplementary Figure S1B, 13 samples overlapped significantly between the TSNE diagrams. After PCA for dimension reduction (Supplementary Figure S1C, D), we select dim = 35 for further analysis.

We clustered cells based on dim = 35 and obtained 17 cell subpopulations. Figure 1A showed t-SNE-maps of adenoma, blood, carcinoma, normal and para-cancer samples. Figure 1B portrayed 17 cell subpopulations after clustering. Then, we annotated the cell subpopulations with some classic markers of immune cells. Supplementary Figure S2 provided TSNE diagram of marker gene expression. Figure 1C showed the clustering characteristics of annotated cell subpopulations. Subpopulations 2, 3, 5, 7, 9, 11, and 12 were epithelial cells expressing EPCAM; Subpopulations 0 and 6 were natural killer (NK) T cells expressing CD3D, KLRD1, and CD8A. Subpopulations 1 and 10 were follicular B cells expressing

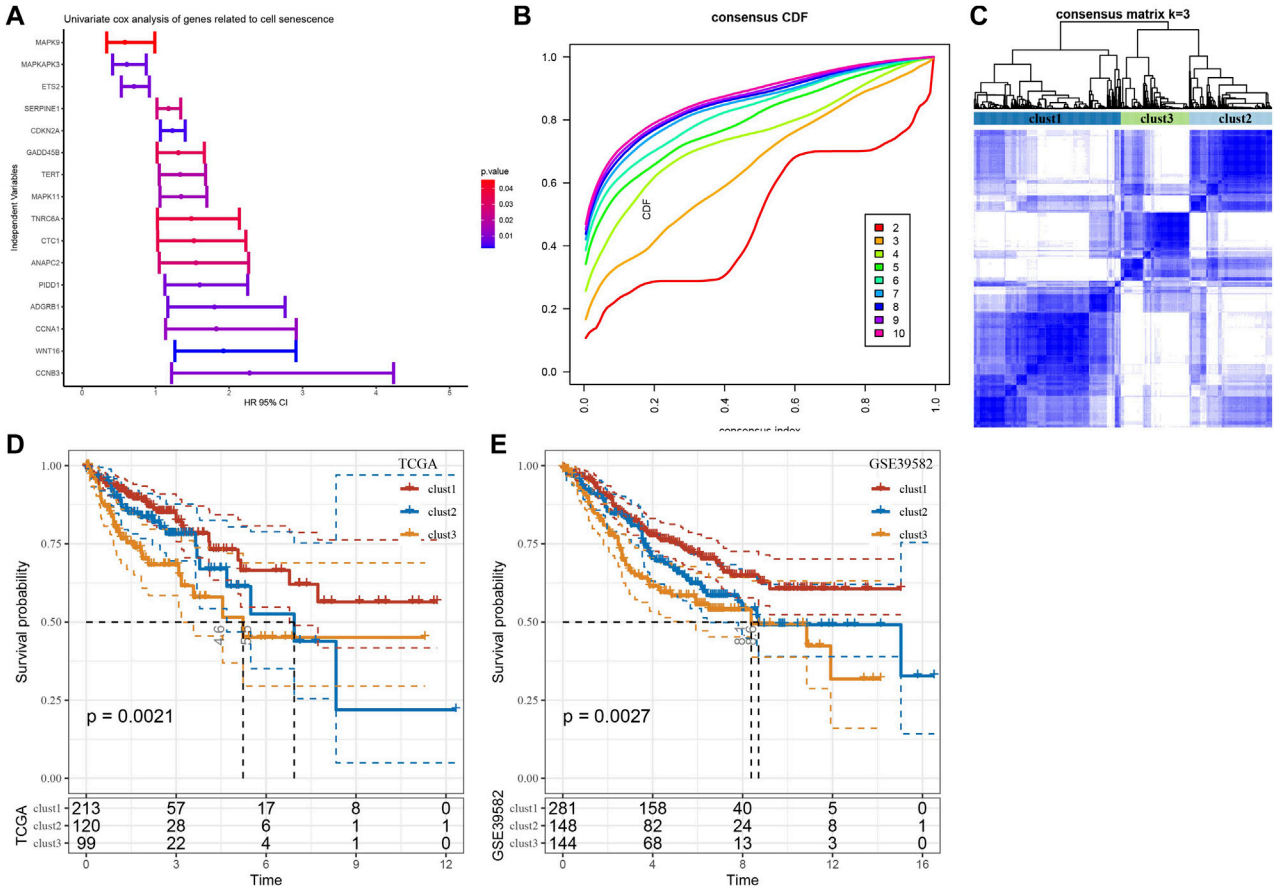


FIGURE 4
Three senescence subtypes were identified. **(A)**, Univariate cox regression analysis of senescence-related genes. **(B)**, Consensus CDF in TCGA cohort. **(C)**, Consensus matrix heatmap defining three clusters (k = 3). **(D-E)**, Kaplan-Meier curves of three subtypes in TCGA cohort and in GSE39582 cohort.

MS4A1; Subpopulation 8 was plasma B cells expressing MZB1; Subpopulation 4 was monocyte derived macrophages (MDMD) expressing CD68, CD14, and FCGR3A; Subpopulation 13 was fibroblasts expressing DCN; Subpopulations 15 and 16 were mast cells expressing KIT. Accordingly, we counted the number of cells of each subpopulation and calculated its proportion in different samples (Figure 1D). The subpopulations epithelial and NK T had larger number of cells than others. Figure 1E showed top 5 maker genes in each subpopulation and the enrichment analysis showed that marker genes were closely associated with human T-cell leukemia virus 1 infection, Th17 cell differentiation, hematopoietic cell lineage, and Th1 and Th2 cell differentiation.

Cellular senescence characteristics in TME

To characterize cellular senescence in TME of single cell, we distinguished aneuploidy and diploid in cell subpopulations. The results revealed that there were 12,362 aneuploid (malignant cells) and 30,833 diploid (non-malignant cells), and their TNSE-maps were shown in Figures 2A, B suggested that there are more malignant cells in cancer tissues, but fewer malignant cells in para-cancer tissues. Further we calculated the cellular senescence-related pathway scores using the ssGSEA method in malignant cells and non-malignant cells. Higher scores of senescence-related pathways were found in malignant cells than that of non-malignant cells ($p < 0.0001$) (Figure 2C).

Cellular senescence was verified based on bulk RNA-seq data

To further verify the cellular senescence characteristics, we evaluated the senescence-related pathways in tumor and para-cancer tissues based on bulk RNA-seq data. GOBP_REPLICATIVE_SENESCENCE, REACTOME_CELLULAR_SENESCENCE, REACTOME_DNA_DAMAGE_TELOMERE_STRESS_INDUCED_SENESCENCE, and KEGG_P53_SIGNALING_PATHWAY were significantly enriched in tumor tissues in TCGA cohort (Figures 3A, B). Through calculating senescence-related pathway scores using the ssGSEA method, several pathways including RGOBP_REPLICATIVE_SENESCENCE, REACTOME_CELLULAR_SENESCENCE, REACT OME_DNA_DAMAGE_TELOMERE_

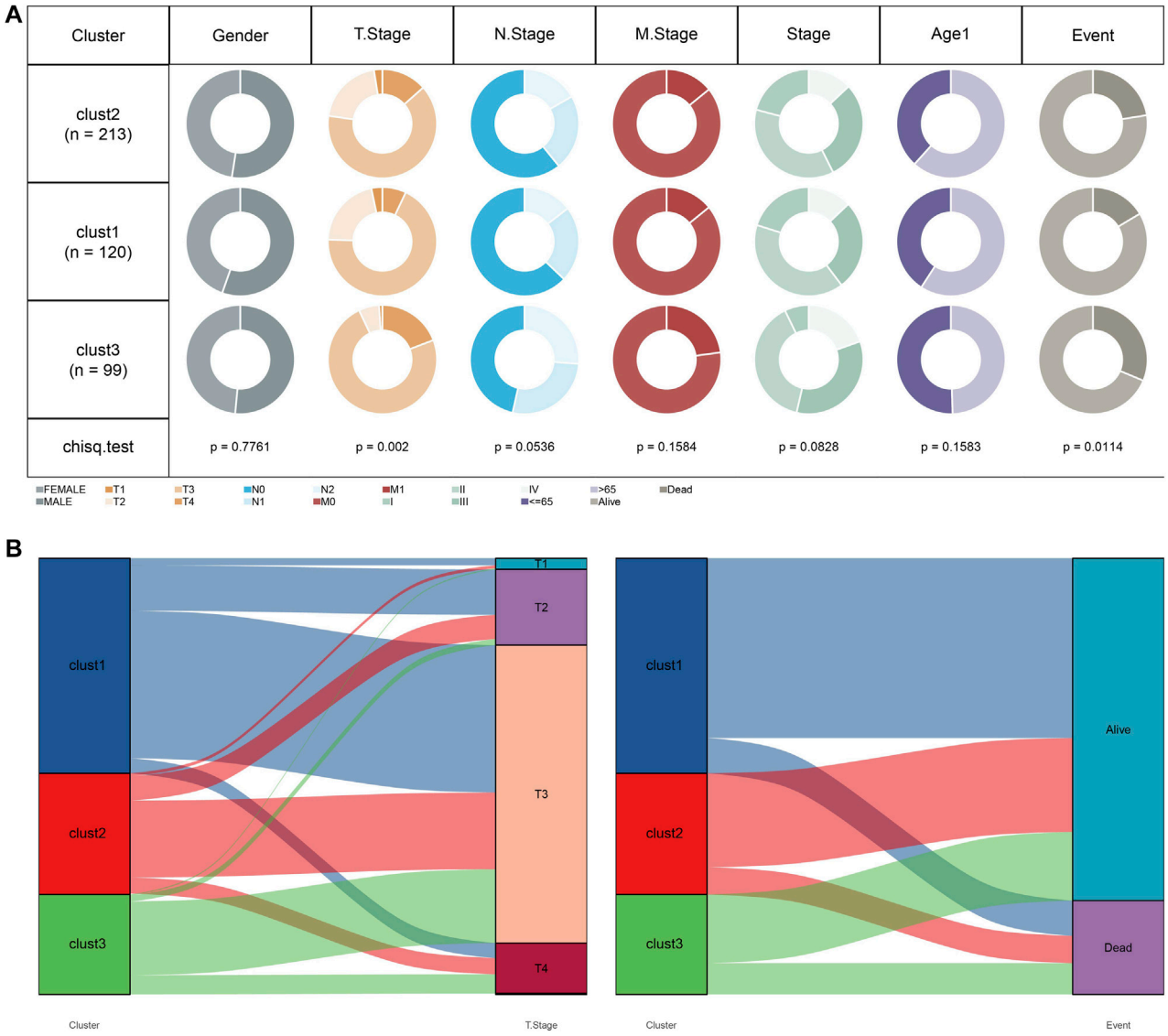


FIGURE 5
Clust3 had poorest prognosis and higher T stage. **(A)**, Distribution of clinicopathologic characteristics among three subtypes in TCGA cohort. **(B)**, Sankey diagram detailed the distribution of three subtypes in T stage and survival status.

STRESS_INDUCED_SENES CENCE, and KEGG_P53_SIGNALING_PATHWAY had higher senescence scores in tumor tissues than that of para-cancer tissues ($p < 0.001$) (Figure 3C).

Three senescence subtypes were identified

We performed univariate Cox regression analysis using genes from these four enriched pathways above (Supplementary Table S1). A total of 16 genes associated with prognosis were identified (Figure 4A). To further identify the subtypes, a consensus clustering analysis was conducted to categorize the 432 TCGA-COAD samples based on the expression profiles of the 16 senescence-related genes. From the results of CDF Delta area, cluster = 3 had a relatively stable clustering effect (Figure 4B). Considering that consensus matrix k = 3 is a preferable choice, we divide the whole cohort into three subtypes (Figure 4C). Next, Kaplan-Meier curves revealed significant variations among the three subtypes, and clust3 had the lowest survival probability while clust1 had the best prognosis in TCGA ($p = 0.0021$) (Figure 4D). Similar results were also observed in GSE39582 cohort ($p = 0.0027$) (Figure 4E).

Clust3 had poorest prognosis and higher T stage

We subsequently compared the distribution of clinicopathologic characteristics (gender, T stage, N stage, M stage, Stage, age, and survival status) among three subtypes in TCGA cohort. The results found

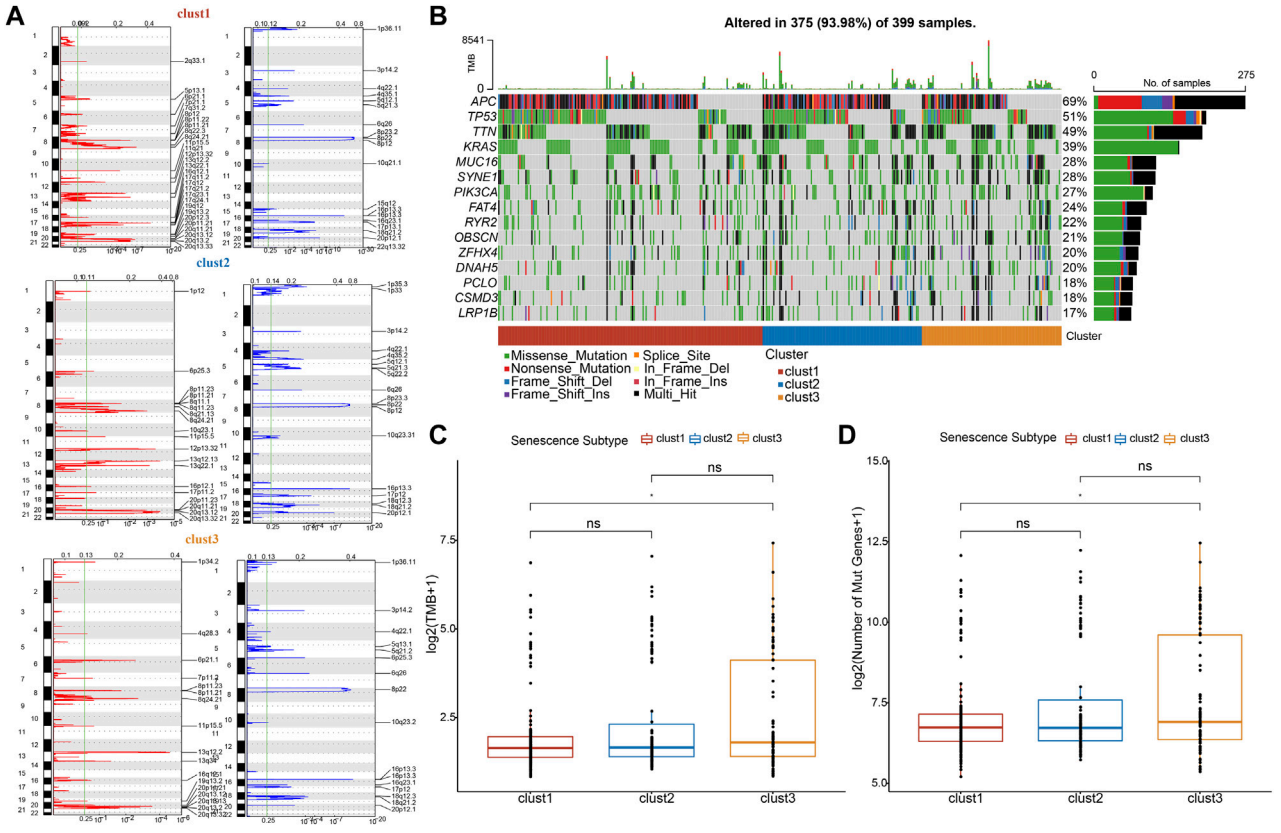


FIGURE 6
Clust3 exhibited higher TMB and mutations. **(A)**, Peaks of CNVs, amplified (red) genes and deleted (blue) genes among three subtypes. **(B)**, Top 15 mutated SNV genes among three subtypes. **(C)**, TMB alterations among three subtypes. **(D)**, The number of mutated genes among three subtypes. Ns represents $p > 0.05$; *$p < 0.05$.

significant differences in T stage and survival status among the three subtypes (Figure 5A). Additionally, Sankey diagram detailed the distribution of three subtypes in T stage and survival status (Figure 5B). The patients with clust3 had poorest prognosis and higher T stage (predominantly in T3 and T4 stage).

Clust3 exhibited higher TMB and mutations

Furthermore, mutation characteristics were further evaluated among senescence subtypes. The CNVs were remarkably changed among three subtypes (Figure 6A). Meanwhile, the results from SNVs showed APC (69%), TP53 (51%), TTIN (49%), and KRAS (39%) exhibited higher mutation frequencies among top 15 mutated genes (Figure 6B). Besides, TMB and the number of mutated genes were both increased in clust3 compared with that of clust1 (Figures 6C, D).

Senescence subtypes were associated with EMT, hypoxia, angiogenesis and tumor-related pathways

In tumors, cellular senescence promotes the extracellular matrix cleavage resulting in growth factors release that can promote epithelial-to-mesenchymal transition (EMT), which leads to tumor metastasis. Hence, we clarified the relationship between senescence subtypes and EMT score. EMT score was distinctly different among senescence subtypes and clust3 has the highest EMT score compared with that of clust1 and clust2 (Figure 7A). At the same time, hypoxia and angiogenesis scores were higher in clust3 than that of clust1 and clust2 (Figures 7B, C). Figure 7D found that 9 tumor-related pathways were significantly altered in the three subtypes, including cell cycle, HIPPO, MYC, NOTCH, NRF1, PI3K, TGF-beta, RAS, TP53 and WNT.

Relationship between senescence subtypes and immune characteristics

Immune infiltration scores including StromalScore, ImmuneScore, and ESTIMATEScore were remarkably different among three subtypes (Figure 8A), and we found that clust3 had a higher degree of immune infiltration. Significant changes in immune cells infiltration were found among the three subtypes (Figure 8B). Clust3 also had higher scores of several inflammation-related pathways such as JAK-STAT signaling pathway, NF-Kappa B signaling pathway, Toll-like receptor signaling pathway, B cell receptor signaling pathway, T cell receptor signaling pathway, and inflammatory response (Figures 8C–H). Furthermore,

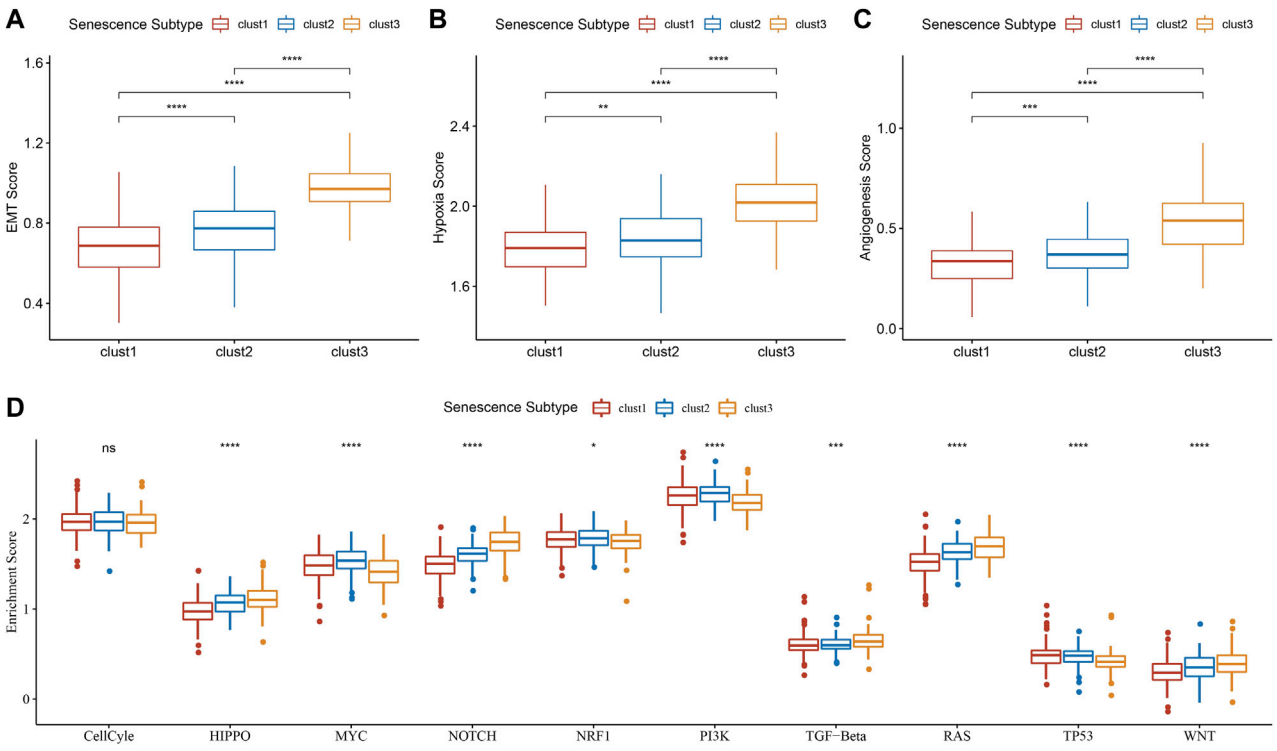


FIGURE 7
Association of senescence subtypes between EMT, hypoxia, angiogenesis and tumor-related pathways. **(A)**, Box plots of EMT score among three subtypes in TCGA cohort. **(B)**, Box plots of hypoxia score among three subtypes in TCGA cohort. **(C)**, Box plots of angiogenesis score among three subtypes in TCGA cohort. **(D)**, Box plots of 10 tumor-related pathways among three subtypes in TCGA cohort. Ns represents $p > 0.05$; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$.

TIDE score was higher in clust3 (Figure 8I), indicating more prone to immune escape of clust3.

Construction and validation of senescence-based risk model

Through differential analysis among the three subtypes, 2,085 DEGs were identified, which were used for univariate Cox regression analysis. 194 genes that have greater impact on prognosis were selected, including 180 risk genes and 14 protective genes (Figure 9A). To reduce the number of genes, LASSO Cox regression was performed. With the gradual increase of lambda, the number of independent variable coefficients tending to zero increased gradually (Figure 9B). 10-fold cross-validation was utilized and the confidence interval under each lambda was shown in (Figure 9C). When lambda = 0.0347, 25 genes were selected for further analysis. Based on stepwise multivariate regression analysis with AIC, 11 genes were finally identified (MFNG, GPRC5B, TNNT1, CCL22, NOXA1, PABPC1L, PCOLCE2, MID2, CPA3, HSPA1A, and CALB1). Subsequently, survival analysis in TCGA cohort revealed that patients with high risk had lower prognosis than that of patients with low risk ($p < 0.0001$) with 1 year AUC of 0.81, 3-year AUC of 0.77, and 5-year AUC of 0.75 (Figure 9D). To validate its robustness, survival analysis was performed in GSE39582 and GSE17537 cohort. High risk patients had lower prognosis than that of low risk patients in GSE39582 ($p < 0.0001$) and in GSE17537 ($p = 0.006$) with good performance in prognosis prediction (Figures 9E, F).

Associations of risk score with clinicopathologic characteristics and biological characteristics

To further clarify the relationship between risk score and clinicopathologic characteristics, we compared the differences of risk score in clinicopathologic characteristics in TCGA cohort and found that patients with higher clinical stage (T stage, N stage, M stage and Stage) had higher risk scores (Figure 10). Besides, patients with clust3 had higher risk score (Figure 10). To evaluate the relationship between risk score and biological characteristics, we performed the correlation analysis of risk score with hypoxia, angiogenesis, and EMT scores. Figures 11A–C showed that hypoxia, angiogenesis, and EMT scores were both positively correlated with risk score. Next, we performed correlation analysis between senescence-based risk score and underlying regulatory KEGG pathways to find risk score-related pathways (Figure 11D) and further statistics

10.3389/fphar.2023.1121634

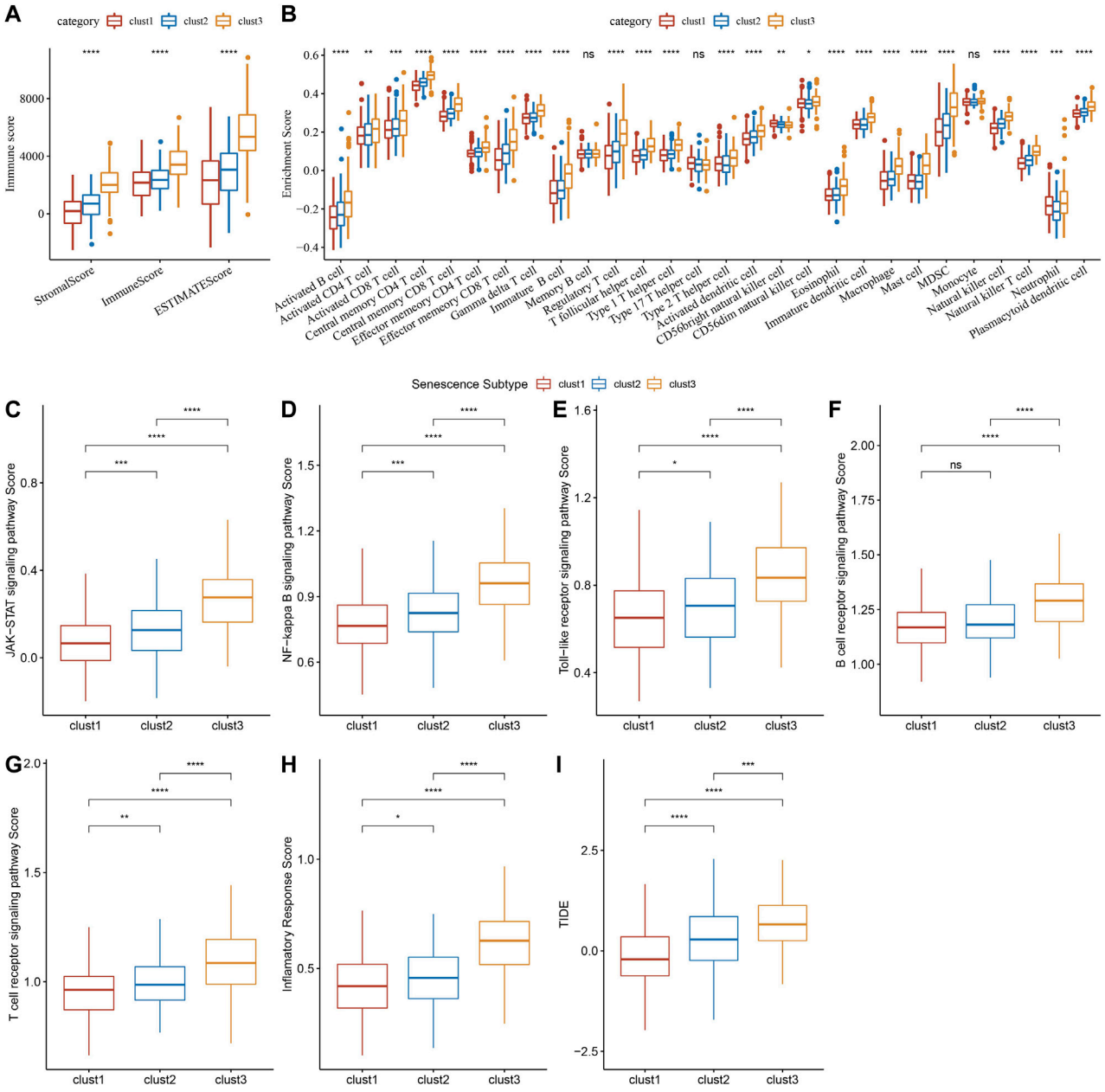


FIGURE 8
Association of senescence subtypes with immune characteristics. **(A)**, Box plots of immune infiltration scores among three subtypes. **(B)**, Box plots of immune infiltration cells among three subtypes. **(C–H)**, Box plots of JAK-STAT signaling pathway, NF-kappa B signaling pathway score, toll-like receptor signaling pathway, B cell receptor signaling pathway, T cell receptor signaling pathway and inflammatory response scores among three subtypes. **(I)**, Alteration of TIDE score among three subtypes. Ns represents $p > 0.05$; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$.

revealed that only some pathways were significant different between high- and low-risk group (Figure 11E). Moreover, we compared the scores of senescence-related pathways between high- and low risk groups. The results showed that high risk patients exhibited higher scores in GOBP_AGING, GOBP_MUSCLE_ATROPHY, and GOBP_NEGATIVE_REGULATION_OF_CELL_AGING; while low risk patients had higher scores in GOBP_STRESS_INDUCED_PREMATURE_SENESCENCE, KEGG_P53_SIGNALING_PATHWAY, and REACTOME_DNA_DAMAGE_TELOMERE_STRESS_INDUCED_SENESCENCE (Supplementary Figure S3).

Prediction of responsiveness to chemotherapy

Furthermore, we assessed the responsiveness to traditional chemotherapy drugs between high- and low-risk groups. As

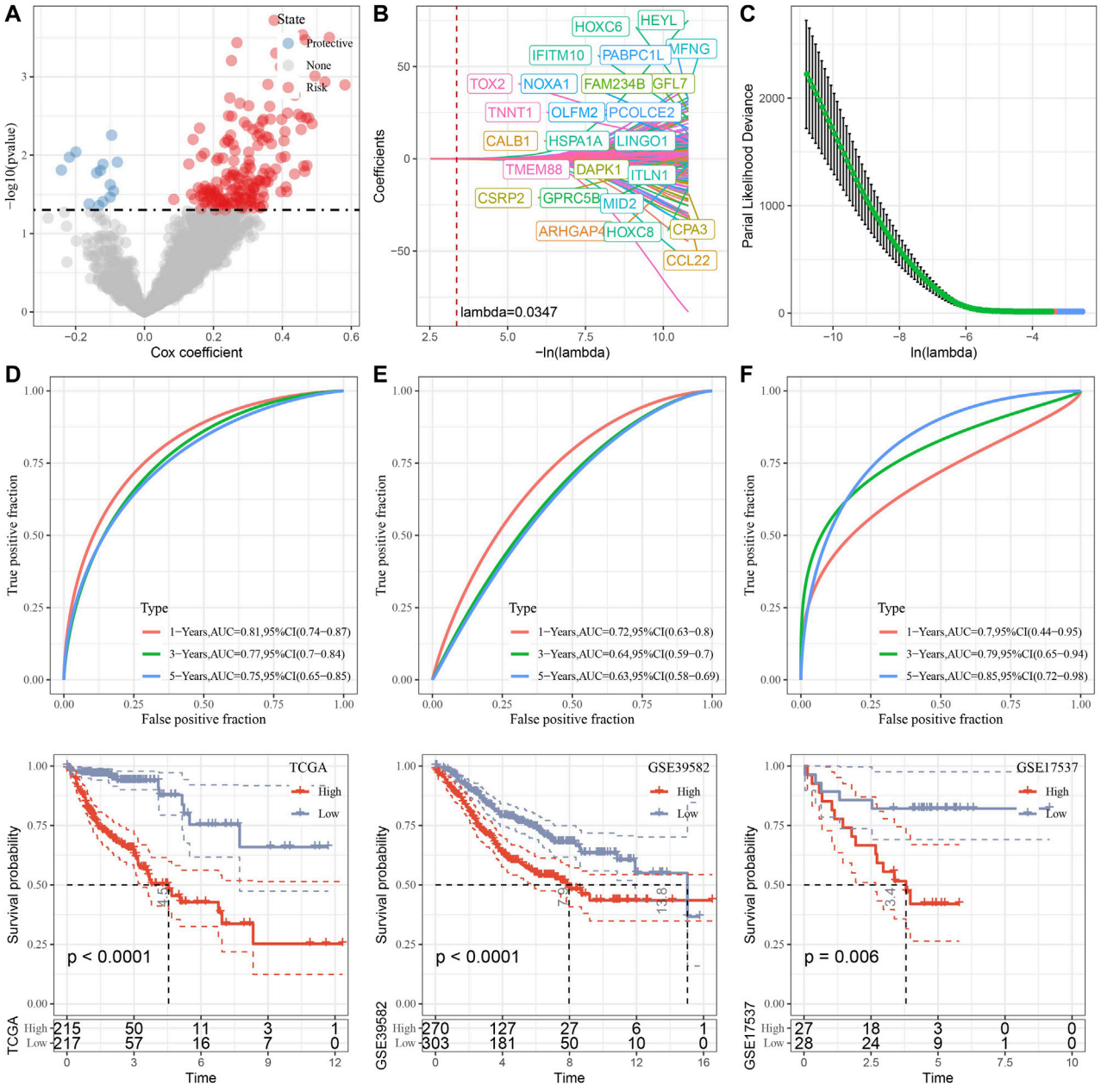


FIGURE 9
Construction and validation of senescence-based risk model. **(A)**, A total of 2,085 promising candidates were identified among the DEGs. **(B)**, Independent variable coefficients changed with lambda increase. When lambda = 0.00347, 25 genes were identified. **(C)**, 10-fold cross validation to determine the confidence interval under each lambda. **(D–F)**, Survival analysis with ROC curves and Kaplan-Meier curves in TCGA cohort, GSE39582 cohort and GSE17537 cohort.

displayed in Figure 12, low risk patients exhibited significant lower IC50 values of Erlotinib ($p < 0.001$), Sunitinib ($p < 0.001$), MG-132 ($p < 0.001$), CGP-082996 ($p < 0.01$), AZ628 ($p < 0.01$), Sorafenib ($p < 0.001$), VX-680 ($p < 0.01$), and Z-LLNle-CHO ($p < 0.01$) than that of high risk patients, which indicated that low risk patients were more sensitive to Erlotinib, Sunitinib, MG-132, CGP-082996, AZ628, Sorafenib, VX-680, and Z-LLNle-CHO.

Relationship between risk gene expression and methylation level

Moreover, we analyzed the methylation level for 11 risk genes in TCGA (Supplementary Figure S4) and performed correlation analysis between risk gene expression and methylation level. Figure 13 displayed that the expression of CALB1, CPA3, NOXA1, and TNNT1 had significant positive

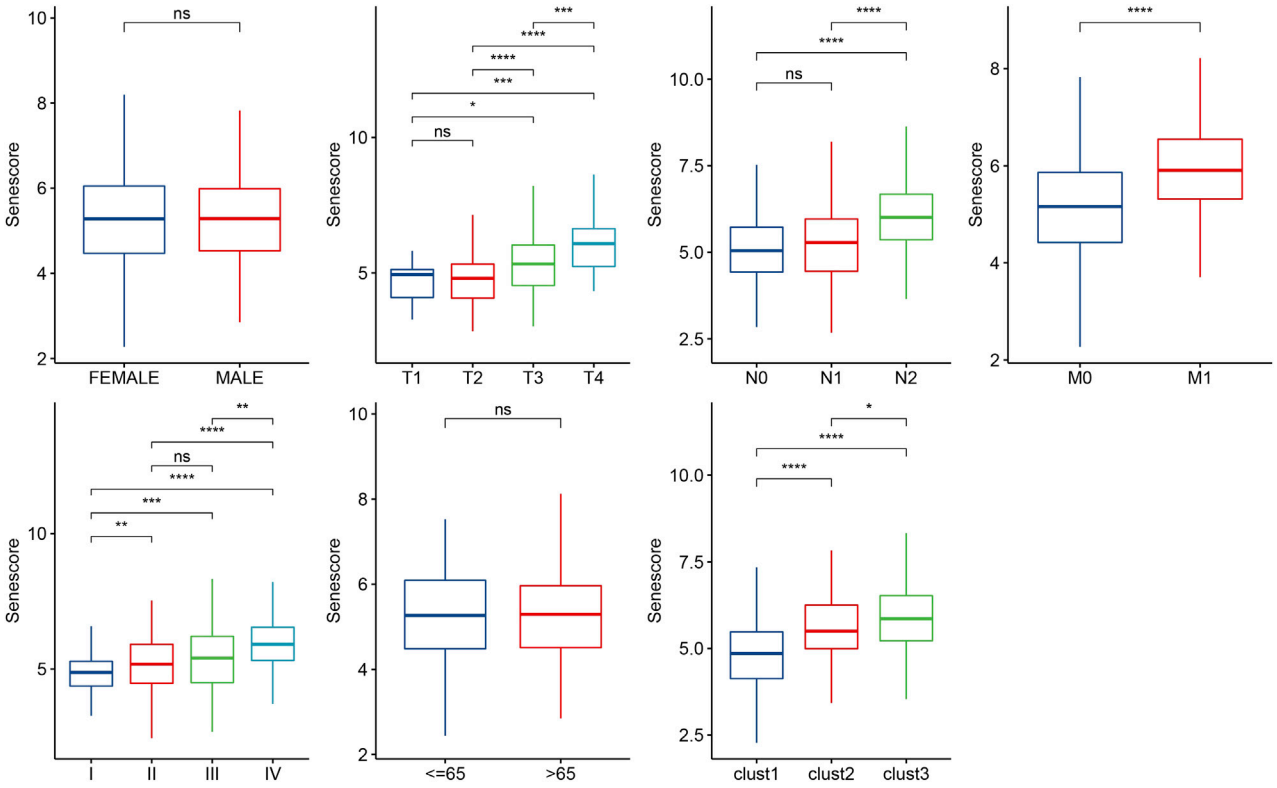


FIGURE 10
Distribution of risk score in different clinicopathologic characteristics. Ns represents $p > 0.05$; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$.

correlation with their methylation level; whereas the expression of CCL22, GPRC5B, HSPA1A, MFNG, PABPC1L, and PCOLCE2 had significant negative correlation with their methylation level.

Construction of nomogram

On basis of univariate and multivariate Cox regression analysis, the results showed that risk score was an independent prognostic factor (Figure 14A). To better quantify the risk assessment and survival probability of COAD patients, we constructed a nomogram to estimate 1-, 3-, and 5-year OS using the risk score and clinicopathological characteristics. Figure 14B revealed that risk score had the most impact on OS of COAD patients. The calibration curves of this nomogram showed high consistency between the observed and predicted values (Figure 14C). To evaluate the reliability of this model, DCA analysis was conducted and confirmed that both nomogram and risk score had the most powerful in predicting prognosis (Figure 14D).

Discussion

Senescent cells are closely related to aging and pathological status. Cellular senescent is crucial in tumorigenesis through the SASPs and the heterogeneity of senescence-associated genes promotes the progression of tumor and its escape from anti-tumor therapy (Junaid et al., 2022). Thus, it is believed that cellular senescence is involved in cancer heterogeneity. A previous study has identified several senescence-associated gene signatures using transcriptome data from TCGA database, which can predict clinical outcomes and responses to immunotherapy in patients with head and neck squamous cell carcinoma (Wang et al., 2022). Another study has developed the senescence-related subtypes, established a prognostic risk model, and further revealed their potential roles in TME in breast cancer only using transcriptome data from GEO cohort (Zhou et al., 2022). In the present study, we downloaded the scRNA-seq data of 13 COAD samples from GEO database and demonstrated that senescence-related pathways were highly expressed in malignant cells than that of non-malignant cells, indicating that cellular senescence was largely associated with heterogeneity of TME in COAD at single cell level. Furthermore, we identified three senescence subtypes and found clust3 had poorest prognosis, manifested with higher T stage, elevated TMB, increased pathway scores (EMT, hypoxia and angiogenesis), activated inflammatory response, and immune cell infiltration as well as immune escape tendency.

Immune cell infiltration serves as an indicator of the immune microenvironment in tumor. Notably, we found that

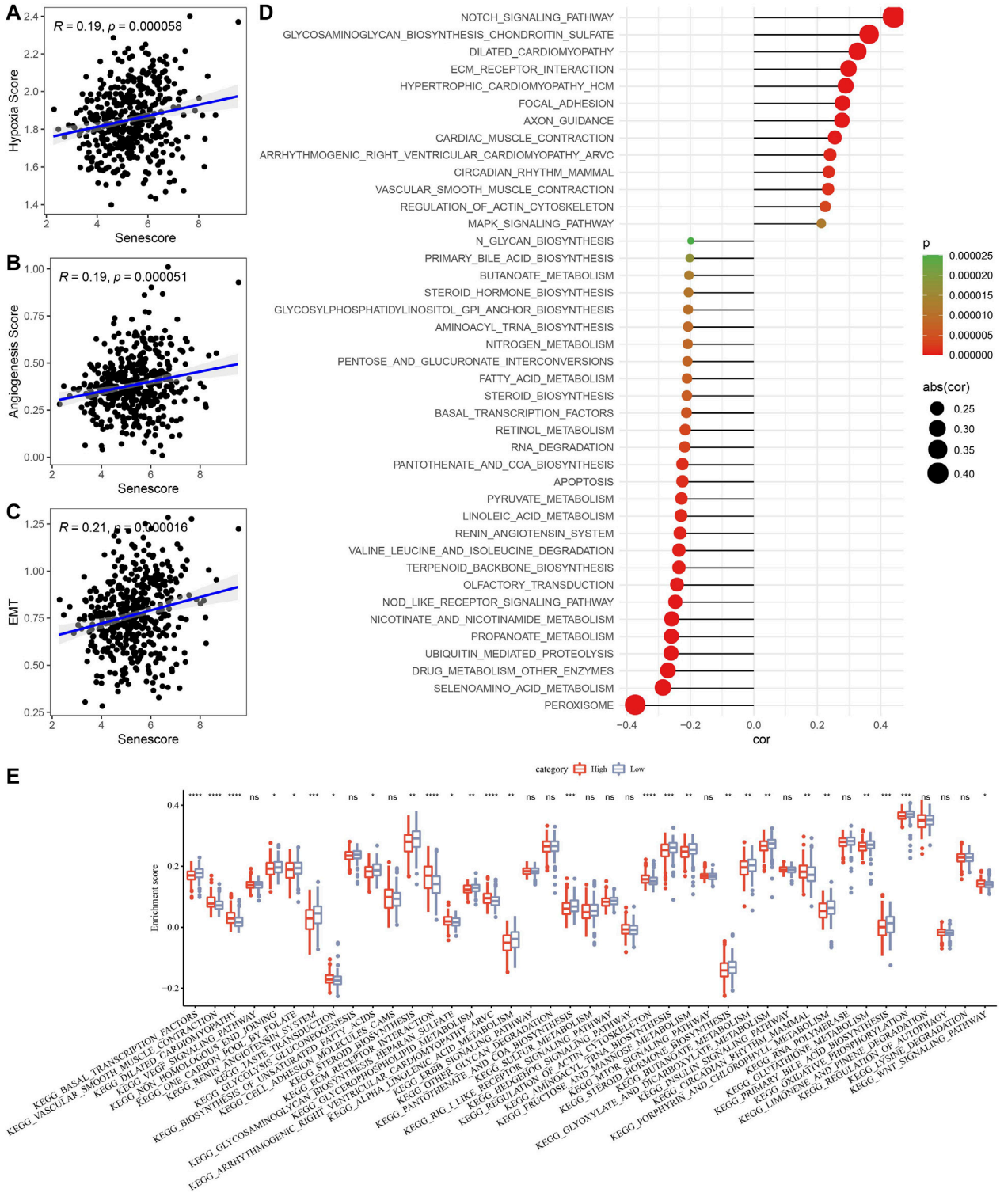


FIGURE 11
Associations of risk score with biological characteristics. **(A–C)**, Correlation analysis of risk score with hypoxia, angiogenesis, and EMT scores. **(D)**, Scatter plots of correlation between risk score and underlying regulatory KEGG pathways. **(E)**, Box plots of underlying regulatory KEGG pathway scores between high- and low-risk group. Ns represents $p > 0.05$; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, and ****$p < 0.0001$.

clust3 patients with poor prognosis had extremely higher expression levels of myeloid-derived suppressor cells (MDSCs) and macrophages than that of other subtypes. It has been reported that MDSCs are immature myeloid cells with heterogeneity, and its accumulation suppress anti-tumor immunity particular suppressing T cells in cancer patients

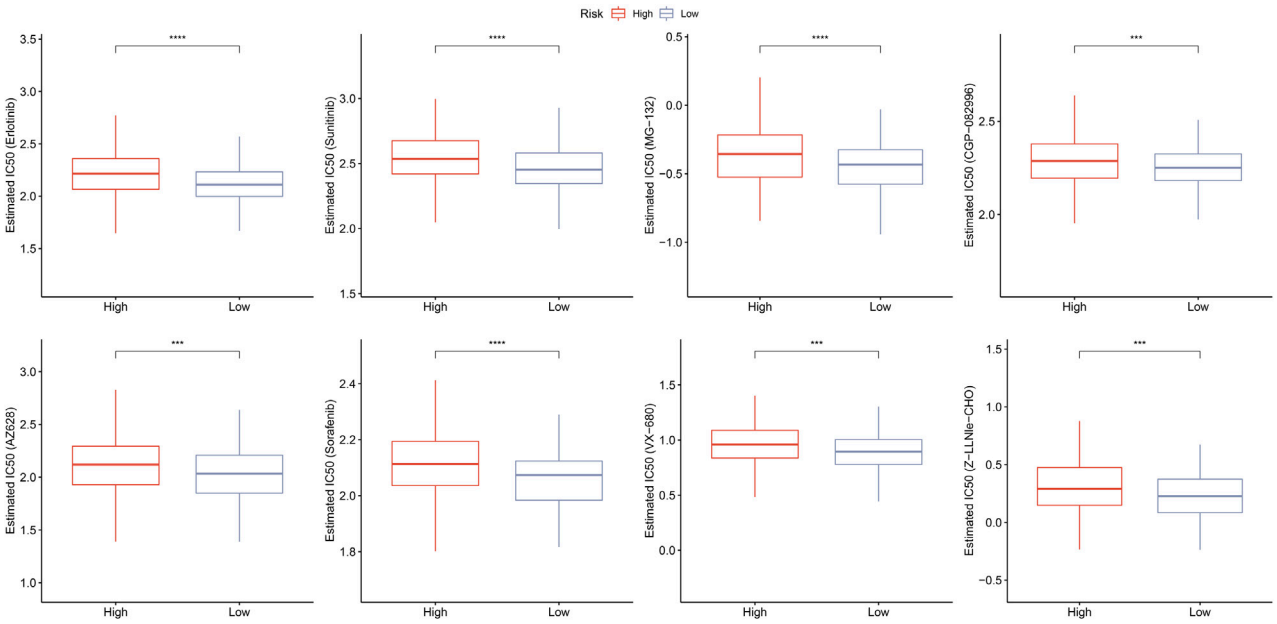


FIGURE 12
Prediction of responsiveness to chemotherapy. Estimated IC50 values for traditional chemotherapy drugs including erlotinib, sunitinib, MG-132, CGP-082996, AZ628, sorafenib, VX-680, and Z-LLNle-CHO. ***$p < 0.001$, and ****$p < 0.0001$.

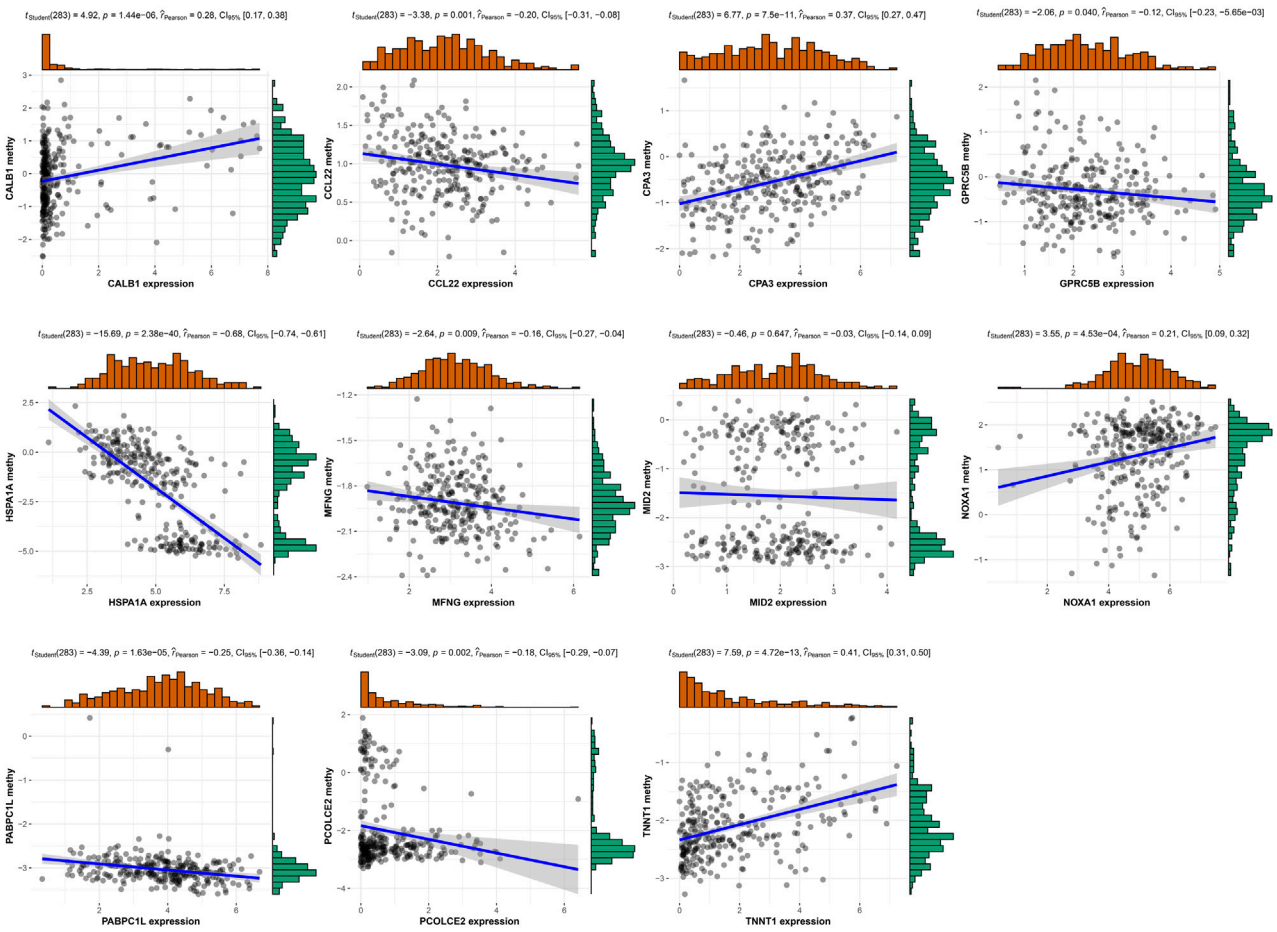


FIGURE 13
Correlation analysis was performed between risk gene expression and methylation level.

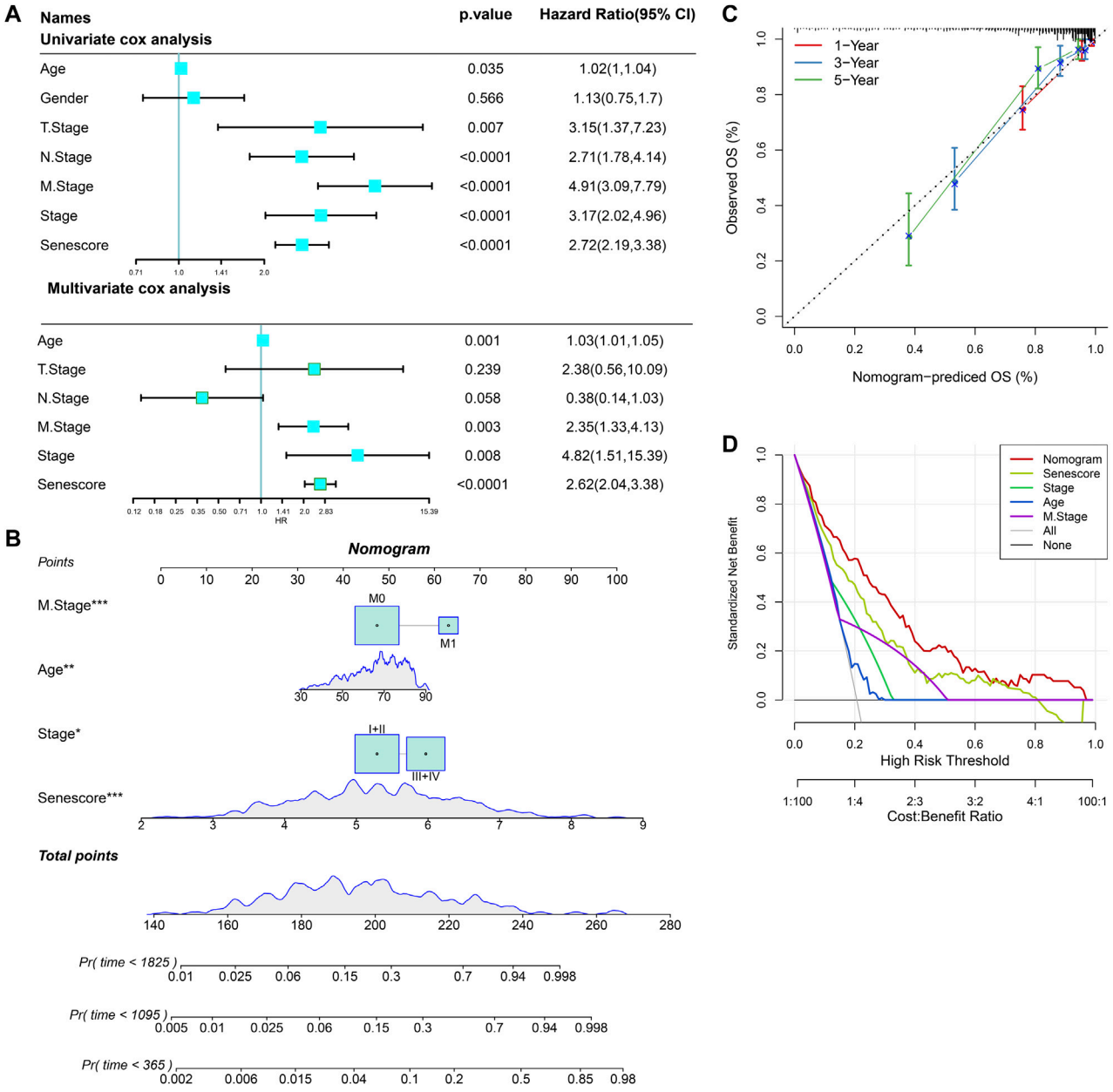


FIGURE 14
Nomogram construction for predicting prognosis of COAD patients. **(A)**, Univariate and multivariate Cox regression analysis of prognostic values of risk score and clinicopathological characteristics. **(B)**, Nomogram for predicting the 1-, 3-, and 5-year OS of COAD patients. **(C)**, Calibration curves for validating the established nomogram. **(D)**, Decision curve analysis o of nomogram. *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

(Ma et al., 2019). Additionally, MDSCs also directly promote tumor growth and metastasis. MDSCs exert these effects mainly through inhibiting T cell proliferation and T cell migration, triggering apoptosis of T cells and NK cells, suppressing immune effector cell functions, and repressing anti-tumor T cell-mediated reactivity by interaction with PD-1 receptor (Umansky et al., 2016). Moreover, macrophage infiltration in solid tumors accounts for poor outcomes and correlates with chemotherapy resistance in most cancers, which contributes to development and progression of cancer *via* provoking angiogenesis, metastasis, and immunosuppression (Cassetta and Pollard, 2018). In this study, highly expressed MDSCs might exert inhibitory effects on T cells and NK cells, inducing a decline of immune function in TME of COAD patients. Besides, macrophage infiltration might induce the suppression of immunity. We also found that patients with poor prognosis had a high TIDE score that represented a low response rate to ICI therapy. Collectively, MDSC infiltration,

macrophage infiltration and low response rate to ICI therapy contribute to poor clinical outcome of COAD patients in clust3. Synergistically, EMT, hypoxia, angiogenesis, and activated inflammatory response were responsible for poor prognosis of COAD patients.

Highly variant tumors are considered to have an increased burden of new antigens that may lead to immunogenicity. It has been recognized that TMB is a potential immune-response marker predicting ICI therapy (Choucair et al., 2020). As a tumor suppressor gene, APC is highly mutated in CRC (about 70%), and its mutation is important in colorectal tumorigenesis (Zhang and Shay, 2017). Recently, a study has analyzed the clinical characterizes and gene mutations in APC-mutant type and APC-wild-type Chinese CRC patients and confirms that APC mutation can be used as a promising biomarker to predict the immunotherapy responsiveness (Feng et al., 2022b). Another tumor suppressor gene TP53, is thought to be a major driver for CRC with approximately 50% mutation frequency (Timar and Kashofer, 2020). It has been reported that TP53 mutation is closely associated with rectum tumor, advanced stage and dismal prognosis of CRC patients (Li, 2019). Additionally, RAS is the most frequent mutated gene in human cancers. Interestingly, KRAS and APC are very common co-mutated (about 80%) and co-mutation of KRAS with TP53 is about 40% in CRC (Timar and Kashofer, 2020). In this study, APC (69%), TP53 (51%), TTIN (49%), and KRAS (39%) exhibited higher mutation frequencies among top 15 mutated genes. The TMB and the number of mutated genes were both increased in clust3 than that of clust1. These results indicated that the higher TMB and large mutations contributed to poor prognosis of COAD patients. Notably, we found that TTIN (49%) was highly mutated except for the known APC, TP53, and KRAS, implying that TTIN mutation is a potential genetic alteration of COAD heterogeneity.

Furthermore, 11-senescence-related gene-based prognostic risk model was established and patients with high risk had lower prognosis. MFNG is a kind of glycosyltransferases that activates Notch signaling and plays an important role in breast cancer, whereas inhibition of MFNG may attenuate the triple-negative breast cancer (Mugisha et al., 2022). A whole transcriptomics analysis has revealed that GPRC5B is elevated in immuno-activated breast cancer cells, while apigenin induces a 94% reduction in GPRC5B expression (Bauer et al., 2019). A recent research of RNA-seq has showed TNNT1 is a representative prognostic mRNAs that is associated with the prognosis of CRC patients (Deng et al., 2022). It has confirmed that high expression of CCL22 is related to a better prognosis in patients with colon cancer, and CCL22 as a prognostic DEG is used to construct a cellular senescence-related risk model in colon cancer (Dai et al., 2022b). NADPH oxidase 1 (NOX1), derived reactive oxygen species and modulated by NOXA1, is crucial in the progression of cancer (Attri et al., 2020). PABPC1L is a key gene in tumor progression and postoperative prognosis, while inhibition of PABPC1L suppresses CRC cell growth and metastasis (Wu et al., 2019). Meanwhile, PABPC1 and FOXC2 bind to cis-regulatory elements and inhibit cellular senescence through downregulating p16INK4a in endothelial cells (Wu et al., 2022). Similarly, PCOLCE2 is a novel senescence-related gene that is used to establish a prognostic model in CRC (Yao et al., 2021). MID2, as a promoter of STAT3, is interacted with protein MORC4, which regulates DNA damage response and gene transcription in breast cancer (Wang et al., 2021). CPA3 belongs to carboxypeptidase family of zinc metalloproteases released by mast cells and has been demonstrated to be involved in endogenous proteins degradation as well as colon cancer prognosis (Fang et al., 2021). The 70 kDa heat shock protein, called HSPA1A, is considered as a potential biomarker for the initiation and development breast cancer (de Freitas et al., 2022). The level of HSPA1A is upregulated after heat stress response, but is downregulated by senescence (Llewellyn et al., 2021). It has been reported that CALB1 can promote the interaction between p53 and MDM2, and alleviates ovarian cancer cell senescence (Cao et al., 2019). Collectively, these results suggest that the prognostic genes in senescence-based signatures may be crucial in cellular senescence and prognosis of COAD.

There are some limitations in this study. Data from TCGA and GEO are collected and used for bioinformatics analysis in this study, and these retrospective data may have selection bias. Thus, prospective studies with large samples are needed to validate these results. Although the robustness of our prognostic risk model has been validated by external GEO datasets, its reliability should be iteratively improved with long-term clinical application. Besides, the regulatory mechanism of MDSCs and macrophages on T cells and NK cells should be further investigated in COAD patients with poorer prognosis and better prognosis according to these senescence-based subtypes.

Conclusion

In conclusion, we developed senescence-based subtypes that could distinguish prognosis, T stage, mutation and immune characteristics, which might guide further mechanism investigation of heterogeneity of COAD. Additionally, we constructed and validated a senescence-based signature and provided a reliable tool for prognosis prediction in COAD 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

All authors contributed to this present work: [JF] and [FF] designed the study, [YN] acquired the data. [YN] drafted the manuscript, [JF] revised the manuscript. All authors read and approved the manuscript.

Funding

This research was funded by National Natural Science Foundation of China (82270577).

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/fphar.2023.1121634/full#supplementary-material

References

Attri, P., Park, J-H., De Backer, J., Kim, M., Yun, J-H., Heo, Y., et al. (2020). Structural modification of NADPH oxidase activator (Noxa 1) by oxidative stress: An experimental and computational study. *Int. J. Biol. Macromol.* 163, 2405–2414. doi:10.1016/j.ijbiomac.2020.09.120

Bauer, D., Mazzio, E., and Soliman, K. F. (2019). Whole transcriptomic analysis of apigenin on TNFα immuno-activated MDA-MB-231 breast cancer cells. *Cancer genomics & proteomics* 16 (6), 421–431. doi:10.21873/cgp.20146

Blanche, P. (2015). TimeROC: Time-dependent ROC curve and AUC for censored survival data. *R. package version* 2.

Cao, L. Q., Wang, Y. N., Liang, M., and Pan, M. Z. (2019). CALB1 enhances the interaction between p53 and MDM2, and inhibits the senescence of ovarian cancer cells. *Mol. Med. Rep.* 19 (6), 5097–5104. doi:10.3892/mmr.2019.10212

Cassetta, L., and Pollard, J. W. (2018). Targeting macrophages: Therapeutic approaches in cancer. *Nat. Rev. Drug Discov.* 17 (12), 887–904. doi:10.1038/nrd.2018.169

Charoentong, P., Finotello, F., Angelova, M., Mayer, C., Efremova, M., Rieder, D., et al. (2017). Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep.* 18 (1), 248–262. doi:10.1016/j.celrep.2016.12.019

Chen, Y-A., Lemire, M., Choufani, S., Butcher, D. T., Grafodatskaya, D., Zanke, B. W., et al. (2013). Discovery of cross-reactive probes and polymorphic CpGs in the illumina infinium HumanMethylation450 microarray. *Epigenetics* 8 (2), 203–209. doi:10.4161/epi.23470

Choucair, K., Morand, S., Stanbery, L., Edelman, G., Dworkin, L., and Nemunaitis, J. (2020). Tmb: A promising immune-response biomarker, and potential spearhead in advancing targeted therapy trials. *Cancer gene Ther.* 27 (12), 841–853. doi:10.1038/s41417-020-0174-y

Dai, L., Wang, X., Bai, T., Liu, J., Chen, B., and Yang, W. (2022). Cellular senescence-related genes: Predicting prognosis in gastric cancer. *Front. Genet.* 13, 909546. doi:10.3389/fgene.2022.909546

Dai, L., Wang, X., Bai, T., Liu, J., Chen, B., Yang, W., et al. (2022). Identification of a novel cellular senescence-related signature for the prediction of prognosis and immunotherapy response in colon cancer. *Front. Genet.* 13, 961554. doi:10.3389/fgene.2022.961554

de Freitas, G. B., Penteado, L., Miranda, M. M., Filassi, J. R., Baracat, E. C., and Linhares, I. M. (2022). The circulating 70 kDa heat shock protein (HSPA1A) level is a potential biomarker for breast carcinoma and its progression. *Sci. Rep.* 12 (1), 13012–13018. doi:10.1038/s41598-022-17414-6

Deng, H., Chen, S., Yuan, X., and Zhang, J. (2022). Transcriptome sequencing analysis of the effect of β-elemene on colorectal cancer from the lncRNA-miRNA-mRNA perspective. *Genet. Res.* 2022, 5896296. doi:10.1155/2022/5896296

Domen, A., Deben, C., De Pauw, I., Hermans, C., Lambrechts, H., Verswyvel, J., et al. (2022). Prognostic implications of cellular senescence in resected non-small cell lung cancer. *Transl. lung cancer Res.* 11 (8), 1526–1539. doi:10.21037/tlcr-22-192

Domen, A., Deben, C., Verswyvel, J., Flieswasser, T., Prenen, H., Peeters, M., et al. (2022). Cellular senescence in cancer: Clinical detection and prognostic implications. *J. Exp. Clin. Cancer Res.* 41 (1), 360–435. doi:10.1186/s13046-022-02555-3

Fang, Z., Xu, S., Xie, Y., and Yan, W. (2021). Identification of a prognostic gene signature of colon cancer using integrated bioinformatics analysis. *World J. Surg. Oncol.* 19 (1), 13–14. doi:10.1186/s12957-020-02116-y

Feng, D-C., Zhu, W-Z., Shi, X., Xiong, Q., You, J., Wei, Q., et al. (2022). Identification of senescence-related molecular subtypes and key genes for prostate cancer. *Asian J. Androl.* 24, 0–7. doi:10.4103/aja202258

Feng, F., Sun, H., Zhao, Z., Sun, C., Zhao, Y., Lin, H., et al. (2022). Identification of APC mutation as a potential predictor for immunotherapy in colorectal cancer. *J. Oncol.* 2022, 6567998. doi:10.1155/2022/6567998

Gao, R., Bai, S., Henderson, Y. C., Lin, Y., Schalck, A., Yan, Y., et al. (2021). Delineating copy number and clonal substructure in human tumors from single-cell transcriptomes. *Nat. Biotechnol.* 39 (5), 599–608. doi:10.1038/s41587-020-00795-2

Gribov, A., Sill, M., Lück, S., Rücker, F., Döhner, K., Bullinger, L., et al. (2010). Seurat: Visual analytics for the integrated analysis of microarray data. *BMC Med. genomics* 3 (1), 21–26. doi:10.1186/1755-8794-3-21

Hänzelmann, S., Castelo, R., and Guinney, J. (2013). Gsva: Gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma.* 14 (1), 7–15. doi:10.1186/1471-2105-14-7

Harrell, F. E., Jr, and Harrell, M. F. E., Jr (2019). *Package 'hmisc'. CRAN2018* 2019, 235–236.

Hastie, T., Qian, J., and Tay, K. (2021). *An Introduction to glmnet.* WU Wien, Austria: CRAN R Repositary.

Hernandez-Segura, A., Nehme, J., and Demaria, M. (2018). Hallmarks of cellular senescence. *Trends Cell Biol.* 28 (6), 436–453. doi:10.1016/j.tcb.2018.02.001

Herranz, N., and Gil, J. (2018). Mechanisms and functions of cellular senescence. *J. Clin. investigation* 128 (4), 1238–1246. doi:10.1172/JCI95148

Hong, J., Wang, X., Yu, L., Li, J., Hu, H., and Mao, W. (2022). Identification and development of an age-related classification and signature to predict prognosis and immune landscape in osteosarcoma. *J. Oncol.* 2022, 5040458. doi:10.1155/2022/5040458

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

Junaid, M., Lee, A., Kim, J., Park, T. J., and Lim, S. B. (2022). Transcriptional heterogeneity of cellular senescence in cancer. *Mol. Cells* 45 (9), 610–619. doi:10.14348/molcells.2022.0036

Kaufman, L., and Rousseeuw, P. J. (1990). Partitioning around medoids (program pam). *Find. groups data Introd. Clust. analysis* 344, 68–125.

Kuksin, M., Morel, D., Aglave, M., Danlos, F-X., Marabelle, A., Zinovyev, A., et al. (2021). Applications of single-cell and bulk RNA sequencing in onco-immunology. *Eur. J. Cancer* 149, 193–210. doi:10.1016/j.ejca.2021.03.005

Lähnemann, D., Köster, J., Szczurek, E., McCarthy, D. J., Hicks, S. C., Robinson, M. D., et al. (2020). Eleven grand challenges in single-cell data science. *Genome Biol.* 21 (1), 31–35. doi:10.1186/s13059-020-1926-6

Li, H. (2019). *Mutational landscape of cancer-related genes and gain of function p53 mutants in colorectal cancer*. The Chinese University of Hong Kong: Hong Kong.

Lightbody, G., Haberland, V., Browne, F., Taggart, L., Zheng, H., Parkes, E., et al. (2019). Review of applications of high-throughput sequencing in personalized medicine: Barriers and facilitators of future progress in research and clinical application. *Briefings Bioinforma.* 20 (5), 1795–1811. doi:10.1093/bib/bby051

Llewellyn, J., Mallikarjun, V., Appleton, E., Osipova, M., Gilbert, H. T., Richardson, S. M., et al. (2021). *Senescence inhibits the chaperone response to thermal stress*. New York: bioRxiv, Cold Spring Harbor Laboratory.

Ma, P., Beatty, P. L., McKolanis, J., Brand, R., Schoen, R. E., and Finn, O. J. (2019). Circulating myeloid derived suppressor cells (MDSC) that accumulate in premalignancy share phenotypic and functional characteristics with MDSC in cancer. *Front. Immunol.* 10, 1401. doi:10.3389/fimmu.2019.01401

Marisa, L., de Reyniès, A., Duval, A., Selves, J., Gaub, M. P., Vescovo, L., et al. (2013). Gene expression classification of colon cancer into molecular subtypes: Characterization, validation, and prognostic value. *PLoS Med.* 10 (5), e1001453. doi:10.1371/journal.pmed.1001453

Masiero, M., Simões, F. C., Han, H. D., Snell, C., Peterkin, T., Bridges, E., et al. (2013). A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis. *Cancer Cell* 24 (2), 229–241. doi:10.1016/j.ccr.2013.06.004

Mayakonda, A., Lin, D-C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118

Mugisha, S., Di, X., Wen, D., Zhao, Y., Wu, X., Zhang, S., et al. (2022). Upregulated GATA3/miR205-5p Axis inhibits MFNG transcription and reduces the malignancy of triple-negative breast cancer. *Cancers* 14 (13), 3057. doi:10.3390/cancers14133057

Muñoz-Espín, D., and Serrano, M. (2014). Cellular senescence: From physiology to pathology. *Nat. Rev. Mol. Cell Biol.* 15 (7), 482–496. doi:10.1038/nrm3823

Munro, M. J., Wickremesekera, S. K., Peng, L., Tan, S. T., and Itinteang, T. (2018). Cancer stem cells in colorectal cancer: A review. *J. Clin. pathology* 71 (2), 110–116. doi:10.1136/jclinpath-2017-204739

Papalexi, E., and Satija, R. (2018). Single-cell RNA sequencing to explore immune cell heterogeneity. *Nat. Rev. Immunol.* 18 (1), 35–45. doi:10.1038/nri.2017.76

Punt, C. J., Koopman, M., and Vermeulen, L. (2017). From tumour heterogeneity to advances in precision treatment of colorectal cancer. *Nat. Rev. Clin. Oncol.* 14 (4), 235–246. doi:10.1038/nrclinonc.2016.171

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids Res.* 43 (7), e47–e. doi:10.1093/nar/gkv007

Sagaert, X., Vanstapel, A., and Verbeek, S. (2018). Tumor heterogeneity in colorectal cancer: What do we know so far? *Pathobiology* 85 (1-2), 72–84. doi:10.1159/000486721

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., Luna, A., La, K. C., et al. (2018). Oncogenic signaling pathways in the cancer genome atlas. *Cell* 173 (2), 321–337. doi:10.1016/j.cell.2018.03.035

Therneau, T. M., and Lumley, T. (2015). Package 'survival. *R. Top. Doc.* 128 (10), 28–33.

Timar, J., and Kashofer, K. (2020). Molecular epidemiology and diagnostics of KRAS mutations in human cancer. *Cancer Metastasis Rev.* 39 (4), 1029–1038. doi:10.1007/s10555-020-09915-5

Umansky, V., Blattner, C., Gebhardt, C., and Utikal, J. (2016). The role of myeloid-derived suppressor cells (MDSC) in cancer progression. *Vaccines* 4 (4), 36. doi:10.3390/vaccines4040036

Wang, H., Zhang, L., Luo, Q., Liu, J., and Wang, G. (2021). MORC protein family-related signature within human disease and cancer. *Cell Death Dis.* 12 (12), 1112–1211. doi:10.1038/s41419-021-04393-1

Wang, J., Zhou, C. C., Sun, H. C., Li, Q., Hu, J. D., Jiang, T., et al. (2022). Identification of several senescence-associated genes signature in head and neck squamous cell carcinoma. *J. Clin. Laboratory Analysis* 36 (7), e24555. doi:10.1002/jcla.24555

Wilkerson, M., Waltman, P., and Wilkerson, M. M. (2013). *Package 'ConsensusClusterPlus'*.

Wu, T., Jiang, D., Zou, M., Sun, W., Wu, D., Cui, J., et al. (2022). Coupling high-throughput mapping with proteomics analysis delineates cis-regulatory elements at high resolution. *Nucleic acids Res.* 50 (1), e5–e. doi:10.1093/nar/gkab890

Wu, Y-Q., Ju, C-L., Wang, B-J., and Wang, R-G. (2019). PABPC1L depletion inhibits proliferation and migration via blockage of AKT pathway in human colorectal cancer cells. *Oncol. Lett.* 17 (3), 3439–3445. doi:10.3892/ol.2019.9999

Xi, Y., and Xu, P. (2021). Global colorectal cancer burden in 2020 and projections to 2040. *Transl. Oncol.* 14 (10), 101174. doi:10.1016/j.tranon.2021.101174

Xiao, C., Zhao, X., Li, X., Zhu, S., Cao, J., Chen, H., et al. (2022). Classification of colorectal cancer into clinically relevant subtypes based on genes and mesenchymal cells. *Clin. Transl. Oncol.* 25, 491–502. doi:10.1007/s12094-022-02964-y

Yao, H., Li, C., and Tan, X. (2021). An age stratified analysis of the biomarkers in patients with colorectal cancer. *Sci. Rep.* 11 (1), 22464–22512. doi:10.1038/s41598-021-01850-x

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, T-J., Ma, D., Liu, Y-Y., Xiao, Y., Gong, Y., Jiang, Y-Z., et al. (2021). Bulk and single-cell transcriptome profiling reveal the metabolic heterogeneity in human breast cancers. *Mol. Ther.* 29 (7), 2350–2365. doi:10.1016/j.ymthe.2021.03.003

Zhang, L., and Shay, J. W. (2017). Multiple roles of APC and its therapeutic implications in colorectal cancer. *JNCI J. Natl. Cancer Inst.* 109 (8), djw332. doi:10.1093/jnci/djw332

Zheng, X., Song, J., Yu, C., Zhou, Z., Liu, X., Yu, J., et al. (2022). Single-cell transcriptomic profiling unravels the adenoma-initiation role of protein tyrosine kinases during colorectal tumorigenesis. *Signal Transduct. Target. Ther.* 7 (1), 60–14. doi:10.1038/s41392-022-00881-8

Zhou, Y., Xiao, L., Long, G., Cao, J., Liu, S., Tao, Y., et al. (2022). Identification of senescence-related subtypes, establishment of a prognosis model, and characterization of a tumor microenvironment infiltration in breast cancer. *Front. Immunol.* 13, 921182. doi:10.3389/fimmu.2022.921182

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 09 March 2023
DOI 10.3389/fphar.2023.1086309

OPEN ACCESS

EDITED BY
Fangfang Tao,
Zhejiang Chinese Medical University, China

REVIEWED BY
Zhixiang Yu,
Xijing Hospital, Fourth Military Medical University, China
Yunfeng Jin,
Fudan University, China

*CORRESPONDENCE
Tiansheng Cao,
caotiansheng2088@sina.com
Tengfei Ji,
243477587@qq.com

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 01 November 2022
ACCEPTED 06 February 2023
PUBLISHED 09 March 2023

CITATION
Cao T, Wu H and Ji T (2023), Bioinformatics-based construction of prognosis-related methylation prediction model for pancreatic cancer patients and its application value.
Front. Pharmacol. 14:1086309.
doi: 10.3389/fphar.2023.1086309

Bioinformatics-based construction of prognosis-related methylation prediction model for pancreatic cancer patients and its application value

Tiansheng Cao*, Hongsheng Wu and Tengfei Ji*

Department of Hepatobiliary Surgery, Affiliated Huadu Hospital, Southern Medical University (People's Hospital of Huadu District), Guangzhou, China

Objective: Pancreatic adenocarcinoma (PAAD) is a highly malignant gastrointestinal tumor with almost similar morbidity and mortality. In this study, based on bioinformatics, we investigated the role of gene methylation in PAAD, evaluated relevant factors affecting patient prognosis, screened potential anti-cancer small molecule drugs, and constructed a prediction model to assess the prognosis of PAAD.

Methods: Clinical and genomic data of PAAD were collected from the Tumor Genome Atlas Project (TCGA) database and gene expression profiles were obtained from the GTEX database. Analysis of differentially methylated genes (DMGs) and significantly differentially expressed genes (DEGs) was performed on tumorous samples with KRAS wild-type and normal samples using the "limma" package and combined analysis. We selected factors significantly associated with survival from the significantly differentially methylated and expressed genes (DMEGs), and their fitting into a relatively streamlined prognostic model was validated separately from the internal training and test sets and the external ICGC database to show the robustness of the model.

Results: In the TCGA database, 2,630 DMGs were identified, with the largest gap between DMGs in the gene body and TSS200 region. 318 DEGs were screened, and the enrichment analysis of DMGs and DEGs was taken to intersect DMEGs, showing that the DMEGs were mainly related to Olfactory transduction, natural killer cell mediated cytotoxicity pathway, and Cytokine -cytokine receptor interaction. DMEGs were able to distinguish well between PAAD and paraneoplastic tissues. Through techniques such as drug database and molecular docking, we screened a total of 10 potential oncogenic small molecule compounds, among which felbamate was the most likely target drug for PAAD. We constructed a risk model through combining three DMEGs (S100P, LY6D, and WFDC13) with clinical factors significantly associated with prognosis, and confirmed the model robustness using external and internal validation.

Conclusion: The classification model based on DMEGs was able to accurately separate normal samples from tumor samples and find potential anti-PAAD drugs by performing gene-drug interactions on DrugBank.

KEYWORDS

pancreatic cancer, methylation, prognosis, prediction model, drugs

1 Background

Pancreatic adenocarcinoma (PAAD) is the 14th most common cancer globally (Kocarnik et al., 2022), with an estimated 458,918 confirmed pancreatic cancer cases and 432,242 death cases each year all over the world (Ferlay et al., 2019). The incidence of PAAD varies widely by country, as Europe and North America showed the highest age-standardized incidence, which was the lowest in South-Central Asia and Africa (Ilic and Ilic, 2016). Incidence rate of PAAD is generally higher in developed countries compared to developing countries, with a standardized incidence rate of 4.9/100,000 and 3.6/100,000 for men and women, respectively. In the United States, 5-year survival rate of PAAD is 9.3%, and it is the fourth leading factor resulting in cancer-related mortality (Gandhi et al., 2018). Apart from smoking, diabetes, alcohol drinking, obesity, occupational exposure and genetic factors, PAAD is as well an epigenetic disease (Goral, 2015; Midha et al., 2016; Hu et al., 2021). Abnormal DNA methylation patterns are a common human tumorous feature (Kulis and Esteller, 2010). From precancerous lesions to PAAD, epigenetic changes play an important role in the multistage carcinogenesis (Xu et al., 2019).

Epigenetics are changes in gene expression but not in DNA sequence, and the major epigenetic alteration leading to PAAD progression is DNA methylation (Wang et al., 2016). To detect epigenetic abnormalities in PAAD, it is necessary to identify genome-wide patterns of DNA methylation. Nones et al. (2014) used high-density arrays to capture 167 untreated PAAD sample methylation and compared it with normal tissue adjacent to the cancerous one and identified 3,522 abnormally methylated genes. In addition, partial methylation of CDKN1C promoter CpG islands and reduced expression of protein products are observed when comparing PAAD precursor cells methylation expression to normal pancreatic duct epithelial cells (Sato et al., 2008). Basic studies have shown that in PAAD precursor cells, CDKN1C gene is under-expressed and there is reduced expression of protein products and partial methylation of CDKN1C promoter CpG islands. The above evidence supports that aberrant DNA hypo/hypomethylation occurs in PAAD precursor lesions leading to further progression to PAAD (13).

As research continues, aberrant methylation of DNA CpG islands has become a prominent feature of PAAD and a potential diagnostic marker and therapeutic target for PAAD. However, the results of clinical trials were disappointing, probably due to the low level of epigenetic specificity (Matsubayashi et al., 2006; Marabelle et al., 2020). Therefore, in order to use methylation as a future therapeutic tool for PAAD, an in-depth understanding of the methylation expression profile and supporting pathways of PAAD is needed. According to the mutation and gene expression profile data of PAAD patients and gene expression profiles of normal pancreas from GTEX, this study screened differentially methylated and expressed genes (DMEGs), and confirmed that methylation was a reliable prognostic marker for PAAD and a potential oncogenic drug target for PAAD.

2 Materials and methods

2.1 Acquisition of clinical data, gene expression profiles and data processing

Methylation data, clinical follow-up data, and gene expression profiles of PPAD came from TCGA (https://portal.gdc.cancer.gov/) by means of UCSC Xena. The gene expression profiles of normal pancreas samples were obtained from the GTEX (http://www.gtexportal.org/home/index.html) databases using UCSC Xena.

For sample data reliability, we set the following inclusion criteria (Kocarnik et al., 2022): only normal samples and primary PAAD samples were retained (Ferlay et al., 2019); PAAD samples with wild-type KRAS gene were retained (Ilic and Ilic, 2016); PAAD samples with complete clinical data were retained. A total of 182 samples were obtained from TCGA, including 178 tumor samples, 70 KRAS wild-type tumor samples and 4 normal samples. A total of 167 normal pancreas samples were obtained from the GTEX database. In order to homogenize the data, the "sva" R package was applied to remove the batch effect from the combined data of the two datasets, and a total of 19,593 protein-coding genes were retained by ENSG conversion of gene symbols using genecode V35.

2.2 Analysis of differentially methylated genes (DMGs)

The Illumina HumanMethylation450 BeadChip matrix contained 380,097 probes of around 99% (n = 26,081) of the RefSeq genes. For each probe, the raw gene methylation intensity was expressed as a beta value. To identify differentially methylated CpG sites (DMS), PAAD tumor samples were compared with paracancer samples using the "limma" R package (Ritchie et al., 2015). The Benjamini and Hochberg (BH) method adjusted p-value of each methylation site to FDR (false discovery rate) (Ghosh, 2012). Statistical thresholds were set for FDR <0.01 and |delta β-value|> 0.1.

The CpG locus to gene match files were downloaded from the Illumina website (https://www.illumina.com/). In different regions (TSS200, TSS1500, Gene body, 5′-UTR, 3′-UTR, transcription start site, integration region), the average β-values of genes were calculated with the correspondence. Using the "limma" R package, the differentially methylated regions were calculated, where FDR <0.01, delta β-values < -0.1 were the demethylated regions, FDR <0.01, delta β > 0.1 were hypermethylated regions.

2.3 Analysis of differentially expressed genes, differentially methylated genes and pathways

Differentially expressed genes (DEGs) were analyzed for normal and tumor samples in the TCGA-PAAD cohort using the "limma" R package, and p values were adjusted using the Benjamini and Hochberg (BH) method, where FDR >0.01 and log2FC > 2 were up-regulated genes, and FDR >0.01 and log2FC < −2 were down-regulated genes.

TABLE 1 Criteria for grouping DMEGs.

Groups	Methylation cut-off	Expression cut-off
HypoUp	FDR <0.01 and delta β-value < −0.1	FDR <0.01 and log2FC > 2
HypoDown	FDR <0.01 and delta β-value < −0.1	FDR <0.01 and log2FC < -2
HyperUp	FDR< 0.01 and delta β-value >0.1	FDR <0.01 and log2FC > 2
HyperDown	FDR <0.01 and delta β-value >0.1	FDR <0.01 and log2FC < -2

FDR:false discovery rate; log2FC: log2 fold change.

To identify the relationship between gene methylation and gene expression profiles, we took the intersection of differentially methylated genes and DEGs to obtain differentially methylated and expressed genes (DMEGs) and classified them into four groups: HyperDown, HyperUp, HypoDown, HypoUp (Table 1). Then, we used Gene Ontology (GO) functional enrichment analysis and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database through the "clusterProfiler," "org.Hs.eg.db," "enrichplot" and "ggplot2" R software packages (Wu et al., 2021), and FDR <0.05 was used as the screening condition to perform enrichment analysis of DEGs to discover the main biological characteristics of DEGs and plot the bubble map.

2.4 Marker evaluation of PAAD methylation and expression profiles

DEGs were proposed as tumor markers for the diagnosis of PAAD, and 50% of the expression profile data of DMEGs and methylation data of DMEGs for PAAD were the training set and 50% as the test set. The training set data were analyzed by principal component analysis (PCA) with the "prcomp" R function (Luu et al., 2017) to clarify the eigenvector weights of the principal components and construct a diagnostic model of PAAD, which was plotted and visualized using the "ggplot2" R software package (Maag, 2018). Finally, to evaluate the diagnostic advantage of PCA model for PAAD, the receiver operating characteristic (ROC) curves of the PCA model were plotted by the "pROC" R software package and the area under curve (AUC) was calculated for the training set and test set (Robin et al., 2011), where AUC showed a low accuracy at 0.5–0.7, higher accuracy at 0.7–0.9, and high accuracy at AUC above 0.9.

2.5 The prediction of DMEGs and target drugs

The use of key genes as potential therapeutic targets is a cornerstone in the development of therapeutic agents for sepsis. We determined PAAD and drug proximity based on drug-target pairs from the drugbank database (https://go.drugbank.com/) and the Protein-Protein interaction (PPI) network (threshold score of 400). Here, given distance d (s,t) as the shortest path between node s and node t (where s ∈ S, PAAD-related genes; t ∈ T, drug target genes), D (degree of related gene set nodes in PPI), T (set of drug target genes), S (PAAD-related genes), and the calculation is as follow:

$$d(S,T) = \frac{1}{|T|}\sum_{t\in T} min_{s\in S}(\mathrm{d}(\mathrm{s},\mathrm{t}) + \omega) \quad (1)$$

where ω, the weight of the target gene, was calculated as ω = -ln (D+1) if the target gene was a gene in the PAAD-related gene set, otherwise ω = 0.

Next, between these simulated drug targets and the key gene set, we calculated the distance d (S,R), and generated the simulated reference distributions after performing random repetitions for 10,000 times, at the same time we the observed distances corresponding to the actual were scored using the mean and standard deviation of the μd (S,R) and σd (S,R) reference distributions and converted into a normalized scoring, i.e., the proximity z.

$$z(S,T) = \frac{d(S,T) - \mu_{d(S,R)}}{\sigma_{d(S,R)}} \quad (2)$$

Finally, a gene set distance density score map was constructed by normalized distance scoring.

2.6 Molecular docking

A technique for designing drugs based on receptor features and the way that drug molecules interact with receptors is called molecular docking. In the realm of computer-aided drug development, it is a theoretical modeling technique that primarily investigates the interaction between molecules (such as ligands and receptors) and forecasts their binding mechanism and affinities. (Lohning et al., 2017; Saikia and Bordoloi, 2019). Autodock Vina software was used in molecular docking (Trott and Olson, 2010). To prepare input files, we applied AutoDockTools 1.5.6. The pdb file of the protein came from Protein Data Bank (Velankar et al., 2021) with PDB ID 6SUK. The Polar hydrogens were added to the solution after all water molecules, potassium ions, and protein B chains had been eliminated. The zinc ion's charge was modified in the receptor protein's PDBQT file to +2.0, and the grid's coordinates in each XYZ direction were −19.5, 74.5, and 34.8 during molecular docking. The lengths were 20 in each XYZ direction. The Lamarckian approach was utilized to determine the ligand molecule's strongest binding mode. The maximum number of output conformations was set to 10, the exhaustiveness was set to 8, and the allowable energy range was set to a maximum of 3 kcal/mol. With the aid of Pymol, the output maps were processed.

2.7 Dynamics simulation

In this study, the binding stability of the receptor-ligand complex was assessed by performing molecular dynamics simulations of 100 ns (Zhou et al., 2022) using the Gromacs2019 package. In the molecular dynamics simulations the CHARMm36 force field was employed. With the aid of the CHARMM Common Force Field (CGenFF) software, the str files for the ligands were acquired. The system was dissolved in TIP3P water molecules in a dodecahedral box. At a concentration of 0.154 M, sodium and chloride ions were introduced to the system to neutralize its charge. Using a cutoff of 5,000 steps and the steepest descent algorithm, the solventized system's energy was minimized. The LINCS method was used to restrict the length of covalent bonds. Using the PME technique, the total electrostatic interactions were determined.

10.3389/fphar.2023.1086309

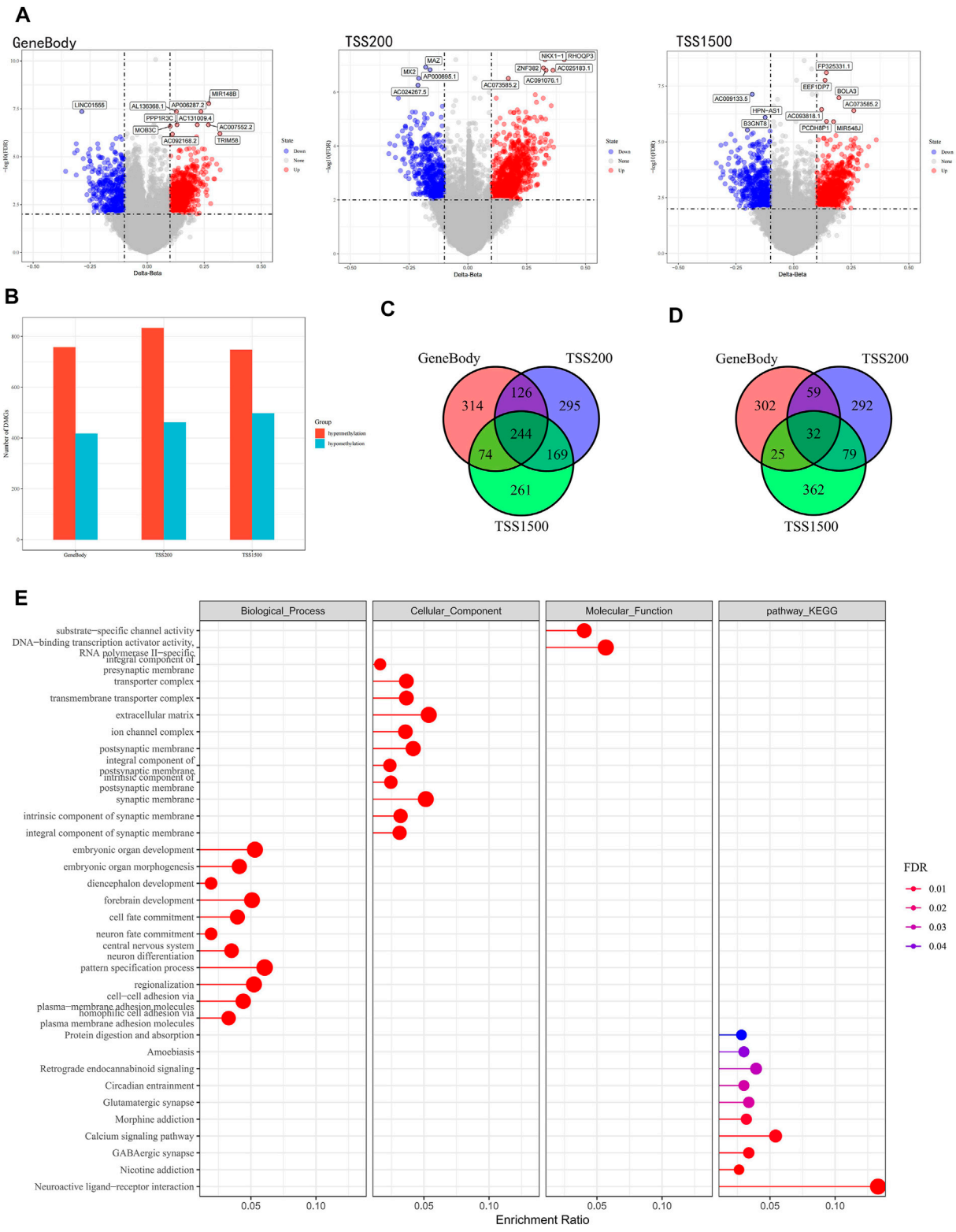


FIGURE 1
Analysis of PAAD differentially methylated genes. **(A)** Volcano plot of differentially methylated within the gene body, TSS200 and TSS1500 regions. **(B)** Histograms of differentially methylated genes within the three regions. **(C)** Venn diagram of hypermethylation within three different regions. **(D)** Venn diagrams of demethylated genes within three different regions. **(E)** KEGG and GO functional enrichment analysis of differentially methylated genes, where the color from blue to red indicates that the FDR is from large to small, and the dots from small to large represent the increasing number of enriched genes, left: hypermethylation, right: demethylation.

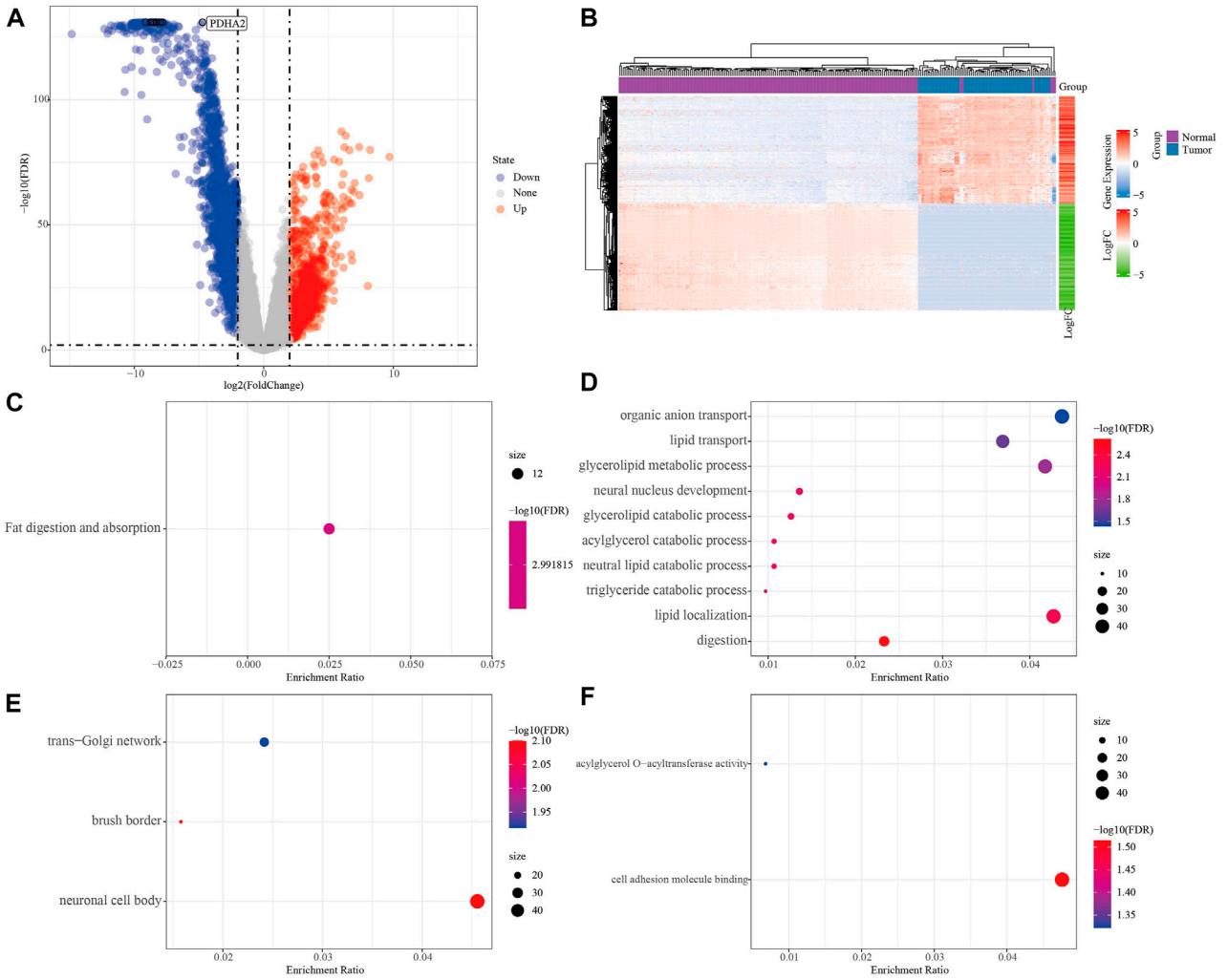


FIGURE 2
Analysis of PAAD differential genes. **(A)** Volcano plot of differentially expressed genes in expression profile. **(B)** Heat map of differentially expressed genes. **(C)** Results of differential gene KEGG enrichment. **(D)** Results of differential gene GO BP enrichment. **(E)** Differential gene GO CC enrichment results. **(F)** Differential gene GO MF enrichment results, the color from blue to red in CDEF represents FDR from large to small, the size of the dot represents the number of enriched to genes, a larger dot indicates more enriched genes.

At constant temperature (300 K) and pressure (1 bar), NVT and NPT simulations were then run for 100 ps, with the compound's confined atoms re-establishing the system's equilibrium at its initial coordinates. Finally, a 100 ns long Pruduct MD run with a 2 fs time step was completed. The Gromacs built-in tool was used to determine the ligands' root mean square deviation (RMSD) values.

2.8 Development and verification of the prognostic gene signature model associated with DMEG

In the TCGA-PAAD dataset, we first randomly and equally divided 241 KRAS wild samples into training (Train) and validation (Test) groups according to the ratio of 1:1, and then reduced the associated genes (Tibshirani, 1997) by Least absolute shrinkage and selection operator (Lasso) regression method. In regression analysis, by compressing some coefficients at the same time setting some coefficients to zero, Lasso regression can better solve multicollinearity. We choose the number of factors when the coefficients of independent variables tended to zero with the gradual increase of lambda. Then, we used the AIC deficit pool information criterion through stepwise regression that takes the statistical fit of the model and parameter numbers into account. A better model of smaller value indicated a sufficient fit of the model with fewer number of parameters (Zhang, 2016). After that, the "survminer" package was used to find the best cutoff (Niu et al., 2021) of the gene signature in the Train dataset of TCGA, and the PAAD was divided into two groups based on the cutoff value, and finally the log-rank test was used to compare the survival differences between the two groups.

To verify the robustness of the gene signature model, we first used the same model and the same coefficients as the training set in the validation set, and then compared the survival differences between the two groups by log-rank test. After that, we

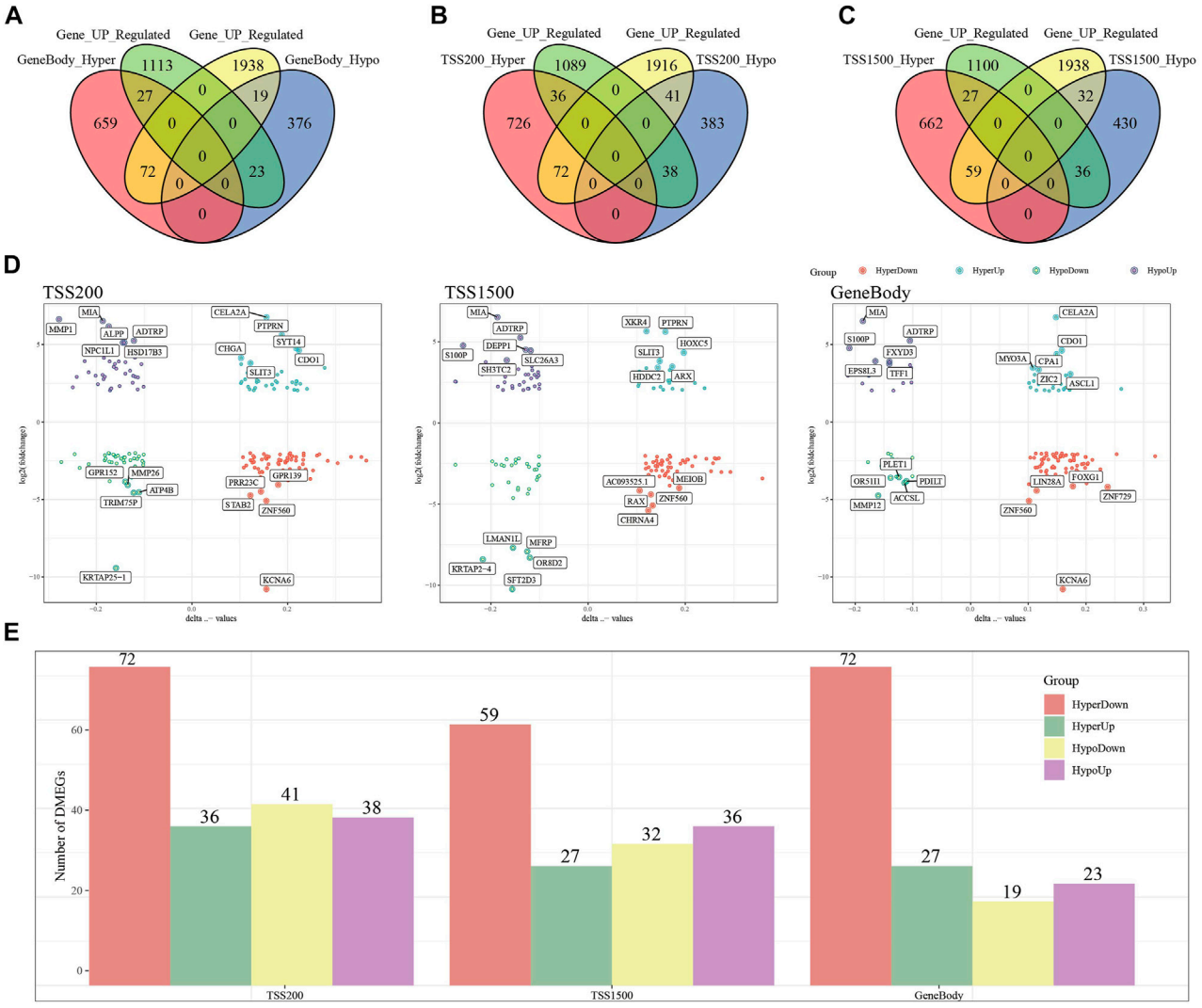


FIGURE 3
Joint analysis of differentially expressed genes and differentially methylated genes. **(A)** Venn diagram of differentially expressed genes with differentially methylated genes in the GeneBody region. **(B)** Venn diagram of differentially expressed genes with differentially methylated genes in the TSS200 region. **(C)** Venn diagrams of differentially expressed genes versus differentially methylated genes within the TSS1500 region. **(D)** Quadrant plots of differentially expressed genes versus differentially methylated genes within the TSS200, TSS1500, and GeneBody regions. **(E)** Histogram of four regulatory patterns of differentially expressed genes and differentially methylated genes in TSS200, TS1500, and GeneBody.

downloaded the expression profiles of PAAD from the ICGC database as well as clinical information, and then used the model constructed above to calculate each score separately and obtain the best cutoff, and then performed the survival curve analysis in the external dataset for the high- and low-risk groups.

2.9 Statistical analysis

All statistical analyses were operated in R software (version 4.1.2, https://cran.r-project.org/doc/manuals/R-lang.html). The optimal threshold of gene expression or score was selected for risk grouping of PAAD using the surv_cutpoint function of the "survminer" package. The Kaplan-Meier assessment method was used to assess the survival differences between the low-risk and high-risk groups, and the Log-rank test was used for comparison. Unless otherwise stated, all statistical tests were two-sided and $p < 0.05$ was considered statistically significant. Comparisons between multiple groups were performed and plotted using the "ggpubr" and "ggplot2" packages, and the statistical significance of box plots was assessed using the Mann-Whitney U or Kruskal–Wallis tests.

3 Results

3.1 Analysis of differentially methylated genes in PAAD

To identify differential gene methylation in PAAD, we first performed a comparative analysis of methylation data from

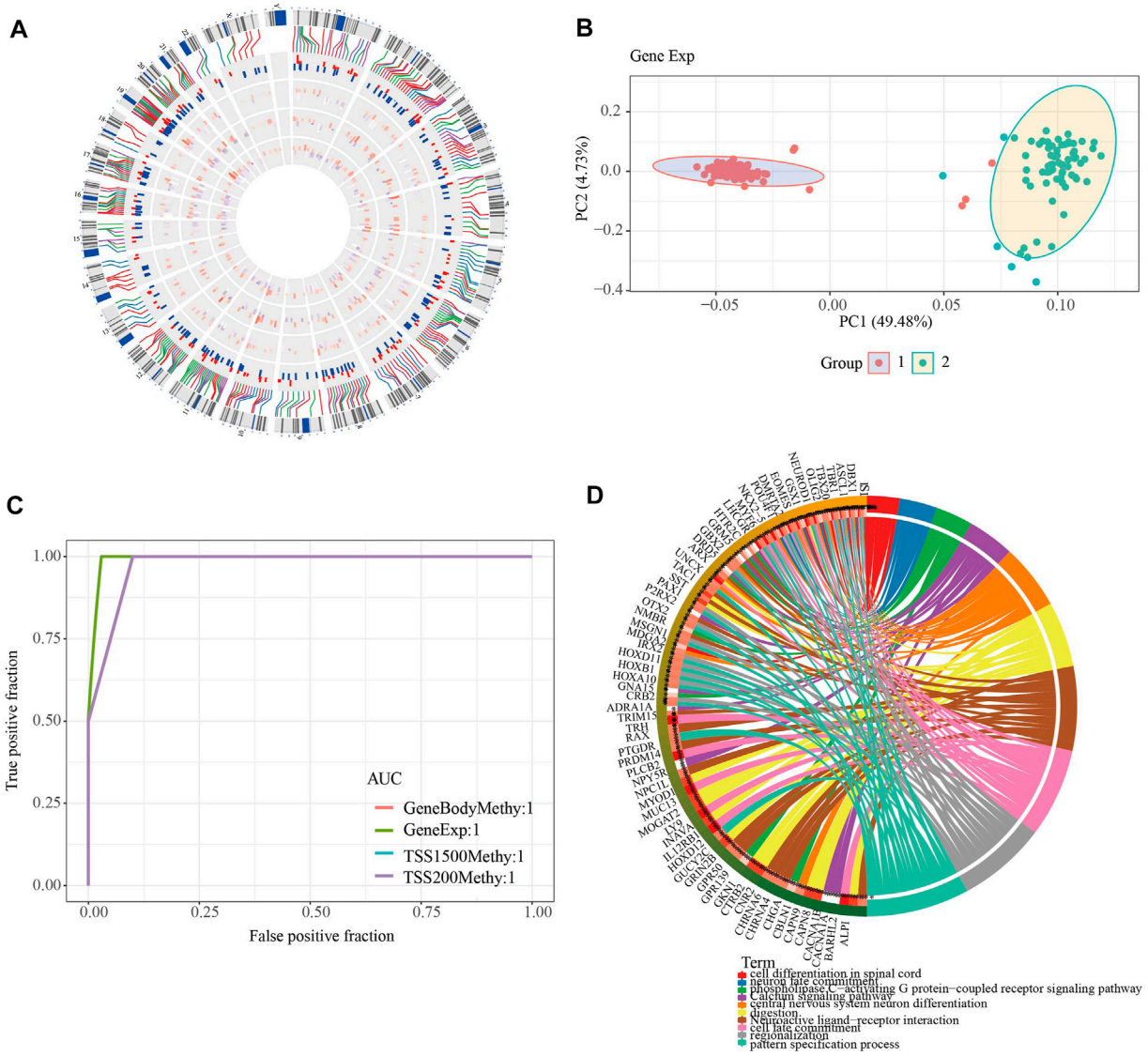


FIGURE 4
Analysis of DMEGs. **(A)** Distribution of DMEGs on the genome. From inside to outside, there are DMGs in the TSS1500 region, DMGs in the TSS200 region, DMGs in the genebody region, DEGs, and corresponding values. The outermost circle is the corresponding chromosome position. **(B)** PCA analysis could distinguish tumor from normal samples based on the gene expression and methylation of DMEGs. **(C)** ROC curves of tumor and normal samples based on a linear discriminant model using the expression profiles and methylation of DMEGs. **(D)** Results of KEGG and GO enrichment analysis of DMEGs, where different colors represent different pathways and connecting lines represent the existence of association between genes and pathways.

185 KRAS wild-type PAAD samples and 10 normal samples, and identified a total of 2,630 differentially methylated genes (FDR <0.01, |delta β-values| > 0.1, Figure 1A), within the Gene body region, 758 genes were hypermethylated and 418 genes were demethylated. 834 genes were hypermethylated and 462 genes were demethylated in TSS20; 748 genes were hypermethylated and 498 genes were demethylated in TSS1500. We found that the number of hypermethylation in the three regions was slightly larger than that of hypermethylation overall (Figure 1B). In the Gene body and TSS200 regions, the difference between hypermethylation and demethylation was the largest, with a ratio of about 1.8:1. Among the hypermethylated genes, 244 genes appeared in all three regions of Gene body, TSS20 and TSS1500, 369 genes appeared in two of them, and the remaining 870 genes appeared in only one region (Figure 1C). Among the demethylated genes, only 32 genes appeared in all three regions, and 163 genes appeared in two of them. These differentially methylated genes were mainly associated with GABAergic synapse, Neuroactive ligand-receptor interaction, Nicotine addiction, and other pathways, as shown by GO and KEGG functional enrichment analysis (Figure 1D). (Figure 1E). The above results confirmed that PAAD methylation was region-specific.

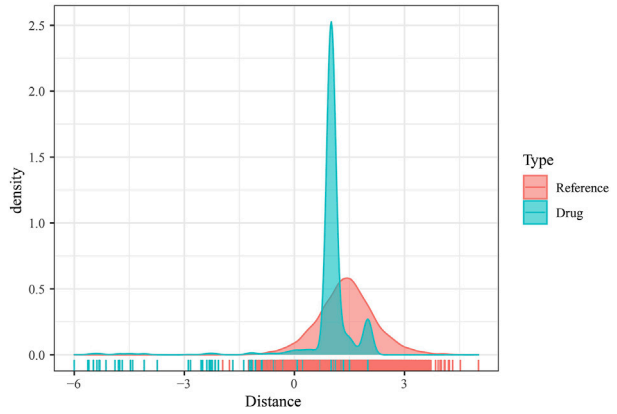


FIGURE 5
DMEGs and potential target therapeutic drugs. Density fractionation plot of drug to DMEGs gene set distance.

3.2 Analysis of differential genes in PAAD and combined analysis of differentially metaylated genes

To screen the differential genes between normal and KRAS wild-type PAAD samples, we analyzed the differential genes between 171 normal samples and 70 KRAS wild-type tumor samples using the "limma" package, and obtained a total of 2,928 significantly DEGs, of which 2029 were down-regulated and 1,163 were up-regulated in tumors (FDR <0.01, |log2FC| > 2, Figure 2A). A total of 2,928 significantly DEGs were obtained, of which 1,163 were up-regulated and 2029 were down-regulated in tumors (FDR <0.01, |log2FC| > 2, Figure 2A). Then, unsupervised hierarchical clustering of these significantly differentially expressed genes revealed that the differential genes could clearly screen tumor samples from the normal ones (Figure 2B). KEGG study showed that the significant differential genes were mainly related to Fat digestion and absorption (Figure 2C). Biological process (BP) enrichment study demonstrated that the differential genes were largely correlated with Lipid transport, Lipid localization and other pathways; cellular components (CC) showed that the differential genes were associated with neural cell body, trans-Golgi. The results of Molecular Function (MF) showed that the differential genes were related to Cytokine-cytokine receptor interaction, natural killer cell-mediated cytotoxicity, Olfactory transduction, and other such pathways that have been previously reported to be associated with PAAD occurrence Figures 2D–F (Malchiodi and Weiner, 2021; Hu et al., 2022).

To search for genes more critical for PAAD occurrence, differentially methylated and expressed genes (DEMGs) were obtained by intersection analysis of DMGs and DEGs. In Gene body, TSS200 and TSS1500, 141, 187 and 154 DEMGs were obtained, respectively (Figures 3A–C). The methylation ploidy and expression difference ploidy of these DMEGs are shown in Figure 3D, and each graph shows the 22 genes with the largest expression difference ploidy. Next, we counted DMEGs in the three regions and identified a total of 318 DMEGs, including 56 in HyperUp, 112 in HyperDown, 69 in HypoUp, and 81 in HypoDwon (Figure 3E).

3.3 Analysis of DMEGs genes in PPAD

To further investigate the role of DMEGs in PAAD, we first used the "circlize" package to map the distribution of 318 DMEGs on chromosomes, with chromosomes chr11 and chr12 having the largest number of 26 DMEGs, chr10, chr12, chr17, chr16, chr2, chr19, chr3, chr20, chr5, chr4, chr7, chr6, and chr6. Chr17, chr16, chr2, chr19, chr3, chr20, chr5, chr4, chr8, chr7, chr6 chromosomes also possessed more than 10 DMEGs each (Figure 4A). We constructed a linear judgment classification model using the gene expression profiles of DMEGs and methylation data from GeneBody, TSS200 and TSS1500, respectively, to evaluate the difference of DNA methylation patterns and gene expression between PAAD tumors and normal samples, and also performed PCA and ROC analyses. The results of PCA showed that DMEGs could classify PAAD and normal samples effectively (Figure 4B), and the AUC values were all 1, suggesting an excellent performance in classification (Figure 4C). GO and KEGG enrichment analysis showed that DMEGs were mainly associated with cell differentiation in spinal cord, neuron fate commitment, calcium signaling pathway, phospholipase C-activating G protein-coupled receptor signaling pathway, digestion, central nervous system neuron differentiation, neuroactive ligand-receptor interation, cell fate commitment, regionalization, and pattern specification process (Figure 4D).

3.4 DMEGs and potential target therapeutic agents

As mentioned previously, DMEGs may be the key genes causing PAAD, and therefore targeting DMEGs is a potential target for the treatment of PAAD. To this end, we calculated the proximity of DMEGs to PAAD according to Formula 1 and converted the observed distances into normalized scores according to Formula 2. We found that either with our randomly selected gene set as a sample or DEMGs as a sample, using the random data acquired for multiple hypothesis testing and selecting drugs with a distance set distributed around 0 to 3 and FDR <0.01, a total of 78 potential target drugs were obtained, and Figure 5 shows the distance density fraction of drugs to DMEGs.

3.5 Molecular docking and pharmacokinetic simulation

Currently, the ADRA1A protein does not have any resolved crystal structure. We used the AlphaFold Protein Structure Database website (https://www.alphafold.ebi.ac.uk/) website for ADRA1A homology modeling to obtain the 3D structure of the ADRA1A protein and the Deepsite (https://www.playmolecule.com/deepsite/) website to predict the protein activity of ADRA1A (32). In addition, the Gromacs2019 software package was used to predict potential small molecule compounds, and a total of 10 small molecule compounds were identified by calculating RMSD values, namely DB06201, DB12733, DB00610, DB00450, DB00699, DB06706, DB06711, DB06764 DB00949, and DB08954 (see Table 2).

Taken together, DB0094 (Felbamate) 9 had the highest molecular docking score and therefore had a higher potential to

TABLE 2 Molecular docking scores of compounds and proteins and the important interactions generated.

Compound	Compound	Target	Docking score	H-bond interactions
DB06201	Rufinamide	GRM5	−4.874	SER143, SER145, THR168
DB12733	Dipraglurant	GRM5	−4.348	SER145, THR168
DB00610	Metaraminol	ADRA1A	−5.158	MET1, GLU87
DB00450	Droperidol	ADRA1A	−5.819	GLU87
DB00699	Nicergoline	ADRA1A	−2.137	MET1
DB06706	Isometheptene	ADRA1A	−2.752	GLU87
DB06711	Naphazoline	ADRA1A	−5.167	—
DB06764	Tetryzoline	ADRA1A	−5.246	—
DB00949	**Felbamate**	**GRIN2B**	**−10.586**	GLU106, SER132
DB08954	Ifenprodil	GRIN2B	−6.821	GLU106, ARG115, ALA135

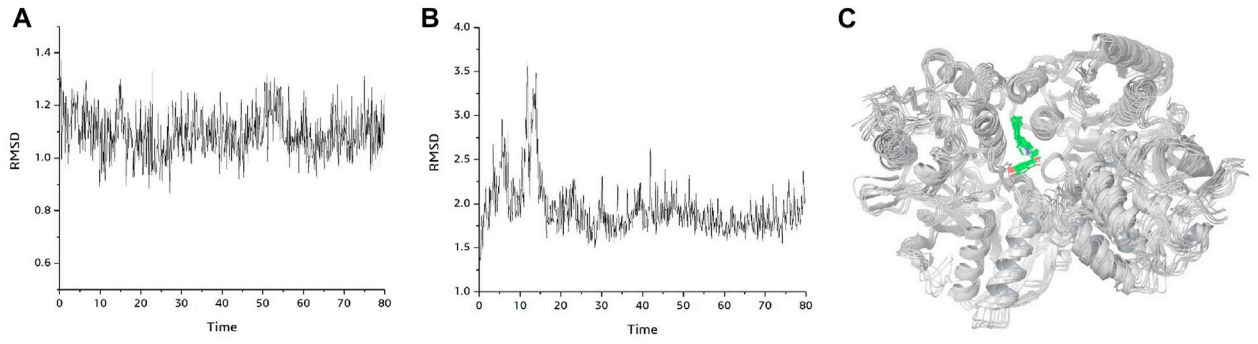


FIGURE 6
Binding mode plot of GRIN2B protein with compound DB00949(Felbamate). **(A)** RMSD diagram of GRIN2B protein during 80 ns molecular dynamics simulation. **(B)** RMSD values of compound DB00949 (Felbamate) during 80 ns molecular dynamics simulation. **(C)** Plot of the dynamic binding pattern of GRIN2B protein with compound DB00949 (Felbamate) during 80 ns molecular dynamics simulation.

be a potential inhibitor of GRIN2B protein. Compound DB0094 interacted with GRIN2B protein, and the RMSD value of compound DB0094 was relatively stable overall (basically stable at around 3 Å) (Figure 5). The compound was able to produce hydrogen bonding interactions with SER132 and GLU106 of GRIN2B protein, and favorable hydrophobic interactions with ILE111, PRO78, ALA107, PRO177 and ALA135, as well as with TYR109, PHE114 and PHE176. Compound DB00949 (Felbamate) showed a high molecular docking score that many favorable interactions with GRIN2B protein were produced.

Figure 6A shows the changes of RMSD values of the D-protein backbone of GRIN2B protein bound to compound DB00949 (Felbamate) during the molecular dynamics simulation at 80 ns As can be seen from the figure, the conformation of the GRIN2B protein was very stable during the molecular dynamics simulation at 80 ns, which also indicated to some extent that the protein structure generated based on homology modeling was relatively reasonable (Figure 6B). In addition, Figure 6C gives the RMSD values of the molecular backbone of compound DB00949 (Felbamate) binding to GRIN2B protein during molecular dynamics (MD) simulation of 80 ns The results demonstrated that compound DB00949(Felbamate)'s RMSD value fluctuated relatively large by an obvious increasing trend during the first 20 ns The stability was basically achieved when it reached 20 ns It remained comparatively constant in the subsequent 60 ns Since the molecular docking was semi-flexible in this experiment, it is understandable that the RMSD values of the ligand's molecular backbone fluctuated moderately in the initial stage of the dynamics simulation. Overall, compound DB00949 (Felbamate) was relatively stable when binding to GRIN2B protein, which further suggested that compound DB00949 (Felbamate) had a high potential to be a potential inhibitor of GRIN2B protein.

3.6 Establishment of prognostic gene signature associated with DMEG

To explore the role of DMEG gene expression in PAAD prognosis, we first randomly divided 241 KRAS wild samples into two groups, one as the training set (n = 121) and one as the validation set (n = 120). We used 10-fold cross-validation to execute

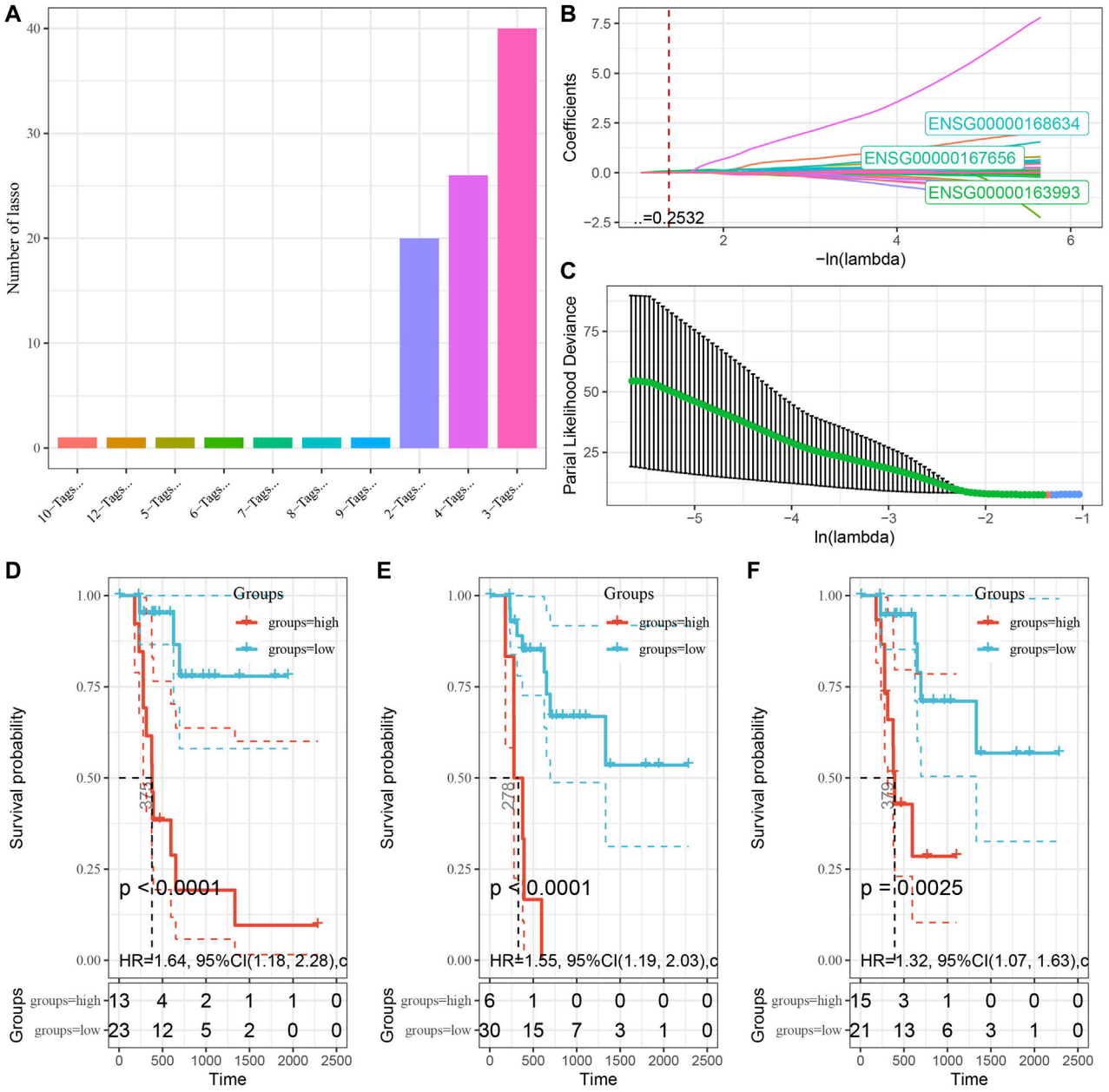


FIGURE 7
Establishment of prognostic gene signature associated with DMEG. **(A)** Frequency of individual gene combinations for one thousand lasso regressions. **(B)** Coefficient change trajectories of individual genes under different lambda. **(C)** Standard deviation distribution of the models under different lambda. **(D)** Prognostic KM curves of S100P in high and low expression groups. **(E)** Prognostic KM curves of LY6D in high and low expression groups. **(F)** Prognostic KM curves of WFDC13 in high and low expression groups.

1,000 Lasso regression analysis on the expression and clinical survival data of these 318 DMEGs genes, and we counted the appearances of each probe 100 times (Figure 7A). 3 probes (S100P, LY6D, and WFDC13) appeared the most frequently, and these 3 genes showed the highest frequency with different coefficient of variation trajectories of lambda as Figure 7B, standard deviation distributions of different lambda as Figure 7C. K-M survival curve results indicated that these three genes were able to distinguish more significantly between the two risk groups (Figures 7D–F). Finally, the risk score formula was obtained as follow:

$$\text{RiskScore} = 0.44*\text{S100P} + 0.147*\text{LY6D} + 0.29*\text{WFDC13}$$

According to the expression level of the sample, we calculated the risk score for PAAD samples, and the RiskScore distribution is shown in Figure 8A. From the results of survival analysis, samples with high risk scores showed a significantly worse overall survival (OS) ($p < 0.001$).

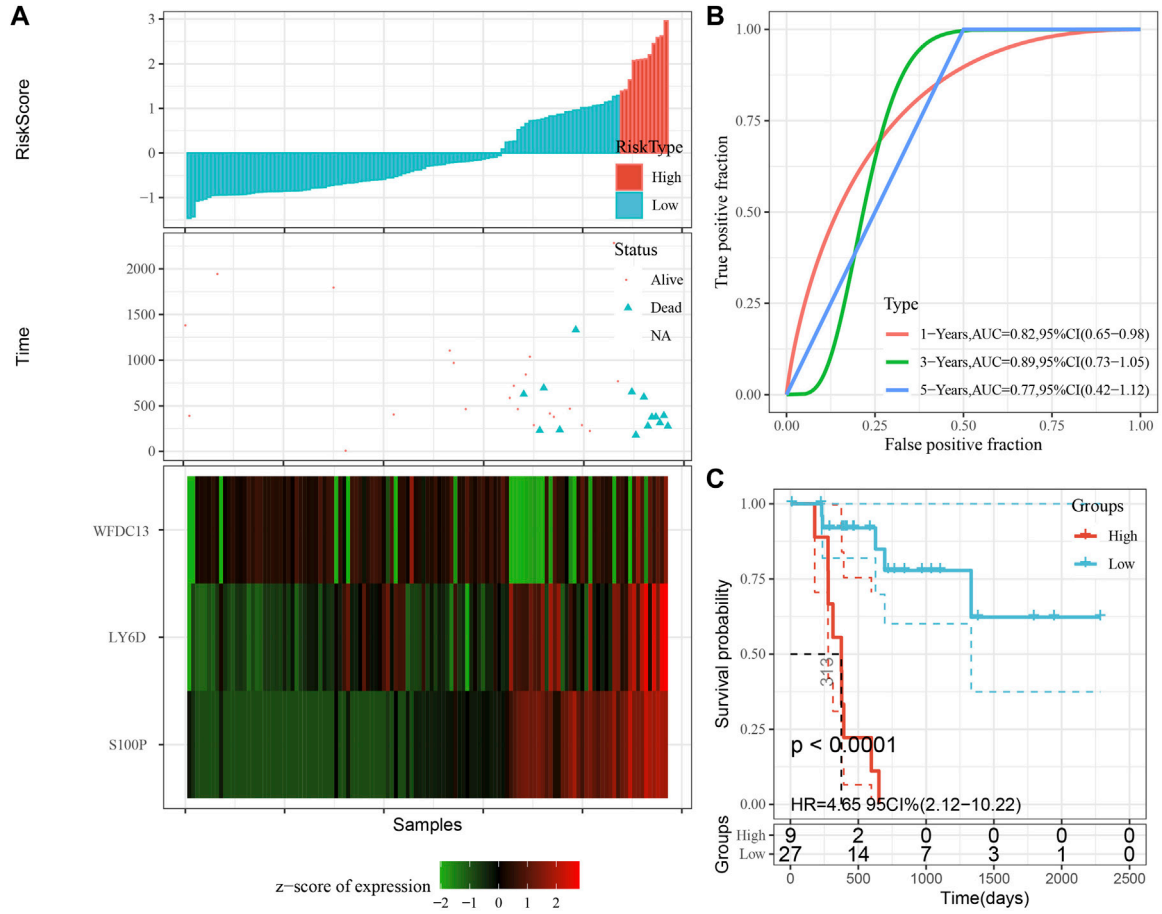


FIGURE 8
Performance of the prognostic gene signature in training set. **(A)** Risk score, survival time and survival status and expression of the 3 genes in training set. **(B)** ROC curve and AUC of the 3-gene signature classification. **(C)** Distribution of KM survival curves of the 3-gene signature in training set.

Then, we used the "timeROC" package to perform ROC analysis for prognostic classification of RiskScore, and the AUCs of predictive classification efficiency were 0.82, 0.89, and 0.77 for one-, three-, and five-year, respectively (Figure 8B), suggesting a good predictive performance. Finally, we performed zscore for Riskscore and determined the cut-off value, divided the sample into high-risk and low-risk groups, and plotted K-M curves. The low-risk group showed significantly better prognosis than that in the high-risk group (Figure 8C, log rank $p < 0.0001$).

3.7 Validation of the prognostic gene signature associated with DMEG

The model was validated further by using the same coefficients and model in the training set as in the validation set. The risk score of each sample was calculated using the same method, and the RiskScore distribution is shown in Figure 9A. Similarly, the AUCs of the classification efficiency of the one-year, three-year, and five-year prognostic predictions were 0.53, 0.86, and 0.85, respectively (Figure 9B), and the OS of the high-risk-score samples was significantly worse than that of the low-risk-score samples (Figure 9C, log rank p = 5e x10-4, HR = 2.42). Next, we used the same coefficients and model in the TCGA-PAAD cohort KRAS wild-type group samples as in the training set. We also calculated risk scores for each sample separately based on the expression level of the samples, and the RiskScore distribution is shown in Figure 10A, with AUCs of 0.72, 0.88, and 0.85 for the prognostic predictive classification efficiency at one, three, and 5 years, respectively (Figure 10B). Survival analysis showed that the OS of the high-risk score samples was significantly smaller than that of the low-risk score samples (Figure 10C, log-rank $p = 0.00039$, HR = 3.78). Finally, we performed the same validation in the ICGC-PAAD external data cohort, and the RiskScore distributions for each sample are shown in Figure 11A. The AUCs for prognostic predictive classification efficiency at one, three, and 5 years were 0.85, 0.85, and 0.91, respectively (Figure 11B), and survival analysis showed that the OS of the high-risk score sample was significantly worse than that of the low-risk score sample (Figure 11C, log rank $p = 0.00024$, HR = 1.68).

4 Discussion

PAAD as one of the most lethal and aggressive malignancies has a 5-year survival rate of less than 10%, (Jiménez et al., 2017), and is now

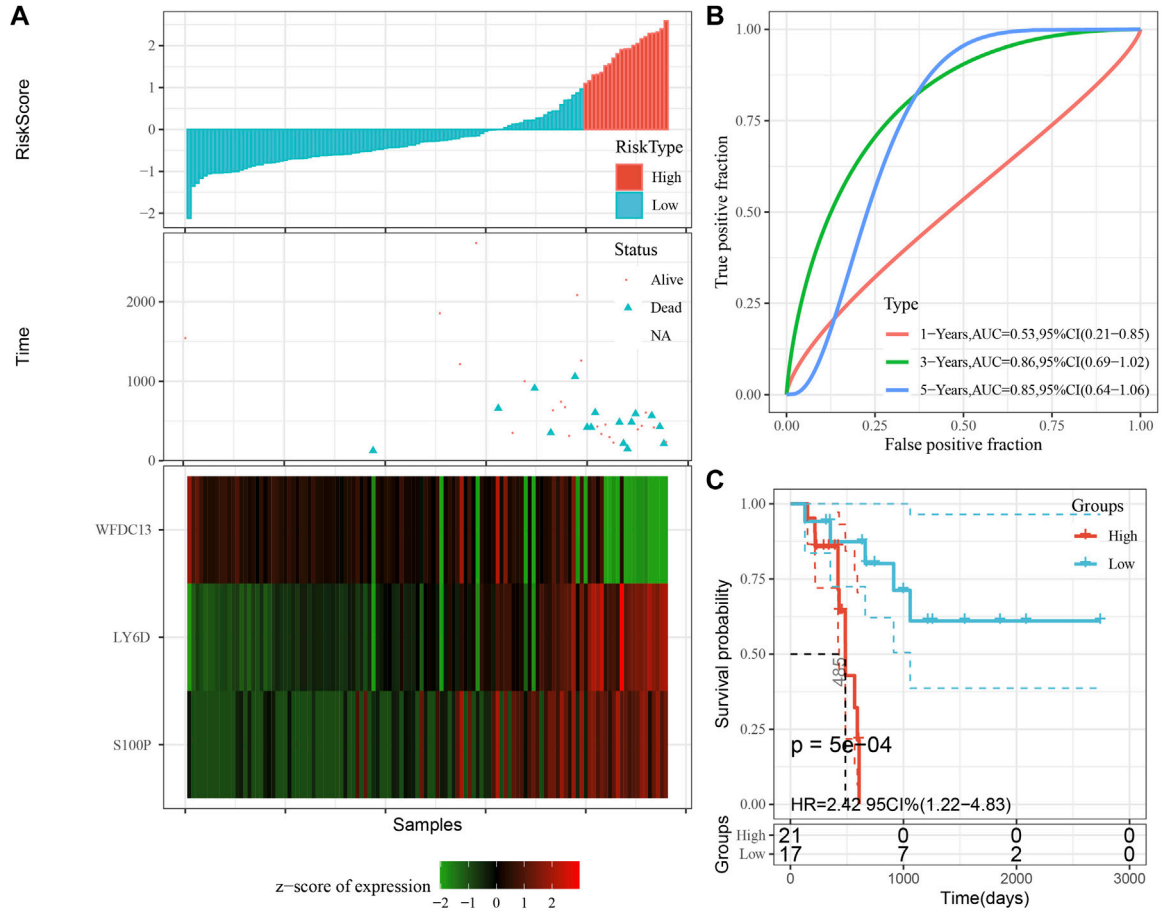


FIGURE 9
Validation of the prognostic gene signature in validation set. **(A)** Risk score, survival time and survival status and expression of the 3 genes. **(B)** ROC curve and AUC of the 3-gene signature classification. **(C)** Distribution of KM survival curves of the 3-gene signature in the validation set.

among the top four leading causes resulting in tumor-associated death (Kleeff et al., 2016). The median age of onset of PAAD is 71 years, and with the aging of the population, its morbidity and mortality will increase rapidly. By 2030, PAAD is estimated as a second cause to tumor mortality (Rahib et al., 2014). The cause of pancreatic cancer is still unclear, and only 5%–10% of pancreatic cancer patients can be attributed to genetic factors (Siegel et al., 2022), although the mutation rate of KRAS reaches 95%, but a single KRAS gene mutation does not lead to the development of pancreatic cancer. Epigenetic alterations are more closely related to environmental and age factors than genetics. Past studies have found that epigenetic alterations occur in the early stages of tumor and are cumulative with tumor development (Nebbioso et al., 2018). In this study, we first DEGs and DMGs in normal samples versus tumor samples without KRAS wild-type based on expression profiling data of pancreatic cancer, and performed functional analysis. Then a classification model was constructed, which can accurately separate normal samples from tumor samples. Finally, we used DMEGs to perform gene-drug interactions on DrugBank to find some potential anti-PAAD drugs, which provides new ideas and potential targets for understanding the role of methylation in PAAD and treating PAAD.

In the early 20th century, Fukushima N and other scholars extensively studied the methylation of different genes in PAAD and its precancerous lesions (intra-epithelial neoplasia (PanIN), and found abnormal methylation of ppENK and p16 (13). Next, it was shown that the incidence of aberrant methylation was 7.3%–7.7% in PanIN-1 patients, 22.7% in PanIN-2 patients, and 46.2% in PanIN-3 patients, a phenomenon that suggests that the incidence of aberrant methylation increases with a more advanced PanIN grade, but the exact mechanism is not clear (Fukushima et al., 2002). Our study, by screening for differentially methylated genes, initially it was found that methylation genes were mediated through Cytokine-cytokine receptor interaction, Natural killer cell-mediated cytotoxicity, Olfactory transduction, and some other pathways leading to the development of PAAD. To further confirm the pathway correlation between PAAD and gene methylation, the intersection of differentially genes and differentially methylated genes was taken and performed enrichment analysis again, and the results demonstrated that methylation led to PAAD by affecting cytokine receptor, NK cell-mediated cytotoxicity pathway.

Illumina human methylation 450 k bead array provides a better technical platform for further study of DNA methylation, therefore, we focused on methylation genes within the three regions of Gene

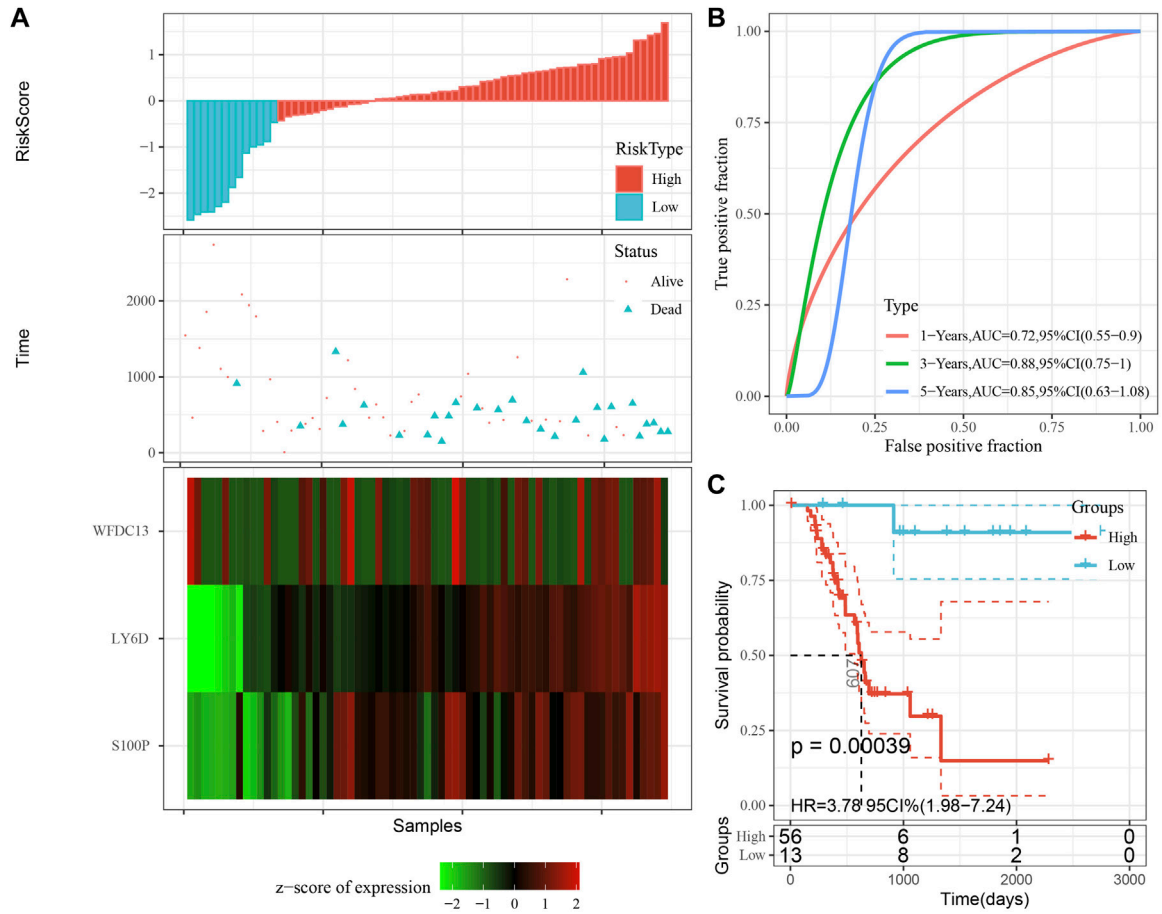


FIGURE 10
Validation of prognostic gene signatures in KRAS wild-type PAAD samples. **(A)** Risk score, survival time and survival status and expression of the 3 genes in KRAS wild-type samples; **(B)** ROC curves and AUC of the 3-gene signature classification; **(C)** Distribution of KM survival curves of 3-gene signature in TCGA KRAS wild-type samples.

body, TSS1500, and TSS200. A total of 758 hypermethylated genes and 418 demethylated genes were identified within the Gene body region, which was consistent with the incidence of PAAD hypomethylation reported in previous studies, and hypomethylation was mainly associated with cell cycle cycling, cell differentiation, and cell surface antigen/cell adhesion (Pedersen et al., 2011; Schäfer et al., 2021; Zhu et al., 2021). TSS1500 is a functional element belonging to differential methylation and is located between 1.5 kb and 200 bp upstream of the transcription start site. Previous studies identified the TSS1500 region as an oncogenic cofactor variable in lung adenocarcinoma and squamous carcinoma by differential methylation probes, and extensive analysis showed that gene probes outside the TSS1500 region could act as potential pathogenic players by affecting the activity of phosphatidylinositol-3,4,5-trisphosphate (Cao et al., 2022). Our study likewise demonstrated an expression imbalance between hypermethylation and hypomethylation in the TSS1500 region, and by using genes in the TSS1500 region, we were able to construct a classification model to distinguish PAAD from normal tissue, providing a useful tool to identify PAAD. tSS200 also belongs to the transcription factor repressor functional element, and methylation in the TSS200 region is not only related to tumor development, but also involved in the acceleration of epigenetic mutational load and epigenetic age, providing a new perspective for our understanding of the age of DNA methylation (Yan et al., 2020).

In 2005, the European Palliative Care Research Collaborative (EPCRC) network working group screened important clinical markers for survival prediction in patients with end-stage cancer based on decades of clinical evidence and recommended a variety of prognostic tools. On this basis, researchers have successively validated and derived several relevant prediction models according to cancer types, and PAAD prognostic models have emerged, which can be broadly classified into traditional manual prediction and statistical-based bioinformatics modeling, with the latter being the majority at present, but they all share common problems such as small sample size, low specificity, and poor predictive performance (Yuan et al., 2021) (Wang et al., 2021; Zhao et al., 2021). Compared with previous PAAD models, we performed model improvement by combining methylation genes (S100P, LY6D, and WFDC13) with clinical factors in prognostic

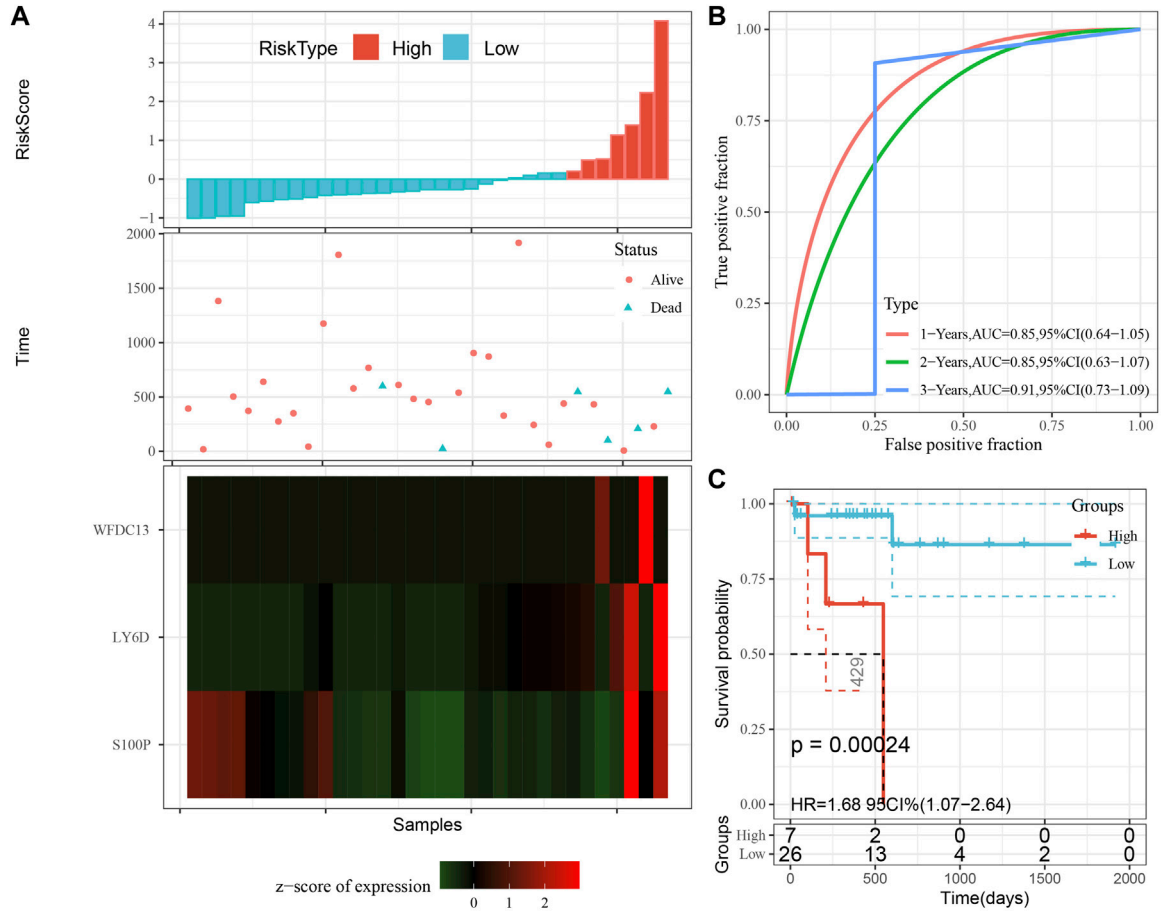


FIGURE 11
Validation of prognostic gene signatures in external datasets. **(A)** Risk score, survival time vs. survival status and expression of the 3 genes; **(B)** ROC curve and AUC for the 3-gene signature classification; **(C)** Distribution of KM survival curves of 3-gene signature in ICGC-PAAD samples.

factors and confirmed the model robustness by external and internal validation. S100P is a member of the S100 protein family containing 2 EF-hand calcium-binding motifs. s100 is localized in the cytoplasm and/or nucleus of a variety of cells and is involved in cell cycle progression and cell differentiation. Meta-analysis showed that S100P is a highly sensitive and highly specific tool for the diagnosis of PAAD (AUC = 0.93) (Hu et al., 2014; Camara et al., 2020). LY6D is mainly involved in lymphoid differentiation and cell surface activity, and the study showed that LY6D is significantly highly expressed in PAAD and is a valid predictor of PAAD, a result consistent with our study (Wang et al., 2020; Xu et al., 2021). WFDC13 belongs to the telomere cluster family of genes, and there are relatively few studies on WFDC13 in PAAD. Our data indicated that WFDC13 was a potential prognostic gene for PAAD and was implicated in the methylation process of PAAD, which provided new ideas for future basic experiments. However, our study was still inadequate and further basic experiments to elucidate the mechanism of the role of this methylation gene in PAAD are required.

There are some limitations in this study. Although the results showed that 3-DMEGs-based signature could distinguish tumor samples and normal samples, the model reliability should be improved with long-term clinical application. Additionally, we downloaded expression profiles and methylation data of PAAD from public databases. Thus, further prospective data should be collected to validate the results. Besides, experimental studies and clinical trials should be performed to verify the results of molecular docking in this study.

5 Conclusion

With the gene expression profile data of PAAD, we identified DEGs and DMGs between normal samples and tumor samples with KRAS wild type; the classification model based on DMEGs was able to accurately separate normal samples from tumor samples, and the gene-drug interactions were performed on DrugBank to find some potential anti PAAD drugs.

Data availability statement

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

Author contributions

All authors contributed to this present work: TJ designed the study and revised the manuscript, HW acquired the data. TC drafted the manuscript. All authors read and approved the manuscript.

Funding

This project is supported by the Key Disciplines of Guangzhou Huadu District People's Hospital Project (2022–2025) (Digestive Diseases Department of Guangzhou Huadu District People's Hospital).

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

Camara, R., Ogbeni, D., Gerstmann, L., Ostovar, M., Hurer, E., Scott, M., et al. (2020). Discovery of novel small molecule inhibitors of S100P with *in vitro* anti-metastatic effects on pancreatic cancer cells. *Eur. J. Med. Chem.* 203, 112621. doi:10.1016/j.ejmech.2020.112621

Cao, L., Ma, X., Rong, P., Zhang, J., Yang, M., and Wang, W. (2022). Comprehensive analysis of DNA methylation and transcriptome to identify PD-1-negative prognostic methylated signature in endometrial carcinoma. *Dis. markers* 2022, 3085289. doi:10.1155/2022/3085289

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., 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

Fukushima, N., Sato, N., Ueki, T., Rosty, C., Walter, K. M., Wilentz, R. E., et al. (2002). Aberrant methylation of preproenkephalin and p16 genes in pancreatic intraepithelial neoplasia and pancreatic ductal adenocarcinoma. *Am. J. pathology* 160 (5), 1573–1581. doi:10.1016/S0002-9440(10)61104-2

Gandhi, N. S., Feldman, M. K., Le, O., and Morris-Stiff, G. (2018). Imaging mimics of pancreatic ductal adenocarcinoma. *Abdom. Radiol. (New York)* 43 (2), 273–284. doi:10.1007/s00261-017-1330-1

Ghosh, D. (2012). Incorporating the empirical null hypothesis into the Benjamini-Hochberg procedure. *Stat. Appl. Genet. Mol. Biol.* 11 (4), 1735. doi:10.1515/1544-6115.1735

Goral, V. (2015). Pancreatic cancer: Pathogenesis and diagnosis. *Asian Pac. J. cancer Prev.* 16 (14), 5619–5624. doi:10.7314/apjcp.2015.16.14.5619

Hu, B., Wu, C., Mao, H., Gu, H., Dong, H., Yan, J., et al. (2022). Subpopulations of cancer-associated fibroblasts link the prognosis and metabolic features of pancreatic ductal adenocarcinoma. *Ann. Transl. Med.* 10 (5), 262. doi:10.21037/atm-22-407

Hu, H., Zhang, Q., Huang, C., Shen, Y., Chen, X., Shi, X., et al. (2014). Diagnostic value of S100P for pancreatic cancer: A meta-analysis. *Tumour Biol. J. Int. Soc. Oncodevelopmental Biol. Med.* 35 (10), 9479–9485. doi:10.1007/s13277-014-2461-4

Hu, J. X., Zhao, C. F., Chen, W. B., Liu, Q. C., Li, Q. W., Lin, Y. Y., et al. (2021). Pancreatic cancer: A review of epidemiology, trend, and risk factors. *World J. gastroenterology* 27 (27), 4298–4321. doi:10.3748/wjg.v27.i27.4298

Ilic, M., and Ilic, I. (2016). Epidemiology of pancreatic cancer. *World J. gastroenterology* 22 (44), 9694–9705. doi:10.3748/wjg.v22.i44.9694

Jiménez, J., Doerr, S., Martínez-Rosell, G., Rose, A. S., and De Fabritiis, G. (2017). DeepSite: Protein-binding site predictor using 3D-convolutional neural networks. *Bioinforma. Oxf. Engl.* 33 (19), 3036–3042. doi:10.1093/bioinformatics/btx350

Kleeff, J., Korc, M., Apte, M., La Vecchia, C., Johnson, C. D., Biankin, A. V., et al. (2016). Pancreatic cancer. *Nat. Rev. Dis. Prim.* 2, 16022. doi:10.1038/nrdp.2016.22

Kocarnik, J. M., Compton, K., Dean, F. E., Fu, W., Gaw, B. L., Harvey, J. D., et al. (2022). Cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life years for 29 cancer groups from 2010 to 2019: A systematic analysis for the global burden of disease study 2019. *JAMA Oncol.* 8 (3), 420–444. doi:10.1001/jamaoncol.2021.6987

Kulis, M., and Esteller, M. (2010). DNA methylation and cancer. *Adv. Genet.* 70, 27–56. doi:10.1016/B978-0-12-380866-0.60002-2

Lohning, A. E., Levonis, S. M., Williams-Noonan, B., and Schweiker, S. S. (2017). A practical guide to molecular docking and homology modelling for medicinal chemists. *Curr. Top. Med. Chem.* 17 (18), 2023–2040. doi:10.2174/1568026617666170130110827

Luu, K., Bazin, E., and Blum, M. G. (2017). pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Resour.* 17 (1), 67–77. doi:10.1111/1755-0998.12592

Maag, J. L. V. (2018). gganatogram: An R package for modular visualisation of anatograms and tissues based on ggplot2. *F1000Research* 7, 1576. doi:10.12688/f1000research.16409.2

Malchiodi, Z. X., and Weiner, L. M. (2021). Understanding and targeting natural killer cell-cancer-associated fibroblast interactions in pancreatic ductal adenocarcinoma. *Cancers* 13 (3), 405. doi:10.3390/cancers13030405

Marabelle, A., Le, D. T., Ascierto, P. A., Di Giacomo, A. M., De Jesus-Acosta, A., Delord, J. P., et al. (2020). Efficacy of pembrolizumab in patients with noncolorectal high microsatellite instability/mismatch repair-deficient cancer: Results from the phase II KEYNOTE-158 study. *J. Clin. Oncol. official J. Am. Soc. Clin. Oncol.* 38 (1), 1–10. doi:10.1200/JCO.19.02105

Matsubayashi, H., Canto, M., Sato, N., Klein, A., Abe, T., Yamashita, K., et al. (2006). DNA methylation alterations in the pancreatic juice of patients with suspected pancreatic disease. *Cancer Res.* 66 (2), 1208–1217. doi:10.1158/0008-5472.CAN-05-2664

Midha, S., Chawla, S., and Garg, P. K. (2016). Modifiable and non-modifiable risk factors for pancreatic cancer: A review. *Cancer Lett.* 381 (1), 269–277. doi:10.1016/j.canlet.2016.07.022

Nebbioso, A., Tambaro, F. P., Dell'Aversana, C., and Altucci, L. (2018). Cancer epigenetics: Moving forward. *PLoS Genet.* 14 (6), e1007362. doi:10.1371/journal.pgen.1007362

Niu, C., Wu, D., Li, A. J., Qin, K. H., Hu, D. A., Wang, E. J., et al. (2021). Identification of a prognostic signature based on copy number variations (CNVs) and CNV-modulated gene expression in acute myeloid leukemia. *Am. J. Transl. Res.* 13 (12), 13683–13696.

Nones, K., Waddell, N., Song, S., Patch, A. M., Miller, D., Johns, A., et al. (2014). Genome-wide DNA methylation patterns in pancreatic ductal adenocarcinoma reveal epigenetic deregulation of SLIT-ROBO, ITGA2 and MET signaling. *Int. J. cancer* 135 (5), 1110–1118. doi:10.1002/ijc.28765

Pedersen, K. S., Bamlet, W. R., Oberg, A. L., de Andrade, M., Matsumoto, M. E., Tang, H., et al. (2011). Leukocyte DNA methylation signature differentiates pancreatic cancer patients from healthy controls. *PloS one* 6 (3), e18223. doi:10.1371/journal.pone.0018223

Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., and Matrisian, L. M. (2014). Projecting cancer incidence and deaths to 2030: The unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* 74 (11), 2913–2921. doi:10.1158/0008-5472.CAN-14-0155

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinforma.* 12, 77. doi:10.1186/1471-2105-12-77

Saikia, S., and Bordoloi, M. (2019). Molecular docking: Challenges, advances and its use in drug discovery perspective. *Curr. drug targets* 20 (5), 501–521. doi:10.2174/1389450119666181022153016

Sato, N., Fukushima, N., Hruban, R. H., and Goggins, M. (2008). CpG island methylation profile of pancreatic intraepithelial neoplasia. *Mod. pathology official J. U. S. Can. Acad. Pathology, Inc.* 21 (3), 238–244. doi:10.1038/modpathol.3800991

Schäfer, D., Tomiuk, S., Küster, L. N., Rawashdeh, W. A., Henze, J., Tischler-Höhle, G., et al. (2021). Identification of CD318, TSPAN8 and CD66c as target candidates for

CAR T cell based immunotherapy of pancreatic adenocarcinoma. *Nat. Commun.* 12 (1), 1453. doi:10.1038/s41467-021-21774-4

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *CA a cancer J. Clin.* 72 (1), 7–33. doi:10.3322/caac.21708

Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Statistics Med.* 16 (4), 385–395. doi:10.1002/(sici)1097-0258(19970228)16:4<385:aid-sim380>3.0.co;2-3

Trott, O., and Olson, A. J. (2010). AutoDock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31 (2), 455–461. doi:10.1002/jcc.21334

Velankar, S., Burley, S. K., Kurisu, G., Hoch, J. C., and Markley, J. L. (2021). The protein Data Bank archive. *Methods Mol. Biol. Clift. NJ)* 2305, 3–21. doi:10.1007/978-1-0716-1406-8_1

Wang, J., Fan, J., Gao, W., Wu, Y., Zhao, Q., Chen, B., et al. (2020). LY6D as a chemoresistance marker gene and therapeutic target for laryngeal squamous cell carcinoma. *Stem cells Dev.* 29 (12), 774–785. doi:10.1089/scd.2019.0210

Wang, L., Zhang, S., Li, H., Xu, Y., Wu, Q., Shen, J., et al. (2021). Quantification of m6A RNA methylation modulators pattern was a potential biomarker for prognosis and associated with tumor immune microenvironment of pancreatic adenocarcinoma. *BMC cancer* 21 (1), 876. doi:10.1186/s12885-021-08550-9

Wang, Y. P., Zhou, W., Wang, J., Huang, X., Zuo, Y., Wang, T. S., et al. (2016). Arginine methylation of MDH1 by CARM1 inhibits glutamine metabolism and suppresses pancreatic cancer. *Mol. Cell.* 64 (4), 673–687. doi:10.1016/j.molcel.2016.09.028

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. Camb. (Mass))* 2 (3), 100141. doi:10.1016/j.xinn.2021.100141

Xu, D., Wang, Y., Zhang, Y., Liu, Z., Chen, Y., and Zheng, J. (2021). Systematic analysis of an invasion-related 3-gene signature and its validation as a prognostic model for pancreatic cancer. *Front. Oncol.* 11, 759586. doi:10.3389/fonc.2021.759586

Xu, R., Xu, Q., Huang, G., Yin, X., Zhu, J., Peng, Y., et al. (2019). Combined analysis of the aberrant epigenetic alteration of pancreatic ductal adenocarcinoma. *BioMed Res. Int.* 2019, 9379864. doi:10.1155/2019/9379864

Yan, Q., Paul, K. C., Lu, A. T., Kusters, C., Binder, A. M., Horvath, S., et al. (2020). Epigenetic mutation load is weakly correlated with epigenetic age acceleration. *Aging* 12 (18), 17863–17894. doi:10.18632/aging.103950

Yuan, H., Liu, J., Zhao, L., Wu, P., Chen, G., Chen, Q., et al. (2021). Prognostic risk model and tumor immune environment modulation of m5C-related LncRNAs in pancreatic ductal adenocarcinoma. *Front. Immunol.* 12, 800268. doi:10.3389/fimmu.2021.800268

Zhang, Z. (2016). Variable selection with stepwise and best subset approaches. *Ann. Transl. Med.* 4 (7), 136. doi:10.21037/atm.2016.03.35

Zhao, S., Nicolle, R., Augustin, J., Svrcek, M., de Mestier, L., Le Corre, D., et al. (2021). Prognostic relevance of pancreatic adenocarcinoma whole-tumor transcriptomic subtypes and components. *Clin. cancer Res. official J. Am. Assoc. Cancer Res.* 27 (23), 6491–6499. doi:10.1158/1078-0432.CCR-21-1907

Zhou, W., Liu, Q., Wang, W., Yuan, X. J., Xiao, C. C., and Ye, S. D. (2022). Comprehensive network analysis reveals the targets and potential multitarget drugs of type 2 diabetes mellitus. *Oxidative Med. Cell. Longev.* 2022, 8255550. doi:10.1155/2022/8255550

Zhu, J., Zhou, Y., Zhu, S., Li, F., Xu, J., Zhang, L., et al. (2021). circRNA circ_102049 implicates in pancreatic ductal adenocarcinoma progression through activating CD80 by targeting miR-455-3p. *Mediat. Inflamm.* 2021, 8819990. doi:10.1155/2021/8819990

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 20 March 2023
DOI 10.3389/fphar.2023.1135626

OPEN ACCESS

EDITED BY
Zhi-Qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Zhenyao Chen,
Fudan University, China
Yan Wang,
Beijing Obstetrics and Gynecology Hospital, Capital Medical University, China
Zigang Zhao,
Hainan Hospital of PLA General Hospital, China

*CORRESPONDENCE
Jianhua Zeng,
300541@hospital.cqmu.edu.cn

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 01 January 2023
ACCEPTED 07 March 2023
PUBLISHED 20 March 2023

CITATION
Liang H, Xiang L, Wu H, Liu Y, Tian W and Zeng J (2023), Anoikis-related long non-coding RNA signatures to predict prognosis and small molecular drug response in cervical cancer.
Front. Pharmacol. 14:1135626.
doi: 10.3389/fphar.2023.1135626

Anoikis-related long non-coding RNA signatures to predict prognosis and small molecular drug response in cervical cancer

Hao Liang[1], Lan Xiang[2], Huan Wu[1], Yang Liu[1], Wei Tian[3] and Jianhua Zeng[1]*

[1]Department of Obstetrics and Gynecology, The Second Affiliated Hospital, Chongqing Medical University, Chongqing, China, [2]Department of Hepatobiliary Surgery, The First Affiliated Hospital, Chongqing Medical University, Chongqing, China, [3]Department of General Surgery, The Second Affiliated Hospital of Tianjin, University of Traditional Chinese Medicine, Tianjin, China

Background: Cervical cancer (CC) is a major health threat to females, and distal metastasis is common in patients with advanced CC. Anoikis is necessary for the development of distal metastases. Understanding the mechanisms associated with anoikis in CC is essential to improve its survival rate.

Methods: The expression matrix of long non-coding RNAs (lncRNAs) from cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) patients was extracted from The Cancer Genome Atlas (TCGA), and highly relevant anoikis-related lncRNAs (ARLs) were identified by the single sample gene set enrichment analysis (ssGSEA) method. ARLs-related molecular subtypes were discerned based on prognosis-related ARLs. ARLs-related prognostic risk score (APR_Score) was calculated and risk model was constructed using LASSO COX and COX models. In addition, we also assessed immune cell activity in the immune microenvironment (TME) for both subtypes and APR_Score groups. A nomogram was utilized for predicting improved clinical outcome. Finally, this study also discussed the potential of ARLs-related signatures in predicting response to immunotherapy and small molecular drugs.

Results: Three ARLs-related subtypes were identified from TCGA-CESC (AC1, AC2, and AC3), with AC3 patients having the highest ARG scores, higher angiogenesis scores, and the worst prognosis. AC3 had lower immune cell scores in TME but higher immune checkpoint gene expression and higher potential for immune escape. Next, we constructed a prognostic risk model consisting of 7-ARLs. The APR_Score exhibited a greater robustness as an independent prognostic indicator in predicting prognosis, and the nomogram was a valuable tool for survival prediction. ARLs-related signatures emerged as a potential novel indicator for immunotherapy and small molecular drug selection.

Conclusion: We firstly constructed novel ARLs-related signatures capable of predicting prognosis and offered novel ideas for therapy response in CC patients.

Abbreviations: ARLs, Anoikis-related lncRNAs; APR_Score, Anoikis prognostic risk model; ARGs, Anoikis-related genes; AUC, area under ROC curve; CC, Cervical cancer; FC, fold change; GSVA, Gene set variation analysis; GSEA, Gene set enrichment analysis; HR, hazard ratio; IC50, the half maximal inhibitory concentration; IDE, integrated development environment; K-M, Kaplan-Meier; LASSO, least absolute shrinkage and selection operator; MSigDB, Molecular Signatures Database; OS, overall survival; ROC, receiver operating characteristic analysis; ssGSEA, single-sample gene set enrichment analysis; TCGA, The Cancer Genome Atlas; TIDE, Tumor Immune Dysfunction and Exclusion; TME, Tumor microenvironment.

KEYWORDS

TCGA, cervical cancer, anoikis, LncRNAs, prognosis, immunotherapy, small molecular drug

Introduction

Cervical cancer (CC) is a serious health threat to females, and according to WHO data, CC is the fourth leading cause of cancer deaths among females, accounting for 342,000 deaths in 2020 (Sung et al., 2021). The most common histological type of CC is squamous cell carcinoma that accounts for about 75% of all cases, followed by the second most common type of adenocarcinoma that accounts for 20% of all cases (Meng et al., 2021). Unlike other cancers among females, effective measures including effective early screening and HPV vaccination could significantly reduce the incidence of CC and early targeted treatment (Canfell et al., 2020). Current mainstream treatment options for CC patients including systemic surgical treatment and radiotherapy could improve the 5-year survival rate of early CC patients to 91.5% (Bhatla et al., 2018). However, survival outcomes for patients with advanced CC due to postoperative recurrence and distant metastases are unsatisfactory (Marth et al., 2017). Therefore, there was an immediate demand for monitoring biomarkers to evaluate the CC metastasis risk.

Anoikis, a defense mechanism to prevent abnormal cell migration, is mainly manifested as the loss of cell adhesion to ECM and cell apoptosis and death in the process of migration (Paoli et al., 2013). Anoikis plays an essential component in inhibiting tumor cell migration and is important in delaying cancer progression (Adeshakin et al., 2021). Studies have found that tumor cells secrete growth factors and activate EMT signaling pathway to resist Anoikis, which also became a prerequisite for tumor cell metastasis (Kim et al., 2012). Anti-anoikis has become a landmark event in the occurrence of remote cancer metastasis (Kim et al., 2012). However, few studies focused on the correlation between anoikis and distal metastasis of CC. In view of the inability to accurately evaluate the prognosis of advanced CC patients, it was indispensable to explore the mechanism of anoikis in CC to improve its treatment.

Therefore, we developed an assessment system for prognostic risk stratification and constructed an anoikis-related long non-coding RNAs (ARLs)-related prognostic model in CC. We further investigated the relationship among prognostic indicators and immune microenvironment (TME), immunotherapeutic response, and chemotherapeutic drug suitability. The purpose of this study was to formulate a novel prognostic scoring system for CC, which was designed to accurately guide the prognosis of CC and improve its treatment options.

Materials and methods

Dataset download and processing

RNA-sequencing data and corresponding clinical information of cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) samples were available in The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/) database using TCGA GDC API. We also downloaded mutect2-processed single-nucleotide variants (SNVs) data from TCGA. The expression matrix with TPM format was transformed into log 2 (TPM+1). Samples were filtered using the Sangerbox (http://sangerbox.com/) (Shen et al., 2022) database as follows: 1) removing samples without follow-up information; 2) retaining samples with survival time greater than 0; 3) remove samples without Status. A total of 291 samples were included after screening. RNA annotation file in GENCODE (https://www.gencodegenes.org/) was used to obtain mRNAs and lncRNAs expression matrix. Anoikis-related genes (ARGs) were available in GeneCards (https://www.genecards.org/).

Recognition of anoikis-related lncRNAs (ARLs) and molecular subtypes

Based on the expression profiles of ARGs, the ARG scores of samples were attained *via* the single sample gene set enrichment

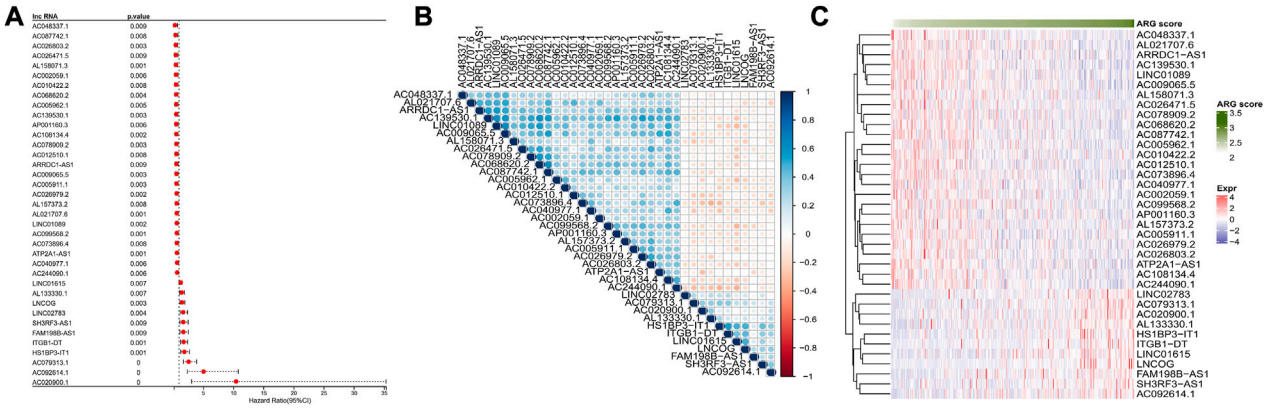


FIGURE 1
Identification of ARLs TCGA-CESC **(A)** Forest map of ARLs affected OS **(B)** Heatmap of ARLs correlation affected OS **(C)** 37 Expression heatmap OS related ARLs.

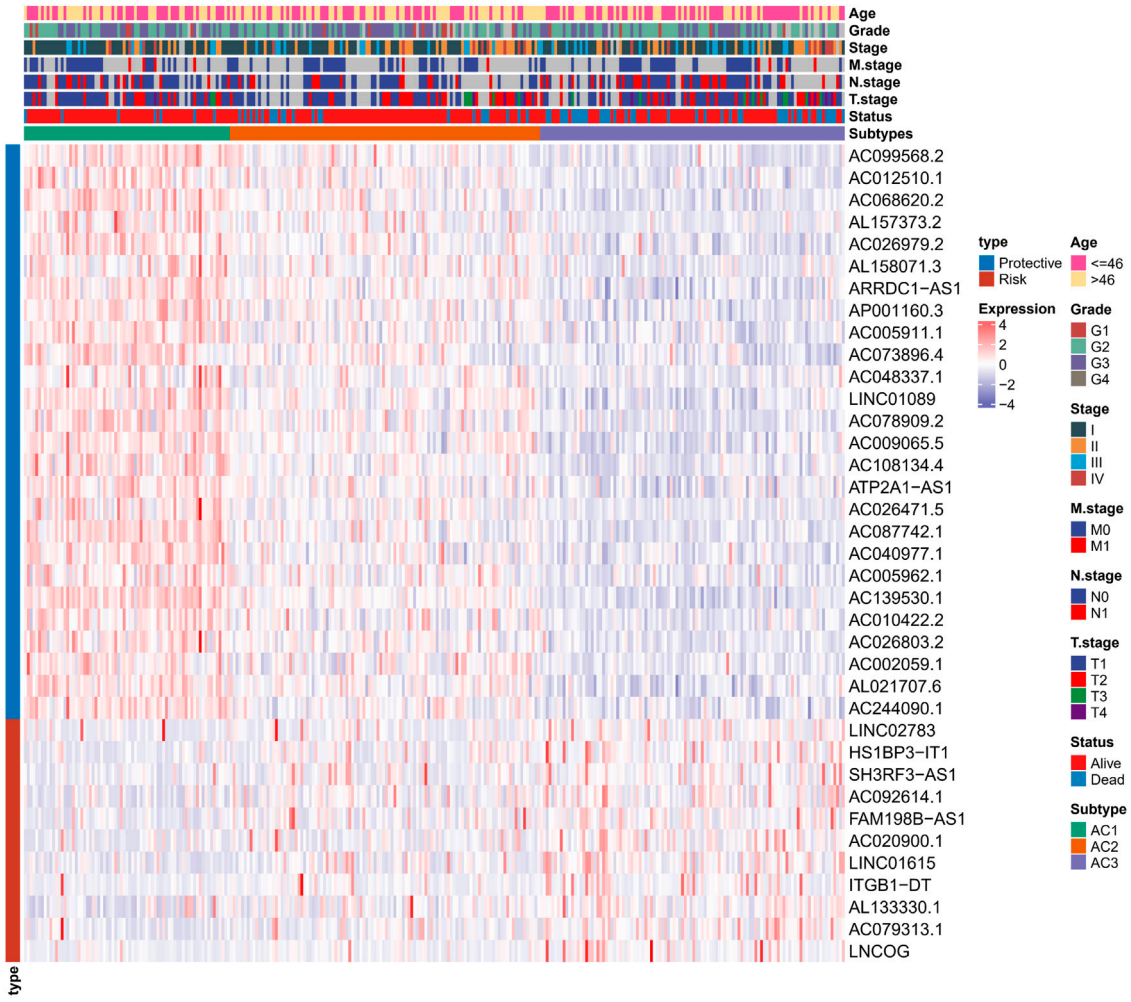


FIGURE 2
Heat map shows the expression of 37 ARLs in ARLs-related molecular subtypes in CC patients and clinical information.

analysis (ssGSEA) algorithm (Barbie et al., 2009). ARLs were identified using the rcorr function in the Hmisc package under screening thresholds of |cor|>0.3 and $p < 0.01$ (Hmisc.pdf). Univariate COX models were performed on the ARLs to screen for prognostically relevant ARLs under the threshold of $p < 0.01$. A consistency clustering analysis was carried out on samples in TCGA-CESC to identify molecular subtypes, according to the methods of Wilkerson et al. (Wilkerson and Hayes, 2010).

Gene set variation analysis (GSVA) and gene mutation landscape in ARLs-related subtypes

For patients in different subtype groups, hallmark gene sets were captured from the Molecular Signatures Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/) and GSVA was conducted using the GSVA package to explore differences in biological pathway variants across groups (Hanzelmann et al., 2013). Next, oncogenic pathway signatures were obtained from Sanchez-Vega et al. (Sanchez-Vega et al., 2018) and differences in oncogenic pathway scores were assessed in subtypes using the ssGSEA method. SNV data from CC were processed in The Genome Analysis Toolkit (GATK) software in the mutect2 plugin. Genes with mutation frequencies ≥3 were screened (fisher test, $p < 0.05$), and the mutation landscape was mapped using the maftools package (Mayakonda et al., 2018).

Construction of ARLs-related prognostic risk model (APR_Score)

The TCGA-CESC samples were clustered into training and test sets at 1:1 ratio for analysis. In the training set, LASSO and multivariate COX models were executed on prognosis-related ARLs to discriminate ARLs significantly affecting CC prognosis. ARLs-related prognostic risk model was constructed based on the following formula (Simon et al., 2011).

$$\text{APR\_Score} = \sum Expression_{\text{ARLs}} * \beta_{ARLs}$$

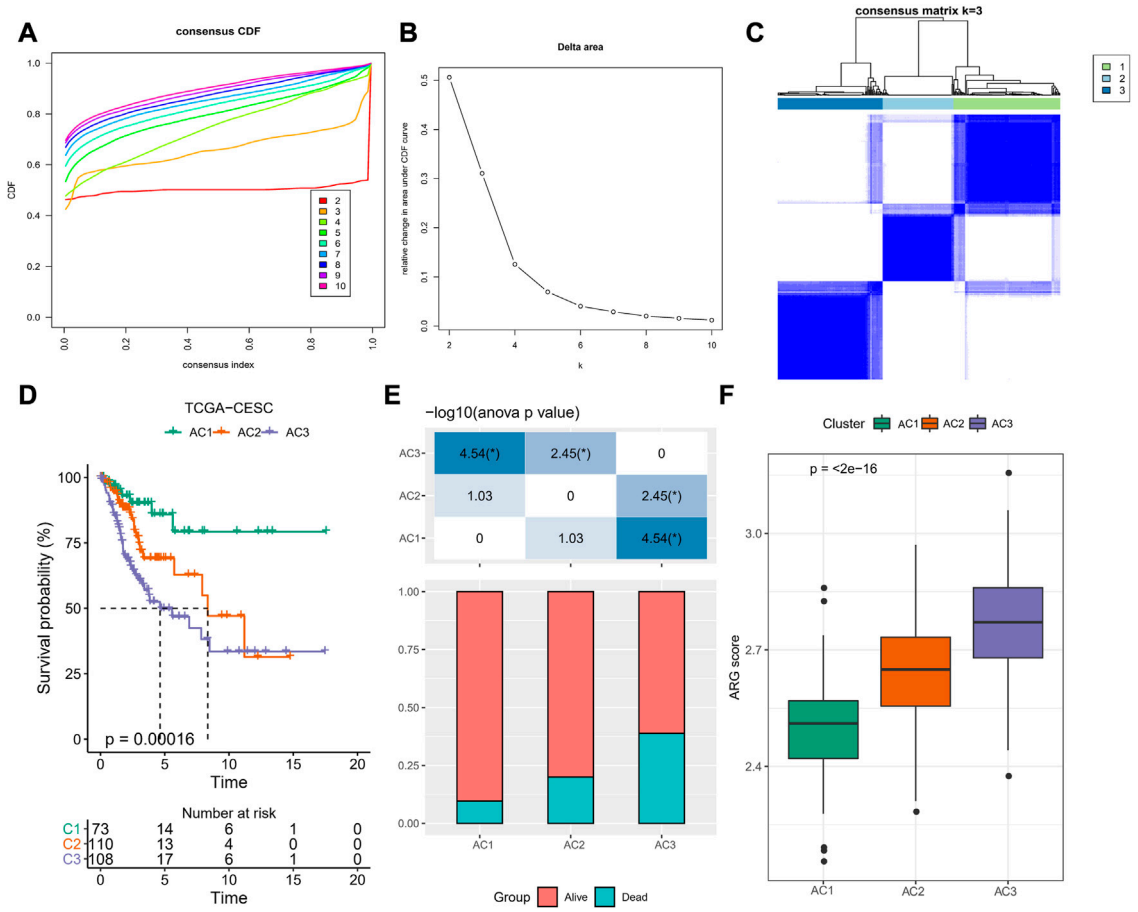


FIGURE 3
ARLs-related molecular subtypes in TCGA-CESC. **(A)** Cumulative distribution function (CDF) of consensus clustering **(B)** CDF Delta area curve for k = 2–10 **(C)** Heatmap of sample clusters shows consensus matrix k=3 **(D)** K-M curves of AC1-3 patients in TCGA-CESC **(E)** Statistical graph of patient survival status for patients in AC1-3 **(F)** Distribution of ARG score amongst clusters AC1-3. *p < 0.05.

In the formula, expression of ARLs represented the expression data of prognosis-related ARLs, and βARLs indicated the COX regression coefficients after normalization. We calculated the ARLs-related prognostic risk score (APR_Score) of each patient in TCGA. Patients were assigned to the high-APR_Score and low-APR_Score groups according to the optimal p-value based on survminer package.

Clinical meaning of APR_Score and association with prognosis

We analyzed the variation in APR_Score and prognosis amongst the molecular subtypes. In this study, for different molecular subtypes and APR_Score groups, K-M curves were evaluated in the training set and the validation set to assess prognostic differences, and ROC curves were used to assess the accuracy of APR_Score. Furthermore, univariate and multivariate COX model analyses were performed on the TCGA-CESC cohort.

Gene set enrichment analysis (GSEA)

For patients in different APR_Score groups, GSEA was performed to explore differences in the biological pathways involved (p < 0.05, FDR<0.25).

Relationship between CC patients and tumor microenvironment (TME)

For ARLs-related molecular subtypes and APR_Score groups, we quantified the relative abundance of 22 immune cell species in the TME of CC patients using the CIBERSORT algorithm (Chen et al., 2018). Next, we further assessed the immune scores of 10 immune cell species in the TME using the MCP-Count method (Becht et al., 2016). Finally, 47 classes of immune checkpoint genes were obtained from the study by Danilova et al. (Danilova et al., 2019) to assess their expression levels in subtypes of CC patients.

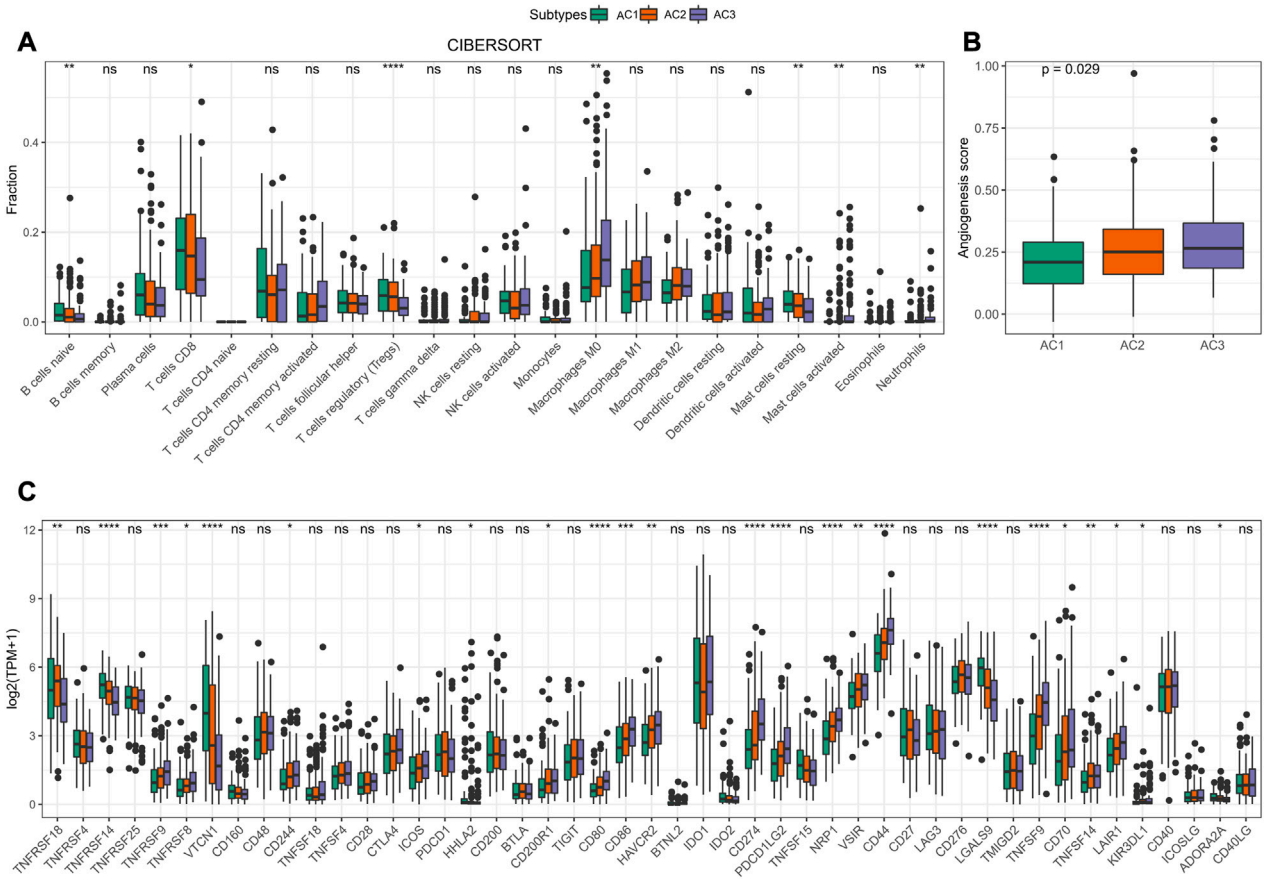


FIGURE 4
Features of TME in AC1-3 subtypes. **(A)** Distribution of immune cell scores estimated by CIBERSORT amongst three subtypes **(B)** Differences of Angiogenesis scores amongst three subtypes **(C)** Differences of immune checkpoint gene expression amongst three subtypes. ns $p > 0.05$; *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$; ****$p < 0.0001$.

Construction of nomogram

In this study, we constructed a nomogram for predicting 1-year, 3-year and 5-year survival rates of patients using the APR_Score (rms.pdf), and calibration curves were applied to evaluate the predictive accuracy of the nomogram. Finally, the ROC curves were exploited to validate the clinical use of APR_Score and nomogram.

Immunotherapy and small molecular drug sensitivity analysis

We assessed TIDE scores in the Tumor Immune Dysfunction and Exclusion (TIDE, http://tide.dfci.harvard.edu/) website between the APR_Score groups. In addition, we computed the half maximal inhibitory concentration (IC50) values of commonly used small molecular drugs in the pRRophetic package (Geeleher et al., 2014).

Statistical analysis

The packages included in this study were downloaded from R (version 4.1.1, https://www.r-project.org/) and analyzed using R Studio, an integrated development environment (IDE) for the R language. The packages included Hmisc, ConsensusClusterPlus, GSVA, survival, survminer, glmnet, maftools, timeROC, rms, and pRRophetic. Sangerbox was deployed for sample screening and data processing. In this study, $p < 0.05$ was considered statistically significant.

Results

Identification of ARLs in TCGA-CESC

Based on annotation file in GENCODE, a total of 14,176 lncRNAs were obtained. Firstly, we computed the ARG score for samples of TCGA-CESC *via* ssGSEA, and 574 ARLs (|cor|>0.3 and $p < 0.01$) were screened based on correlation analysis. This was followed by univariate COX model analysis, filtering a total of 37 prognosis-related ARLs (26 risk factors and 11 protective factors, $p < 0.01$) (Figure 1A). Furthermore, we analyzed the correlation among 37-ARLs, and the correlation heat map was illustrated in Figure 1B. Finally, the expression of 37-ARLs in 291 TCGA-CESC samples was counted, and we observed that 37-ARLs were differentially expressed, showing changes of ARG score in TCGA-CESC patients (Figure 1C).

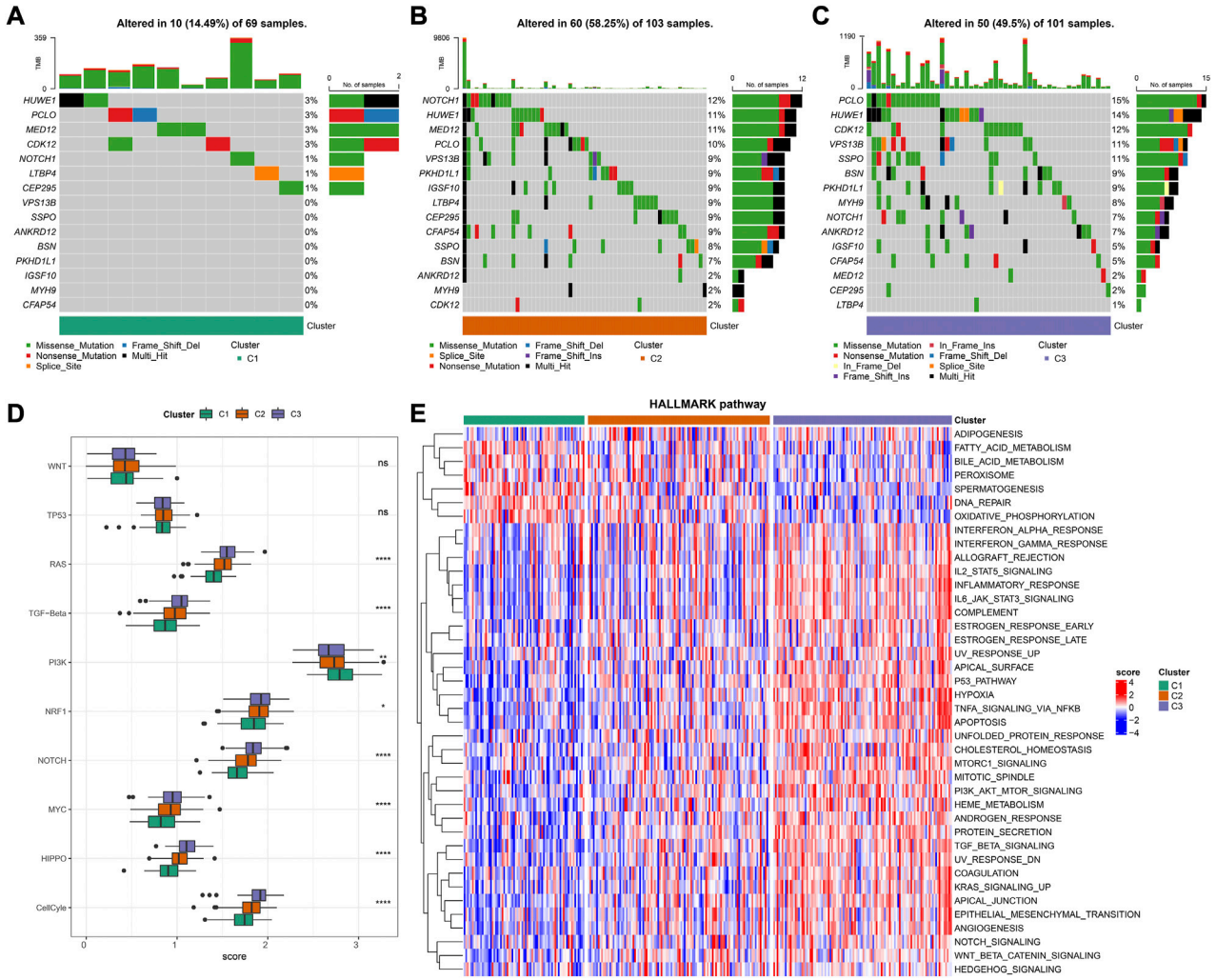


FIGURE 5
Genetic Landscapes and Biological Pathways in AC1-3 subtypes **(A–C)** The mutation frequencies of top 15 genes in clusters AC1-3 **(D)** Distribution of tumorigenesis-related pathways score in AC1-3 **(E)** Heatmap shows pathway characteristic based on GSVA results. Red represents ssGSEA score is greater than 0 while blue represents ssGSEA score is lower than 0. ns $p > 0.05$; *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$; ****$p < 0.0001$.

ARLs-related molecular subtypes in TCGA-CESC

To determine the potential connections among ARG scores and clinical information, this study counted the expression of clinical information and 37-ARLs in TCGA-CESC patients. We found that risk factors were highly expressed in AC3 and protective factors were highly expressed in AC1 (Figure 2). Classification of TCGA-CESC patients was conducted based on consistent clustering of the 37-ARLs expression matrix. We found that the TCGA-CESC patients could be significantly clustered into three clusters (Figures 3A–C), therefore we derived three ARLs-related molecular subtypes, namely, ARLs-Cluster 1 (AC1), ARLs-Cluster 2 (AC2), and ARLs-Cluster 3 (AC3). Furthermore, the K-M curves showed significant prognostic differences between patients with AC1-3, with AC1 patients having the optimal prognosis and AC3 having the poorest prognosis (Figure 3D). Survival status statistics also showed more deaths among AC3 patients ($p < 0.05$, Figure 3E). We examined the ARG scores of patients in AC1-3 and found that the highest ARG scores were in AC3 patients (Figure 3F).

Features of TME in AC1-3 subtypes

To determine the relationship between ARG score and TME in CC, we assessed the relative infiltration abundance of 22 immune cell species in TME using the CIBERSORT method. We determined that the infiltration abundance of B cells naïve, T cells CD8, T cells regulatory (Tregs), and Mast cells resting was significantly lower in AC3 than in AC1 and AC2. In contrast, the infiltration risk of Macrophages M0, Mast cells activated, and Neutrophils was remarkably higher in AC3 than in AC1 and AC2

TABLE 1 Clinical information of CC patients in the training and validation sets.

Characteristics	Train(*N* = 146)	Test (*N* = 145)	Total (*N* = 291)	pvalue	FDR
Status				0.76	1
Alive	112 (38.49%)	108 (37.11%)	220(75.60%)		
Dead	34 (11.68%)	37 (12.71%)	71 (24.40%)		
T.stage				0.49	1
T1	70 (24.05%)	67 (23.02%)	137 (47.08%)		
T2	30 (10.31%)	37 (12.71%)	67 (23.02%)		
T3	9 (3.09%)	7(2.41%)	16(5.50%)		
T4	3 (1.03%)	7 (2.41%)	10 (3.44%)		
Uknown	34 (11.68%)	27 (9.28%)	61 (20.96%)		
N.stage				0.91	1
N0	63 (21.65%)	65 (22.34%)	128 (43.99%)		
N1	29 (9.97%)	26 (8.93%)	55 (18.90%)		
Uknown	54 (18.56%)	54 (18.56%)	108 (37.11%)		
M.stage				0.99	1
M0	53 (18.21%)	54 (18.56%)	107 (36.77%)		
M1	5 (1.72%)	5 (1.72%)	10 (3.44%)		
Uknown	88 (30.24%)	86 (29.55%)	174 (59.79%)		
Stage				0.77	1
I	84 (28.87%)	75(25.77%)	159 (54.64%)		
II	29 (9.97%)	35 (12.03%)	64 (21.99%)		
III	20 (6.87%)	21 (7.22%)	41 (14.09%)		
IV	11 (3.78%)	10 (3.44%)	21 (7.22%)		
Uknown	2 (0.69%)	4 (1.37%)	6 (2.06%)		
Grade				0.86	1
G1	8 (2.75%)	10 (3.44%)	18 (6.19%)		
G2	65 (22.34%)	64 (21.99%)	129 (44.33%)		
G3	59 (20.27%)	57 (19.59%)	116 (39.86%)		
G4	0 (0.0e+0%)	1 (0.34%)	1 (0.34%)		
Uknown	14 (4.81%)	13 (4.47%)	27 (9.28%)		
Age1				0.86	1
<=46	76 (26.12%)	73 (25.09%)	149 (51.20%)		
>46	70 (24.05%)	72 (24.74%)	142 (48.80%)		

(Figure 4A). Accumulating evidence indicated that angiogenesis could influence tumor development and metastasis (Detmar, 2000; Parmar and Apte, 2021). Therefore, we computed the Angiogenesis score in AC1-3 and clearly observed that the poor prognosis AC3 subtype had the highest Angiogenesis score ($p < 0.05$, Figure 4B). In addition, we found that 17 immune checkpoint genes were highly expressed in clusters AC3 (Figure 4C). These results suggested that a higher ARG score was associated with immunosuppressive activity in TME of CC, which could lead to higher Angiogenesis scores, and that the

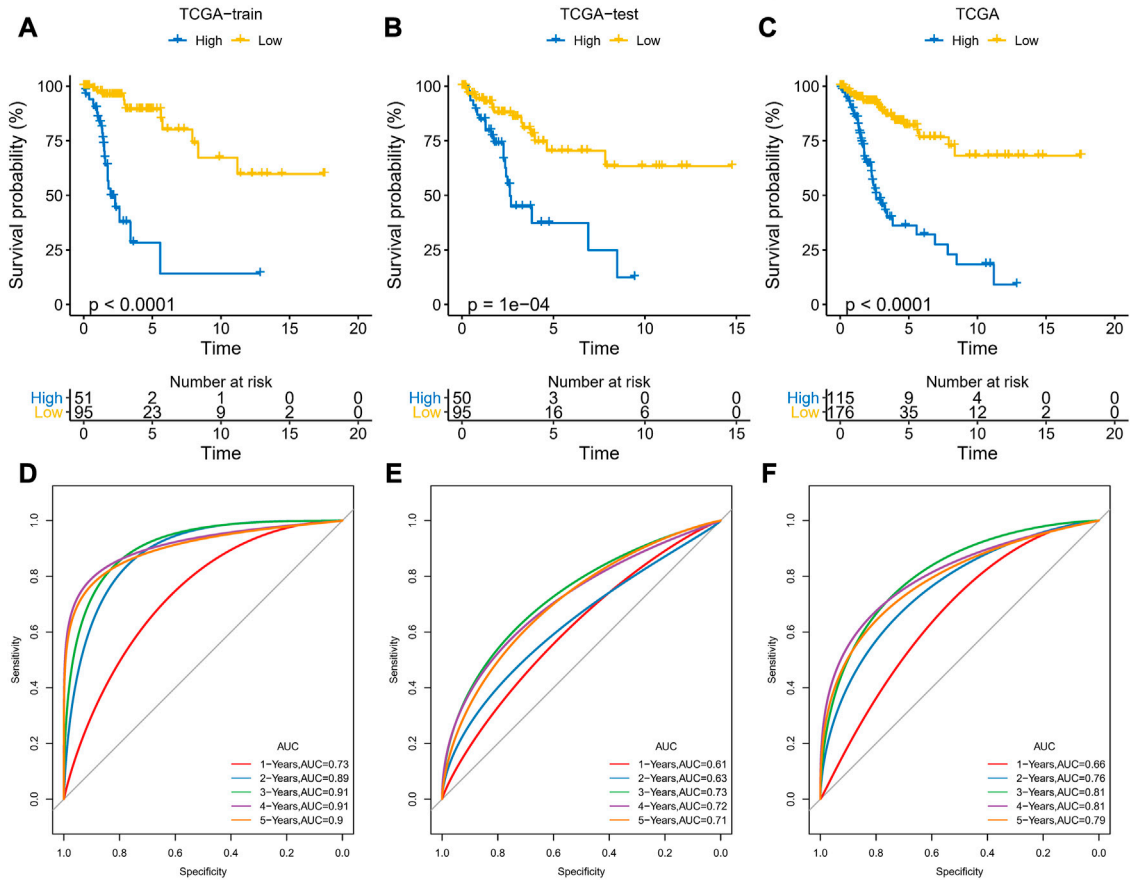


FIGURE 6
Development and validation of APR_Score **(A–C)** K-M curves for the 7 ARLs in the training set, validation set and TCGA-CESC cohort **(D–F)** ROC curves of the training set, validation set and TCGA-CESC cohort.

development and metastasis of CC might be inextricably linked to immune escape.

Genetic landscapes and biological pathways in AC1-3 subtypes

Based on the genomic mutational landscape, we found variation among AC1-3 subtypes. AC3 showed markedly higher mutation frequencies than in AC1 and AC2, with the highest mutation frequencies for PCLO, HDWE1 and CDK12 in AC3 (Figures 5A–C). Based on the results of ssGSEA analysis, we noted that AC3 was substantially enriched in tumorigenesis-related pathways containing RAS, TGF-Beta, NRF1, NOTCH, MYC, HIPPO, CellCyle (Figure 5D). Similarly, GSVA results further validated the result because TGF BETA SIGNALING, PI3K AKT MTOR SIGNALING, TNFA SIGNALING VIA NFKB, WNT BETA CATENIN SIGNALING and NOTCH SIGNALING were remarkably activated in AC3 but remarkably inhibited in AC1 (Figure 5E). These results further supported a higher risk of metastasis in patients with CC in AC3, which in turn could lead to the negative prognosis.

Construction and validation of APR_Score

The prognostic risk model was constructed based on ARLs by randomly dividing patients in TCGA-CESC into training and validation set at a ratio of 1:1, with the clinical information of the samples in each group shown in Table 1. In the training set, the LASSO COX model was applied to optimize the model, and 11 ARLs were identified based on the penalty parameter lambda and the model trajectory change curve (Supplementary Figure S1). Seven ARLs affecting prognosis were selected based on the multivariate COX model, namely AC092614.1, AL158071.3, AC016394.2, LINC02749, MIR100HG, AC079313.1, and ATP2A1-AS1. According to APR_Score=1.847*AC092614.1+(−0.88*AL158071.3)+(−0.675*AC016394.2) +(−1.099*LINC02749)+0.514*MIR100HG+0.766*AC079313.1+(-0.323*ATP2A1-AS1), the prognostic risk model was assessed. Patients in the training and validation sets were classified into high-APR_Score and low-APR_Score groups by the optimal *p*-value in the survminer. Based on the K-M curves, we noted that high APR_Score scores predicted poorer OS in the training set, validation set and total TCGA-CESC cohort (Figures 6A–C). In the training set, the AUCs were 0.73, 0.89, 0.91, 0.91 and 0.9 for 1, 2, 3, 4 and 5 year(s) of survival, respectively (Figure 6D). While in the validation set and the total TCGA-CESC cohort, APR_Score still showed

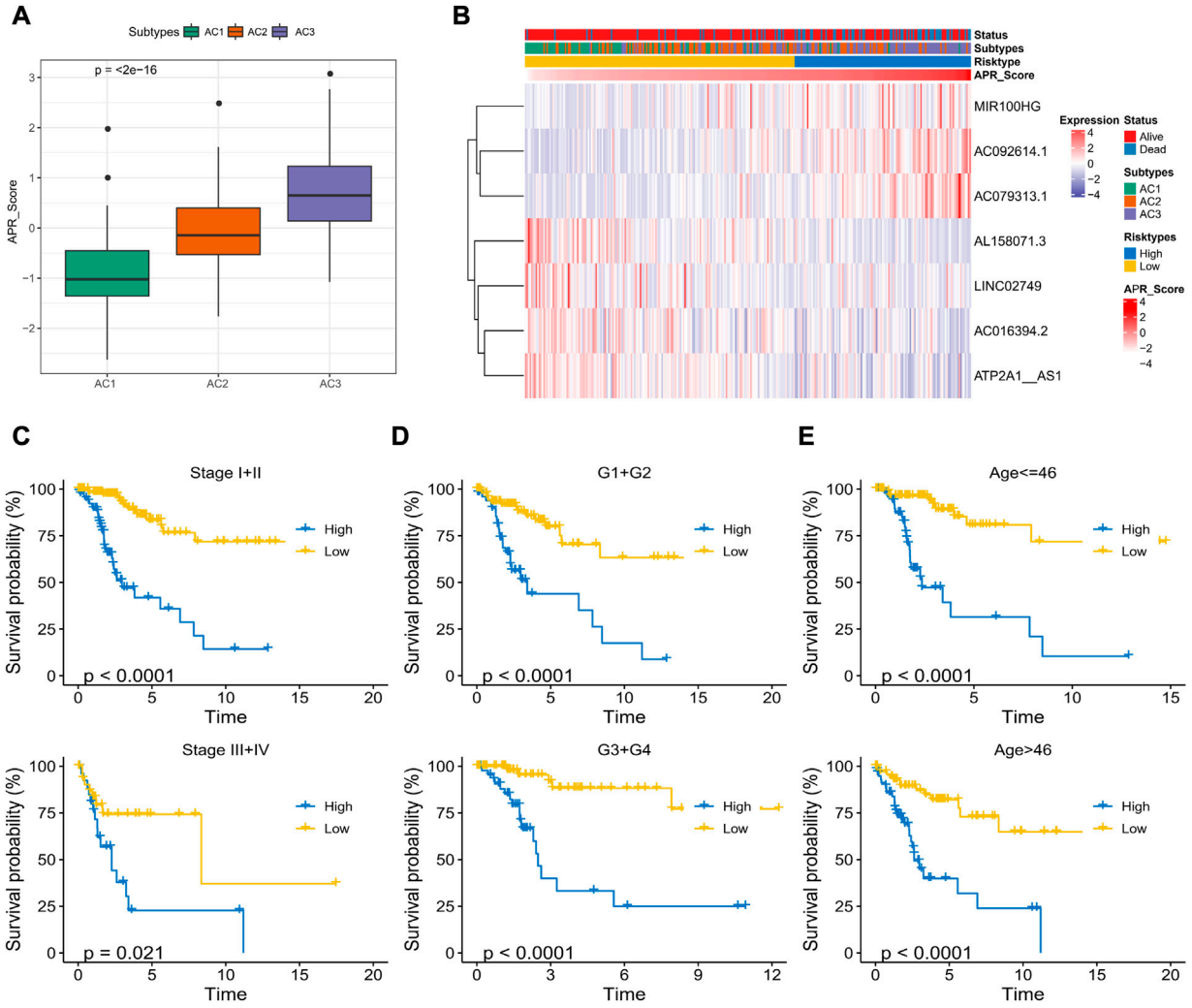


FIGURE 7
Correlation between APR_Score and clinical characteristics of CC **(A)** APR_Score in AC1-3 subtypes **(B)** Heatmap showing the expression of 7 OS-related ARLs in ARLs-related molecular subtypes and APR_Score groupings **(C–E)** K-M curves for patients with high APR_Score and low APR_Score in Stage, Grade, and Age subgroups.

higher AUC values in predicting 1-, 2-, 3-, 4-, and 5-year OS in CC patients (Figures 6E, F). These results suggested that APR_Score had a high accuracy in predicting OS risk in CC patients and might be a novel prognostic indicator.

Correlation between APR_Score and clinical characteristics

To determine the relationship between APR_Score and clinical characteristics, this study discussed the interaction between APR_Score and different clinical features. First, we found that the APR_Score was the highest in the poor prognosis AC3 subtype, with more death cases in the high APR_Score group (Figures 7A, B). Expression differences of AC092614.1, AL158071., AC016394.2, LINC02749, MIR100HG, AC079313.1, ATP2A1- AS1 in different types of patients are shown in Figure 7B. In addition, to compare the meaning of APR_Score in clinicopathological subgroups for CC prognosis, K-M survival analysis was performed on patients with high APR_Score and low APR_Score in clinicopathological subgroups. The results showed that patients with high APR_Score had markedly poorer OS than those with low APR_Score in Stage, Grade and Age subgroups ($p < 0.0001$, Figures 7C–E). Additionally, we compared the differences of clinical characteristics (N satge and M stage) between risk groups in TCGA-CESC cohort. The patients with metastasis (including lymph node metastasis and distal metastasis) in the high-APR_Score group were more than those in the low-APR_Score group (Supplementary Figure S2).

Nomogram for predicting survival rate of CC

To further discuss the clinical value of APR_Score in CC patients, univariate and multivariate COX model analyses were

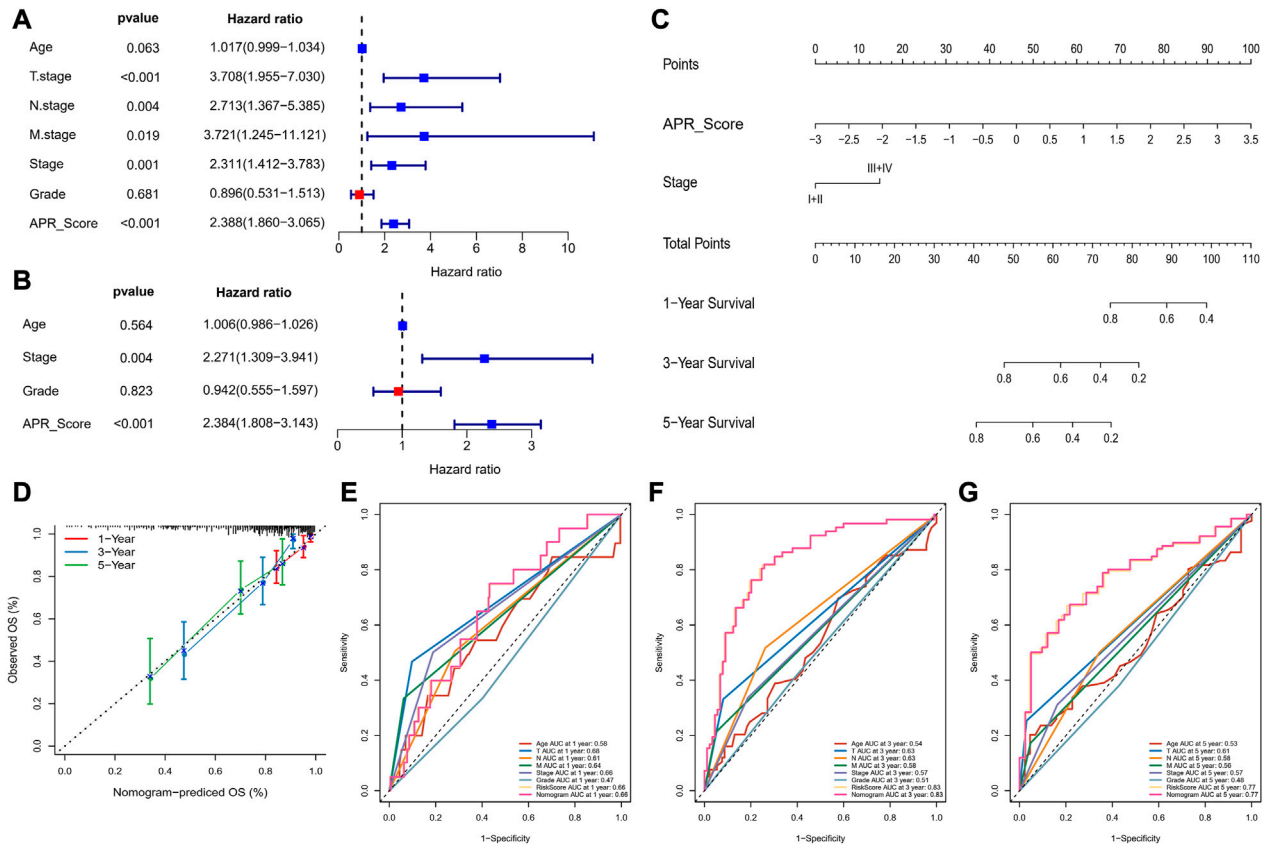


FIGURE 8
Nomogram for predicting survival rate of CC **(A, B)** Forest plots for univariate and multivariate COX models for APR_Score **(C)** Nomogram of APR_Score and Stage **(D)** Calibration curves for 1-year, 3-years and 5-years survival **(E–G)** The ROC curves of a variety of clinical features for overall OS at 1-year, 3-year and 5-year.

performed, and we determined that APR_Score and Stage were clinically significant for CC prognosis ($p < 0.05$, Figures 8A, B). Due to a close correlation between APR_Score and Stage and prognosis, we created a nomogram based on APR_Score and Stage to assess OS of CC patients (Figure 8C). The calibration curves showed a great overlap between Nomogram predicted survival at 1-, 3- and 5-year OS with actual observations, indicating that nomogram was a reliable tool for predicting OS (Figure 8D). We also found that the nomogram and APR_Score showed higher accuracy in predicting 1-year, 3-year and 5-year OS when compared to Age, TNM Stage, Stage and Grade (Figures 8E–G).

TME activity assessment in APR_Score groups

The CIBERSORT and MCP-count algorithms were applied to assess the abundance of immune cell infiltration and immune score in TME of APR_Score groups. CIBERSORT results, as presented in Figure 9A, showed that T cells CD8, Tregs infiltration abundance was lower in high-APR_Score group. MCP-count results also showed lower immune score of CD8 T cells in high-APR_Score group (Figure 9B). Then, to discuss the potential connection between APR_Score and immunotherapy response, we assessed TIDE scores and Exclusion scores in the APR_Score groups. TIDE scores and Exclusion scores were slightly higher in the high APR_Score group than in the low APR_Score group, suggesting that patients with high APR_Score were more likely to experience immune escape and less responsive to immunotherapy (Figures 9C, D).

Association between APR_Score and chemotherapy drug sensitivity

Until immunotherapy was proposed as an alternative treatment for CC, conventional resection and radiotherapy were the dominant treatments (Serkies and Jassem, 2018). In this study, to discuss the potential of APR_Score as therapeutic response marker for predicting chemotherapeutic agents, we evaluated the IC50 values of 20 chemotherapeutic agents in TCGA-CESC. Initially, correlations between APR_Score and drug IC50 values were computed, and highly correlated drugs were selected for comparison in the APR_Score groups (Supplementary Figure S3). We found positive responses to small molecular drugs including Rapamycin, KIN001-135, Roscovitine, Phenformin treatments in the low APR_Score

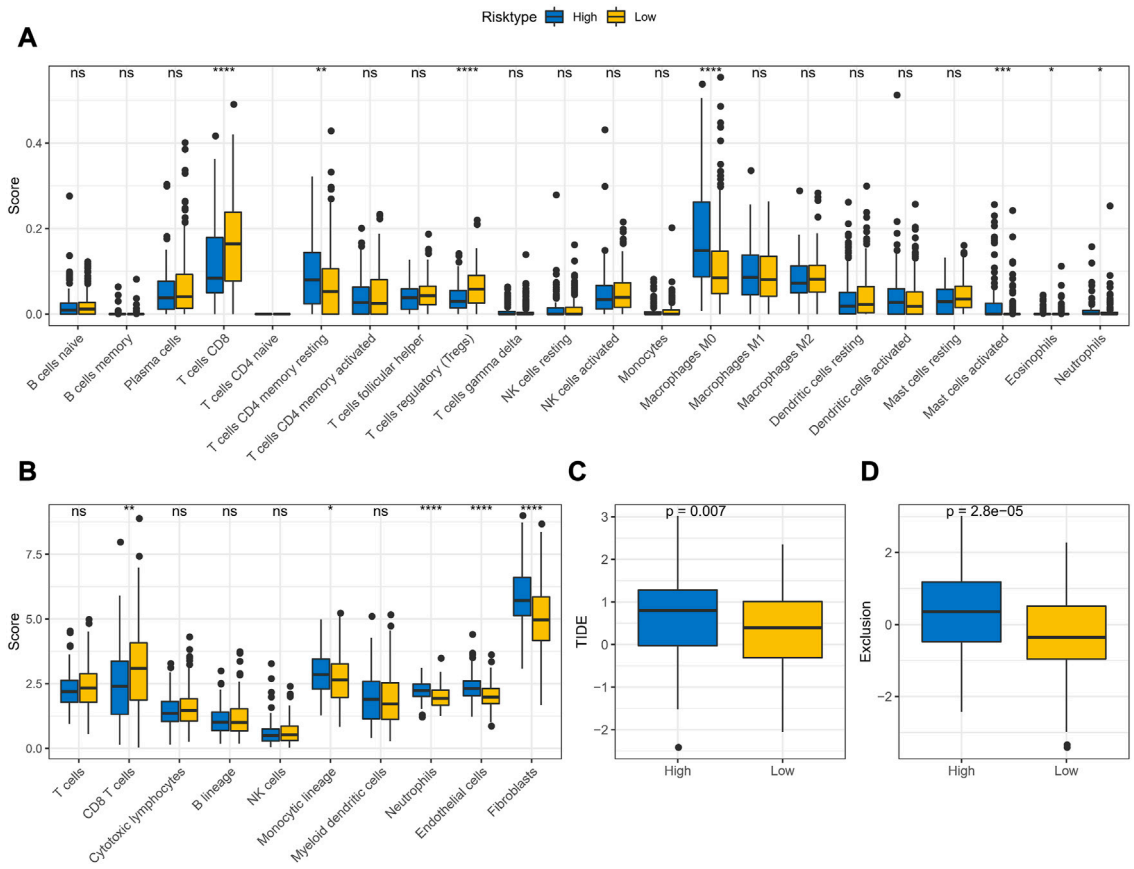


FIGURE 9
TME activity assessment in APR_Score groups **(A)** Results of immune cell scores estimated by CIBERSORT between APR_Score groups in TCGA-CESC cohort **(B)** Results of immune cell scores using MCP-count between APR_Score groups in TCGA-CESC cohort **(C)** Distribution of TIDE score between APR_Score groups in TCGA-CESC cohort **(D)** Exclusion scores in high-APR_Score group and low-APR_Score group. ns $p > 0.05$; *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$; ****$p < 0.0001$.

group, while the high APR_Score group responded positively to Sunitinib, MG-132, Paclitaxel, AZ628, Sorafenib, Saracatinib, Dasatinib, CGP-60474, A-770041, WH-4-023, WZ-1 -84, CMK, Bortezomib, Lapatinib, Midostaurin, Embelin treatments (Figure 10). Overall, the APR_Score was correlated with the sensitivity to small molecular drugs.

Biological pathway characterization of APR_Score groups

Moreover, to discuss the biological pathway variation in APR_Score groups, we evaluated the markedly enriched pathways in distinct APR_Score groups using GSEA methods *via* h.all.v7.5.1.symbols.gmt signatures. We noted that TGF BETA SIGNALING, APICAL JUNCTION, HYPOXIA, EPITHELIAL MESENCHYMAL TRANSITION, APOPTOSIS, UV RESPONSE DN, KRAS SIGNALING UP, TNFA SIGNALING VIA NFKB, PROTEIN SECRETION, ANGIOGENESIS were markedly enriched in high-APR_Score group ($p < 0.05$, FDR<0.25, Figure 11).

Discussion

Anoikis is a protective mechanism *via* which the organism self-corrects abnormal disorders in the presence of disorders and damage, and is essential for the normal growth and development process of living organisms (Zhong and Rescorla, 2012). Advanced cancer patients are affected by tumor metastasis, which occurs in large tumors and could result in death (Steeg, 2016). The prerequisite for tumor cells to migrate would be their own resistance to Anoikis, otherwise the self-correcting capacity of the organism could prevent the formation of metastatic lesions (Simpson et al., 2008; Kim et al., 2012). In fact, the existing studies related to the principles of the anoikis mechanism in CC remained limitations.

In this study, we defined three potential subtypes based on anoikis-related lncRNAs in CC that exhibited remarkable prognostic variation as well as immunoreactivity. Notably, AC3 patients exhibited higher ARG scores and angiogenesis scores. Angiogenesis is essential and important in the development of malignant tumor (Viallard and Larrivee, 2017). In contrast, tumor cell anoikis resistance was

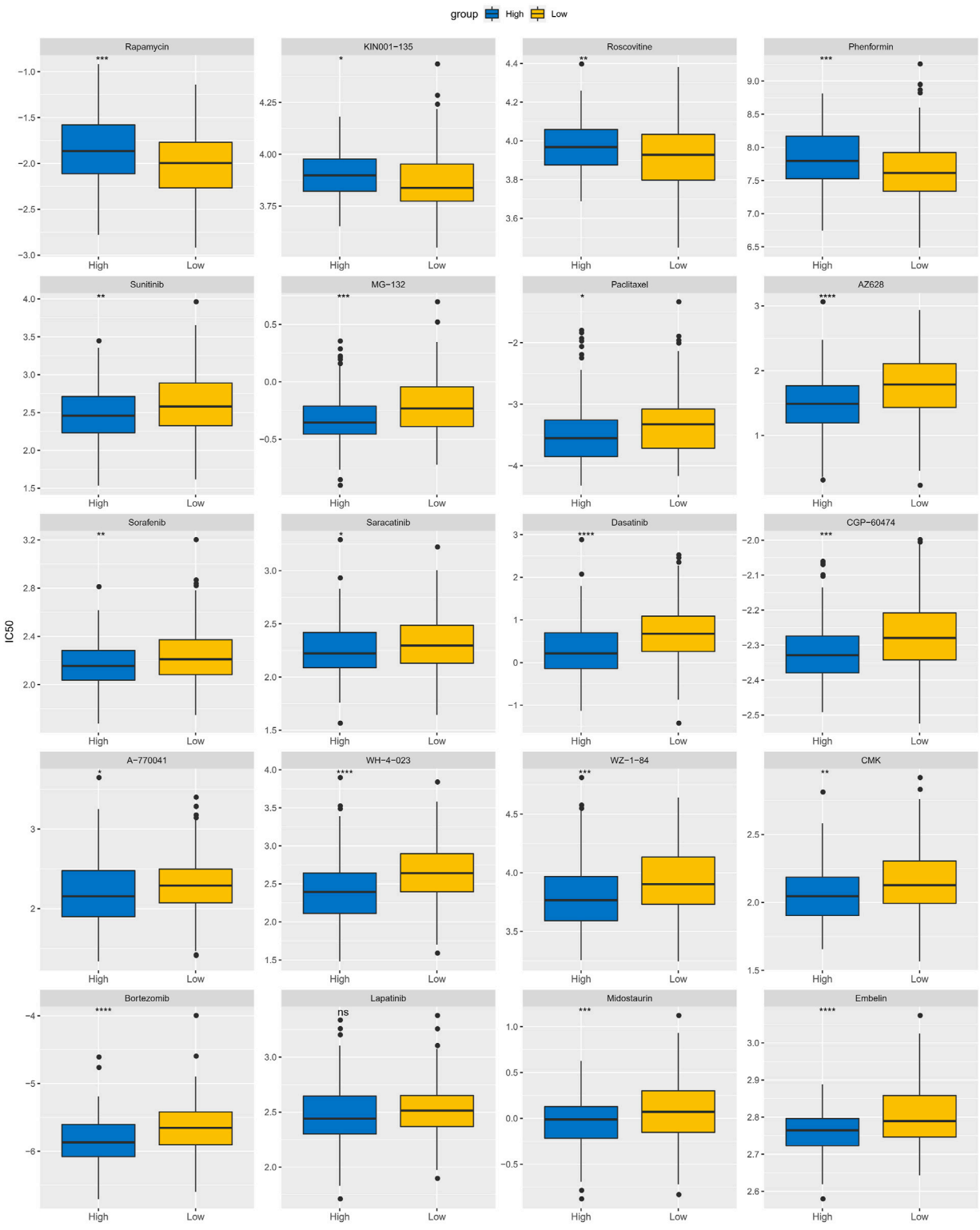


FIGURE 10
Differences of IC50 values for 20 chemotherapy drugs between APR_Score groups in TCGA-CESC cohort. ns $p > 0.05$; *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$; ****$p < 0.0001$.

sufficiently necessary for the formation of new blood vessels and for migration (Kim et al., 2012). The interaction between anoikis and tumor angiogenesis was reported. Gao and colleagues showed that osteosarcoma cells resist through anoikis by activating Src kinase, which in turn activates JNK/ERK/VEGF-A to promote the formation of tumor metastases (Gao

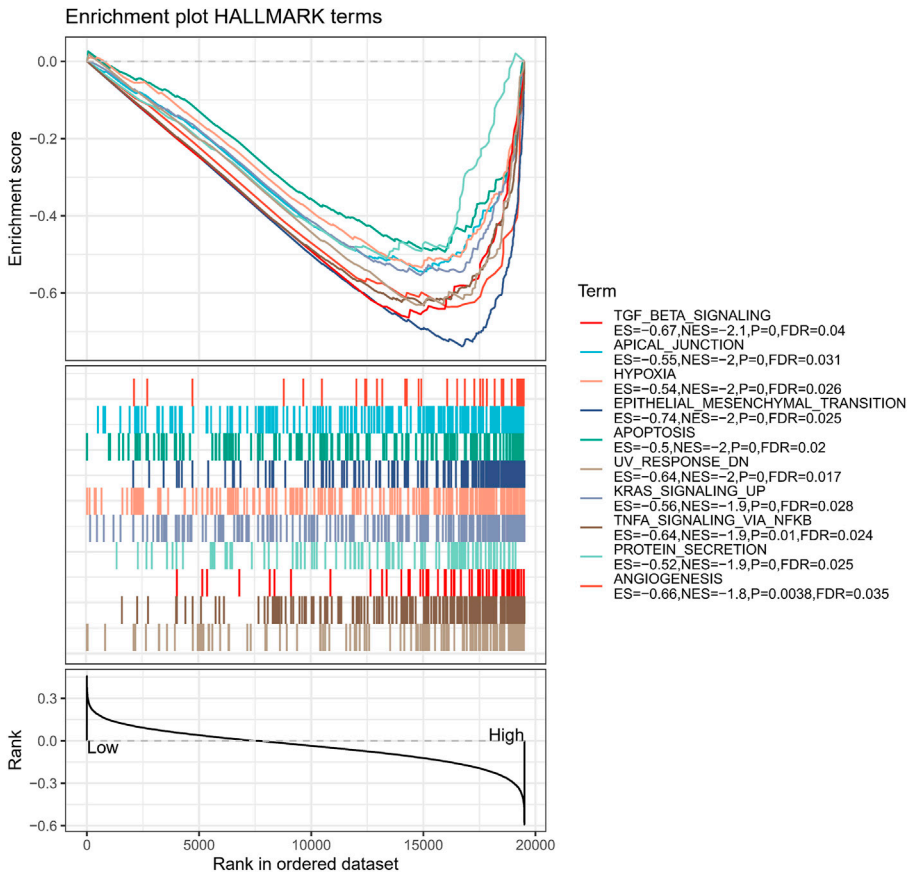


FIGURE 11
Biological pathway characterisation of APR_Score groups *via* GSEA.

et al., 2019). In the present study, we also found marked activation of angiogenesis and NOTCH signaling in the AC3 subtype (Figure 5E). The survival rate of AC3 patients was not satisfactory. This allowed us to speculate that AC3 patients might have a higher probability of postoperative recurrence and higher prognostic risk. However, distal metastases occurred in the majority of cases at the time of diagnosis in CC patients at present (Dereje et al., 2020; Sengayi-Muchengeti et al., 2020; Ryzhov et al., 2021). Accurate determination of disease staging would be crucial for patient treatment modalities and clinical outcomes. Therefore, the ARLs-related subtypes defined in this study might assist M Stage and more accurately calculate the risk of metastasis in CC patients.

Here, we have constructed prognostic risk stratification models for CC based on the expression of ARLs. We explored the potential utility of APR_Score in predicting patient survival and prognosis. In addition, this study also explored the potential function of APR_Score in guiding immunotherapy and chemotherapy. Immunotherapy offered promising opportunities for patients as an emerging option for the rehabilitation of advanced CC patients (Duska et al., 2020; Colombo et al., 2021). Pembrolizumab was approved by the Food and Drug Administration (FDA) as an immunotherapeutic agent for CC (De Felice et al., 2021), however, the current treatment with Pembrolizumab for CC remained limited due to the absence of efficacy biomarkers to determine treatment benefit (Marret et al., 2019). In this study, there were discrepancies in immune cell scores between high- and low-APR_Score patients and in the TIDE scores. We observed higher T cell CD8, T cell CD4 memory resting and TIDE scores in patients with high APR_Score. Tumor cells blocked T cell killing by expression of PD-1/PD-L1, causing immune escape of tumor cells (Pardoll, 2012; Chabanon et al., 2016). Another study concluded that CC cells had a greater tendency of immune escape and were accompanied by increased T cells (Mortezaee, 2020). Moreover, acquisition of anoikis-resistance enhances the abilities of invasiveness, escaping from immune surveillance and therapeutic agents in cancer cells. Fanfone and colleagues have demonstrated that mechanically stressed and anoikis-resistant cancer cells had increased cell motility and escape from killing by natural killer cells (Fanfone et al., 2022), suggesting that anoikis was strongly associated with immune escape of tumor cells. We constructed a prognostic model based on ARLs, and patients with high APR_Score had suboptimal immunotherapy benefit and were more suitable for taking conventional chemotherapy, suggesting that ARLs-based prognostic risk model was a promising predictor in immunotherapy and chemotherapy. Furthermore, the COX model demonstrated that APR_Score was an independent prognostic indicator, and that the nomogram developed in conjunction with APR_Score was a valid prognostic tool. Thus, this finding confirmed the positive utility of ARLs-related signatures in guiding the prognosis and treatment selection, which might provide a

theoretical basis for the development of precise and personalized treatment guidelines for CC.

ARLs-related prognostic risk model consisted of AC092614.1, AL158071.3, AC016394.2, LINC02749, MIR100HG, AC079313.1, ATP2A1-AS1. We found that MIR100HG and ATP2A1-AS1 were correlated with multiple cancers and the remaining ARLs were the first identified tumor prognostic markers. Several studies suggested that aberrant expression of MIR100HG was associated with poor clinical outcome and pathological features, and that it was involved in multiple pathways related to tumorigenesis (Wu et al., 2022). aTP2A1-AS1 is a potential prognostic biomarker for CC (Feng et al., 2021). The mechanism of anoikis in CC was not completely elucidated, and the ARLs identified in this study were important for the elucidation of the molecular mechanisms of CC. However, limitations of this study should be equally noted. In other datasets, there are few cervical cancer-related data sets and they lack clinical information (survival time). Due to the difficulty in annotating lncRNAs in other databases, only the TCGA database was used and randomly divided into training set and validation set to construct and validate the risk model in this study. The accuracy of APR_Score should be validated in the future using clinical samples or sequencing data from multiple centers. Finally, the mechanism of 7-ARLs in CC still required further study, which was our follow-up research plan.

Conclusion

In this study, we defined three molecular subtypes of anoikis-related lncRNAs and generated ARLs-related signatures for assessing the prognosis of patients with CC. The molecular subtypes contributed to a better understanding of the mechanism of CC metastasis and the signatures held potential clinical value in predicting response to therapy.

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

All authors contributed to this present work: HL and LX designed the study and acquired the data. HW drafted the manuscript, YL and JZ revised the manuscript. All authors read and approved the manuscript.

Funding

The subject was funded by the General Program of Chongqing Natural Science Foundation: Study on the Role of CircRNA-0008193 Targeting miR-1226-5p in Regulating PD-L1 in Immune Escape of Cervical Cancer(cstc2021jcyj-msxmX0219). We are grateful for the support of the Tianjin Municipal Education Commission Scientific Research Program Project (2018KJ036).

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/fphar.2023.1135626/full#supplementary-material

References

Adeshakin, F. O., Adeshakin, A. O., Afolabi, L. O., Yan, D., Zhang, G., and Wan, X. (2021). Mechanisms for modulating anoikis resistance in cancer and the relevance of metabolic reprogramming. *Front. Oncol.* 11, 626577. doi:10.3389/fonc.2021.626577

Barbie, D. A., Tamayo, P., Boehm, J. S., Kim, S. Y., Moody, S. E., Dunn, I. F., et al. (2009). Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 462 (7269), 108–112. doi:10.1038/nature08460

Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 17 (1), 218. doi:10.1186/s13059-016-1070-5

Bhatla, N., Aoki, D., Sharma, D. N., and Sankaranarayanan, R. (2018). Cancer of the cervix uteri. *Int. J. Gynaecol. Obstet.* 143 (2), 22–36. doi:10.1002/ijgo.12611

Canfell, K., Kim, J. J., Brisson, M., Keane, A., Simms, K. T., Caruana, M., et al. (2020). Mortality impact of achieving WHO cervical cancer elimination targets: A comparative modelling analysis in 78 low-income and lower-middle-income countries. *Lancet* 395 (10224), 591–603. doi:10.1016/S0140-6736(20)30157-4

Chabanon, R. M., Pedrero, M., Lefebvre, C., Marabelle, A., Soria, J. C., and Postel-Vinay, S. (2016). Mutational landscape and sensitivity to immune checkpoint blockers. *Clin. Cancer Res.* 22 (17), 4309–4321. doi:10.1158/1078-0432.CCR-16-0903

Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol. Biol.* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Colombo, N., Dubot, C., Lorusso, D., Caceres, M. V., Hasegawa, K., Shapira-Frommer, R., et al. (2021). Pembrolizumab for persistent, recurrent, or metastatic cervical cancer. *N. Engl. J. Med.* 385 (20), 1856–1867. doi:10.1056/NEJMoa2112435

Danilova, L., Ho, W. J., Zhu, Q., Vithayathil, T., De Jesus-Acosta, A., Azad, N. S., et al. (2019). Programmed cell death ligand-1 (PD-L1) and CD8 expression profiling identify an immunologic subtype of pancreatic ductal adenocarcinomas with favorable survival. *Cancer Immunol. Res.* 7 (6), 886–895. doi:10.1158/2326-6066.CIR-18-0822

De Felice, F., Giudice, E., Bolomini, G., Distefano, M. G., Scambia, G., Fagotti, A., et al. (2021). Pembrolizumab for advanced cervical cancer: Safety and efficacy. *Expert Rev. Anticancer Ther.* 21 (2), 221–228. doi:10.1080/14737140.2021.1850279

Dereje, N., Gebremariam, A., Addissie, A., Worku, A., Assefa, M., Abraha, A., et al. (2020). Factors associated with advanced stage at diagnosis of cervical cancer in addis ababa, Ethiopia: A population-based study. *BMJ Open* 10 (10), e040645. doi:10.1136/bmjopen-2020-040645

Detmar, M. (2000). Tumor angiogenesis. *J. Investig. Dermatol Symp. Proc.* 5 (1), 20–23. doi:10.1046/j.1087-0024.2000.00003.x

Duska, L. R., Scalici, J. M., Temkin, S. M., Schwarz, J. K., Crane, E. K., Moxley, K. M., et al. (2020). Results of an early safety analysis of a study of the combination of pembrolizumab and pelvic chemoradiation in locally advanced cervical cancer. *Cancer* 126 (22), 4948–4956. doi:10.1002/cncr.33136

Fanfone, D., Wu, Z., Mammi, J., Berthenet, K., Neves, D., Weber, K., et al. (2022). Confined migration promotes cancer metastasis through resistance to anoikis and increased invasiveness. *Elife* 11, e73150. doi:10.7554/eLife.73150

Feng, Q., Wang, J., Cui, N., Liu, X., and Wang, H. (2021). Autophagy-related long non-coding RNA signature for potential prognostic biomarkers of patients with cervical cancer: A study based on public databases. *Ann. Transl. Med.* 9 (22), 1668. doi:10.21037/atm-21-5156

Gao, Z., Zhao, G. S., Lv, Y., Peng, D., Tang, X., Song, H., et al. (2019). Anoikis-resistant human osteosarcoma cells display significant angiogenesis by activating the Src kinase-mediated MAPK pathway. *Oncol. Rep.* 41 (1), 235–245. doi:10.3892/or.2018.6827

Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One* 9 (9), e107468. doi:10.1371/journal.pone.0107468

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

Harrell Jr, F. E., and Harrell Jr, M. F. E. (2019). Package 'hmisc'[J]. *CRAN2018*, 235–236.

Harrell Jr, F. E., Harrell Jr, M. F. E., and Hmisc, D. (2017). Package 'rms'[J]. *Vanderbilt University* 17 (4), 229.

Kim, Y. N., Koo, K. H., Sung, J. Y., Yun, U. J., and Kim, H. (2012). Anoikis resistance: An essential prerequisite for tumor metastasis. *Int. J. Cell Biol.* 2012, 306879. doi:10.1155/2012/306879

Marret, G., Borcoman, E., and Le Tourneau, C. (2019). Pembrolizumab for the treatment of cervical cancer. *Expert Opin. Biol. Ther.* 19 (9), 871–877. doi:10.1080/14712598.2019.1646721

Marth, C., Landoni, F., Mahner, S., McCormack, M., Gonzalez-Martin, A., Colombo, N., et al. (2017). Cervical cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 28 (4), iv72–iv83. doi:10.1093/annonc/mdx220

Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118

Meng, Y., Chu, T., Lin, S., Wu, P., Zhi, W., Peng, T., et al. (2021). Clinicopathological characteristics and prognosis of cervical cancer with different histological types: A population-based cohort study. *Gynecol. Oncol.* 163 (3), 545–551. doi:10.1016/j.ygyno.2021.10.007

Mortezaee, K. (2020). Immune escape: A critical hallmark in solid tumors. *Life Sci.* 258, 118110. doi:10.1016/j.lfs.2020.118110

Paoli, P., Giannoni, E., and Chiarugi, P. (2013). Anoikis molecular pathways and its role in cancer progression. *Biochim. Biophys. Acta* 1833 (12), 3481–3498. doi:10.1016/j.bbamcr.2013.06.026

Pardoll, D. M. (2012). The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* 12 (4), 252–264. doi:10.1038/nrc3239

Parmar, D., and Apte, M. (2021). Angiopoietin inhibitors: A review on targeting tumor angiogenesis. *Eur. J. Pharmacol.* 899, 174021. doi:10.1016/j.ejphar.2021.174021

Ryzhov, A., Corbex, M., Pineros, M., Barchuk, A., Andreasyan, D., Djanklich, S., et al. (2021). Comparison of breast cancer and cervical cancer stage distributions in ten newly independent states of the former soviet union: A population-based study. *Lancet Oncol.* 22 (3), 361–369. doi:10.1016/S1470-2045(20)30674-4

Sanchez-Vega, F., Mina, M., Armenia, J., Chatila, W. K., Luna, A., La, K. C., et al. (2018). Oncogenic signaling pathways in the cancer Genome Atlas. *Cell* 173 (2), 321–337. e10. doi:10.1016/j.cell.2018.03.035

Sengayi-Muchengeti, M., Joko-Fru, W. Y., Miranda-Filho, A., Egue, M., Akele-Akpo, M. T., N'da, G., et al. (2020). Cervical cancer survival in sub-saharan africa by age, stage at diagnosis and human development index: A population-based registry study. *Int. J. Cancer* 147 (11), 3037–3048. doi:10.1002/ijc.33120

Serkies, K., and Jassem, J. (2018). Systemic therapy for cervical carcinoma - current status. *Chin. J. Cancer Res.* 30 (2), 209–221. doi:10.21147/j.issn.1000-9604.2018.02.04

Shen, W., Song, Z., Zhong, X., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 1 (3), e36. doi:10.1002/imt2.36

Simon, N., Friedman, J., Hastie, T., and Tibshirani, R. (2011). Regularization paths for cox's proportional hazards model via coordinate descent. *J. Stat. Softw.* 39 (5), 1–13. doi:10.18637/jss.v039.i05

Simpson, C. D., Anyiwe, K., and Schimmer, A. D. (2008). Anoikis resistance and tumor metastasis. *Cancer Lett.* 272 (2), 177–185. doi:10.1016/j.canlet.2008.05.029

Steeg, P. S. (2016). Targeting metastasis. *Nat. Rev. Cancer* 16 (4), 201–218. doi:10.1038/nrc.2016.25

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

Viallard, C., and Larrivee, B. (2017). Tumor angiogenesis and vascular normalization: Alternative therapeutic targets. *Angiogenesis* 20 (4), 409–426. doi:10.1007/s10456-017-9562-9

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

Wu, Y., Wang, Z., Yu, S., Liu, D., and Sun, L. (2022). LncmiRHG-MIR100HG: A new budding star in cancer. *Front. Oncol.* 12, 997532. doi:10.3389/fonc.2022.997532

Zhong, X., and Rescorla, F. J. (2012). Cell surface adhesion molecules and adhesion-initiated signaling: Understanding of anoikis resistance mechanisms and therapeutic opportunities. *Cell Signal* 24 (2), 393–401. doi:10.1016/j.cellsig.2011.10.005

<Hmisc.pdf>.

<rms.pdf>.

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 30 March 2023
DOI 10.3389/fphar.2023.1130399

OPEN ACCESS

EDITED BY
Zhi-Qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Dapeng Lei,
Qilu Hospital, Shandong University, China
Lei Tao,
Fudan University, China
Yuxia Fan,
First Affiliated Hospital of Zhengzhou University, China

*CORRESPONDENCE
Xudong Zhao,
zhaoxdent@hotmail.com

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal *Frontiers in Pharmacology*

RECEIVED 23 December 2022
ACCEPTED 06 March 2023
PUBLISHED 30 March 2023

CITATION
Zhang W, Liu T, Li X, Li T, Ma X, Zhao D, Liu Y, Zheng X and Zhao X (2023), Identification of novel immune-related molecular subtypes and a prognosis model to predict thyroid cancer prognosis and drug resistance. *Front. Pharmacol.* 14:1130399. doi: 10.3389/fphar.2023.1130399

Identification of novel immune-related molecular subtypes and a prognosis model to predict thyroid cancer prognosis and drug resistance

Wei Zhang[1], Ting Liu[2], Xinyi Li[2], Tianshu Li[2], Xiangchi Ma[2], Dongxu Zhao[2], Yueyang Liu[2], Xueke Zheng[2] and Xudong Zhao[2]*

[1]Department of Endocrinology, Shengjing Hospital of China Medical University, Shenyang, China, [2]Department of Otolaryngology-Head and Neck Surgery, Shengjing Hospital of China Medical University, Shenyang, China

Background: Thyroid cancer is a common malignant tumor of the endocrine system that has shown increased incidence in recent decades. We explored the relationship between tumor-infiltrating immune cell classification and the prognosis of thyroid carcinoma.

Methods: RNA-seq, SNV, copy number variance (CNV), and methylation data for thyroid cancer were downloaded from the TCGA dataset. ssGSEA was used to calculate pathway scores. Clustering was conducted using ConsensusClusterPlus. Immune infiltration was assessed using ESTIMATE and CIBERSORT. CNV and methylation were determined using GISTIC2 and the KNN algorithm. Immunotherapy was predicted based on TIDE analysis.

Results: Three molecular subtypes (Immune-enrich(E), Stromal-enrich(E), and Immune-deprived(D)) were identified based on 15 pathways and the corresponding genes. Samples in Immune-E showed higher immune infiltration, while those in Immune-D showed increased tumor mutation burden (TMB) and mutations in tumor driver genes. Finally, Immune-E showed higher CDH1 methylation, higher progression-free survival (PFS), higher suitability for immunotherapy, and higher sensitivity to small-molecule chemotherapeutic drugs. Additionally, an immune score (IMScore) based on four genes was constructed, in which the low group showed better survival outcome, which was validated in 30 cancers. Compared to the TIDE score, the IMScore showed better predictive ability.

Conclusion: This study constructed a prognostic evaluation model and molecular subtype system of immune-related genes to predict the thyroid cancer prognosis of patients. Moreover, the interaction network between immune genes may play a role by affecting the biological function of immune cells in the tumor microenvironment.

KEYWORDS

thyroid cancer, immune, prognosis, methylation, molecular subtype, chemotherapy drug

Introduction

Among endocrine tumors, thyroid cancer is a malignant tumor with the highest incidence and manifests low mortality and a relatively favorable prognosis (Wang et al., 2020). However, for locally advanced or recurrent and metastatic thyroid cancer, the existing treatment methods cannot effectively improve patient prognosis. Therefore, novel therapeutic approaches such as immunotherapy targeting the molecular mechanisms of thyroid cancer initiation and progression are under exploration (Farkona et al., 2016). Immune-related genes can be used to predict the prognosis of patients with thyroid cancer and can also serve as therapeutic targets (Gunda et al., 2018). The tumor microenvironment (TME) includes the various cell types (immune cells, fibroblasts, endothelial cells, etc.) and extracellular components (growth factors, cytokines, extracellular matrix, hormones, etc.) surrounding cancer cells (Wu and Dai, 2017). Recent studies have shown that different types of immune cells affect the tumor progression of various cancer types, reflecting TME heterogeneity (Zhang et al., 2020; Chen et al., 2021). Therefore, it is important to understand the role of immune cells and immune genes in the thyroid cancer microenvironment.

Immune checkpoint inhibitors have achieved great efficacy in the treatment of a variety of tumors (Branchoux et al., 2019). In papillary thyroid carcinoma, BRAF V600E mutation is positively correlated with the expression of programmed death ligand 1/programmed death receptor 1 in tumor tissues and immune checkpoint inhibitors can effectively kill thyroid tumor cells (Bai et al., 2018). Gnjatic et al. (2017) found that the number and distribution of tumor-infiltrating immune cells (TIICs) could affect the treatment response in patients with cancer and that TIICs are a potential drug target to further improve patient survival. Inflammation and immune cell infiltration are closely involved in thyroid cancer initiation and development; therefore, the exploration of immune infiltration patterns is needed to evaluate patient treatment response and prognosis (Mould et al., 2017). Immune genes as prognostic molecular markers and potential targets for thyroid cancer immunotherapy have attracted attention (Ma et al., 2020; Zhi et al., 2020; Qin et al., 2021). The current AJCC TNM staging and risk stratification of recurrence for patients with differentiated thyroid cancer are used to guide individualized treatment and are formulated based on the clinicopathological data on patients with thyroid cancer without molecular detection. As the AJCC TNM system is still not sufficiently accurate to classify patients with cancer with different prognoses, patients must be classified at the RNA level.

This study applied bioinformatic methods to identify immune molecular subtypes and construct prognostic models and risk-scoring systems. We evaluated the prognosis of thyroid cancer at the gene and molecular levels and further analyzed the immune gene regulatory network of thyroid cancer to provide new ideas for the study of the immune-related mechanisms of thyroid cancer and the development of immune-targeted drugs.

TABLE 1 Pathological types of thyroid cancer.

Var1	Freq
Other, specify	9
Thyroid papillary carcinoma—classical/usual	355
Thyroid papillary carcinoma—follicular (≥99% follicular patterned)	101
Thyroid papillary carcinoma—tall cell (≥50% tall cell features)	36

Materials and methods

Raw dataset

RNA-seq, clinical data, transcriptome data, SNV, CNV, and methylation data on patients with thyroid cancer were downloaded from The Cancer Genome Atlas on 23 April 2022. For RNA-seq data, samples without clinical follow-up

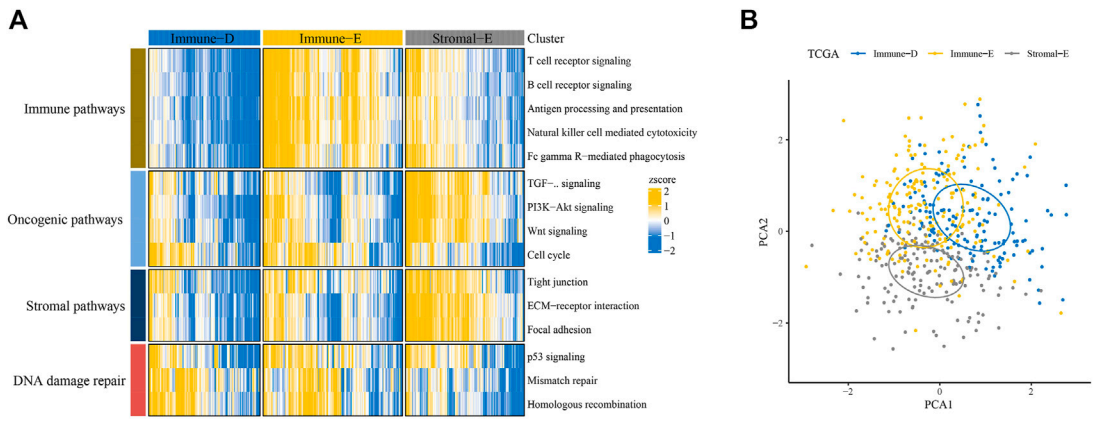


FIGURE 1
Identification of three molecular subtypes in thyroid cancer. **(A)** Identification of Immune-Enrich, Stromal-Enrich, and Immune-Deprived subtypes in the TCGA dataset. **(B)** Principal component analyses of the three molecular subtypes.

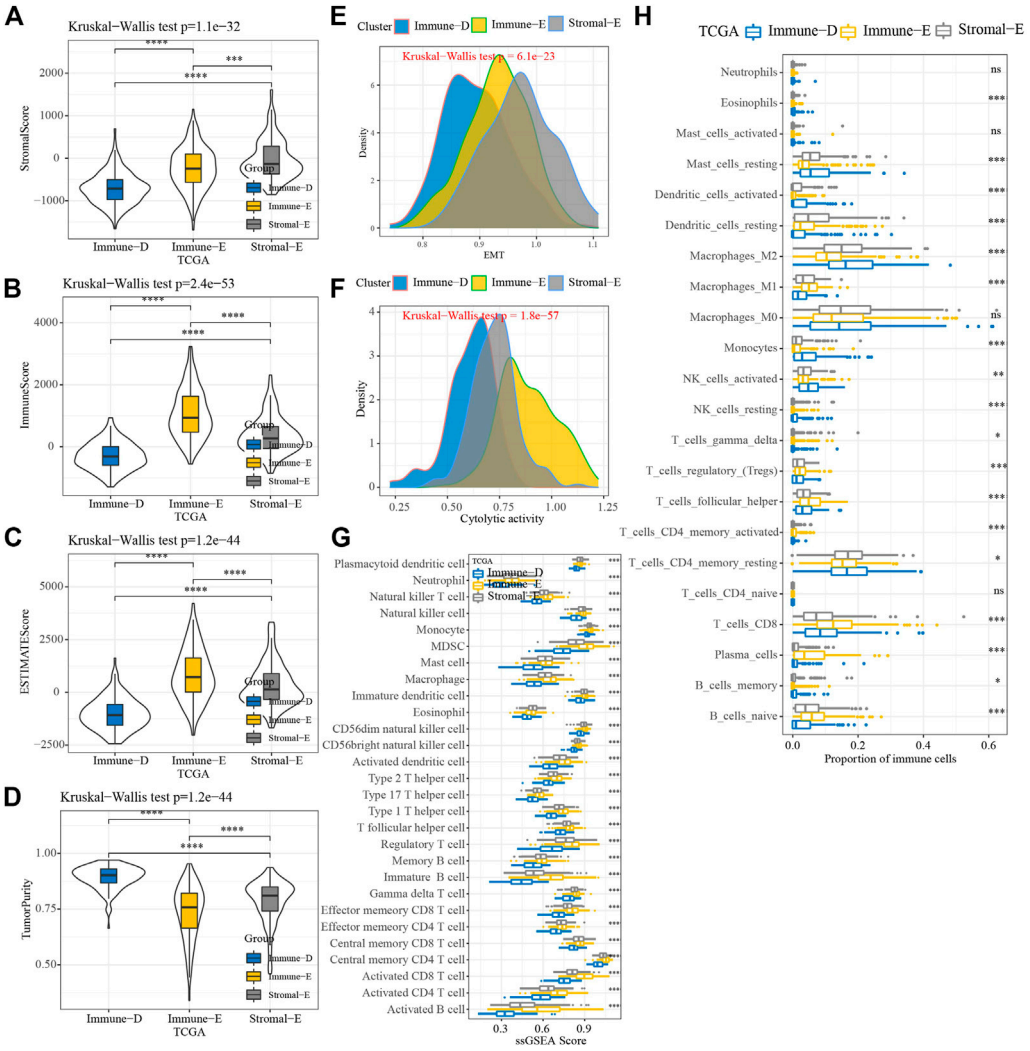


FIGURE 2
Immune infiltration analysis among three molecular subtypes. **(A)** Distributions of StromalScore among the three molecular subtypes. **(B)** Distributions of ImmuneScore among the three molecular subtypes. **(C)** Distributions of ESTIMATEscores among the three molecular subtypes. **(D)** Distributions of TumorPurity among the three molecular subtypes. **(E)** Differences in EMTscores among the three molecular subtypes. **(F)** Differences in cytolytic activity among the three molecular subtypes. **(G)** Differences in the scores for 28 kinds of immune cells among the three molecular subtypes. **(H)** Differences in the scores for 22 kinds of immune cells among the three molecular subtypes. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns: no significance.

information, survival time, or status were removed; Ensembl was converted into gene symbol, and the median expression of multiple GeneSymbols was used. The pathological types of thyroid cancer are shown in Table 1.

Data on a total of 15 pathways (immune, stromal, DNA damage repair, and oncogenic) and their corresponding genes were obtained from a previous study (Li and Wang, 2021).

ssGSEA

ssGSEA analysis was used to calculate the scores of the 15 pathways, EMT pathways (HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION), and cytolytic activity (Rooney et al., 2015) using the R package "GSVA".

Clustering analysis

Molecular subtyping was performed separately for the TCGA dataset samples *via* ConsensusClusterPlus 1.52.0 using the scores for the 15 pathways (Wilkerson and Hayes, 2010). A total of 500 bootstraps were completed with "pam" arithmetic and "pearson" distances. Each bootstrap involved TCGA dataset specimens (≥80%). The cluster number k was set from 2 to 10, and the optimum k was defined as per cumulative distribution function (CDF) and AUC. Differences in survival (KM) curves were analyzed according to the molecular subtypes. Similarly, the distribution differences in clinical characteristics were compared, and chi-square tests were conducted. $p < 0.05$ was defined as statistically significant. Principal component analysis (PCA) was also performed to test the rationality of the molecular subtype distributions.

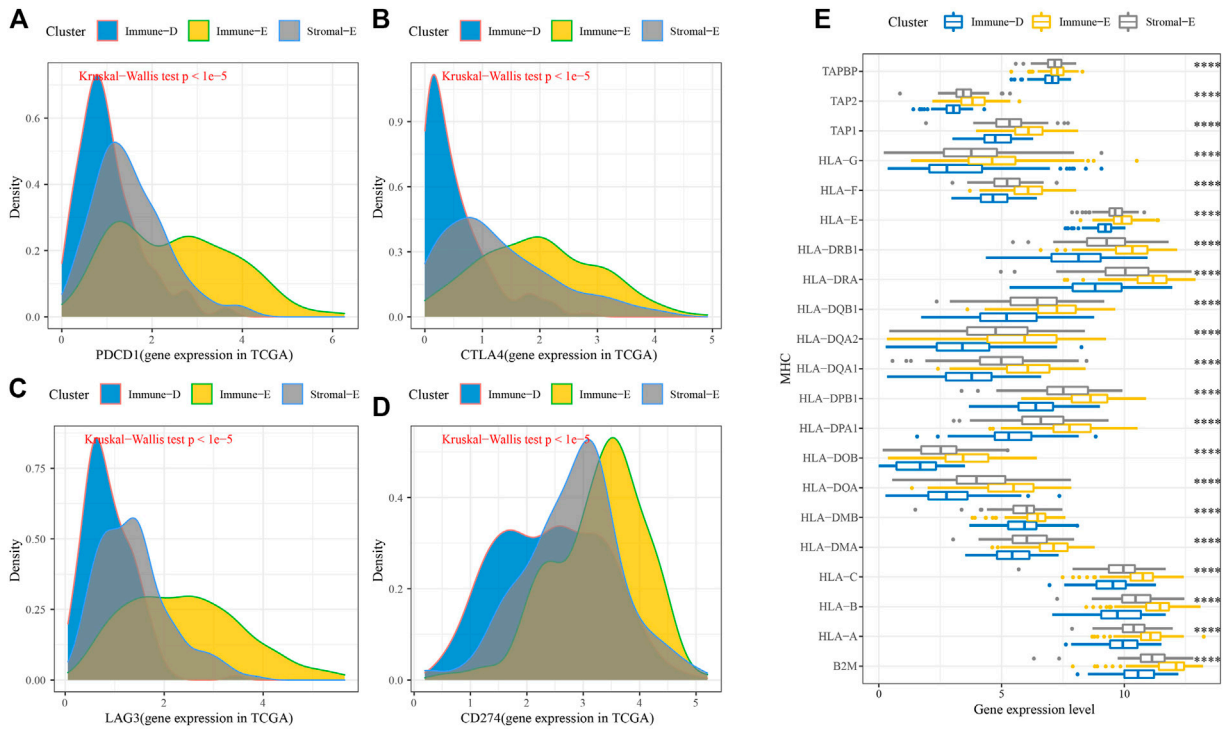


FIGURE 3
Immune checkpoint genes among the three molecular subtypes. **(A)** Differences in *PDCD1* expression among the three molecular subtypes. **(B)** Differences in *CTLA4* expression among the three molecular subtypes. **(C)** Differences in *LAG3* expression among the three molecular subtypes. **(D)** Differences in *CD274* expression among the three molecular subtypes. **(E)** Differences in MHC-related gene expression among the three molecular subtypes. ****$p < 0.0001$.

Immune cell infiltration

The proportions of 22 tumor-infiltrating immune cells (TIICs) were calculated using the CIBERSORT algorithm in all malignant tumor samples. ImmuneScore, StromalScore, ESTIMATEScore, and TumorPurity were determined using the ESTIMATE algorithm. ssGSEA identified scores for 28 kinds of immune cells.

Genetic mutations and epigenetics

For 172 tumor driver genes (159 of which had copy data) (Gao et al., 2013), we used GISTIC2 to analyze the changes in copy number. Those with a ratio >0.2 were considered Gains, while those with a ratio <0.2 were considered Losses; and the rest were considered to be Diploid. SNV was determined using maftools. Methylation of 450K in seven EMT genes and two mismatch repair genes was determined using the KNN algorithm in the impute package.

Tumor immune dysfunction and exclusion (TIDE)

The TIDE (Jiang et al., 2018; Fu et al., 2020) algorithm (http://tide.dfci.harvard.edu) was used to evaluate the exclusion of CTL and dysfunction of tumor infiltration cytotoxic T lymphocytes (CTL) by immunosuppressive factors.

Drug sensitivity analysis

The sensitivity to traditional medicines (IC50 values) was predicted using pRRophetic (Geeleher et al., 2014).

Construction of the IMscore

In the TCGA dataset, thyroid cancer samples were randomly grouped into the training and test datasets in a 1:1 ratio. In the TCGA dataset, we identified pathway genes and pathways with Pearson correlations below the threshold $|R| > 0.4$, $p < 0.05$ to obtain related genes. In the training dataset, univariate Cox analysis was performed to screen genes related to prognosis. LASSO Cox regression in the glmnet package in R language and stepAIC in the MASS package were performed to select the best prognostic genes. A penalty coefficient λ of the optimal value and genes for the model development were determined through 10-fold cross-validation for a total of 1000 iterations. The risk scores for each were calculated using the following formula:

$$IMscore = \sum \beta i \times Expi,$$

where βi refers to the Cox hazard ratio coefficient of mRNA and Expri is the expression level of a gene. Samples in the training dataset were assigned into two groups of high-risk and low-risk based on the optimal segmentation point cutoff, which was determined using the

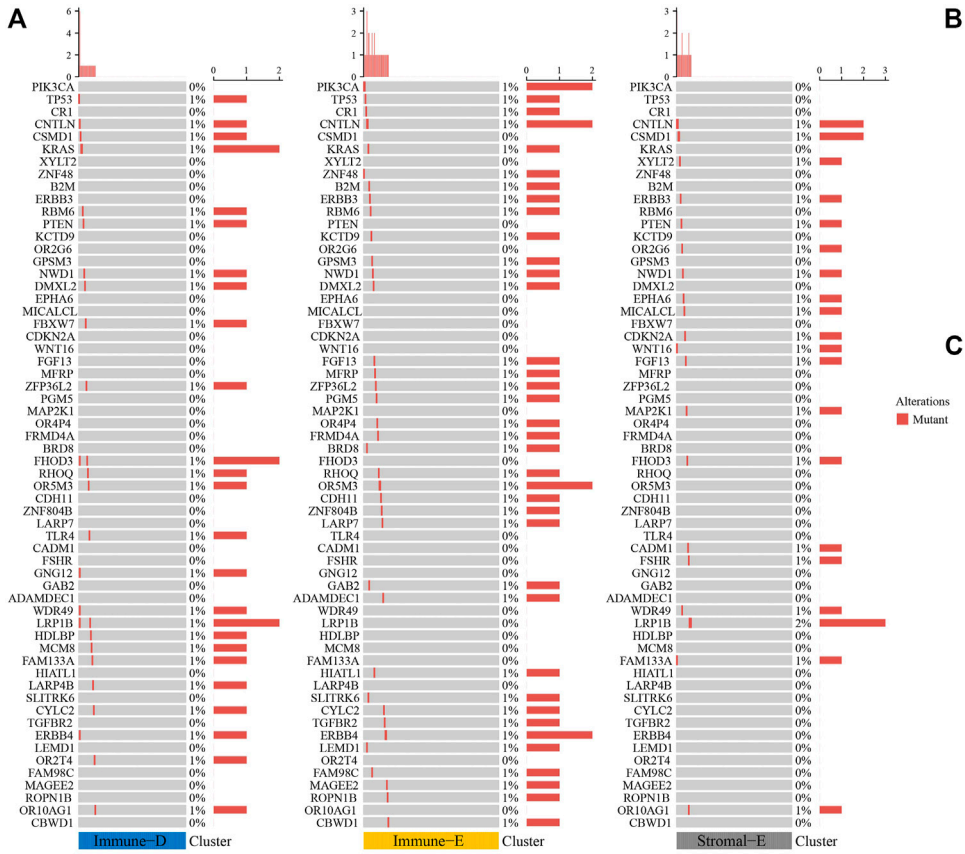


FIGURE 4
Mutation analysis of tumor-driving genes. **(A)** Mutation analysis of tumor-driving genes among the three molecular subtypes. **(B)** Differences in TBM among the three molecular subtypes. **(C)** KM survival curves of CSMD1 and ERBB3 mutants and wildtype. $^{**}p < 0.01$; ns: no significance.

survminer package. Simultaneously, the effectiveness and robustness of the prognostic risk model were validated in test and entire TCGA datasets. Survival differences among the risk groups were evaluated using Kaplan–Meier (KM) curves combined with log-rank tests. The performance of IMscore in pan-cancer, immunotherapy datasets (IMvigor210 and GSE91061) was also evaluated.

Sangerbox assisted with this article (Shen et al., 2022).

Statistical analysis

The software packages used in this study were implemented in R software (version 4.2.2; https://www.r-project.org/). A p-value < 0.05 was considered statistically significant.

Results

Identification of three molecular subtypes in thyroid cancer

Based on scores in 15 pathways, three molecular subtypes (Immune-Enrich (E), Stromal-Enrich (E), and Immune-Deprived (D)) in thyroid cancer were identified by ConsensusClusterPlus for k = 3 (Figure 1A). The PCA results showed that the three molecular subtypes had distinct boundaries, indicating the rationality of the subtype classification (Figure 1B). Samples in Immune-D showed better OS, while those in Immune-D showed better progression-free survival (PFS) (Supplementary Figure S1A). The distribution of clinical features of the three molecular subtypes indicated the significance of the T and N stages (Supplementary Figure S1B).

Immune cell infiltration analysis among molecular subtypes

The results of the ESTIMATE analysis showed higher and lower ImmuneScore, StromalScore, and ESTIMATEScore in Immune-E and Immune-D, respectively (Figures 2A–C). TumorPurity was lower in Immune-E (Figure 2D). A higher EMT score was observed in Stromal-E (Figure 2E). The cytolytic activity score was increased in Immune-E (Figure 2F). In total, 28 kinds of immune cells showed higher scores in Immune-E (Figure 2G), while 18 of 22 immune cells also had higher scores in Immune-E (Figure 2H).

Furthermore, the expression levels of PDCD1, CTLA4, LAG3, and CD274(PD-L1) were upregulated in Immune-E (Figures 3A–D). The expression analysis of MHC genes showed increases

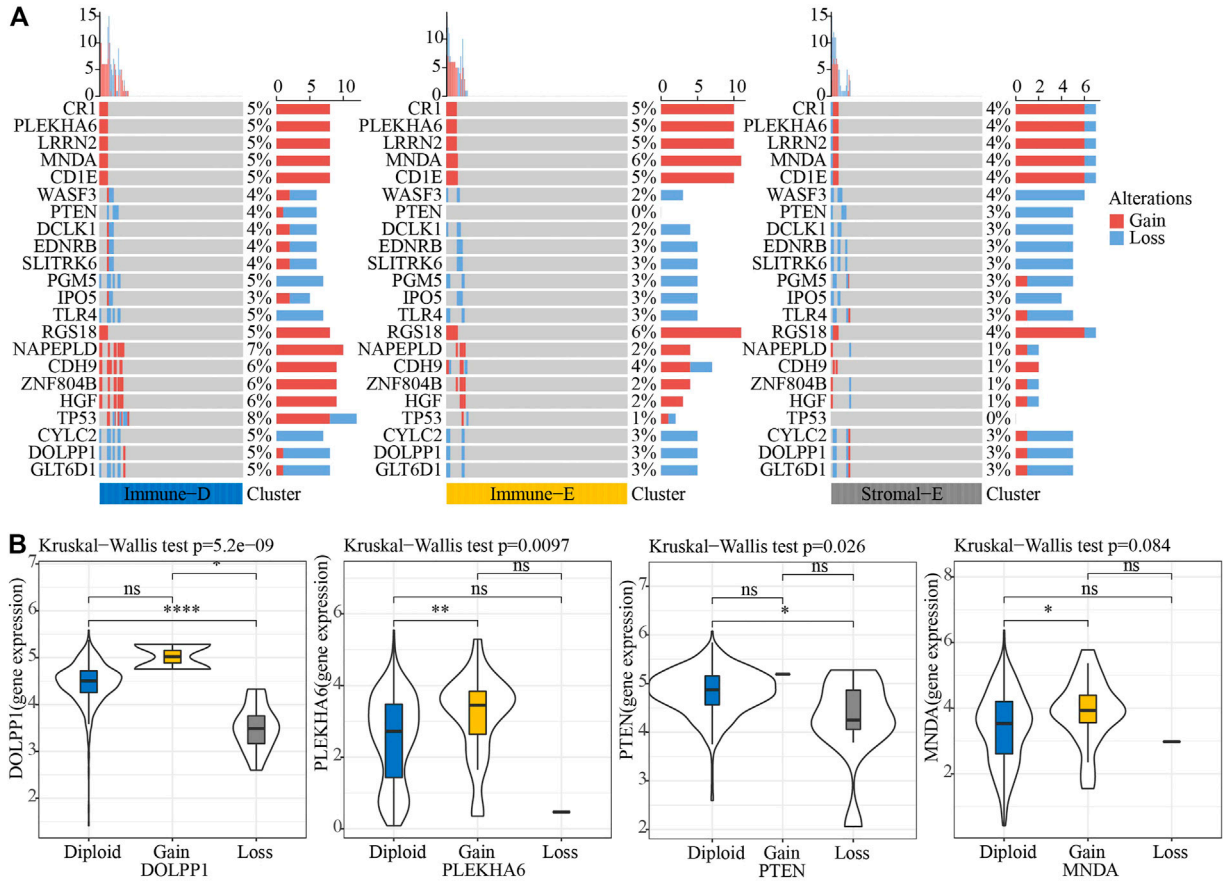


FIGURE 5
CNV analysis of tumor-driving genes. **(A)** CNV analysis of tumor-driving genes among the three molecular subtypes. **(B)** Expression differences of four genes in three CNV groups. *$p < 0.05$; **$p < 0.01$; ****$p < 0.0001$; ns: no significance.

in 21 genes in Immune-E (Figure 3E). These results indicated higher immune infiltration in Immune-E.

Genome mutation analysis of the molecular subtypes

Next, we analyzed the gene mutations among the molecular subtypes. The results demonstrated that 60 genes among 172 tumor-driving genes showed varying degrees of mutation in the three molecular subtypes (Figure 4A). The TMB was higher in Immune-D compared to Stromal-E (Figure 4B). Tumor driver gene mutations and wild-type samples used for KM analysis showed better survival outcomes in samples with *CSMD1* and *ERBB3* wildtype compared to samples with *CSMD1* and *ERBB3* mutations (Figure 4C). CNV analysis of 159 genes showed copy number amplification and deletion in 22 genes in the three molecular subtypes (Figure 5A). Expression analysis of the corresponding genes in CNV groups of *DOLPP1*, *PLEKHA6*, *PTEN*, and *MNDA* demonstrated that the four genes had higher expression levels in the Gain group and low expression in the Loss group (Figure 5B).

A total of seven EMT genes and two mismatch repair genes were used to calculate the 450K beta values. The beta values of *ZEB1*, *TW1ST1*, *CDH2*, *CDH1*, and *MLH1* differed among the three molecular subtypes (Figure 6A). Pearson correlation analysis of gene expressions and beta values showed that *ZEB1*, *VIM*, *CDH2*, *CDH1*, and *CLDN1* expressions were negatively correlated with beta value (Figure 6B). The beta value of the cg probe site in *CDH1* was higher in Immune-E (Figure 6C). Similarly, the beta value of the cg probe site was negatively correlated with CDH1 expression (Figure 6D).

Immunotherapy prediction and drug sensitivity analysis

We used TIDE (http://tide.dfci.harvard.edu/) software to evaluate the potential clinical effect of immunotherapy according to the molecular subtypes. The TIDE score was lower for Immune-E, indicating that Immune-E may be more suitable for immunotherapy. Moreover, 47% of samples in Immune-E showed immunotherapy response, a proportion higher than those in Stromal-E and Immuno-D (Figure 7A). The IC_{50} values for cisplatin, erlotinib, sunitinib, paclitaxel, saracatinib, and dasatinib

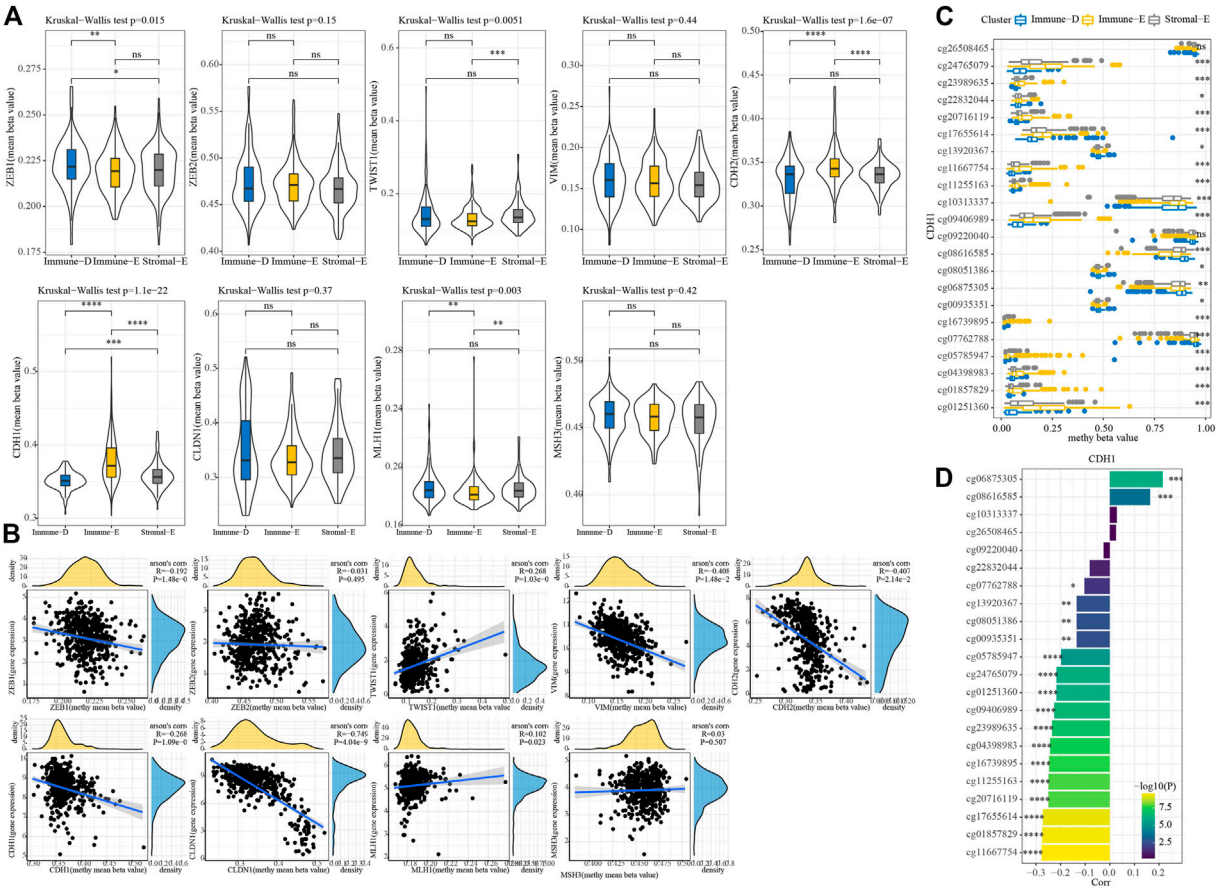


FIGURE 6
Methylation analysis of genes among the three molecular subtypes. **(A)** 450K beta value differences for nine genes among the three molecular subtypes. **(B)** Correlation analysis between 450K beta values and gene expression. **(C)** Distributions of beta in the cg probe site in *CDH1* among the three molecular subtypes. **(D)** Correlation analysis between 450K beta values of CDH1 and *CDH1* expression. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns: no significance.

were lower in Immune-E, suggesting that Immune-E is more sensitive to those chemotherapeutic drugs (Figure 7B).

Construction of the IMScore

In the TCGA dataset, Pearson correlation analysis between genes in pathways and pathways identified 1784 genes with |R| >0.4 and $p < 0.05$. Then, in the TCGA training dataset, seven prognosis genes ($p < 0.05$) for thyroid cancer were screened from 1784 genes using univariate Cox analysis. Finally, four genes were used to construct a prognostic model (IMScore = 0.732*HSPA6+0.917*FLNC 1.083*CLDN2 0.966*E2F1) through lasso analysis and the stepAIC method.

In the TCGA training, testing, and entire TCGA datasets, samples were classified into high and low IMScore groups using the cutoff. KM curve analysis showed that patients in the low group had longer survival times. Moreover, in terms of PFS, DFI, and DSS, the low group showed better PFI ($p = 0.04$) and DSS ($p < 0.0001$) (Figure 8A). The IMScores were higher in the Immune-D and Stromal-E subtypes (Figure 8B).

Performance prediction of the prognostic model

Among 32 cancer types in the TCGA dataset, high IMScore survival times were shorter than low IMScore survival times except for TGCT and UCS (Supplementary Figure S2). We validated the prediction effect of IMScore in the immunotherapy datasets IMvigor210 and GSE91061. In the IMvigor210 dataset, samples with a low IMScore had better survival outcomes, and the 0.5-, 1-, and 1.5-year AUCs were 0.58, 0.64, and 0.65, respectively (Figure 9A). The samples with low TIDE had better survival outcomes, and the 0.5-, 1-, and 1.5-year AUCs were 0.54, 0.57, and 0.57, respectively (Figure 9B). Samples with low PD-L1 also had better survival outcomes, and the 0.5-, 1-, and 1.5-year AUCs were 0.6, 0.6, and 0.59, respectively (Figure 9C). The prediction of the response to treatment showed AUCs of TIDE, PD-L1, and IMScore of 0.58, 0.57, and 0.67, respectively (Figure 9D). In the GSE91061 dataset, samples with a low IMScore had better survival outcomes, and the 0.5-, 1-, and 1.5-year AUCs were 0.59, 0.75, and 0.75, respectively (Figure 9E). The samples with low TIDE had better survival outcomes, and the 1-, 2-, 2.5-year

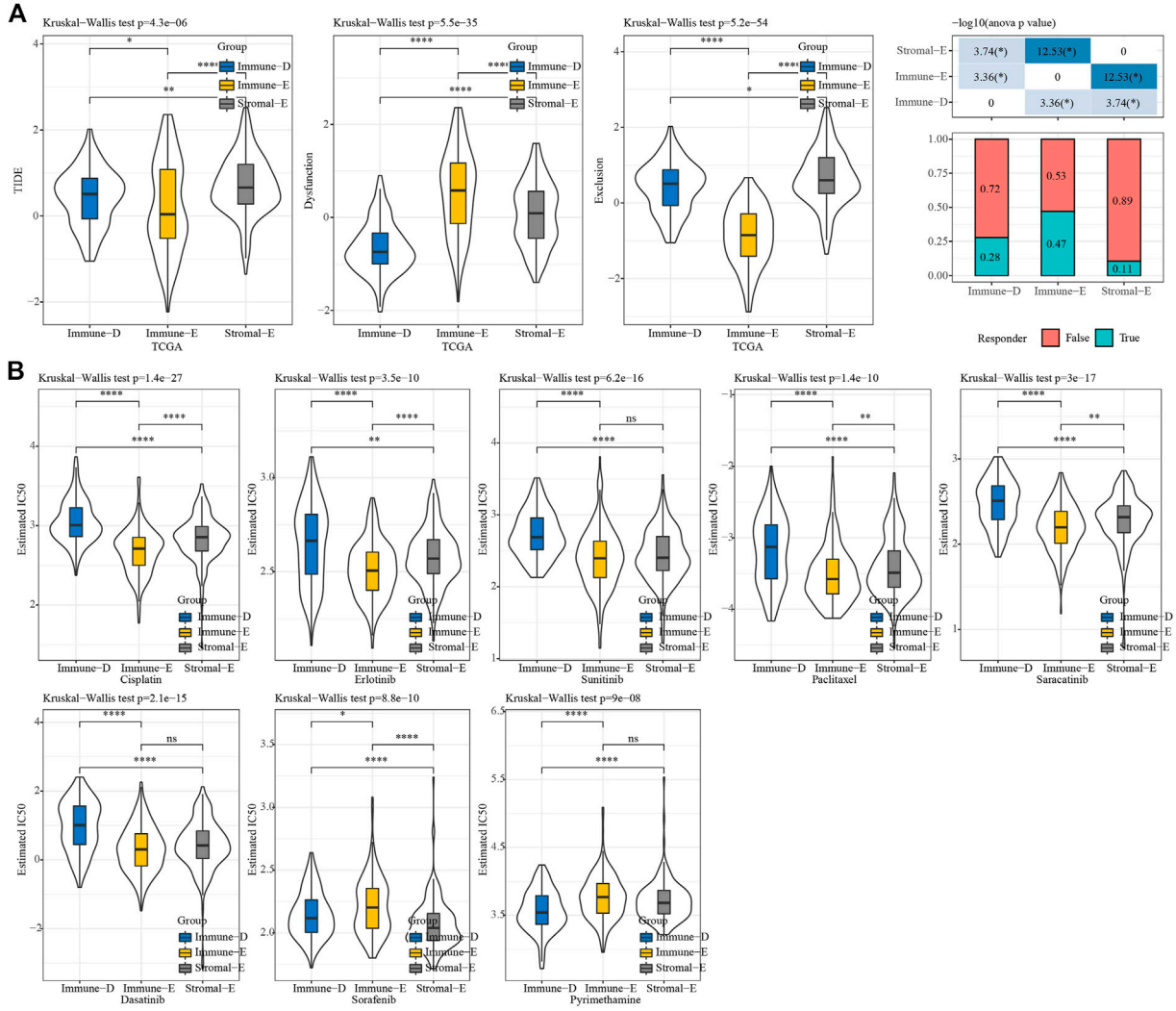


FIGURE 7
TIDE and drug sensitivity analysis. **(A)** TIDE analysis among the three molecular subtypes. **(B)** IC_{50} analysis of eight drugs among the three molecular subtypes. *$p < 0.05$; **$p < 0.01$; ****$p < 0.0001$; ns: no significance.

AUCs were 0.61, 0.58, and 0.59, respectively (Figure 9F). The survival outcomes did not differ significantly between the low- and high- PD-L1 groups, and the 0.5-, 1-, 1.5-year AUCs were 0.54, 0.57, and 0.57, respectively (Figure 9G). The prediction of the response to treatment showed AUCs of TIDE, PD-L1, and IMScore of 0.58, 0.55, and 0.61, respectively (Figure 9H). The results of the aforementioned analyses demonstrated the better prediction effect of the IMScore compared to TIDE.

Nomogram model of thyroid cancer

First, univariate analysis showed that age, gender, TNM stage ($p < 0.001$), stage, and IMScore were significantly associated with a shorter OS in patients with thyroid cancer (Figure 10A). Then, we established a nomogram model that included the important predictors in the Cox analysis to predict the prognosis of thyroid cancer (Figure 10B). The calibration curve showed good concordance between the predicted and observed values of 1-, 3-, and 5-year OS (Figure 10C). The decision curve showed that the nomogram had the best prediction performance for the prognosis of thyroid cancer (Figure 10D).

Discussion

The main obstacle to tumor progression is the immune system, which sees tumors as emerging pathogens that require elimination (Martin et al., 2021). Understanding tumor immunity is critical for improving current immunotherapy regimens. In 2018, Thorsson et al. (2018) developed a new immune classification system comprising six immune subtypes: C1 (wound healing), C2 (IFN-γ phenotype), C3 (inflammatory), C4 (lymphocyte depletion), C5 Type I (immunosilencing), and type C6 (TGF-β dominant). In different tumors, different immune subtypes have different prognoses, and patients with C4 and C6 tumors have worse

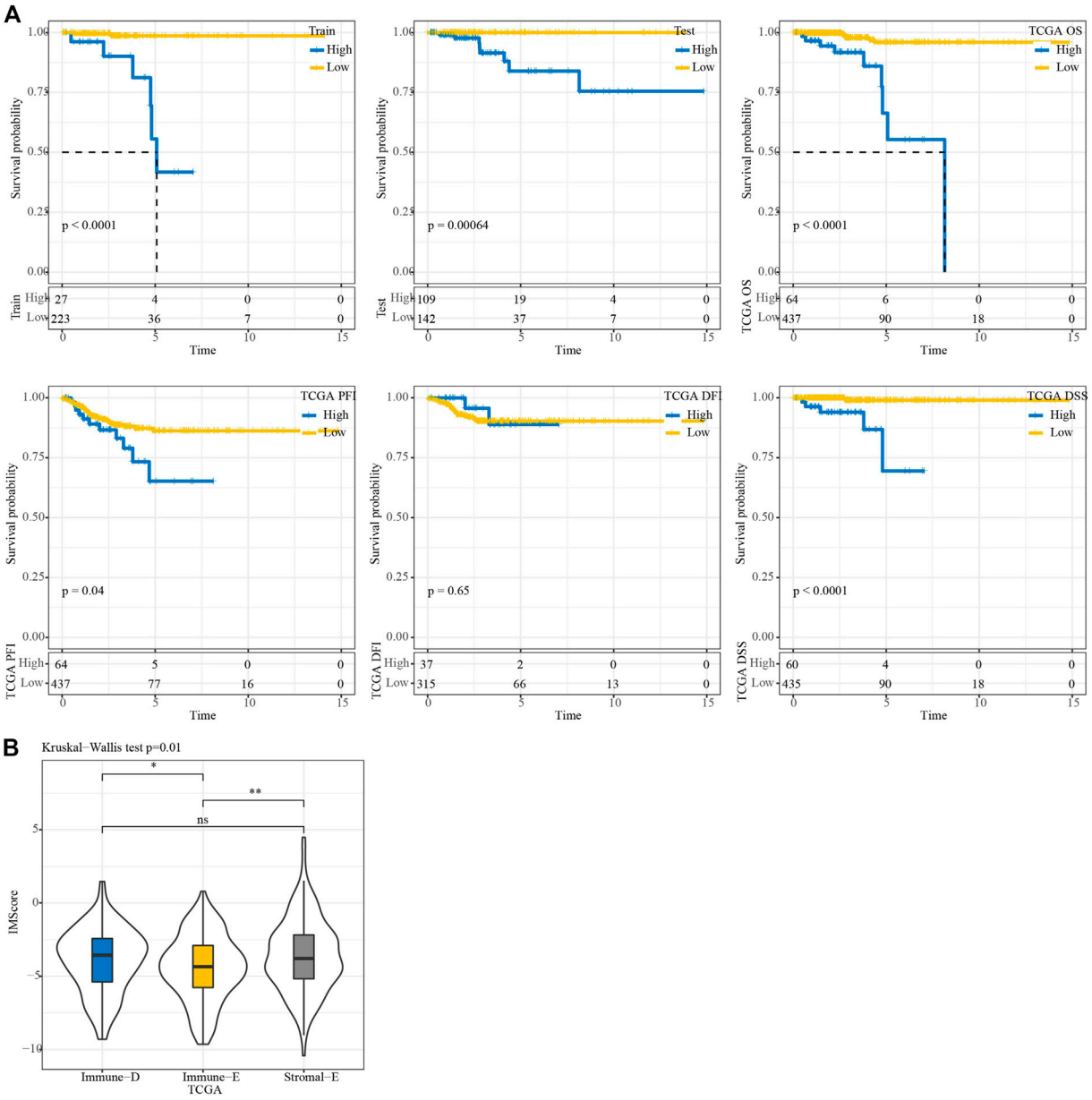


FIGURE 8
KM survival analysis. **(A)** KM survival analysis of the high and low groups in TCGA train, test, entire, TCGA-PFI, TCGA-DFI, and TCGA-DSS datasets. **(B)** Differences in IMScore among the three molecular subtypes.

prognoses. In colorectal cancer, the immune subtypes are mainly types C1 and C2. Different immune subtypes cause different biological differences, which may explain drug heterogeneity in patients with traditional cytotoxic drugs and immunotherapy (Soldevilla et al., 2019). We divided thyroid cancer samples into three immune types based on immune cells: Immune-enrich (E), Stromal-enrich(E), and Immune-deprived(D). Immuno-E showed a high immune cell infiltration but shorter OS, probably because of a small number of dead samples (7.96%).

Cancer is essentially a genomic disease that progresses as mutations including CNVs and SNPs accumulate in somatic cells, as well as epigenomic alterations with or without inherited alterations. CNV is one of the most common markers in the cancer genome, which can lead to oncogene activation and tumor suppressor gene inactivation (Nakagawa and Fujita, 2018). DNA methylation is the most important epigenetic variation in the human genome, and the process of cell carcinogenesis is always accompanied by extensive changes in DNA methylation (Locke et al., 2019; Pan et al., 2021). To further explore the underlying differences in mechanism among the three immune subtypes, we selected methylation and gene copy number. We detected CNV and hypermethylation of tumor driver genes in all three subtypes. The methylation and copy values of genes were negatively and positively correlated with mRNA expression levels, respectively; hence, the differences between thyroid cancer subtypes may be due to changes in gene copy number and methylation.

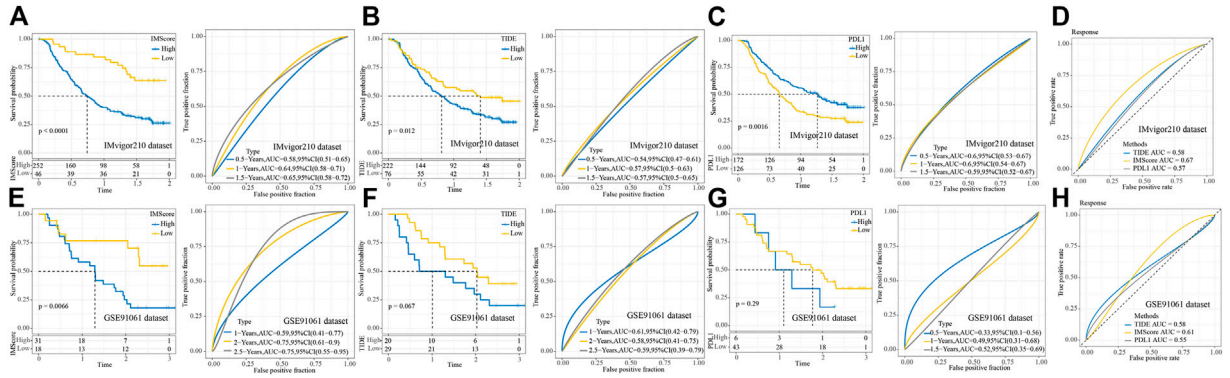


FIGURE 9
Performance of IMScore in immunotherapy datasets. **(A)** KM survival curve and ROC analysis of IMScore in the IMvigor210 dataset. **(B)** KM survival curve and ROC analysis of TIDE in the IMvigor210 dataset. **(C)** KM survival curve and ROC analysis of PD-L1 in the IMvigor210 dataset. **(D)** ROC analysis of IMScore and TIDE in the IMvigor210 dataset. **(E)** KM survival curve and ROC analysis of IMScore in the GSE91061 dataset. **(F)** KM survival curve and ROC analysis of TIDE in the GSE91061 dataset. **(G)** KM survival curve and ROC analysis of PD-L1 in the GSE91061 dataset. **(H)** ROC analysis of IMScore and TIDE in the GSE91061 dataset.

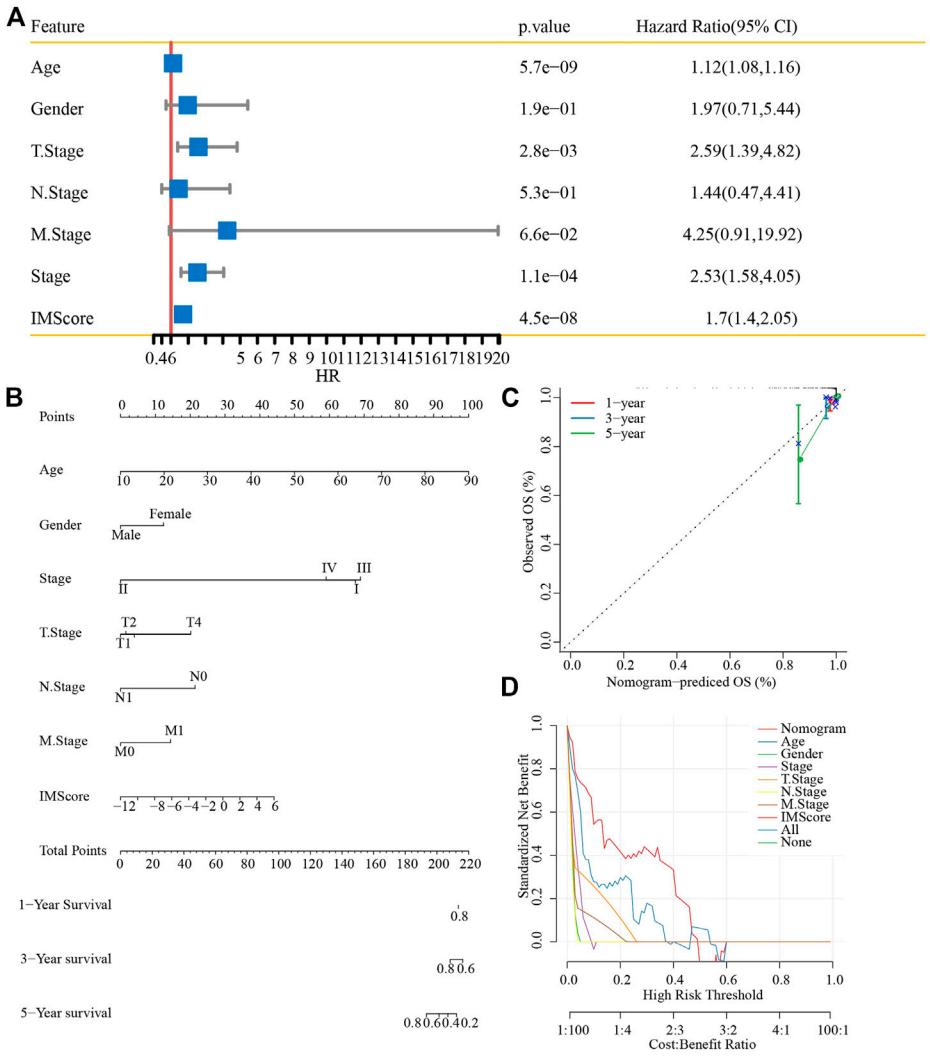


FIGURE 10
Nomogram construction. **(A)** Univariate analysis of the IMScore and clinical features. **(B)** Nomogram that incorporated the IMScore and clinical features was developed. **(C)** Calibration curve. **(D)** Decision curve analysis.

In this study, biological information analysis identified four genes, *HSPA6*, *FLNC*, *CLDN2*, and *E2F1* as candidate biomarkers of thyroid cancer. Recently, *HSPA6* was found to be indispensable in the Withaferin A-mediated inhibition of apoptosis/autophagy or migration in breast cancer cells (Hahm et al., 2021). Alterations in Claudin-2 (*CLDN2*), a component of cellular tight junction, are involved in the progression of a variety of cancer types (Buchert et al., 2010; Tabariès et al., 2011; Tabariès et al., 2012). *E2F1* is a potent oncogene in human cancers, including thyroid cancer, prostate cancer, lung cancer, and colorectal cancer, that can accelerate the invasion, spread, and metastasis of cancer cells and further predict poor prognosis (Bi et al., 2017; Yin et al., 2017; Zhou et al., 2020; Yang et al., 2022).

Although we used bioinformatics methods on large numbers of samples to identify genetic subgroups and develop a prognosis model of thyroid carcinoma that showed significant prognostic differences, this study has several limitations. Future work will place a greater emphasis on research that is both fundamentally experimental and functionally in-depth. Moreover, we were unable to consider other factors because the samples lacked essential clinical follow-up information, such as diagnostic specifics; for instance, whether the patients had other health conditions. These factors may have informed the differentiation of the molecular subtypes.

In conclusion, we identified three immune molecular subtypes and developed a prognostic model based on four prognostic genes, which may provide new targets for the diagnosis and treatment of thyroid cancer. Further studies are needed to confirm the mechanism of prognostic genes, which will provide new opportunities for the diagnosis and treatment of thyroid cancer.

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

WZ and TnL performed the statistical analyses and prepared the draft manuscript. XL, TaL, XM, and DZ edited the manuscript. YL, XdZ, and XkZ provided critical comments on the manuscript. All authors checked and proofread the final version of the manuscript.

Funding

This work was supported by Key Projects of Intergovernmental Cooperation from National Key R&D Programs (grant no. 2022YFE0131800), a Project of the Department of Education of Liaoning Province (grant no. QNZR2020017, the China Postdoctoral Science Foundation (grant no. 2022MD713825), the 345 Talent Project of Shengjing Hospital of China Medical University, a project from the Shenyang Science and Technology Bureau (grant no. RC210316), and the Natural Science Foundation of Liaoning Province (grant no. 2022-MS-235).

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.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2023.1130399/full#supplementary-material

References

Bai, Y., Guo, T., Huang, X., Wu, Q., Niu, D., Ji, X., et al. (2018). In papillary thyroid carcinoma, expression by immunohistochemistry of BRAF V600E, PD-L1, and PD-1 is closely related. *Virchows Archiv : Int. J. pathology* 472 (5), 779–787. doi:10.1007/s00428-018-2357-6

Bi, X. C., Pu, X. Y., Liu, J. M., and Huang, S. (2017). Effect of transcription factor E2F1 expression on the invasion of prostate cancer. *Zhonghua yi xue za zhi* 97 (36), 2856–2859. doi:10.3760/cma.j.issn.0376-2491.2017.36.016

Branchoux, S., Bellera, C., Italiano, A., Rustand, D., Gaudin, A. F., and Rondeau, V. (2019). Immune-checkpoint inhibitors and candidate surrogate endpoints for overall survival across tumour types: A systematic literature review. *Crit. Rev. oncology/hematology* 137, 35–42. doi:10.1016/j.critrevonc.2019.02.013

Buchert, M., Papin, M., Bonnans, C., Darido, C., Raye, W. S., Garambois, V., et al. (2010). Symplekin promotes tumorigenicity by up-regulating claudin-2 expression. *Proc. Natl. Acad. Sci. United States of America* 107 (6), 2628–2633. doi:10.1073/pnas.0903747107

Chen, Y. P., Lv, J. W., Mao, Y. P., Li, X. M., Li, J. Y., Wang, Y. Q., et al. (2021). Unraveling tumour microenvironment heterogeneity in nasopharyngeal carcinoma identifies biologically distinct immune subtypes predicting prognosis and immunotherapy responses. *Mol. cancer* 20 (1), 14. doi:10.1186/s12943-020-01292-5

Farkona, S., Diamandis, E. P., and Blasutig, I. M. (2016). Cancer immunotherapy: The beginning of the end of cancer? *BMC Med.* 14, 73. doi:10.1186/s12916-016-0623-5

Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., et al. (2020). Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med.* 12 (1), 21. doi:10.1186/s13073-020-0721-z

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

Geeleher, P., Cox, N., and Huang, R. S. (2014). pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PloS one* 9 (9), e107468. doi:10.1371/journal.pone.0107468

Gnjatic, S., Bronte, V., Brunet, L. R., Butler, M. O., Disis, M. L., Galon, J., et al. (2017). Identifying baseline immune-related biomarkers to predict clinical outcome of immunotherapy. *J. Immunother. cancer* 5, 44. doi:10.1186/s40425-017-0243-4

Gunda, V., Gigliotti, B., Ndishabandi, D., Ashry, T., McCarthy, M., Zhou, Z., et al. (2018). Combinations of BRAF inhibitor and anti-PD-1/PD-L1 antibody improve survival and tumour immunity in an immunocompetent model of orthotopic murine anaplastic thyroid cancer. *Br. J. cancer* 119 (10), 1223–1232. doi:10.1038/s41416-018-0296-2

Hahm, E. R., Kim, S. H., Singh, K. B., and Singh, S. V. (2021). RNA-seq reveals novel cancer-selective and disease subtype-independent mechanistic targets of withaferin A in human breast cancer cells. *Mol. Carcinog.* 60 (1), 3–14. doi:10.1002/mc.23266

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

Li, L., and Wang, X. (2021). Identification of gastric cancer subtypes based on pathway clustering. *NPJ Precis. Oncol.* 5 (1), 46. doi:10.1038/s41698-021-00186-z

Locke, W. J., Guanzon, D., Ma, C., Liew, Y. J., Duesing, K. R., Fung, K. Y. C., et al. (2019). DNA methylation cancer biomarkers: Translation to the clinic. *Front. Genet.* 10, 1150. doi:10.3389/fgene.2019.01150

Ma, M., Lin, B., Wang, M., Liang, X., Su, L., Okose, O., et al. (2020). Immunotherapy in anaplastic thyroid cancer. *Am. J. Transl. Res.* 12 (3), 974–988.

Martin, T. D., Patel, R. S., Cook, D. R., Choi, M. Y., Patil, A., Liang, A. C., et al. (2021). The adaptive immune system is a major driver of selection for tumor suppressor gene inactivation. *Sci. (New York, NY)* 373 (6561), 1327–1335. doi:10.1126/science.abg5784

Mould, R. C., van Vloten, J. P., AuYeung, A. W. K., Karimi, K., and Bridle, B. W. (2017). Immune responses in the thyroid cancer microenvironment: Making immunotherapy a possible mission. *Endocrine-related cancer* 24 (12), T311–T29. doi:10.1530/ERC-17-0316

Nakagawa, H., and Fujita, M. (2018). Whole genome sequencing analysis for cancer genomics and precision medicine. *Cancer Sci.* 109 (3), 513–522. doi:10.1111/cas.13505

Pan, H., Renaud, L., Chaligne, R., Bloehdorn, J., Tausch, E., Mertens, D., et al. (2021). Discovery of candidate DNA methylation cancer driver genes. *Cancer Discov.* 11 (9), 2266–2281. doi:10.1158/2159-8290.CD-20-1334

Qin, X. J., Lin, X., Xue, G., Fan, H. L., Wang, H. Y., Wu, J. F., et al. (2021). CXCL10 is a potential biomarker and associated with immune infiltration in human papillary thyroid cancer. *Biosci. Rep.* 41 (1). doi:10.1042/BSR20203459

Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160 (1-2), 48–61. doi:10.1016/j.cell.2014.12.033

Shen, W., Song, Z., Xiao, Z., Huang, M., Shen, D., Gao, P., et al. (2022). Sangerbox: A comprehensive, interaction-friendly clinical bioinformatics analysis platform. *iMeta* 1 (3), e36. doi:10.1002/imt2.36

Soldevilla, B., Carretero-Puche, C., Gomez-Lopez, G., Al-Shahrour, F., Riesco, M. C., Gil-Calderon, B., et al. (2019). The correlation between immune subtypes and consensus molecular subtypes in colorectal cancer identifies novel tumour microenvironment profiles, with prognostic and therapeutic implications. *Eur. J. cancer* 123, 118–129. doi:10.1016/j.ejca.2019.09.008

Tabariès, S., Dong, Z., Annis, M. G., Omeroglu, A., Pepin, F., Ouellet, V., et al. (2011). Claudin-2 is selectively enriched in and promotes the formation of breast cancer liver metastases through engagement of integrin complexes. *Oncogene* 30 (11), 1318–1328. doi:10.1038/onc.2010.518

Tabariès, S., Dupuy, F., Dong, Z., Monast, A., Annis, M. G., Spicer, J., et al. (2012). Claudin-2 promotes breast cancer liver metastasis by facilitating tumor cell interactions with hepatocytes. *Mol. Cell. Biol.* 32 (15), 2979–2991. doi:10.1128/MCB.00299-12

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830. doi:10.1016/j.immuni.2018.03.023

Wang, J., Yu, F., Shang, Y., Ping, Z., and Liu, L. (2020). Thyroid cancer: Incidence and mortality trends in China, 2005-2015. *Endocrine* 68 (1), 163–173. doi:10.1007/s12020-020-02207-6

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

Wu, T., and Dai, Y. (2017). Tumor microenvironment and therapeutic response. *Cancer Lett.* 387, 61–68. doi:10.1016/j.canlet.2016.01.043

Yang, J., Ying, Y., Zeng, X., Liu, J., Xie, Y., Deng, Z., et al. (2022)., 2022. Amsterdam, 7081611. doi:10.1155/2022/7081611Transcription factor E2F1 exacerbates papillary thyroid carcinoma cell growth and invasion via upregulation of LINC00152*Anal. Cell. Pathol.*

Yin, J., Fu, W., Dai, L., Jiang, Z., Liao, H., Chen, W., et al. (2017). ANKRD22 promotes progression of non-small cell lung cancer through transcriptional up-regulation of E2F1. *Sci. Rep.* 7 (1), 4430. doi:10.1038/s41598-017-04818-y

Zhang, Y., Song, J., Zhao, Z., Yang, M., Chen, M., Liu, C., et al. (2020). Single-cell transcriptome analysis reveals tumor immune microenvironment heterogenicity and granulocytes enrichment in colorectal cancer liver metastases. *Cancer Lett.* 470, 84–94. doi:10.1016/j.canlet.2019.10.016

Zhi, J., Yi, J., Tian, M., Wang, H., Kang, N., Zheng, X., et al. (2020). Immune gene signature delineates a subclass of thyroid cancer with unfavorable clinical outcomes. *Aging* 12 (7), 5733–5750. doi:10.18632/aging.102963

Zhou, C., Liu, H. S., Wang, F. W., Hu, T., Liang, Z. X., Lan, N., et al. (2020). circCAMSAP1 promotes tumor growth in colorectal cancer via the miR-328-5p/E2F1 Axis. *Mol. Ther. : J. Am. Soc. Gene Ther.* 28 (3), 914–928. doi:10.1016/j.ymthe.2019.12.008

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 31 March 2023
DOI 10.3389/fphar.2023.1131610

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Tengda Wang,
First Affiliated Hospital of Harbin Medical University, China
Qun Zhang,
Nanjing Medical University, China

*CORRESPONDENCE
Qing Yang,
25744194@qq.com
Ting Xia,
xiating@cdutcm.edu.cn

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 25 December 2022
ACCEPTED 23 March 2023
PUBLISHED 31 March 2023

CITATION
Deng Y, Wang F, Wu X, Du K, Yang Q and Xia T (2023), The m6A-regulation and single cell effect pattern in sunitinib resistance on clear cell renal cell carcinoma: Identification and validation of targets.
Front. Pharmacol. 14:1131610.
doi: 10.3389/fphar.2023.1131610

The m6A-regulation and single cell effect pattern in sunitinib resistance on clear cell renal cell carcinoma: Identification and validation of targets

Yanxi Deng[1†], Fang Wang[1†], Xinhui Wu[2], Kangming Du[3], Qing Yang[3*] and Ting Xia[2,4*]

[1]Clinical Laboratory, Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China, [2]Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China, [3]Department of Cardiothoracic Surgery, Hospital of Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China, [4]Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China

Background: Sunitinib is the main target drug for clear cell renal cell carcinoma. However, the effect of sunitinib is often limited by acquired drug resistance.

Methods: The open-accessed data used in this study were obtained from different online public databases, which were analyzed using the R software. The RNA level of specific genes was detected using quantitative Real-Time PCR. Sunitinib-resistant cell lines were constructed based on protocol get from the previous study. Colony formation and Cell Counting Kit-8 assays were applied to detect cell proliferation ability.

Results: In this study, through publicly available data and high-quality analysis, we deeply explored the potential biological mechanisms that affect the resistance of sunitinib. Detailed, data from GSE64052, GSE76068 and The Cancer Genome Atlas were extracted. We identified the IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 were associated with sunitinib resistance. Single-cell analysis, prognosis analysis and m6A regulatory network were conducted to investigate their role. Moreover, the MX2 was selected for further analysis, including its biological role and effect on the ccRCC microenvironment. Interestingly, we noticed that MX2 might be an immune-related gene that could affect the response rate of immunotherapy. Then, *in vitro* experiments validated the overexpression of MX2 in sunitinib-resistance cells. Colony formation assay indicated that the knockdown of MX2 could remarkably inhibit the proliferation ability of 786-O-Res and Caki-1-Res when exposed to sunitinib.

Conclusion: In summary, through publicly available data and high-quality analysis, we deeply explored the potential biological mechanisms that affect the resistance of sunitinib. MX2 was selected for further analysis, including its biological role and effect on the ccRCC microenvironment. Finally, *in vitro* experiments were used to validate its role in ccRCC.

KEYWORDS

epigenetic, sunitinib, ccRCC, MX2, biological

Introduction

Renal cell carcinoma (RCC) is a malignant tumor that arises out of the renal tubular epithelium, which is very common in the world. It is estimated that about 300,000 new cases are created each year, and 130,000 cancer-related deaths are caused at the same time (Cohen and McGovern, 2005). Among them, seventy to eighty percent of all cases of renal cell carcinoma are clear cell (ccRCC) (Wettersten et al., 2017). As a multifactorial disease, the pathogenesis of ccRCC is not completely clear, and it is closely related to age, obesity, smoking, hypertension, genetic factors, and so on (Znaor et al., 2015). In the past decade, the incidence rate of renal tumors worldwide has shown a trend of continuous growth, and the internal microenvironment of tumors is usually accompanied by the reprogramming of metabolic networks and pathways. Through metabolic reprogramming, tumor cells proliferate rapidly, survive under hypoxia and nutrient depletion, and escape immune surveillance (Wettersten et al., 2017). Meanwhile, due to the lack of early clinical symptoms, more than 30%–50% of ccRCC patients missed the best opportunity for surgery, making the diagnosis, treatment, and prevention of it a serious public health problem worldwide.

At present, for early and resectable RCC, radical surgery is still a major treatment. Although surgery can cure most early-stage patients, due to the high blood metastasis rate, about 25% of locally progressed or localized patients will suffer from metastasis (Dudani et al., 2021). Additionally, considering the occult symptoms of RCC, about 20%–25% of patients had distant metastasis at the time of diagnosis and could not undergo radical surgery (Xue et al., 2021). Moreover, for patients with advanced or metastatic stages, the five-year survival rate is only about 23% due to the high heterogeneity and invasiveness of the disease (Atkins and Tannir, 2018). Unfortunately, the treatment of ccRCC with radiotherapy and chemotherapy is not effective, which limits its treatment options to some extent. Sunitinib is a kind of drug that can selectively target tyrosine kinase, which is widely utilized in RCC and has achieved encouraging results (Bex et al., 2019). However, some patients receiving sunitinib treatment will still be limited by acquired drug resistance (McDermott et al., 2018). Considering the practical significance of this problem, researchers have begun to pay attention to and identify the specific biological mechanism of acquired resistance to sunitinib (Broxterman et al., 2009). Zhu and their colleagues found that the ZHX2 can induce sunitinib resistance through the autophagy regulated by MEK/ERK axis (Zhu et al., 2020). Bender and their colleagues noticed that the overexpressed PRKX, TTBK2 and RSK4 can lead to sunitinib resistance (Bender and Ullrich, 2012). The m6A RNA methylation is an epigenetic modification pathway widely existing in the cancer microenvironment. Chen and their colleagues revealed that TRAF1 can contribute to sunitinib resistance based on the METTL14 and m6A modifications (Chen et al., 2022). Consequently, it is of practical clinical significance to identify biological targets that may participate in the resistance of sunitinib.

Access to public data can provide convenience for researchers (Wang et al., 2019a; Wei et al., 2020; Zhang et al., 2022). Here, through publicly available data and high-quality analysis, we deeply explored the potential biological mechanisms that affect the resistance of sunitinib. Detailed, data from GSE64052, GSE76068 and TCGA were extracted. We identified the IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 were associated with sunitinib resistance. Single-cell analysis, prognosis analysis and m6A regulatory network were conducted to investigate their role. Moreover, the MX2 was selected for further analysis, including its biological role and effect on the ccRCC microenvironment. Interestingly, we noticed that MX2 might be an immune-related gene that could affect the response rate of immunotherapy. Then, *in vitro* experiments validated the overexpression of MX2 in sunitinib-resistance cells. Colony formation assay indicated that the knockdown of MX2 could remarkably inhibit the proliferation ability of 786-O-Res and Caki-1-Res when exposed to sunitinib.

Methods

Acquisition and pre-processing of open-accessed data

The open-accessed data used in this study were obtained from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) databases. The GSE64052 and GSE76068 contain the sequence information between the sunitinib-resistant and wild-type RCC cells (Zhang et al., 2015). For the TCGA database, the clinical features and transcription profile information were directly downloaded from the TCGA-KIRC project. Initially, the expression profile file of the individual patient was downloaded in "STAR-Counts" format and converted into TPM format through R code. Before analysis, we adjusted the range of expression values to 1–20 through data preprocessing for all data. The first step is to annotate the probe ID as the corresponding gene symbol through the annotation file (GRCh38. p13). The second part is to complete the missing values in the expression matrix. The third step is to average the expression amount of duplicate gene symbols and remove the part where the mean value is less than 0.1. Limma package was applied to identify the genes differentially expressed between different groups (Ritchie et al., 2015). The genes affected by sunitinib were get from the CTD database. The baseline information of TCGA-KIRC patients was shown in Supplementary Table S1.

Gene ontology (GO) and kyoto encyclopedia of genes and genomes (KEGG)

GO and KEGG analysis can reflect the biological effect based on the input molecules, which was performed using the clusterprofiler package (Yu et al., 2012). Detailed, the "OrgDb" was "org.Hs.eg.db"; the "pvalueCutoff" was 0.05; the "qvalueCutoff" was 0.05; the "ont" was "all".

Single-cell evaluation

Specific gene expression patterns in the ccRCC microenvironment were evaluated using the online website TISCH project, a scRNA-seq database aiming to

characterize tumor microenvironment at single-cell resolution (Sun et al., 2021). Detailed, the database in TISCH projects KIRC_GSE111360, KIRC_GSE121636, KIRC_GSE139555 and KIRC_GSE145281 were selected to illustrate the single-cell expression pattern of MX2 (major-linegae).

Cytoscape software

The co-expression analysis was visualized using the Cytoscape software (Shannon et al., 2003).

Pathway investigation

To identify pathways significantly different between the two groups, gene set enrichment analysis (GSEA) was employed. Reference gene set was "Hallmark". The enriched pathways with false discovery rate (FDR) < 0.25 and p.adjust < 0.05 were regarded as significant (Subramanian et al., 2005). Based on the pathway set, single sample GSEA (ssGSEA) analysis was conducted (Hänzelmann et al., 2013).

Methylation

The list of molecules involved m6A process was collected from the previous study (Lv et al., 2021). The correlation between clinical features and gene methylation was investigated using the MEXPRESS database (https://mexpress.be/).

Tumor microenvironment

Through bioinformatics analysis, the tumor microenvironment can be quantified using specific algorithms. In our study, the tumor microenvironment was quantified using the EPIC, MCPCOUNTER, TIMER, CIBERSORT, QUANTISEQ and XCELL algorithms (Becht et al., 2016; Li et al., 2017; Chen et al., 2018; Racle and Gfeller, 2020).

Tumor immune dysfunction and exclusion (TIDE)

The TIDE score quantified by the TIDE algorithm can reflect the response rate of patients on immunotherapy. Meanwhile, as well as immune dysfunction and immune exclusion levels, the TIDE algorithm quantified cancer-associated fibroblasts, M2 macrophages, and myeloid-derived suppressor cells (Fu et al., 2020).

Immunohistochemistry

In the HPA database, MX2 was immunohistochemically detected in ccRCC tumors and normal tissue (Uhlén et al., 2015).

Establishment of sunitinib-resistant cell lines and cell culture

The 786-O and Caki-1 cell lines were laboratories stored and cultured in RPMI-1640 culture medium added with 10% fetal bovine serum (FBS) under the standard cell culture conditions of 37°C with 5% CO_2. The process to induce the cell lines resistant to sunitinib was followed by a previous study (Sakai et al., 2013; Wei et al., 2021). The IC50 of used cells 786-O/786-O-Res and Caki-1/Caki-1-Res were 27.66/102.1 and 10.26/73.59 nM.

Quantitative Real-Time PCR

Total RNA extraction and cDNA preparation were conducted following the standard process (Wei et al., 2021). The primer used for PCR was: forward, 5′-TGAACGTGCAGCGAGCTT-3′, reverse, 5′-GGCTT GTGGGCCTTAGACAT-3′; GPADH, 5′-CTGGGC TACACTGAGCACC-3'; reverse, 5′-AAGTGGTCGTTGAGG GCAATG-3'.

RNA interference

The plasmids used for cell transfection were purchased from Shanghai GenePharma Co., Ltd., and the sequences were: sh#1: 5′-GCACGATTGAAGACATAAA-3′, sh#2: 5′- GGGACGCCTTCA CAGAATA-3′, sh#3: 5′-GCCAACCAGATCCCATTTA-3'. The processes of cell transfection were conducted following the standard process using the Lipofectamine 3,000 regrant.

Cell Counting Kit-8 (CCK8) and colony formation assays

CCK8 and colony formation assays were conducted following the standard process (Wei et al., 2021).

Statistical analysis

All the analysis were completed in the R, SPSS and GraphPad Prism 8 software. The 0.05 was set as the statistical threshold. Normally distributed data are analyzed using independent T-tests. Non-normally distributed data are analyzed using the Mann-Whitney U tests.

Results

Figure 1 illustrates the flow chart of our study. In this study, we identified the molecules involved in sunitinib resistance through the data from GSE64052, GSE76068 and TCGA-KIRC. Then, the biological enrichment and single-cell analysis based on TISCH project were conducted to investigate the role of identified molecules in ccRCC, as well as their interaction network with m6A regulators. Ultimately, MX2 was identified for further analysis, including expression pattern, prognosis role, biological

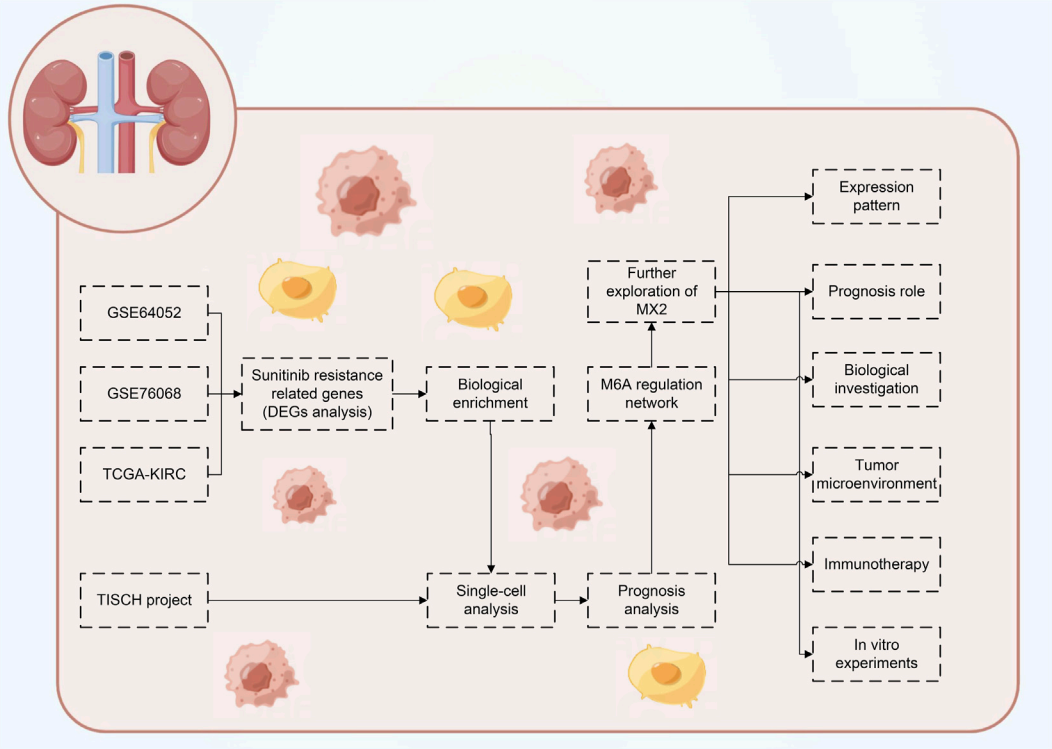


FIGURE 1
The flow chart of the whole study.

investigation, tumor microenvironment, immunotherapy evaluation and *in vitro* experiments validation.

Effect of sunitinib-resistant related genes in ccRCC

The data normalization process of GSE64052 and GSE76068 were shown in Figures 2A–D. We investigated the underlying biological effects of sunitinib on ccRCC cells. For genes positively correlated with sunitinib (Supplementary Material S1), the genes were enriched in cytoplasmic translation, rRNA processing, ribosome biogenesis, ncRNA processing, process utilizing autophaic mechanism and autophagy (Figure 2E, GO-BP); cytosolic ribosome, late endosome, lysosomal membrane, vacuolar membrane and lytic vacuole membrane (Figure 2F, GO-CC); phosphatidylinositol binding, ubiquitin-protein transferase activity and ubiquitin-like protein transferase activity (Figure 2G, GO-MF); biosynthesis of amino acids, HIF-1 signaling pathway, p53 signaling pathway, phosphatidylinositol signaling system, FoxO signaling pathway and mTOR signaling pathway (Figure 2H, KEGG). For genes negatively correlated with sunitinib (Supplementary Material S1), the genes were enriched in DNA-dependent DNA replication, mitotic sister chromatid segregation, mitotic nuclear division, DNA replication, nuclear division, and organelle fission (Figure 2I, GO-BP); centromeric region, condensed chromosome, chromosomal region and spindle (Figure 2J, GO-CC); DNA replication origin binding, structural constituent of muscle, actinin binding, DNA helicase activity, catalytic activity, acting on DNA and actin binding (Figure 2K, GO-MF); DNA replication, cell cycle, cardiac muscle contraction, age-race signaling pathway in diabetic complications, prion disease and parkinsion disease (Figure 2L, KEGG). Through the limma package with the threshold of $|\log FC| > 0.5$ and $p < 0.05$, 280 downregulated and 200 upregulated genes were identified in GSE64052 between the sunitinib-resistant and wild-type RCC cells (Figure 2M); 83 downregulated and 53 upregulated genes were identified in GSE76068 between the sunitinib-resistant and wild-type RCC cells (Figure 2N). Furthermore, we found that six genes were commonly upregulated, while nine genes were commonly downregulated in both GSE64052 and GSE76068 cell lines (Figure 2O). The clinical roles of these six genes were shown in Supplementary Figure S1A–D.

Single-cell analysis

Following this, we evaluated the single-cell level of six commonly upregulated genes in the ccRCC single-cell level (Figures 3A–D; Supplementary Figure S2A–E). Results indicated that IFITM1 was mainly expressed in NK cells, Treg and $CD8^+$ T cells in four ccRCC single-cell cohorts, GSE111360, GSE121636, GSE139555 and GSE145281; the overall expression level of IL6, PCOLEC2 and RSAD2 seems to be very low; MX2 and SLC2A3 are expressed in various cells. KM survival curves were then used to identify the prognosis role of these genes (Figures 3E–G).

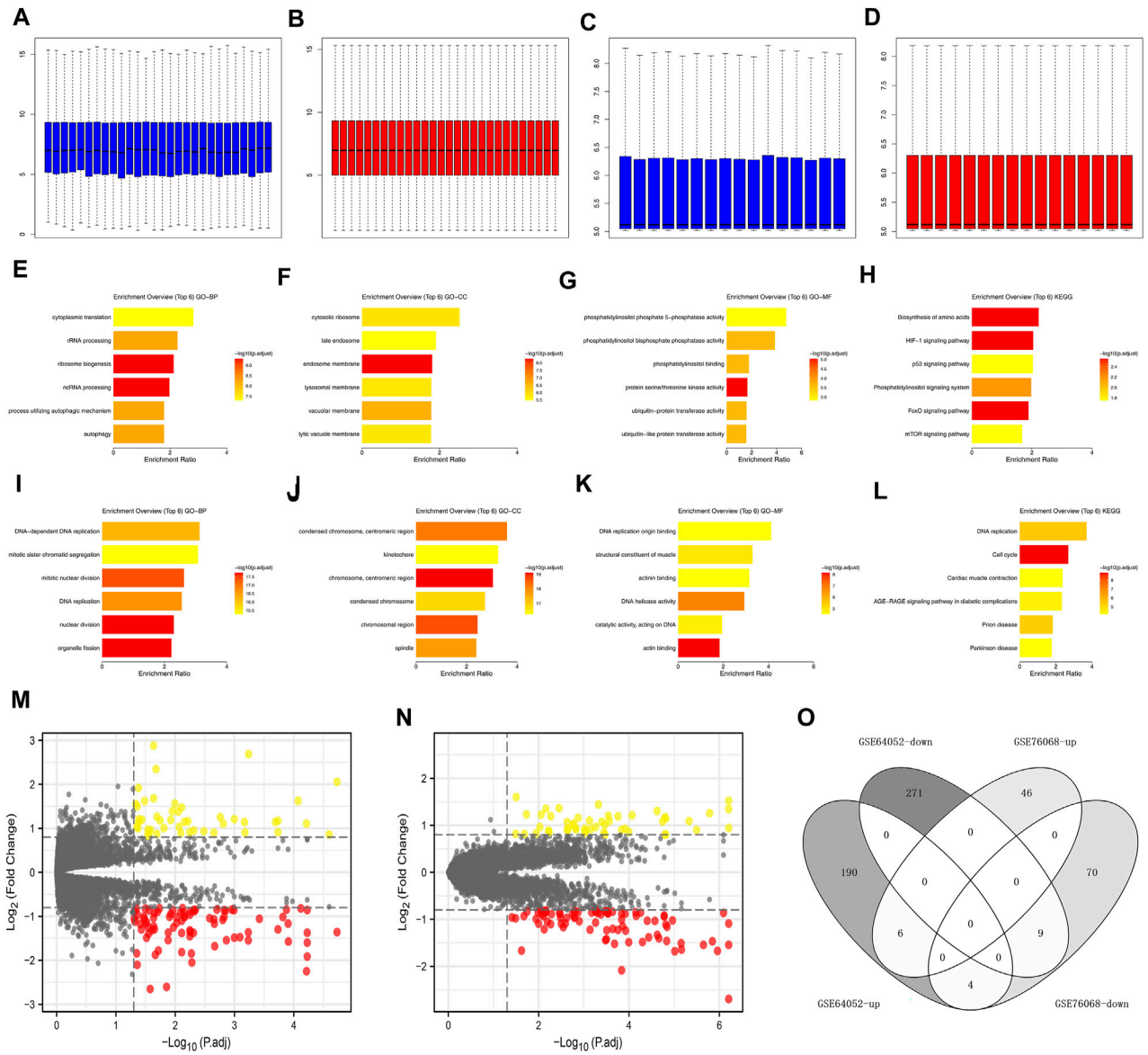


FIGURE 2
Effect of sunitinib-resistant related genes in ccRCC. **(A)** The GO-BP terms of genes positively correlated with sunitinib (CTD database); **(B)** The GO-CC terms of genes positively correlated with sunitinib (CTD database); **(C)** The GO-CC terms of genes positively correlated with sunitinib (CTD database); **(D)** The KEGG terms of genes positively correlated with sunitinib (CTD database); **(E)** The GO-BP terms of genes negatively correlated with sunitinib (CTD database); **(F)** The GO-CC terms of genes negatively correlated with sunitinib (CTD database); **(G)** The GO-CC terms of genes negatively correlated with sunitinib (CTD database); **(H)** The KEGG terms of genes negatively correlated with sunitinib (CTD database); **(I,J)** Data normalization of GSE64052; **(K,L)** Data normalization of GSE76068; **(M)** DEGs analysis of GSE64052; **(N)** DEGs analysis of GSE76068; **(O)** Intersection of DEGs result of GSE76068 and GSE64052.

Results indicated that the MX2 and IL6 are associated with worse survival performance of patients, but the statistical *p*-value of IFITM1, PCOLCE2, RSAD2, and SLC2A3 were not significant.

The m6A-regulation regulatory network of sunitinib-resistant related genes

The m6A modification is an important part of the epigenetic field and has been reported to affect sunitinib resistance (Li et al., 2022). The expression pattern of m6A regulators was shown in Figure 4A. We noticed that the IFITM1 was regulated by YTHDC1, METTL14, RBM15, ALKBH5, WTAP, HNRNPC, YTHDF1, METTL3, ZC3H13, YTHDF2 and FTO (Figure 4B); RSAD2 was regulated by YTHDC2, FTO, ALKBH5, RBM15, ZC3H13, YTHDF2, YTHDF1, WTAP, HNRNPC, YTHDC1, METTL14, METTL3 and YTHDC2 (Figure 4C); PCOLCE2 was regulated by YTHDC2, ZC3H13, RBM15, FTO, HNRNPC, ALKBH5, WTAP, YTHDF1, METTL14, YTHDC1, YTHDF2 (Figure 4D); MX2 was regulated by YTHDC2, FTO,

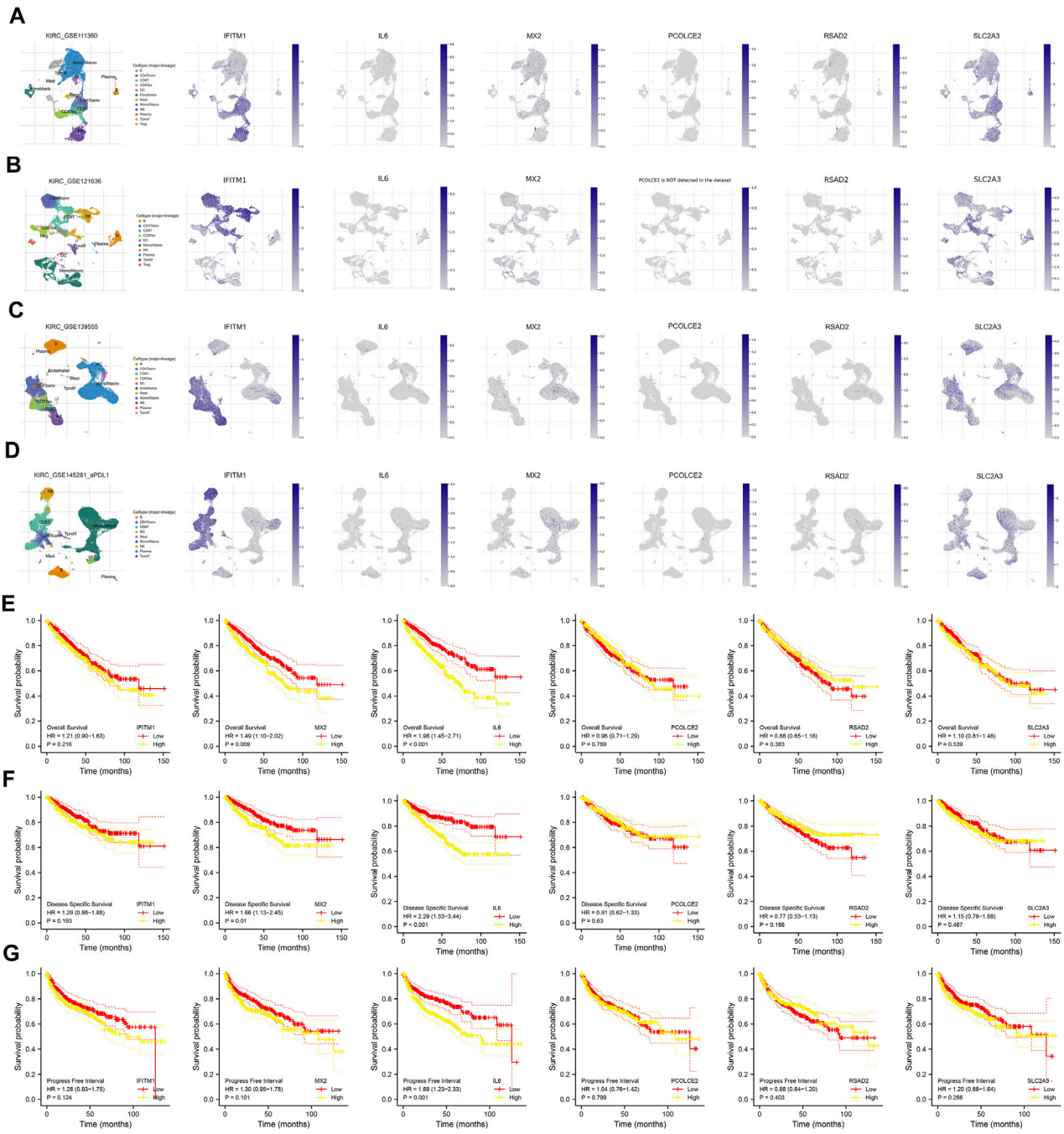


FIGURE 3
Single-cell and prognosis analysis of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3. **(A)** Single-cell analysis of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 in GSE111360; **(B)** Single-cell analysis of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 in GSE121636; **(C)** Single-cell analysis of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 in GSE139555; **(D)** Single-cell analysis of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 in GSE145281; **(E)** Overall survival of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 in ccRCC; **(F)** Disease free survival of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 in ccRCC; **(G)** Progression free survival of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 in ccRCC.

RBM15, YTHDC1, ZC3H13, METTL3, ALKBH5, YTHDF2, HNRNPC, WTAP, METTL14 and YTHDF1 (Figure 4E); SLC2A3 was regulated by ZC3H13, METTL14, RBM15, YTHDF2, ALKBH5, YTHDF1, YTHDC2, FTO, HNRNPC, WTAP, METTL3 and YTHDC1 (Figure 4F). Interestingly, we noticed MX2 was positively correlated with all m6A regulators, including HNRNPC, YTHDF2, METTL3, YTHDF1, YTHDC2, ALKBH5, FTO, YTHDC1, ZC3H13, RBM15, WTAP and METTL14 (Figures 4G–R). Moreover, we noticed that the methylation sites cg00764652, cg05656374, cg152811283, and cg21130374 were negatively correlated with the MX2 expression (Figures 4S–V).

10.3389/fphar.2023.1131610

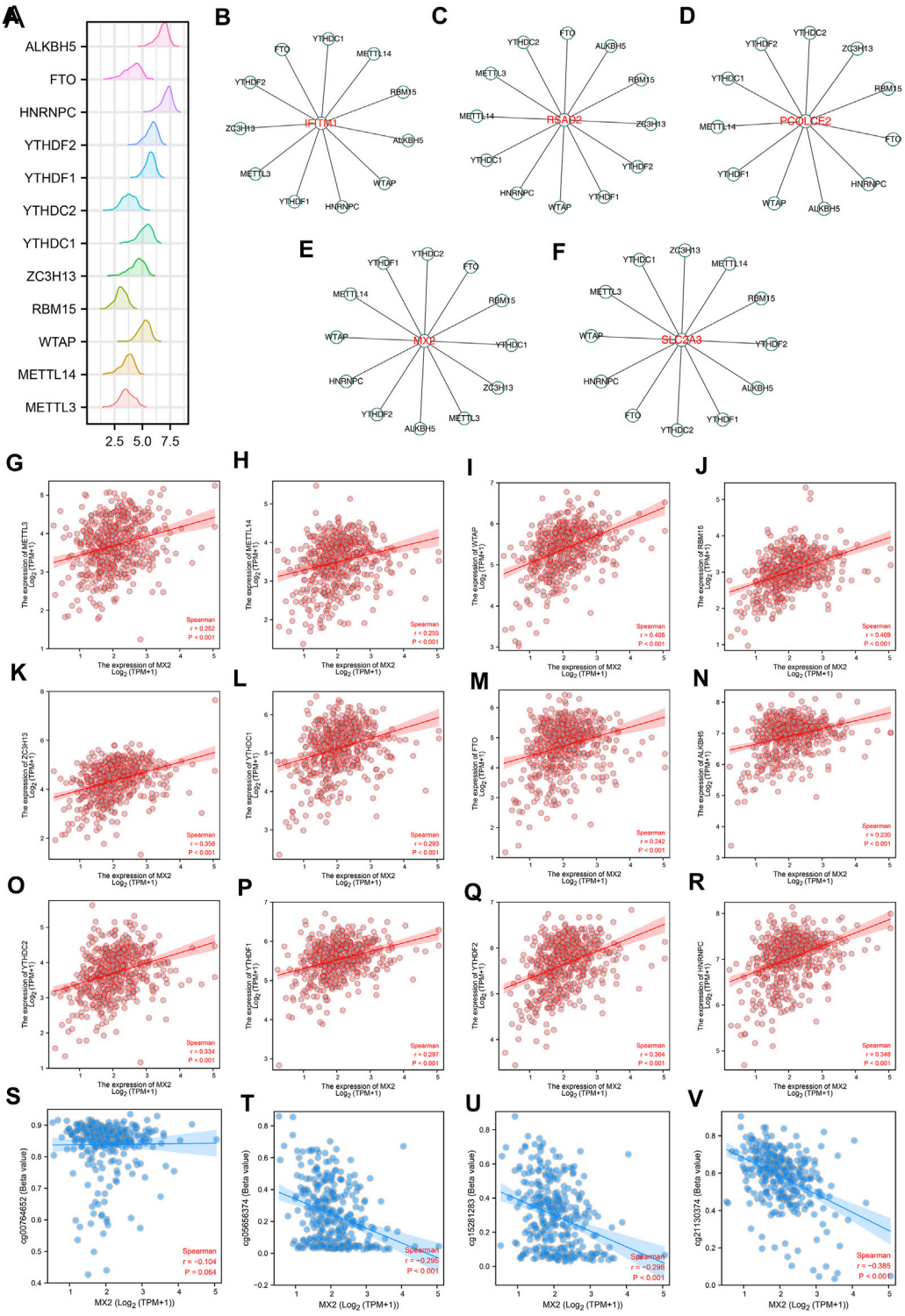


FIGURE 4
The m6A regulatory network of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3. **(A)** The expression pattern of m6A regulators in ccRCC; **(B–F)** The m6A regulatory network of IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3; **(G–R)** Correlation between MX2 and m6A regulators; **(S–V)** Correlation between MX2 and methylation site.

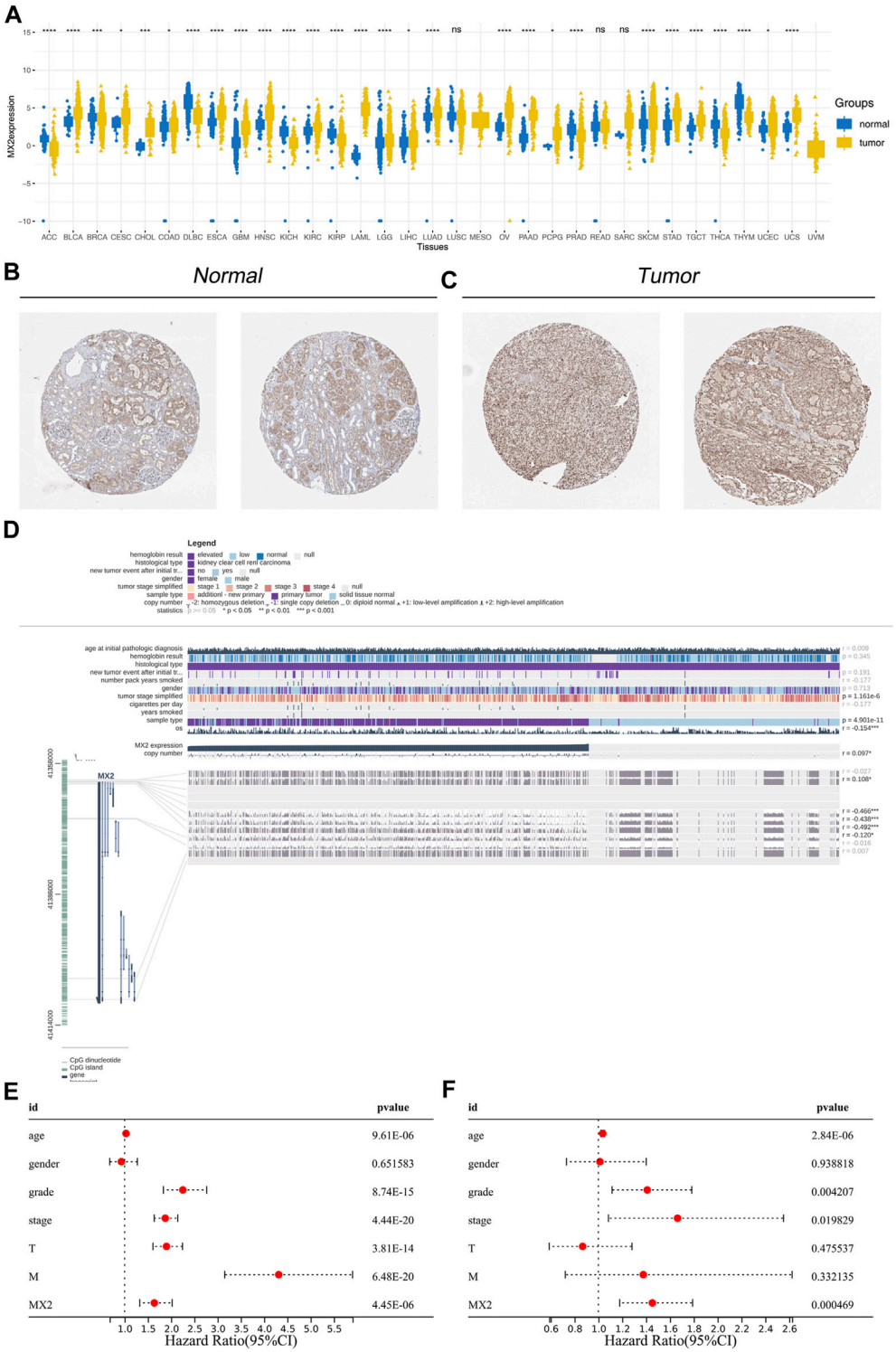


FIGURE 5
Effect pattern of MX2 in ccRCC. **(A)** Pan-cancer analysis of MX2; **(B)** The immunohistochemistry image of MX2 in normal renal tissue, ns = $p > 0.05$, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$; **(C)** The immunohistochemistry image of MX2 in renal cancer tissue; **(D)** Overview of MX2 methylation in ccRCC, * = $p < 0.05$, *** = $p < 0.001$; **(E)** Univariate Cox regression analysis of MX2; **(F)** Multivariate Cox regression analysis of MX2.

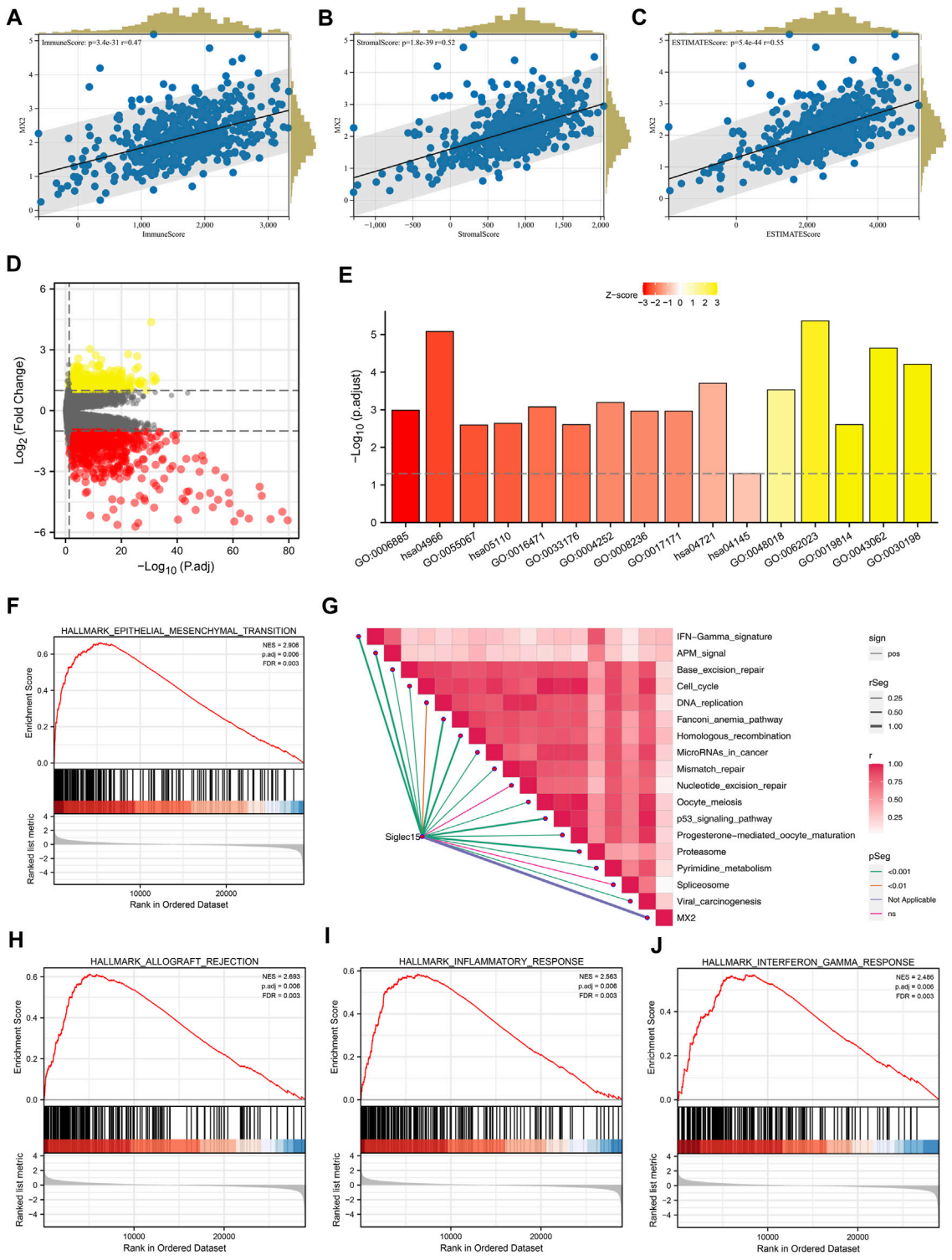


FIGURE 6
Biological investigation of MX2. **(A–C)** Correlation between MX2 and immune score, stromal score and estimate score quantified by estimate package; **(D)** DEGs analysis in patients with high and low MX2 expression; **(E)** GO and KEGG analysis of MX2 in ccRCC; **(F, H–J)** GSEA analysis based on Hallmark gene set; **(G)** ssGSEA algorithm was used to quantify the enrichment score of immune pathways.

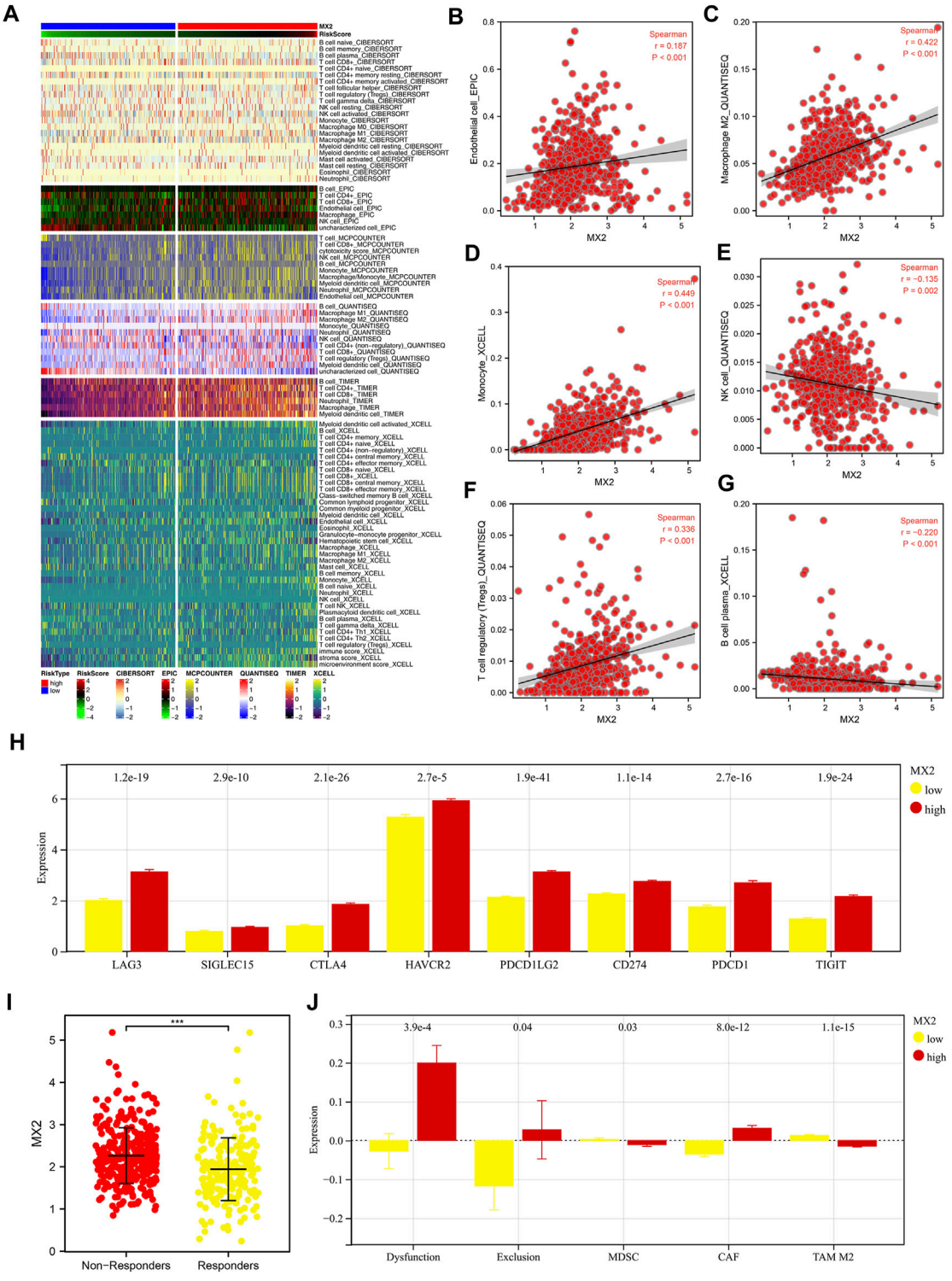


FIGURE 7
Effect of MX2 on ccRCC microenvironment. **(A)** The ccRCC microenvironment was quantified based on multiple algorithms; **(B–G)** Correlation between MX2 and specific cells; **(H)** The expression level of specific immune checkpoints in patients with high and low MX2 expression; **(I)** The expression level of MX2 in immunotherapy responders and non-responders, *** = $p < 0.001$; **(J)** Levels of immune dysfunction, immune exclusion and CAF, MDSC and TAM M2 in patients with high and low MX2 expression.

Subsequent analysis of MX2

Next, we evaluated the expression pattern of MX2 in pan-cancer. Results showed that MX2 was differentially expressed in most cancer (Figure 5A). We noticed a relatively higher protein level of MX2 in ccRCC protein (Figures 5B, C). The overview of the MX2 and methylation site were shown in Figure 5D. Cox regression analysis of single factor and multiple factors showed that MX2 is an independent prognosis factor for ccRCC survival (Figures 5E, F). We also explored the lncRNAs and mRNAs significantly correlated with MX2 expression, which was shown in Supplementary Material S2, S3.

Biological investigation

The Estimate R package was utilized to quantify the tumor microenvironment of the ccRCC microenvironment. In the correlation analysis, the immune score, stromal score, and estimate score were positively correlated with the MX2 (Figures 6A–C). The differentially expressed genes (DEGs) analysis was performed between the patients with high and low MX2 expression (Figure 6D). Based on these DEGs, we found MX2 was mainly enriched in the terms of GO:0006885, hsa04966, GO:0055067, hsa05110, GO:0033176, GO:0004252, GO:0008236, GO:0017171, hsa04145, GO:0048018, GO:0019814, GO:0043062, GO:0030198 (Figure 6E). There was a positive correlation between MX2 and multiple pathways in the ssGSEA analysis (Figure 6G). Using GSEA analysis, it was revealed that the DEGs with a high level of epithelial-mesenchymal transition, allograft rejection and inflammation were enriched in the Hallmark signaling (Figures 6F, H–J).

Effect of MX2 on tumor microenvironment

Multiple algorithms mentioned in the method section were utilized to quantify the tumor microenvironment of ccRCC. From the heatmap, we observed a remarkably different infiltration pattern of quantified cells in patients with high and low MX2 expression (Figure 7A). Correlation analysis showed that MX2 was positively correlated with endothelial cell_EPIC, macrophages M2_QUANTISEQ, monocyte_XCELL, Tregs_QUANTISEQ, yet negatively correlated with the NK cell_QUANTISEQ and B cell plasma_XCELL (Figures 7B–G). Moreover, we noticed that all the key immune checkpoints, including LAG3, SIGLEC15, CTLA4, HAVCR2, PDCD1LG2, CD274, PDCD1 and TIGIT were overexpressed in patients with high MX2 level (Figure 7H). Furthermore, we tried to explore whether MX2 has an impact on the immunotherapeutic response of ccRCC. Results showed that the immunotherapy non-responders had a higher MX2 level (Figure 7I). Meanwhile, patients with higher MX2 expression might have a higher level of immune dysfunction, immune exclusion and CAF, while a lower level of MDSC and TAM M2 (Figure 7J).

MX2 is associated with sunitinib resistance

Through the method mentioned above, we construct two cell lines resistant to sunitinib, named 786-O-Res and Caki-1-Res. The results of IC50 to sunitinib validated the resistance of these cells on sunitinib (Figure 8A, IC50 of 786-O-wild = 27.66, IC50 of 786-O-Res = 102.1; Figure 8B, IC50 of Caki-1-wild = 10.26, IC50 of Caki-1-Res = 73.59). The result of the PCR revealed that MX2 was overexpressed in sunitinib-resistance cell lines (Figure 8C, 786-O-Res and Caki-1-Res). The inhibition efficiency of MX2 in cell lines was validated using the PCR and sh#2 was selected for further experiments (Figure 8D). Colony formation assay indicated that the knockdown of MX2 could remarkably hamper the proliferation ability of 786-O-Res and Caki-1-Res when exposed to sunitinib (Figure 8E).

Discussion

With the change in the comprehensive environment, the incidence rate of ccRCC is increasing year by year (Jonasch et al., 2021). Advanced RCC is mainly treated with drugs, and it is not sensitive to radiotherapy and has a poor effect on chemotherapy (Barata and Rini, 2017). Although non-specific immunotherapy is beneficial to some patients with advanced RCC, it has little clinical benefit in most cases and obvious toxic exposure (Barata and Rini, 2017). Sunitinib has effectively improved the survival performance of patients with RCC, with relatively small side effects, and is currently the main means of drug treatment for renal cancer (Barata and Rini, 2017). However, in practical clinical application, some patients receiving sunitinib treatment often have acquired drug resistance, which limits their therapeutic benefits (Larroquette et al., 2021).

In this study, through publicly available data and high-quality analysis, we deeply explored the potential biological mechanisms that affect the resistance of sunitinib. Detailed, data from GSE64052, GSE76068 and TCGA were extracted. We identified the IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3 were associated with sunitinib resistance. Single-cell analysis, prognosis analysis and m6A regulatory network were conducted to investigate their role. Moreover, the MX2 was selected for further analysis, including its biological role and effect on the ccRCC microenvironment. Interestingly, we noticed that MX2 might be an immune-related gene that could affect the response rate of immunotherapy. Then, *in vitro* experiments validated the overexpression of MX2 in sunitinib-resistance cells. Colony formation assay indicated that the knockdown of MX2 could remarkably inhibit the proliferation ability of 786-O-Res and Caki-1-Res when exposed to sunitinib.

Six genes were identified to induce sunitinib resistance in ccRCC, including IFITM1, IL6, MX2, PCOLCE2, RSAD2 and SLC2A3. Some of these genes have been reported to play an important role in cancer. Provance and their colleagues found that the IFITM1 could be affected by crosstalk between the NF-κB and interferon-alpha and regulated breast cancer progression (Provance et al., 2021). Lee and their colleagues noticed that the IFITM1 affected gastric cancer pathological characteristics through epigenetic regulation (Lee et al., 2012). Yu and their

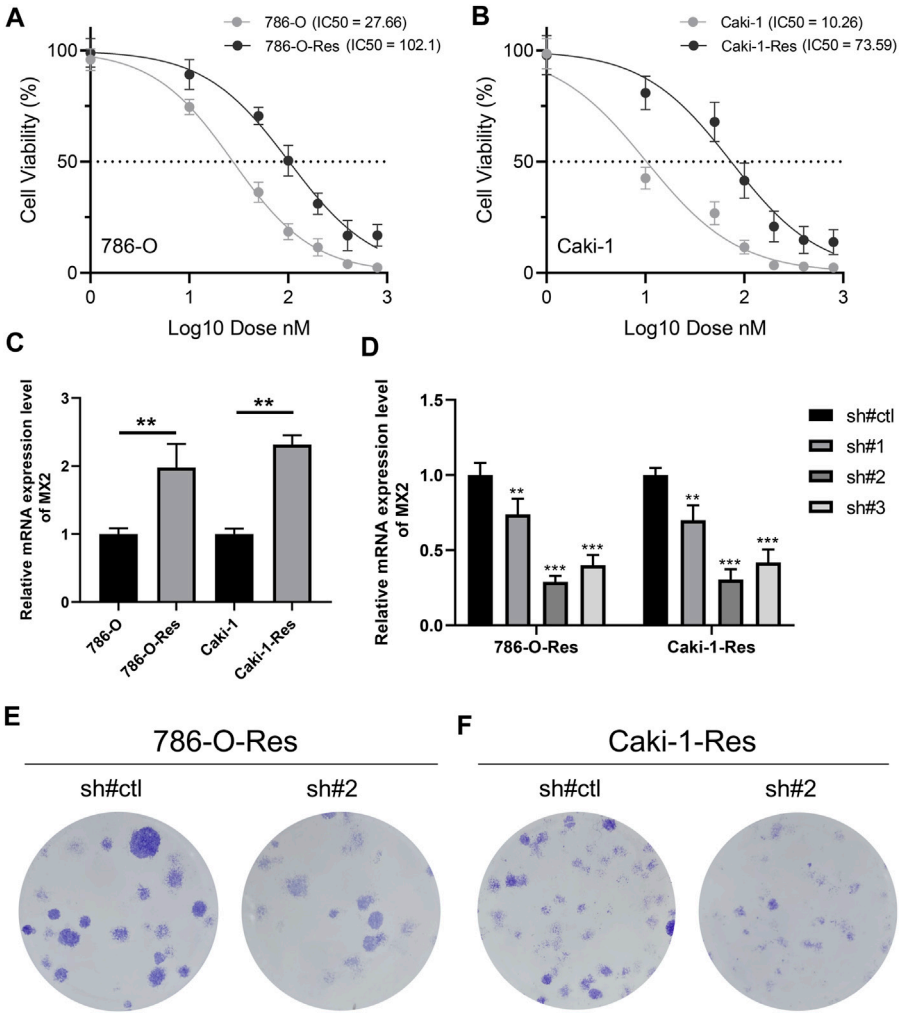


FIGURE 8
MX2 is associated with sunitinib resistance **(A–B)** The IC50 of wild-type and sunitinib-resistant cells (786-O and Caki-1); **(C)** The expression level of MX2 in wild-type and sunitinib-resistant cells, ** = $p < 0.01$; **(D)** PCR was used to validate the knockdown efficiency of MX2, ** = $p < 0.01$, *** = $p < 0.001$; **(E)** Colony formation assay in 786-O-Res and Caki-1-Res when exposed to sunitinib.

colleagues indicated that the IFITM1 could facilitate colon cancer metastasis by regulating CAV-1 (Yu et al., 2015). Yao and their colleagues found that the SLC2A3 could facilitate M2 macrophage infiltration by inducing glycolysis reprogramming (Yao et al., 2020). Liu and their colleagues demonstrated that the SLC2A3 could lead to the reduction of vitamin C uptake, therefore inhibiting leukemia development (Liu et al., 2020). Juraleviciute and their colleagues noticed that the MX2 could regulate the XAF1 and make the melanoma cells sensitive to targeted therapy (Juraleviciute et al., 2021). Wang and their colleagues found that the MX2 could suppress the glioblastoma progression through ERK/P38/NF-κB signaling (Wang et al., 2019b). Our results provide a reference for revealing the mode of action of these genes in ccRCC. Meanwhile, we deeply and comprehensively analyzed the role pattern of MX2 in ccRCC, and validated its influence on sunitinib resistance through *in vitro* experiments, making it a potential clinical target.

We found that these sunitinib-resistant related genes were regulated by multiple m6A regulators. The m6A epigenetic modification has also been reported to be related to sunitinib resistance. Chen and their colleagues noticed that TRAF1 can contribute to sunitinib resistance based on the METTL14 and m6A modifications (Chen et al., 2022). Li and their colleagues noticed that the level of YTHDC1 was downregulated by YY1/HDAC2 and could regulate the sunitinib resistance targeting the ANXA1-MAPK pathway (Li et al., 2022).

Correlation analysis showed that MX2 was positively correlated with endothelial cell_EPIC, macrophages M2_QUANTISEQ, monocyte_XCELL, Tregs_QUANTISEQ, yet negatively correlated with the NK cell_QUANTISEQ and B cell plasma_XCELL. Previous studies have reported the relationship between these cells and the progression of ccRCC. For example, van Hooren and their colleagues noticed that agonistic CD40-antibody could be enhanced by sunitinib through reducing MDSCs, increasing endothelial activation, and enhancing T cell recruitment (van Hooren et al., 2016). Dannenmann and their colleagues found that the tumor-associated

macrophages could destroy the function of T cells and reduce the survival rate of ccRCC (Dannenmann et al., 2013). Xu and their colleagues found that HK3 could facilitate the immune escape of ccRCC by inducing monocyte infiltration (Xu et al., 2021). Our results indicate that MX2 may complete the remodeling of the tumor microenvironment by affecting the infiltration level of these cells and then play its biological role.

Although our article provides a biological explanation for sunitinib resistance, some limitations still need to be noted. Firstly, the result from GSE64052 and GSE76068 was only at the cell level. However, due to the complex regulatory mechanism *in vivo*, our conclusions should be subsequently validated *in vivo* models. Secondly, the deep biological mechanism of MX2 in ccRCC still needs to be explored.

Conclusion

In summary, through publicly available data and high-quality analysis, we deeply explored the potential biological mechanisms that affect the resistance of sunitinib. MX2 was selected for further analysis, including its biological role and effect on the ccRCC microenvironment. Finally, *in vitro* experiments were used to validate its role in ccRCC.

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 authors.

Author contributions

DY, WF, and WX collected the public data. DY and WF performed the experiments. DY, WX, and DK worte the manuscript. YQ and XT designed this work.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphar.2023.1131610/full#supplementary-material

References

Atkins, M. B., and Tannir, N. M. (2018). Current and emerging therapies for first-line treatment of metastatic clear cell renal cell carcinoma. *Cancer Treat. Rev.* 70, 127–137. doi:10.1016/j.ctrv.2018.07.009

Barata, P. C., and Rini, B. I. (2017). Treatment of renal cell carcinoma: Current status and future directions. *CA a cancer J. Clin.* 67 (6), 507–524. doi:10.3322/caac.21411

Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 17 (1), 218. doi:10.1186/s13059-016-1070-5

Bender, C., and Ullrich, A. (2012). Prkx, Ttbk2 and Rsk4 expression causes sunitinib resistance in kidney carcinoma- and melanoma-cell lines. *Int. J. cancer* 131 (2), E45–E55. doi:10.1002/ijc.26486

Bex, A., Mulders, P., Jewett, M., Wagstaff, J., van Thienen, J. V., Blank, C. U., et al. (2019). Comparison of immediate vs deferred cytoreductive nephrectomy in patients with synchronous metastatic renal cell carcinoma receiving sunitinib: The surtime randomized clinical trial. *JAMA Oncol.* 5 (2), 164–170. doi:10.1001/jamaoncol.2018.5543

Broxterman, H. J., Gotink, K. J., and Verheul, H. M. (2009). Understanding the causes of multidrug resistance in cancer: A comparison of doxorubicin and sunitinib. *Drug Resist. Updat. Rev. Comment. Antimicrob. anticancer Chemother.* 12 (4-5), 114–126. doi:10.1016/j.drup.2009.07.001

Chen, B., Khodadoust, M. S., Liu, C. L., Newman, A. M., and Alizadeh, A. A. (2018). Profiling tumor infiltrating immune cells with cibersort. *Methods Mol. Biol. Clift. NJ)* 1711, 243–259. doi:10.1007/978-1-4939-7493-1_12

Chen, Y., Lu, Z., Qi, C., Yu, C., Li, Y., Huan, W., et al. (2022). N(6)-Methyladenosine-Modified Traf1 promotes sunitinib resistance by regulating apoptosis and angiogenesis in a mettl14-dependent manner in renal cell carcinoma. *Mol. cancer* 21 (1), 111. doi:10.1186/s12943-022-01549-1

Cohen, H. T., and McGovern, F. J. (2005). Renal-cell carcinoma. *N. Engl. J. Med.* 353 (23), 2477–2490. doi:10.1056/NEJMra043172

Dannenmann, S. R., Thielicke, J., Stöckli, M., Matter, C., von Boehmer, L., Cecconi, V., et al. (2013). Tumor-associated macrophages subvert T-cell function and correlate with reduced survival in clear cell renal cell carcinoma. *Oncoimmunology* 2 (3), e23562. doi:10.4161/onci.23562

Dudani, S., de Velasco, G., Wells, J. C., Gan, C. L., Donskov, F., Porta, C., et al. (2021). Evaluation of clear cell, papillary, and chromophobe renal cell carcinoma metastasis sites and association with survival. *JAMA Netw. open* 4 (1), e2021869. doi:10.1001/jamanetworkopen.2020.21869

Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., et al. (2020). Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med.* 12 (1), 21. doi:10.1186/s13073-020-0721-z

Hänzelmann, 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

Jonasch, E., Walker, C. L., and Rathmell, W. K. (2021). Clear cell renal cell carcinoma ontogeny and mechanisms of lethality. *Nat. Rev. Nephrol.* 17 (4), 245–261. doi:10.1038/s41581-020-00359-2

Juraleviciute, M., Nsengimana, J., Newton-Bishop, J., Hendriks, G. J., and Slipicevic, A. (2021). Mx2 mediates establishment of interferon response profile, regulates Xaf1, and can sensitize melanoma cells to targeted therapy. *Cancer Med.* 10 (8), 2840–2854. doi:10.1002/cam4.3846

Larroquette, M., Peyraud, F., Domblides, C., Lefort, F., Bernhard, J. C., Ravaud, A., et al. (2021). Adjuvant therapy in renal cell carcinoma: Current knowledges and future perspectives. *Cancer Treat. Rev.* 97, 102207. doi:10.1016/j.ctrv.2021.102207

Lee, J., Goh, S. H., Song, N., Hwang, J. A., Nam, S., Choi, I. J., et al. (2012). Overexpression of Ifitm1 has clinicopathologic effects on gastric cancer and is regulated by an epigenetic mechanism. *Am. J. pathology* 181 (1), 43–52. doi:10.1016/j.ajpath.2012.03.027

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, W., Ye, K., Li, X., Liu, X., Peng, M., Chen, F., et al. (2022). Ythdc1 is downregulated by the yy1/hdac2 complex and controls the sensitivity of ccrcc to sunitinib by targeting the anxa1-mapk pathway. *J. Exp. Clin. cancer Res. CR* 41 (1), 250. doi:10.1186/s13046-022-02460-9

Liu, J., Hong, J., Han, H., Park, J., Kim, D., Park, H., et al. (2020). Decreased vitamin C uptake mediated by Slc2a3 promotes leukaemia progression and impedes Tet2 restoration. *Br. J. cancer* 122 (10), 1445–1452. doi:10.1038/s41416-020-0788-8

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

McDermott, D. F., Huseni, M. A., Atkins, M. B., Motzer, R. J., Rini, B. I., Escudier, B., et al. (2018). Clinical activity and molecular correlates of response to atezolizumab alone or in combination with bevacizumab versus sunitinib in renal cell carcinoma. *Nat. Med.* 24 (6), 749–757. doi:10.1038/s41591-018-0053-3

Provance, O. K., Geanes, E. S., Lui, A. J., Roy, A., Holloran, S. M., Gunewardena, S., et al. (2021). Disrupting interferon-alpha and nf-kappab crosstalk suppresses Ifitm1 expression attenuating triple-negative breast cancer progression. *Cancer Lett.* 514, 12–29. doi:10.1016/j.canlet.2021.05.006

Racle, J., and Gfeller, D. (2020). Epic: A tool to estimate the proportions of different cell types from bulk gene expression data. *Methods Mol. Biol. Clift. NJ)* 2120, 233–248. doi:10.1007/978-1-0716-0327-7_17

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Sakai, I., Miyake, H., and Fujisawa, M. (2013). Acquired resistance to sunitinib in human renal cell carcinoma cells is mediated by constitutive activation of signal transduction pathways associated with tumour cell proliferation. *BJU Int.* 112 (2), E211–E220. doi:10.1111/j.1464-410X.2012.11655.x

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

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, D., Wang, J., Han, Y., Dong, X., Ge, J., Zheng, R., et al. (2021). Tisch: A comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment. *Nucleic acids Res.* 49 (1), D1420–D1430. doi:10.1093/nar/gkaa1020

Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Sci. (New York, NY)* 347 (6220), 1260419. doi:10.1126/science.1260419

van Hooren, L., Georganaki, M., Huang, H., Mangsbo, S. M., and Dimberg, A. (2016). Sunitinib enhances the antitumor responses of agonistic Cd40-antibody by reducing mdscs and synergistically improving endothelial activation and T-cell recruitment. *Oncotarget* 7 (31), 50277–50289. doi:10.18632/oncotarget.10364

Wang, H., Guan, Q., Nan, Y., Ma, Q., and Zhong, Y. (2019b). Overexpression of human Mx2 gene suppresses cell proliferation, migration, and invasion via erk/P38/nf-?b pathway in glioblastoma cells. *J. Cell. Biochem.* 120 (11), 18762–18770. doi:10.1002/jcb.29189

Wang, H., Liu, Y., Shen, K., Dong, Y., Sun, J., Shu, Y., et al. (2019a). A comparison between radiofrequency ablation combined with transarterial chemoembolization and surgical resection in hepatic carcinoma: A meta-analysis. *J. cancer Res. Ther.* 15 (7), 1617–1623. doi:10.4103/jcrt.JCRT_503_19

Wei, X., Dong, Y., Chen, X., Ren, X., Li, G., Wang, Y., et al. (2020). Construction of circrna-based cerna network to reveal the role of circrnas in the progression and prognosis of metastatic clear cell renal cell carcinoma. *Aging* 12 (23), 24184–24207. doi:10.18632/aging.104107

Wei, Y., Chen, X., Ren, X., Wang, B., Zhang, Q., Bu, H., et al. (2021). Identification of Mx2 as a novel prognostic biomarker for sunitinib resistance in clear cell renal cell carcinoma. *Front. Genet.* 12, 680369. doi:10.3389/fgene.2021.680369

Wettersten, H. I., Aboud, O. A., Lara, P. N., Jr., and Weiss, R. H. (2017). Metabolic reprogramming in clear cell renal cell carcinoma. *Nat. Rev. Nephrol.* 13 (7), 410–419. doi:10.1038/nrneph.2017.59

Xu, W., Liu, W. R., Xu, Y., Tian, X., Anwaier, A., Su, J. Q., et al. (2021). Hexokinase 3 dysfunction promotes tumorigenesis and immune escape by upregulating monocyte/macrophage infiltration into the clear cell renal cell carcinoma microenvironment. *Int. J. Biol. Sci.* 17 (9), 2205–2222. doi:10.7150/ijbs.58295

Xue, J., Chen, W., Xu, W., Xu, Z., Li, X., Qi, F., et al. (2021). Patterns of distant metastases in patients with clear cell renal cell carcinoma--a population-based analysis. *Cancer Med.* 10 (1), 173–187. doi:10.1002/cam4.3596

Yao, X., He, Z., Qin, C., Deng, X., Bai, L., Li, G., et al. (2020). Slc2a3 promotes macrophage infiltration by glycolysis reprogramming in gastric cancer. *Cancer Cell Int.* 20, 503. doi:10.1186/s12935-020-01599-9

Yu, F., Xie, D., Ng, S. S., Lum, C. T., Cai, M. Y., Cheung, W. K., et al. (2015). Ifitm1 promotes the metastasis of human colorectal cancer via cav-1. *Cancer Lett.* 368 (1), 135–143. doi:10.1016/j.canlet.2015.07.034

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

Zhang, L., Wang, X., Bullock, A. J., Callea, M., Shah, H., Song, J., et al. (2015). Anti-S1p antibody as a novel therapeutic strategy for vegfr tki-resistant renal cancer. *Clin. Cancer Res.* 21 (8), 1925–1934. doi:10.1158/1078-0432.Ccr-14-2031

Zhang, X., Ren, X., Zhang, T., Zhou, X., Chen, X., Lu, H., et al. (2022). Comprehensive analysis of the association between human non-obstructive azoospermia and plasticisers via single-cell and traditional rna sequencing methods. *Expo. Health* 14 (4), 829–842. doi:10.1007/s12403-021-00460-2

Zhu, L., Ding, R., Yan, H., Zhang, J., and Lin, Z. (2020). Zhx2 drives cell growth and migration via activating mek/erk signal and induces sunitinib resistance by regulating the autophagy in clear cell renal cell carcinoma. *Cell death Dis.* 11 (5), 337. doi:10.1038/s41419-020-2541-x

Znaor, A., Lortet-Tieulent, J., Laversanne, M., Jemal, A., and Bray, F. (2015). International variations and trends in renal cell carcinoma incidence and mortality. *Eur. Urol.* 67 (3), 519–530. doi:10.1016/j.eururo.2014.10.002

frontiers | Frontiers in Pharmacology

TYPE Original Research
PUBLISHED 17 April 2023
DOI 10.3389/fphar.2023.1091378

OPEN ACCESS

EDITED BY
Zhi-qian Zhang,
Southern University of Science and Technology, China

REVIEWED BY
Elena Levantini,
National Research Council (CNR), Italy
Yuyi Hou,
First Affiliated Hospital of Jiamusi University, China
Haitao Xu,
Anqing Hospital affiliated to Anhui Medical University, China

*CORRESPONDENCE
Tengfei Ji,
243477587@qq.com
Tiansheng Cao,
caotiansheng2088@sina.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Pharmacology of Anti-Cancer Drugs, a section of the journal Frontiers in Pharmacology

RECEIVED 07 November 2022
ACCEPTED 21 March 2023
PUBLISHED 17 April 2023

CITATION
Zhang S, Yang J, Wu H, Cao T and Ji T (2023), Establishment of a 7-gene prognostic signature based on oxidative stress genes for predicting chemotherapy resistance in pancreatic cancer. *Front. Pharmacol.* 14:1091378. doi: 10.3389/fphar.2023.1091378

Establishment of a 7-gene prognostic signature based on oxidative stress genes for predicting chemotherapy resistance in pancreatic cancer

Shengmin Zhang†, Jianrong Yang†, Hongsheng Wu, Tiansheng Cao* and Tengfei Ji*

Department of Hepatobiliary Surgery, Affiliated Huadu Hospital, Huadu People's Hospital, Guangzhou, Guangdong, China

Background: Oxidative stress is involved in regulating various biological processes in human cancers. However, the effect of oxidative stress on pancreatic adenocarcinoma (PAAD) remained unclear.

Methods: Pancreatic cancer expression profiles from TCGA were downloaded. Consensus ClusterPlus helped classify molecular subtypes based on PAAD prognosis-associated oxidative stress genes. Limma package filtered differentially expressed genes (DEGs) between subtypes. A multi-gene risk model was developed using Lease absolute shrinkage and selection operator (Lasso)-Cox analysis. A nomogram was built based on risk score and distinct clinical features.

Results: Consistent clustering identified 3 stable molecular subtypes (C1, C2, C3) based on oxidative stress-associated genes. Particularly, C3 had the optimal prognosis with the greatest mutation frequency, activate cell cycle pathway in an immunosuppressed status. Lasso and univariate cox regression analysis selected 7 oxidative stress phenotype-associated key genes, based on which we constructed a robust prognostic risk model independent of clinicopathological features with stable predictive performance in independent datasets. High-risk group was found to be more sensitive to small molecule chemotherapeutic drugs including Gemcitabine, Cisplatin, Erlotinib and Dasatinib. The 6 of 7 genes expressions were significantly associated with methylation. Survival prediction and prognostic model was further improved through a decision tree model by combining clinicopathological features with RiskScore.

Conclusion: The risk model containing seven oxidative stress-related genes may have a greater potential to assist clinical treatment decision-making and prognosis determination.

KEYWORDS

pancreatic cancer, oxidative stress, methylation, molecular subtypes, risk score, small molecule chemotherapeutic drugs, prognosis, tumor immunity

Introduction

Pancreatic adenocarcinoma (PAAD) is one of the most difficult malignancies to treat (Katona et al., 2021), with gallstones, chronic pancreatitis, smoking, alcohol drinking as the most common risk factors for PAAD (Lowenfels et al., 1993). Ductal adenocarcinoma of the pancreas is the predominant histopathological type accounting for 85% of all the PAAD cases. Surgical resection is not available to proximately 80%–85% of patients due to a lack of typical manifestations at the initial stage (Singhi et al., 2019). For those patients with PAAD who have taken surgery, 5-year overall survival probability is only about 20% (Wu et al., 2019). The technology of genome sequencing has further characterized the molecular patterns and genotypic heterogeneity of pancreatic cancer. Given that molecular targeting therapies have become indispensable in treatment, discovering new therapeutic targets is crucial. Hence, for improving the prognostic prediction of PAAD, it is imperative to identify novel prognostic indicators.

Oxidative stress functions importantly in pathogenesis of multiple diseases, including inflammatory diseases, cancer, and immune-mediated (Azmanova and Pitto-Barry, 2022). Oxidative stress induces reactive oxygen species (ROS) that could damage lipids, proteins, DNA, and produce mutagenic metabolites to affect tumor biological behaviors and transform malignant phenotype (Sosa et al., 2013). Tumor microenvironment consists of surrounding tissue components and interacting tumor cells, with the latter favoring biological behaviors of tumor cells. ROS has a complex and multifaceted role in tumor microenvironment. A study found that the non-classical glutamine pathway promotes the development of pancreatic cancer with dysregulation of oxidative stress (Son et al., 2013). ROS inhibits the arginine methylation enzyme CARM1, which in turn inhibits MDH1 activity. Thus, ROS could activate non-classical glutamine metabolism to promote pancreatic cancer cell growth (Son et al., 2013). Glutamine and asparagine are two key nutrients affecting pancreatic cancer cell development, moreover, these two are one of the bases of protein synthesis in pancreatic cancer cells to promote resistance to oxidative stress and are essential for pancreatic cancer cell growth and proliferation. Pathria et al. showed that simultaneous inhibition of asparagine metabolic pathway and MAPK pathway inhibited pancreatic cancer development (Pathria et al., 2019). Methionine residues has been found to serve as a reversible redox switch in controlling different signaling outcomes. To control tumor metastasis, MSRA-PKM2 axis is a regulatory bridge between cancer metabolism and redox biology (He et al., 2022). Therefore, future studies on the role of oxidative stress in PAAD and the impact on TME are needed to optimize immunotherapy or develop new therapeutic strategies.

Consistent clustering screened stable molecular subtypes utilizing genes of oxidative stress pathway. We also compared immune features, mutational, clinical pathway features among the subtypes. Finally, we identified genes associated with oxidative stress phenotypes using differential expression analysis and LASSO. Moreover, a risk model and clinical prognostic model was developed for facilitating personalized PAAD treatment.

Materials and methods

Data collection and processing

We used TCGA GDC API to download the mutation data and RNA-seq data [transcripts per million (TPM)] of TCGA-PAAD. A total of 176 primary tumor samples were finally obtained after screening. We downloaded transcriptomic data of samples from the pancreatic cancer-Australia (PACA-AU) and pancreatic cancer-Canada (PACA-CA) cohorts in the International Cancer Genome Consortium (ICGC) database (https://dcc.icgc.org/projects), with each cohort containing 267 and 215 pancreatic cancer samples, respectively. Oxidative stress-related genes were obtained e oxidative stress pathway "GOBP_RESPONSE_TO_OXIDATIVE_STRESS" in MSigDB database.

Data pre-processing

The RNA-seq data from TCGA were preprocessed as follows.

1) Removing samples that did not contain clinical information of follow-up;
2) Removing samples that did not show survival time;
3) Removing samples that did not show status;
4) Conversion of Ensembl to Gene symbol;
5) Mean value taken for expression in the cases of multiple Gene Symbols.

Molecular subtyping of oxidative stress-related genes

Clustering and subtyping of the samples were achieved using ConsensusClusterPlus (Wilkerson and Hayes, 2010). To obtain molecular subtypes, expression of cellular senescence-correlated genes were utilized. "KM" algorithm and "1—Pearson correlation" was the metric distance in performing 500 bootstraps. Each bootstrapping contained 80% training set patients. Cluster number was from 2 to 10. Molecular subtypes as well as the optimal classification were obtained through calculation of consistency matrix and consistency cumulative distribution function.

Risk model

1) Among subtypes, differentially expressed genes (DEGs) were identified by the previously identified molecular subtypes, and we used the Limma package to calculate genes differentially expressed between C1 vs. Other, C2 vs. Other and C3 vs. Other in the TCGA-PAAD cohort (Ritchie et al., 2015).
2) Selection of differentially expressed genes of prognostic significance ($p < 0.01$).

3) Furthermore, genes were reduced by lasso regression (Tibshirani, 1997) to obtain prognostically significant genes associated with the oxidative stress phenotype.
4) Risk modeling, the formula RiskScore = Σβi × Expi, where Expi is gene expression of the prognostic-related gene with features of the oxidative stress phenotype, and β is corresponding gene lasso cox regression coefficient, was used to calculate the risk score for each patient. Then zscore was performed, and patient classification into low- and high- RiskScore groups was conducted under the threshold "0". To draw raw curves for prognostic analysis, KM method was used and the log-rank test determined difference significance.

Gene set enrichment analysis (GSEA)

To explore pathways of different biological processes, we used "GSEA" based on all the candidate gene sets in Hallmark database for pathway analysis in different subtypes (Liberzon et al., 2015). Significant enrichment was when false discovery rate (FDR) < 0.05. Ferroptosis pathway were from "WP_FERROPTOSIS" in MSigDB database; autophagy pathway were from "GOBP_REGULATION_OF_AUTOPHAGY" in MSigDB database; from Liu et al. (Liu et al., 2020), we obtained inflammatory signature-related gene set; angiogenesis-related gene set were from Masiero et al. (Masiero et al., 2013).

Protein-protein interaction (PPI) analysis

PPI networks were produced. The DEGS between subtypes were entered into the STRING online tool (https://string-db.org/), and in Cytoscape (version 3.9.1) software visualization of the PPI networks were done. Next, module analysis of the PPI networks was performed using the Molecular Complex Detection (MCODE) tool of Cytoscape software (Bader and Hogue, 2003).

Calculation of TME cell invasion

In PAAD, CIBERSORT algorithm (https://cibersort.stanford.edu/) was introduced to quantify relative abundance of 22 immune cells (Newman et al., 2015). ESTIMATE software was applied for the calculation of immune cells proportion, followed by comparison of immune cell infiltration using Wilcoxon test (Runa et al., 2017).

Correlation analysis of risk score and drug sensitivity

Drug sensitivity data of about 1,000 cancer cell lines were retrieved from Genomics of Drug Sensitivity in Cancer (GDSC) (http://www.cancerrxgene.org) (Yang et al., 2013). Area under ROC curve (AUC) for each antitumor drug served as an indicator for drug response in cancer cell lines. To calculate the association of AMrs scores with drug sensitivity, Spearman correlation analysis was carried out, we considered | Rs |> 0.1. FDR was adjusted by Benjamini and Hochberg, a significant correlation was defined when FDR was less than 0.01.

Results

Molecular subtyping based on oxidative stress-associated genes

The expression pattern of oxidative stress-related genes pancreatic cancer samples in the TCGA-PAAD and PACA-AU datasets with clinical information was determined *via* univariate Cox regression. A total of 27 oxidative stress genes showing significant prognosis in both pancreatic cancer datasets were screened. Univariate cox analysis of these 27 genes in TCGA-PAAD and PACA-AU filtered 19"risk genes" and 8"protective genes" (Figures 1A,B). Next, we classified patients by consistent clustering based on 27 prognostically significant oxidative stress gene expression profiles, and according to the cumulative distribution function (CDF), determined the optimal number of clusters. From CDF Delta area curve, we could see that the Cluster selection of 3 had more stable clustering results (Figures 1C,D), and three molecular subtypes (C1, C2, C3) were categorized under k = 3 (Figure 1E). Furthermore, we analyzed their prognostic characteristics and significant differences in prognosis (Figure 1F). Generally, the prognosis of C3 was better in contrast to a worse prognosis of C1. Also, this result was validated in the PACA-AU cohort (Figure 1G).

Analysis on the "oxidative stress ssGSEA scores" for each pancreatic cancer patient in the TCGA-PAAD cohort showed that the C1 subtype had higher "oxidative stress ssGSEA scores" and it was the lowest in C3 (Figure 1H), noticeably, C1 presented activated oxidative stress. We also compared the expression differences of 27 oxidative stress genes in different molecular subtypes (Figure 1J). C1 subtype showed an overall high-expressed "Risk" genes, while in the C3 subtype, the "Protective" gene was high-expressed. This phenomenon was also observed in the PACA-AU cohort (Figures 1I,K).

Genomic landscape between molecular subtypes

To further investigate the potential molecular mechanisms underlying the classification of oxidative stress subtypes, we explored genomic alteration differences among these three TCGA cohort molecular subtypes. Here, information of molecular signature of TCGA-PAAD was acquired from a previous pan-cancer study (Thorsson et al., 2018). The Anenploidy Score, loss of heterogeneity (LOH), tumor mutation burden (TMB), Homologous Recombination Defects all differed greatly among the three subtypes. It has been observed that the C1 subtype had higher levels of these four indicators (Figure 2A). In addition, in this study, according to 160 different immune signatures, five molecular subtypes of PADD were categorized, among which the most favorable prognosis was the immunoassay subtype C3. Then, comparison of the current molecular subtypes were compared with the five immune molecular subtypes showed that our

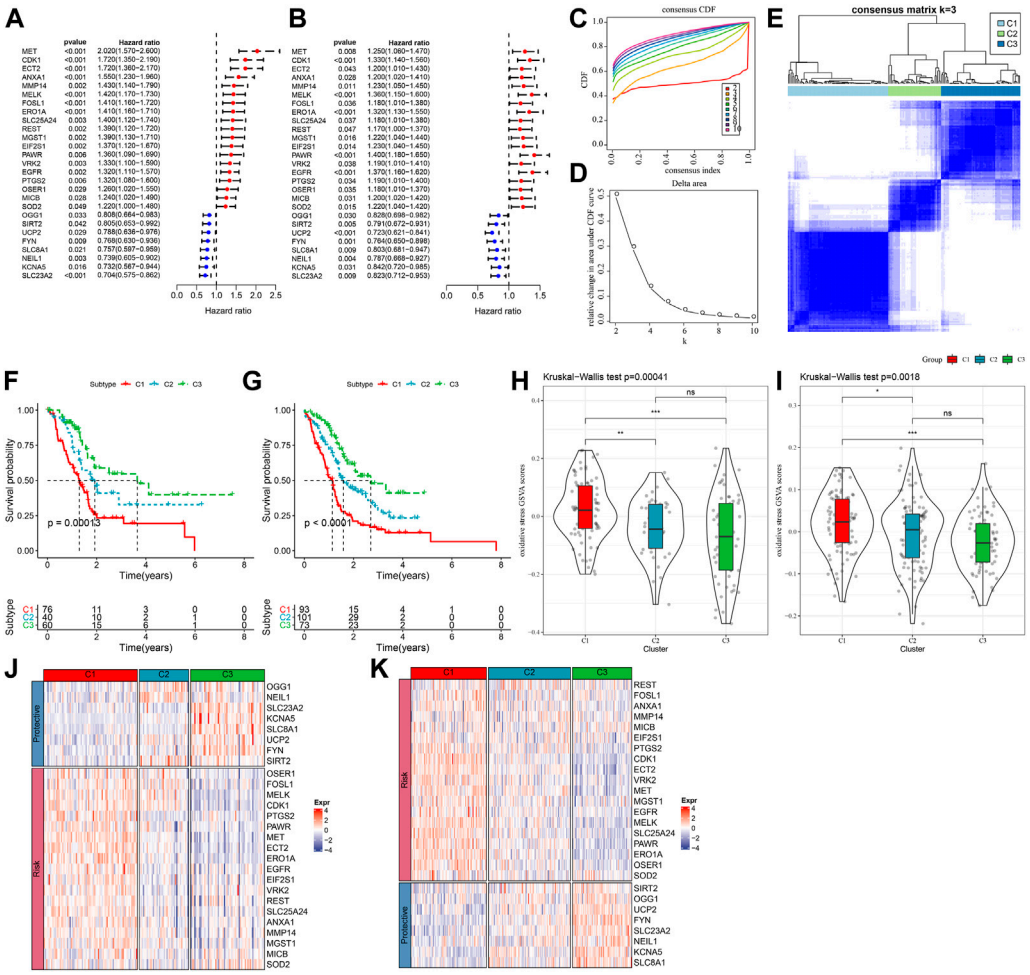


FIGURE 1
Three pancreatic cancer molecular subtypes based on oxidative stress-related genes. **(A)** In the TCGA-PAAD cohort, the forest plot of 27 prognostically significant oxidative stress genes; **(B)** The forest plot of 27 prognostically significant oxidative stress genes in the PACA-AU cohort; **(C)** CDF curves of TCGA-PAAD samples; **(D)** CDF Delta area curves of the samples, with the horizontal axis indicating the number of categories k and the vertical axis indicating the relative change in area under the CDF curve; **(E)** At consensus k = 3, the heat map of clustered samples; **(F)** KM curve of the prognosis of three subtypes of the TCGA-PAAD samples; **(G)** KM curve of the prognosis of three subtypes of the PACA-AU samples; **(H)** Differences in "oxidative stress ssGSEA scores" among the TCGA-PAAD molecular subtypes; **(I)**Differences in "oxidative stress ssGSEA scores" among the PACA-AU molecular subtypes; **(J)** Heat map of prognostic significant oxidative stress-related genes in TCGA-PAAD subtypes; **(K)** Heat map of prognostic significant oxidative stress-related genes in PACA-AU subtypes.

C3 subtype was more occupied by the immune molecular subtype C3, which coincided with the most favorable prognosis of our molecular subtype C3 (Figure 2B). The top significant 20 genes were shown (Figure 2C). It could be seen that genes such as KRAS and TP53 had significantly different mutation frequencies between the three molecular subtypes.

Immune characteristics between molecular subtypes and differences in immunotherapy/chemotherapy

Between different molecular subtypes, differences in the PAAD immune microenvironment was further explored by assessing immune cell infiltration in patients in TCGA-PAAD and PACA-AU cohorts based on gene expression in the immune cells. Relative abundance of 22 immune cell types was determined using CIBERSORT, and in the TCGA-PAAD cohort we found that six immune cell types (Mcrophages, CD8 T cells, naive B cells, Monocytes, memory CD4 T cells, regulatory T cells, Macrophages M0) differed significantly between subtypes, and T_cells_CD8 and Monocytes were enriched in the C3 subtype (Figure 3A). Immune cell infiltration was assessed using ESTIMATE. In the TCGA cohort, the three subtypes differed significantly in distribution in the StromalScore, ImmuneScore and ESTIMATEScore, and the highest score was in the C3 subtype but the lowest was in the C1 subtype (Figure 3B). We also analyzed the PACA-AU data set and found that eight immune cell types, including resting memory CD4 T cells, M0 Macrophages, naïve CD4 T cells, M1 Macrophages, CD8 T cells, helper follicular T cells, Monocytes, Neutrophils, in the PACA-AU cohort differed significantly between subtypes

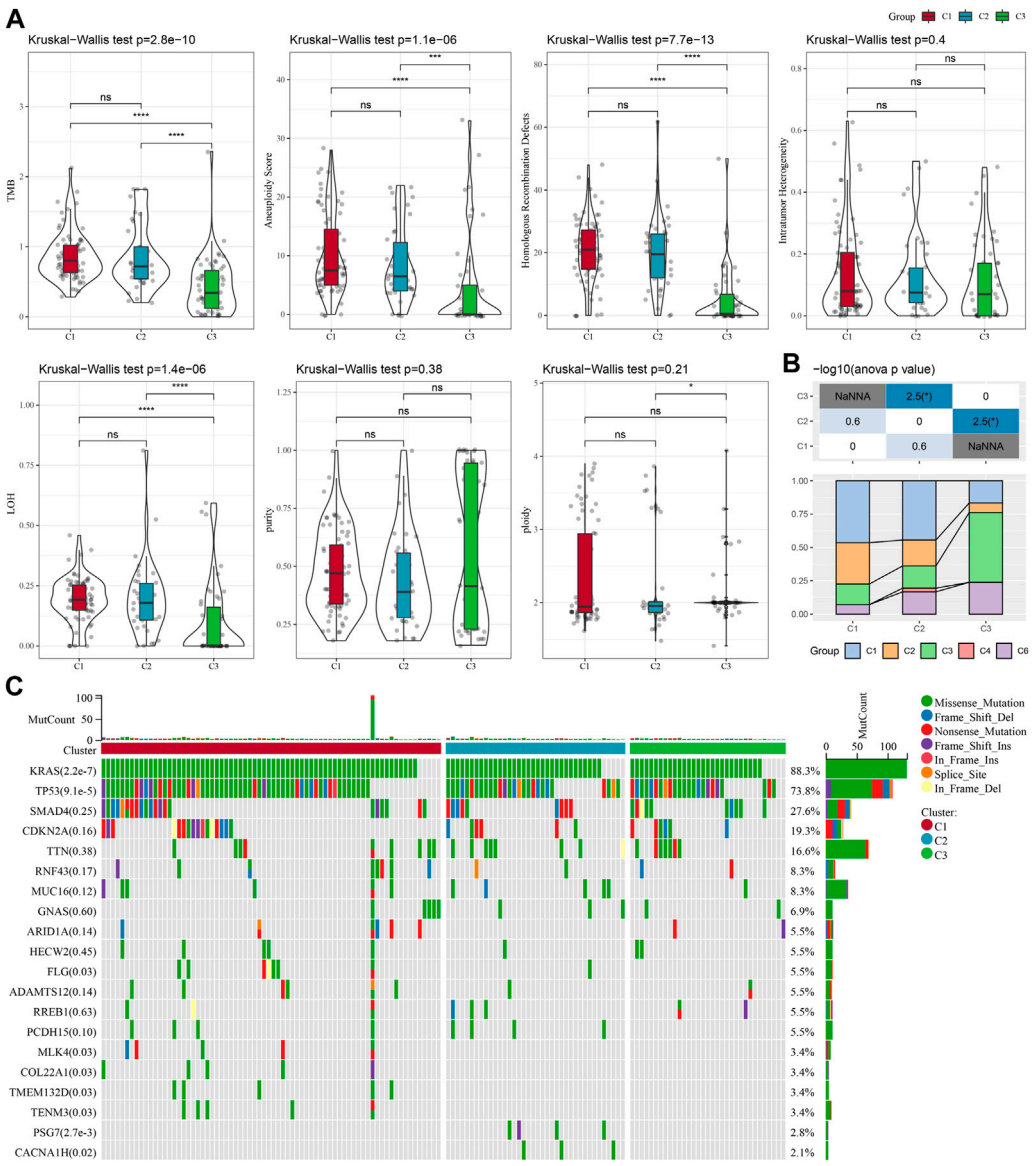


FIGURE 2
In TCGA-PAAD cohort genomic alterations of molecular subtypes. **(A)** Comparison of Aneuploidy Score, LOH, tumor 25 burden, Intratumor Heterogeneity, ploidy, Homologous Recombination Defects, purity in TCGA-PAAD subtypes; **(B)** Comparing our molecular subtypes to the other six existing immune molecular subtypes; **(C)** Chi-square test on the somatic mutations in the three molecular subtypes. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; and ****$p < 0.0001$.

(Figure 3C). Furthermore, immune cell infiltration in the PACA-AU cohort was consistent with the TCGA-PAAD cohort (Figure 3D).

Some studies have reported that oxidative stress and inflammation are intertwined processes in disease progression and response to therapy by interfering with multiple signaling pathways. Here, the enrichment scores of seven metagenes clusters were greatly different in the three molecular subtypes, with the exception of Interferon, STAT1, and the remaining five metagenes clusters, and overall, the C3 subtype had higher inflammatory activity (Figure 3E). In addition, it has been reported that ferroptosis from oxidative stress and inflammation plays a key role in the pathogenesis of cardiovascular diseases (Yu et al., 2021), such as stroke, vascular sclerosis, heart failure, ischemia-reperfusion injury. Thus, comparison on the differences in ferroptosis scores between the three subtypes has demonstrated significant distributional differences between C1 and C2 subtypes, with C2 subtype having a lower ferroptosis score (Figure 3F). In addition, a study reported that inflammation stimulates excessive autophagy or severe oxidative stress could result in autophagy-dependent cell death (Cai et al., 2018). The autophagy scores for the subtypes (Figure 3G) were significantly different between the three subtypes, with the C3 subtype having a higher autophagy score. Additionally, we found statistically significant differences in angiogenesis between C2 and C3 subtypes and between C1 and

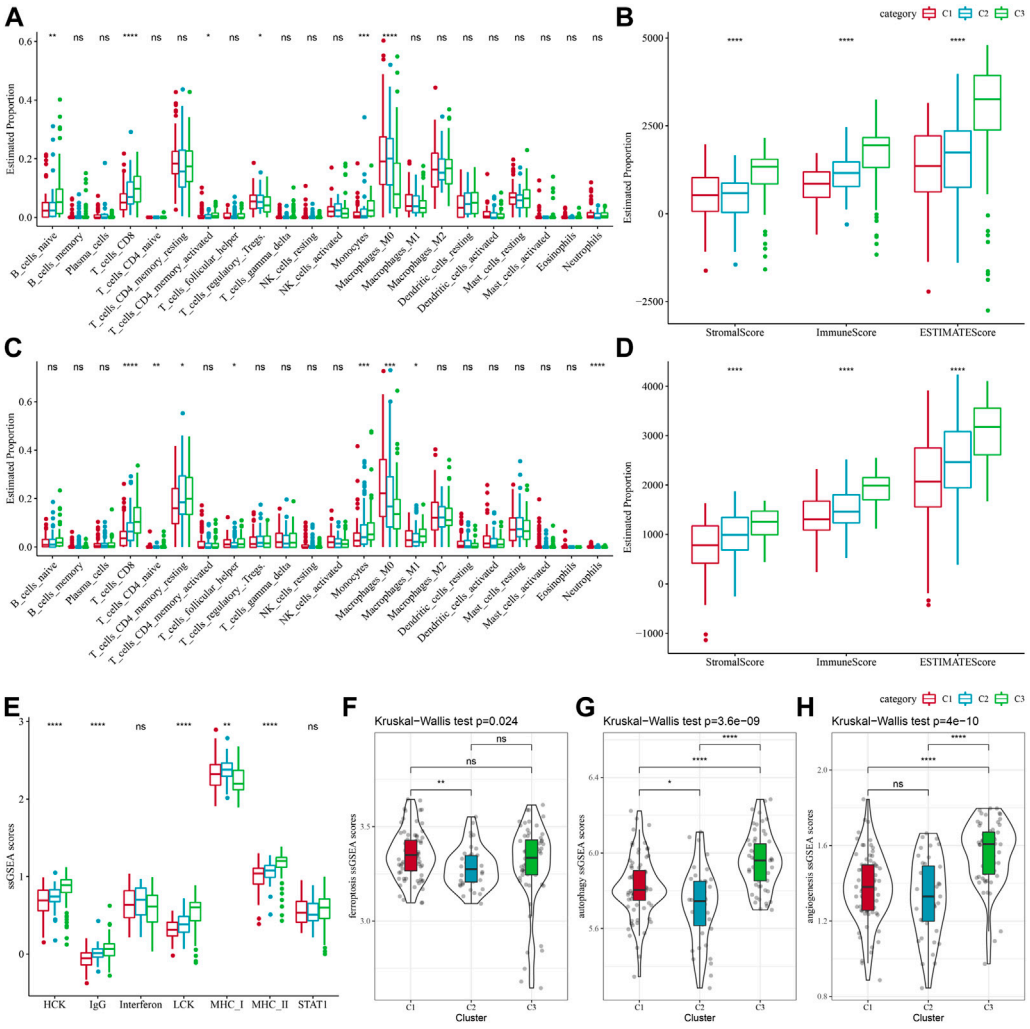


FIGURE 3
Immune characteristics of different subtypes. **(A)** TCGA-PAAD molecular subtypes varied in the differences of 22 immune cell scores; **(B)** TCGA-PAAD molecular subtypes varied in the differences of ESTIMATE immune infiltration; **(C)** PACA-AU molecular subtypes varied in the differences of 22 immune cell scores; **(D)** PACA-AU molecular subtypes varied in the differences of ESTIMATE in immune infiltration; **(E)** TCGA-PAAD molecular subtypes varied in the differences of scores of seven inflammation-related gene clusters; **(F)** TCGA-PAAD molecular subtypes varied in the differences of ferroptosis pathway; **(G)** TCGA-PAAD molecular subtypes varied in the differences in scores of autophagy pathway; **(H)** TCGA-PAAD molecular subtypes varied in the differences in scores of angiogenesis-related genes; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; and ****$p < 0.0001$.

C3 subtypes, with C3 subtype having the highest score (Figure 3H).

Immunotherapy and drug sensitivity differences between molecular subtypes

Some sample compounds were examined because immune checkpoint blockade (ICB) cancer treatment works by suppressing important immune checkpoints. Among the three subtypes, CTLA4 and PD-1 were differentially distributed, and C3 was significantly more high-expressed, while PD-L1 was not differentially expressed (Figure 4A). We also applied the "T-cell-inflamed GEP score" to assess the predictive potential of different molecular subtypes in immunotherapy for cancers. It could be observed from Figure 4B, the C3 subtype had a noticeably higher "T-cell-inflamed GEP score". Considering that IFN-γ is a cytokine with a key role in anti-cancer immunity and immunomodulation (Rydyznski Moderbacher et al., 2022), our analysis revealed that in the C3 subtype the IFN-γ response was significantly enhanced (Figure 4C). Additionally, we also found that Cytolytic activity (CYT) scores, which reflects cytotoxic effects, were significantly higher in C3 subtypes compared with other subtypes (Figure 4D).

Additionally, response of the molecular subtypes in the TCGA-PAAD cohort to the conventional chemotherapeutic agents such as Gemcitabine, Erlotinib, Cisplatin, 5-Fluorouracil were analyzed, and found that C1 was more

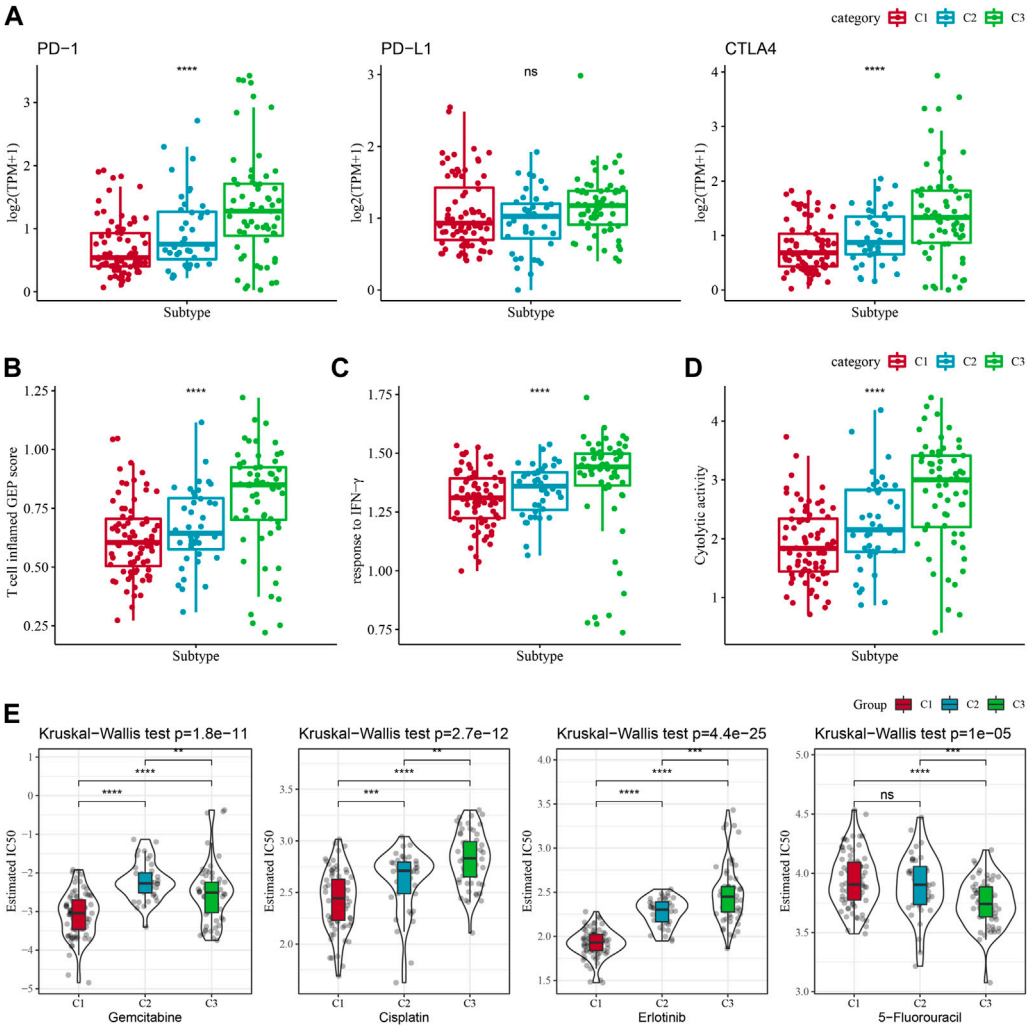


FIGURE 4
Differences in treatment sensitivity among the molecular subtypes. Among different molecular subtypes, **(A)** Differences in "T cell inflamed GEP score"; **(B)** differences in "response to IFN-γ"; **(C)** Differences in "response to IFN-γ". **(B)** Differences in "response to IFN-γ"; **(C)** Differences in immune checkpoint gene expression; **(D)** Differences in "Cytolytic activity"; **(E)** Box plots of IC50 of cisplatin, gemcitabine, 5-fluorouracil, erlotinib in TCGA-PAAD; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; and ****$p < 0.0001$.

sensitive to Gemcitabine, Erlotinib, Cisplatin (Figure 4E), while C3 was more sensitive to 5-Fluorouracil.

Differential functional analysis between the molecular subtypes

Limma package was used to determine DEGs. There is 64 DEGs in C1vs. other in TCGA-PAAD dataset and PACA-AU dataset, 137 DEGs in C3 vs. other in TCGA-PAAD dataset and PACA-AU dataset. After union analysis, 144 DEGs were obtained. Functional enrichment analysis was conducted on the DEGs among the subtypes. The enrichment results of GO and KEGG pathways on the DEGs in the "C1" subtype demonstrated that the DEGs had significant enrichment in some biological functions than cellular communication (Supplementary Figure S1A). However, in "C3" subtype these DEGs were significantly enriched to some immune-related biological pathways (Supplementary Figure S1B). To better investigate the interactions among these DEGs, the STRING online tool for developing a PPI network (Supplementary Figure S1C) was applied. In addition, two important modules in the PPI network were determined based on the module analysis (Supplementary Figure S1D).

Identification of key genes for the oxidative stress phenotype

Next, we performed univariate COX regression analysis on 144 DEGs among the subtypes and identified 61 genes showing great prognostic significance ($p < 0.01$), including 31"Risk" and

10.3389/fphar.2023.1091378

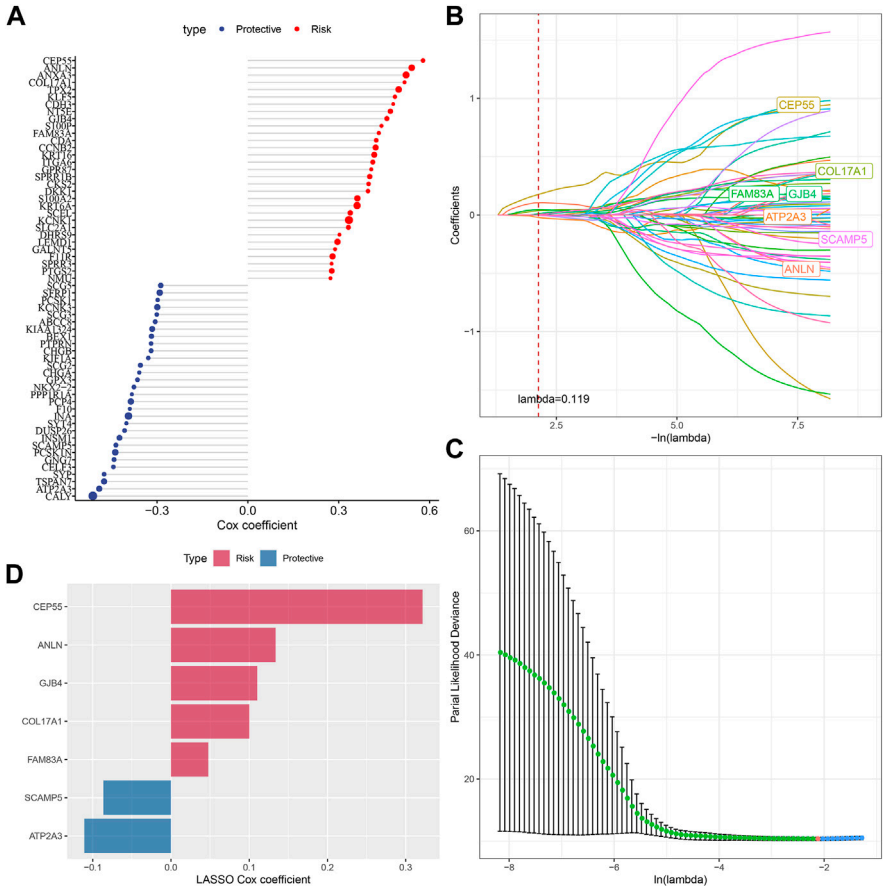


FIGURE 5
Screening of genes correlated with oxidative stress that affect prognosis. **(A)** Totally 61 candidates were screened from all the DEGs; **(B)** With the change of lambda, trajectory of each independent variable was shown; **(C)** Confidence interval under lambda; **(D)** Oxidative stress-related prognostic gene markers and the distribution of LASSO coefficients.

30"Protective" genes (Figure 5A). PPI network analysis indicated that these genes are related to each other (Supplementary Figure S2). To further compress these 61 genes in the risk model, Lasso regression was used. Independent variable's trajectory is shown in Figure 5B. The number of independent variable coefficients close to zero likewise showed a progressive increase as the lambda gradually increased. Moreover, 10-fold cross-validation was applied to develop a model and to analyze confidence intervals under each lambda (Figure 5C). When lambda = 0.119, the model was optimized, therefore, 7 genes at lambda = 0.119 were determined in this study as the genes related to oxidative stress to affect patients prognosis (Figure 5D). These genes included ATP2A3, ANLN, GJB4, FAM83A, CEP55, COL17A1, and SCAMP5. The formula as followed: RiskScore = + 0.134*ANLN-0.086*SCAMP5+0.048*FAM83A-0.111*ATP2A3+0.322*CEP55 + 0.11*GJB4+0.1*COL17A1 Single-cell division TISCH2 (http://tisch.comp-genomics.org/home/) analyzed the expression distribution of seven genes in multiple single-cell data of pancreatic cancer, and the results showed that the expression of COL17A1 and FAM83A genes in malignant cells was significantly higher than that in other cell types (Supplementary Figure S3).

The performance and validation of prognostic model

The expression and coefficients of seven prognostic genes were used to construct a clinical prognostic model and for calculating and ranking the risk values of TCGA-PAAD samples. According to the cut-off, we divided 81 samples into "Low-risk" group and 95 samples were in the "High-risk" group. The prognosis prediction at 1, 2, and 3 years (s) was further analyzed for its classification efficiency (Figure 6B), respectively. The model demonstrated a high area under the AUC line (1-Year, AUC = 0.73; 2-Year, AUC = 0.75; 3-Year, AUC = 0.79). KM curves were plotted and a highly significant difference was shown between the two RiskScore groups ($p < 0.0001$), with the "Low-risk" group showing a significantly better prognostic outcome than "High-risk" group (Figure 6C). To confirm the robustness of the clinical prognostic model, we performed validation in 2 additional independent pancreatic cancer cohorts (PACA-AU, PACA-CA), and it can be seen that in the validation cohort showed similar results to the training set, with the "Low-risk" group showing a significantly better prognostic outcome than "High-risk" group (Figures 6D–G).

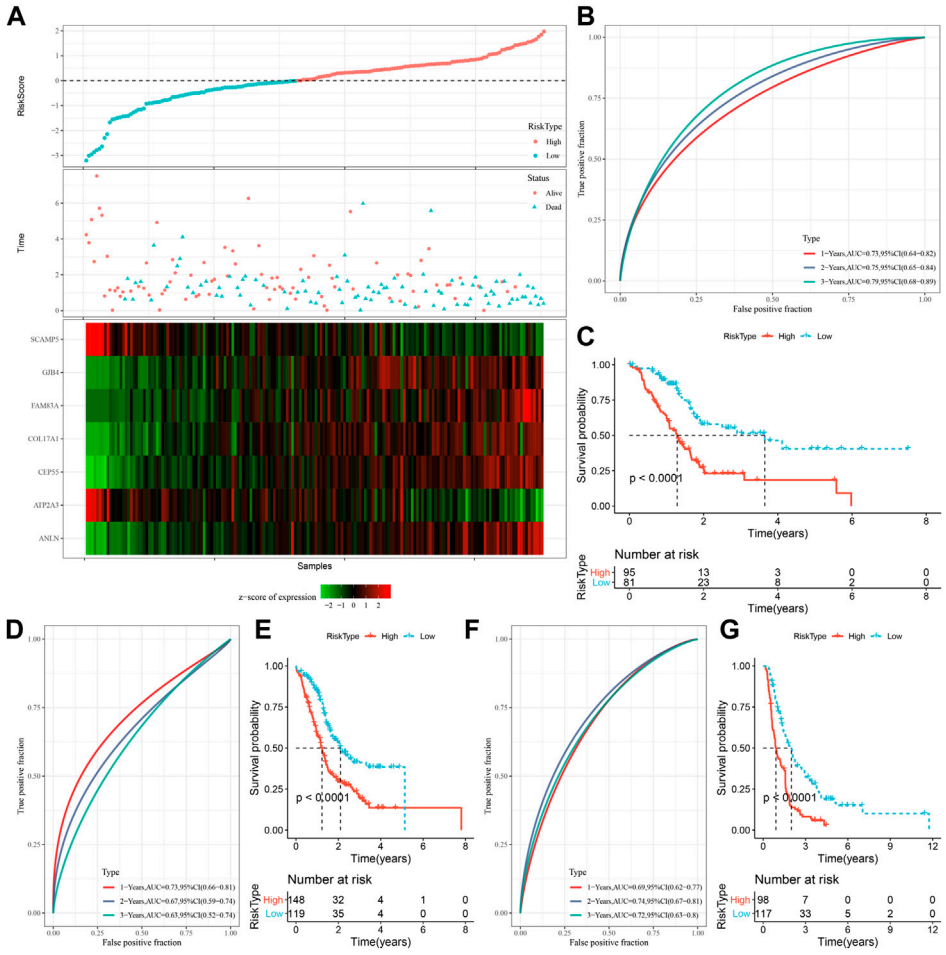


FIGURE 6
Generation and evaluation of risk score models using 7 genes related to oxidative stress. **(A)** RiskScore, expression of oxidative stress-related prognostic genes and survival time and status in TCGA dataset; **(B)** RiskScore classification in TCGA dataset and ROC and AUC curves; **(C)** Distribution of KM survival curve of RiskScore in TCGA dataset; **(D, E)** ROC curve and KM survival curve distribution of RiskScore in PACA-AU cohort; **(F, G)** ROC curves and KM survival curves of RiskScore in the PACA-CA cohort.

The RiskScore on the subtypes and various clinicopathological features

To assess the correlation of PAAD clinical features with RiskScore, the differences in RiskScore between different TNM grades and clinical stages in the TCGA-PAAD and PACA-AU datasets were studied. Samples with higher clinical grades showed higher RiskScore. Also, C1 subtypes had the highest RiskScore but C3 subtypes had the lowest RiskScore (Supplementary Figures S4A,C). Additionally, difference comparison between RiskScore groups and molecular subtypes was conducted, showing a majority of "C1" or "C2" patients in the "high-risk" group (Supplementary Figures S4B,D). Moreover, the prognosis of TCGA-PAAD between the low- and high-risk groups in relation to clinicopathological characteristics was explored, and our risk grouping was equally effective across clinical subgroups, with the "Low-risk" group showing a significantly better prognosis, demonstrating the reliability of our risk grouping (Supplementary Figures S4E). In addition, the correlation analysis of spearman between expression levels and methylation on 6 genes except ANLN gene showed a negative phenomenon but had a positive with SCAMP5 gene (Supplementary Figures S5).

Immune infiltration/pathway characteristics between RiskScore subgroups

Differences in immune microenvironment in the RiskScore subgroups were studied, we used ESTIMATE to assess immune cell infiltration (Figure 7A), and observed that the "Low-risk" group was significantly higher in immune cell infiltration. The most significant top 10 pathways showing differences between the high- and low-risk groups are shown in Figure 7B. It can be seen that high-RiskScore was significantly enriched to some cell cycle-related pathways such as G2M_CHECKPOINT, MTOTIC

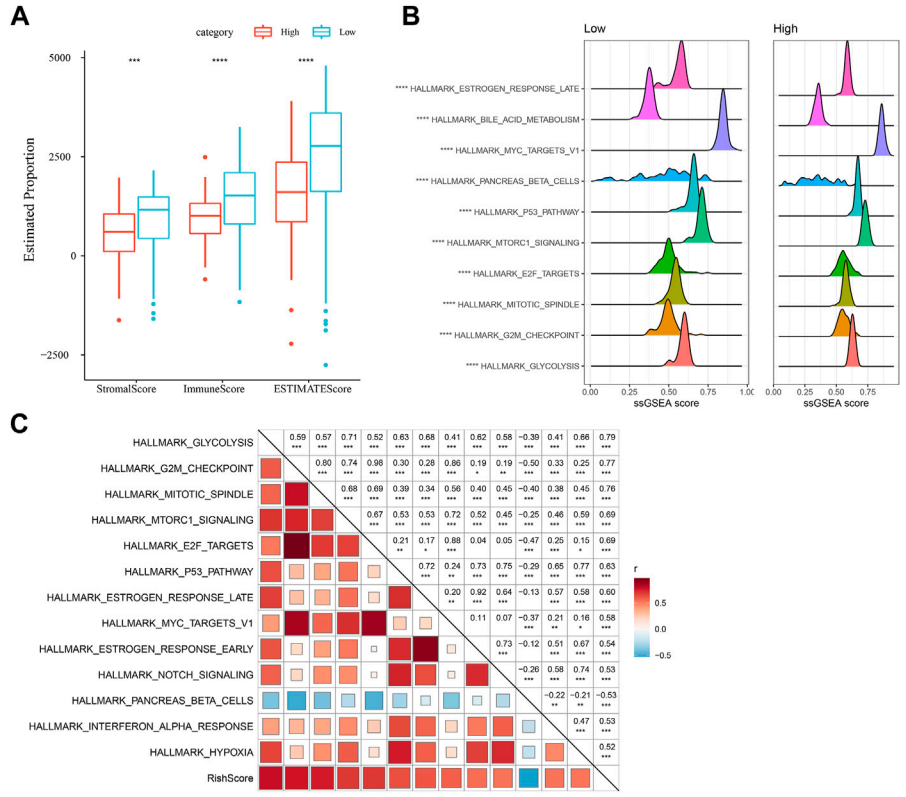


FIGURE 7
Immunology and pathway between different RiskScore subgroups. **(A)** ESTIMATE software was applied to determine immune cell components in the TCGA database; **(B)** The top 10 pathways showing the greatest significant differences between Low-risk and high-risk groups; **(C)** Correlation analysis results on the RiskScore and KEGG pathways scored greater than 0.5; *$p < 0.05$; **$p < 0.01$; * **$p < 0.001$; and ****$p < 0.0001$.

SIGNSLING, and E2F TARGETS. Furthermore, association of the RiskScore with the enrichment scores of these functions was analyzed, with the functions showing a correlation greater than 0.5 being identified. Figure 7C manifests a positive correlation of the RiskScore with cell cycle-related pathways.

Differences in chemotherapy/immunotherapy among the RiskScore subgroups

First, we used the "T-cell-inflamed GEP score" to assess the prediction potential of the different RiskScore subgroups in cancer immunotherapy (Figure 8A), and the results showed that in the low-RiskScore group the "T-cell-inflamed GEP score" was significantly higher. Further analysis on the response to IFN-γ in both groups revealed that the response to IFN-γ was significantly enhanced in the low-RiskScore group (Figure 8B). Moreover, the CYT score had cytotoxic effect, and it was significantly lower in the high-RiskScore group (Figure 8C). Some representative immune checkpoint molecules were significantly high-expressed CTLA4, PD-1 in the low-RiskScore group, while PD-L1 was not differentially expressed between molecular subtypes (Figure 8D).

The effect of RiskScore on drug response was analyzed based on the relationship between RiskScore and cancer cell lines' response to drugs. There were 11 drug sensitivities in the GDSC database showing significant correlation with RiskScore. There were three drug sensitivities showing a negative correlation with RiskScore, namely, KU-55933, Tozasertib, Dasatinib (Figure 8E). Furthermore, we also analyzed the signaling pathways of the genes targeted by these drugs, which mainly target the SRC pathway (Figure 8F).

Moreover, response degree of the TCGA-PAAD subtypes to chemotherapeutic agents (Gemcitabine, Erlotinib, Cisplatin, 5-Fluorouracil, and Dasatinib) was studied. We found that the High-risk group responded to Gemcitabine, Cisplatin, Erlotinib and Dasatinib. Overall, High-risk group showed a higher sensitivity to Gemcitabine, Cisplatin, Erlotinib and Dasatinib. Low-risk group had higher sensitivity to 5-Fluorouracil (Figure 8G).

RiskScore in combination with clinicopathological features to improve survival prediction and the prognostic models

For the TCGA-PAAD cohort (Figure 9A), only RiskType, and Age, N Stage remained in the decision tree that was originally

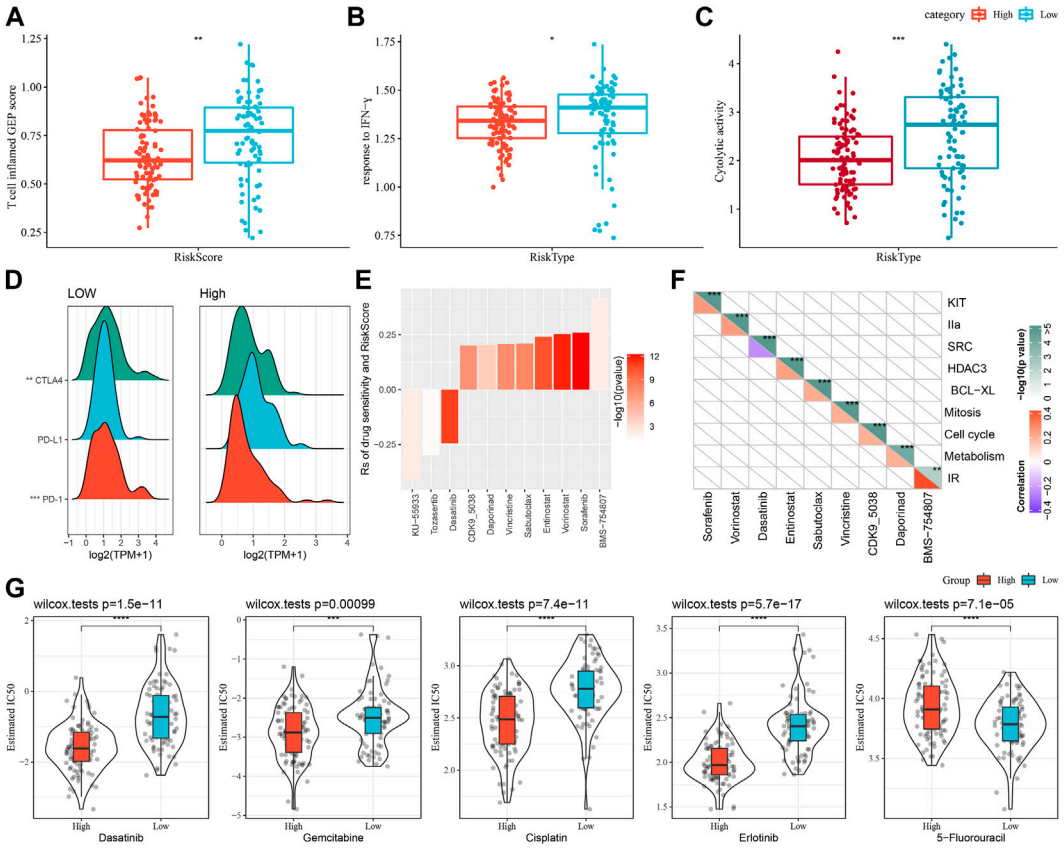


FIGURE 8
The prognostic risk models in predicting patients' benefit from immunization/chemotherapy. **(A)** Differences in "T cell inflamed GEP score" between subgroups; **(B)** Differences in "response to IFN-γ" between subgroups; **(C)** Differences in "C-γ" between subgroups. **(C)** Difference of "Cytolytic activity" between different subgroups; **(D)** Difference of immune checkpoint gene expression between different subgroups; **(E)** Spearman analysis was conducted for correlation analysis on drug sensitivity and RiskScore, with each column representing a type of drug. Correlation significance is reflected in color brightness. The correlation of RiskScore with drug sensitivity (Rs < 0) or drug resistance (Rs > 0) was reflected in the height of a column. **(F)** The horizontal axis is the drug name and the vertical axis is the signaling pathway targeted by the drug. The signaling pathway targeted by the drug is sensitive to RiskScore (blue); **(G)** Box plots of IC50 estimates for dasatinib, gemcitabine, cisplatin, erlotinib and 5-fluorouracil in TCGA-PAAD; *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; and ****$p < 0.0001$.

established with TNM Stage, gender, pathology information and RiskScore, patient age, and it identified four different risk subgroups (Lowest, Low, Mediate, High). Among them RiskType was the parameter of the greatest impact. The four risk subgroups showed significant difference in overall survival, with the "Lowest" group having the optimal prognosis and the "High" group having the worst prognosis (Figure 9B). Patients in the risk subgroups were all "Low","Lowest", and "Mediate" Low-risk group patients (Figure 9C). In addition, "High-risk" group showed more distribution of our defined molecular subtypes C1 and C2 (Figure 9D). From Figures 9E,F, the most significant prognostic factor was the RiskScore. To quantify survival probability of PAAD patients and the risk assessment, other clinicopathological characteristics were combined with RiskScore for nomogram development (Figure 9G), here, the RiskScore demonstrated the greatest impact on predicting patients' survival. Furthermore, we evaluated the model prediction accuracy with calibration curve (Figure 9H). At the calibration points of 1, 2, 3 years (s), the prediction calibration curves almost completely encircled the standard curve, indicating good prediction accuracy. We also assessed model reliability by decision curve analysis, and in comparison to the extreme curves, benefit of both RiskScore and Nomogram was significantly greater and the two showed a stronger survival prediction (Figure 9I).

Herein, we selected three risk models from previous studies (5-gene signature (Yan) (Yan et al., 2022), 3-gene signature (Yang) (Yang et al., 2022) and 9-gene signature (Liu) (Liu et al., 2022)) to compare with our model. In order to make the model have a certain comparability, the same method is used to calculate the sample risk score according to the corresponding genes in the three models, and zscore is performed for Riskscore. After zscore, the samples with Riskscore greater than zero are divided into high-risk group and those with riskscore less than zero are divided into low-risk group.

ROC analysis showed that the AUC value was lower than that in our model (Supplementary Figures S6A–C). The C-index in our model was higher than that in other 3 models (Supplementary Figure S6D).

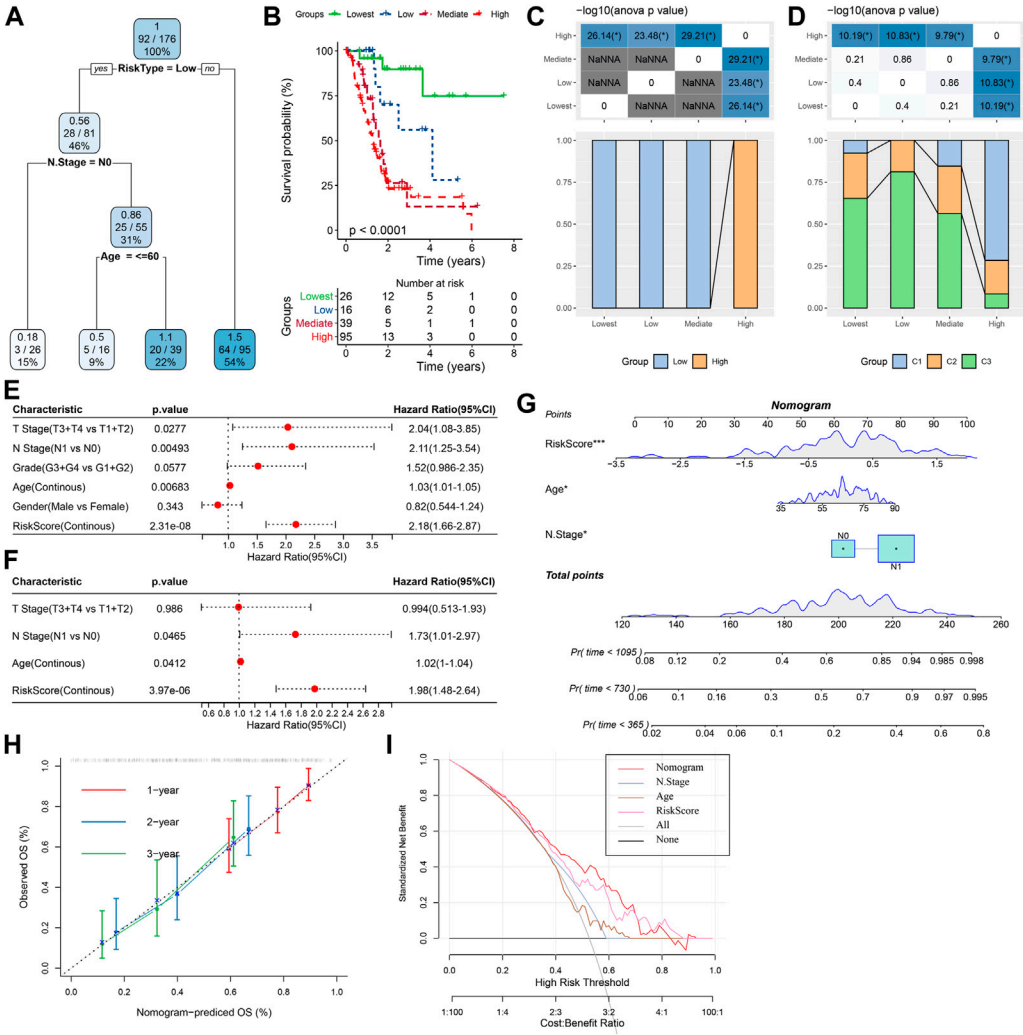


FIGURE 9
The nomogram of prognostic risk models with clinicopathological features. **(A)** To optimize risk stratification, patients with full-scale annotations including TNM Stage, age, gender, and RiskScore were enrolled for developing a survival decision tree; **(B)** Risk subgroups showed significant overall survival differences; **(C, D)** Comparative analysis between different subgroups; **(E, F)** RiskScore and clinicopathological **(E, F)** Univariate and multifactorial Cox analysis on clinicopathological features and RiskScore; **(G)** The nomogram model; **(H)** Calibration curves for 1, 3, and 5 years (s) of the nomogram; (H:Decision curve for the columnar graph; **(I)** The most powerful capacity of the nomogram for survival prediction when compared with other clinicopathological features. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$; and ****$p < 0.0001$.

Discussion

Pancreatic cancer shows a significantly poor prognosis with a 5-year survival chance of approximately 5% (Ilic and Ilic, 2016). Accurate prognostic evaluation enables patients suffering from PAAD to benefit more from effective treatments such as more intensive surgery, targeted molecular therapies, neoadjuvant therapy, immunotherapy, radiotherapy, chemotherapy. Thus, treatment could be personalized to individual patient for improving prognosis. In the early diagnosis of highly heterogeneous PAAD, molecular prognostic markers are potentially valuable, which at the same time could help overcome the impediment of heterogeneity. Multiple molecular markers increase the accuracy than single molecular markers in reflecting pancreatic cancer prognosis, the progression of which is in a complex network involving different signaling pathways. Wu et al. identified a 9-gene signature and also developed a prognostic nomogram that could reliably predict PAAD overall survival (Wu et al., 2019). Weng et al. established a multi-omics perspective consisting of 3 mRNAs, 3 miRNAs, 60 lncRNAs related to PAAD prognosis, and constructed a classifier based on 14 mRNAs with a good predictive function in the cohort and helped to predict PAAD prognosis (Weng et al., 2020). Based on 14 necroptosis-associatedgenes, Wu et al. developed a prognostic model for the diagnosis, prognosis of PAAD and its treatment (Wu et al., 2022). Our work identified the molecular subtypes of PAAD based on oxidative stress due to the non-negligible regulatory impact from oxidative stress plays.

First, we employed oxidative stress-related genes to consistently cluster three stable molecular subtypes, each of which has its own unique prognostic, route, clinical, and immunological properties. Our analysis demonstrated a better prognosis of C3 and a worse one of C1. ROS could cause several types of DNA damage (Ebrahimi et al., 2020). Persistent DNA damage is resulted from the continuous production of ROS and an inflammatory cascade that triggers genomic changes and a tendency to increase epigenetic alterations. The development of cancer may be facilitated by the accumulation of epigenetic changes that disrupt genome-wide cell signaling system and promote malignant transformation (Kgatle et al., 2017). Our comparative analysis of genomic alterations in the three subtypes revealed that the C1 subtype showed higher "TMB", "Homologous Recombination Defects", "Aneuploidy Score", "Intratumor Heterogeneity", "LOH". TMB is a sensitive biomarker for screening sensitive responders to immunotherapy and has been shown to be correlated with more significantly with response, with higher blockade of PD-L1 and PD-1 than PD-1 or PD-L1 expression. Mechanistically, high TMB provides more opportunities for "non-me" neoantigen production and activation of immune cell enrichment. Nevertheless, these theories have only been confirmed in some places for immunotherapy of certain tumors, but they may not be applicable to tumors such as pancreatic cancer (Strickler et al., 2021). For immune microenvironment differences, significantly higher immune cell infiltration and "ImmuneScore" were observed in C2 and C3. In addition, as oxidative stress is closely associated with multiple physiological activities, and our analysis demonstrated that C3 subtypes had higher inflammatory activity and autophagy scores and the lowest angiogenic scores. Though many achievements have been made in the immunotherapy of cancer, not all patients can benefit from immunotherapy. Our analysis showed significantly enhanced IFN-γ response and higher "T-cell-inflamed GEP score" in the C3. Additionally, the CYT score, which reflects cytotoxic effects, was noticeably higher in C3 than in other subtypes. In the multimodal treatment of pancreatic cancer, chemotherapy is an important component. Adjuvant chemotherapy can significantly improve disease-free survival and overall survival after curative resection (Springfeld et al., 2019). Analysis on conventional chemotherapeutic drug response showed C1 was more sensitive to Erlotinib, Cisplatin, and Gemcitabine.

Seven key genes (GJB4, CEP55, SCAMP5, ANLN, FAM83A, ATP2A3, COL17A1) associated with oxidative stress phenotypes, were identified. CEP55 plays an important role in cytoplasmic division, tumor stage, aggressiveness, metastasis and poor prognosis in many tumor types such as breast, lung, colon and liver cancers (Jeffery et al., 2016). In many malignancies ANLN is an upregulated actin-binding protein. Wang et al. found that in pancreatic cancer tissues and cell lines, ANLN expression is upregulated and is predictive of a poor PAAD prognosis. ANLN-mediated pancreatic cancer invasion and migration, colony formation, cell proliferation may involve EZH2/miR-218-5p/LASP1 signaling axis (Wang et al., 2019). The gap junction β-4 protein is an integral membrane protein member involved in tumorigenesis and may play a role as a tumor promoter (Liu et al., 2019). Moreover, in lung cancer, it has also been found to induce chemoresistance and metastasis *via* Src activation (Lin et al., 2019). As an important component of type I hemibridges (HD), COL17A1 encodes collagen XVII (COL17) (Yodsurang et al., 2017), and has been identified as a marker for pancreatic cancer by Shen et al. (Shen et al., 2017). In a variety of human tumors, family with sequence similarity 83 member A was initially identified by bioinformatics methods as a potential tumor-specific gene with overexpression, including in bladder, lung, testicular, breast cancers, etc. Chen et al. found that in pancreatic cancer FAM83A shows significant overexpression, which promotes CSC-like features by activating Wnt/β-catenin and TGF-β pathways. Therefore they concluded that FAM83A has the potential of acting as a therapeutic target for patients with pancreatic cancer (Chen et al., 2017). SCAMP functions as a post-Golgi transporter protein in all mammalian cells and is an effective prognostic and diagnostic biomarker for pancreatic cancer (Mao et al., 2021). ATP2A3 is a significantly upregulated gene that encodes a Ca2+ -ATPase localized to the ER membrane and is involved in Ca2+ transport (Zhang et al., 2019). Eduardod et al. showed that resveratrol upregulates the expression of the ATP2A3 gene in breast cancer cell lines through an epigenetic mechanism (Izquierdo-Torres et al., 2019).

In spite of this, there are some limitations in this study, which should be verified by PCR and immunohistochemical experiments. We did not consider other factors because the samples lacked necessary clinical follow-up information, especially diagnostic details.

Conclusion

This paper first identified a novel prognostic risk model consisting of 7 oxidative stress-related genes that well predict PAAD prognosis of PAAD. The 7 genes demonstrated complex molecular functions that remained to be explored further. In addition, this work highlighted the correlation between of the prognosis of PAAD with oxidative stress-related genes. The current findings facilitate personalized treatment for PAAD patients.

Data availability statement

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

Author contributions

All authors contributed to this present work: TJ designed the study and revised the manuscript, HW acquired the data. TC drafted the manuscript. All authors read and approved the manuscript.

Funding

This project is supported by the Program of Huadu District Science and Technology (No.21-HDWS-042) and the Key Disciplines of Guangzhou Huadu District People's Hospital Project (2022–2025) (Digestive Diseases Department of Guangzhou Huadu District People's Hospital).

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/fphar.2023.1091378/full#supplementary-material

References

Azmanova, M., and Pitto-Barry, A. (2022). Oxidative stress in cancer therapy: Friend or enemy? *Chembiochem* 23 (10), e202100641. doi:10.1002/cbic.202100641

Bader, G. D., and Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinforma.* 4, 2. doi:10.1186/1471-2105-4-2

Cai, X., She, M., Xu, M., Chen, H., Li, J., Chen, X., et al. (2018). GLP-1 treatment protects endothelial cells from oxidative stress-induced autophagy and endothelial dysfunction. *Int. J. Biol. Sci.* 14 (12), 1696–1708. doi:10.7150/ijbs.27774

Chen, S., Huang, J., Liu, Z., Liang, Q., Zhang, N., and Jin, Y. (2017). FAM83A is amplified and promotes cancer stem cell-like traits and chemoresistance in pancreatic cancer. *Oncogenesis* 6 (3), e300. doi:10.1038/oncsis.2017.3

Ebrahimi, S. O., Reiisi, S., and Shareef, S. (2020). miRNAs, oxidative stress, and cancer: A comprehensive and updated review. *J. Cell. Physiol.* 235 (11), 8812–8825. doi:10.1002/jcp.29724

He, D., Feng, H., Sundberg, B., Yang, J., Powers, J., Christian, A. H., et al. (2022). Methionine oxidation activates pyruvate kinase M2 to promote pancreatic cancer metastasis. *Mol. Cell.* 82, 3045–3060.e11. doi:10.1016/j.molcel.2022.06.005

Ilic, M., and Ilic, I. (2016). Epidemiology of pancreatic cancer. *World J. Gastroenterol.* 22 (44), 9694–9705. doi:10.3748/wjg.v22.i44.9694

Izquierdo-Torres, E., Hernandez-Oliveras, A., Meneses-Morales, I., Rodriguez, G., Fuentes-Garcia, G., and Zarain-Herzberg, A. (2019). Resveratrol up-regulates ATP2A3 gene expression in breast cancer cell lines through epigenetic mechanisms. *Int. J. Biochem. Cell. Biol.* 113, 37–47. doi:10.1016/j.biocel.2019.05.020

Jeffery, J., Sinha, D., Srihari, S., Kalimutho, M., and Khanna, K. K. (2016). Beyond cytokinesis: The emerging roles of CEP55 in tumorigenesis. *Oncogene* 35 (6), 683–690. doi:10.1038/onc.2015.128

Katona, B. W., Brand, R. E., Canto, M. I., Chak, A., Farrell, J. J., Kastrinos, F., et al. (2021). Screening for pancreatic ductal adenocarcinoma: Are we asking the impossible?-letter. *Cancer Prev. Res. (Phila)* 14 (10), 973–974. doi:10.1158/1940-6207.CAPR-21-0233

Kgatle, M. M., Spearman, C. W., Kalla, A. A., and Hairwadzi, H. N. (2017). DNA oncogenic virus-induced oxidative stress, genomic damage, and aberrant epigenetic alterations. *Oxid. Med. Cell. Longev.* 2017, 3179421. doi:10.1155/2017/3179421

Liberzon, A., Birger, C., Thorvaldsdottir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell. Syst.* 1 (6), 417–425. doi:10.1016/j.cels.2015.12.004

Lin, Y. P., Wu, J. I., Tseng, C. W., Chen, H. J., and Wang, L. H. (2019). Gjb4 serves as a novel biomarker for lung cancer and promotes metastasis and chemoresistance via Src activation. *Oncogene* 38 (6), 822–837. doi:10.1038/s41388-018-0471-1

Liu, G., Pang, Y., Zhang, Y., Fu, H., Xiong, W., and Zhang, Y. (2019). GJB4 promotes gastric cancer cell proliferation and migration via Wnt/CTNNB1 pathway. *Onco Targets Ther.* 12, 6745–6755. doi:10.2147/OTT.S205601

Liu, Q., Cheng, R., Kong, X., Wang, Z., Fang, Y., and Wang, J. (2020). Molecular and clinical characterization of PD-1 in breast cancer using large-scale transcriptome data. *Front. Immunol.* 11, 558757. doi:10.3389/fimmu.2020.558757

Liu, T., Liu, Q., Wang, Y., Yang, R., and Tian, F. (2022). Cuproptosis scoring model predicts overall survival and assists in immunotherapeutic decision making in pancreatic carcinoma. *Front. Genet.* 13, 938488. doi:10.3389/fgene.2022.938488

Lowenfels, A. B., Maisonneuve, P., Cavallini, G., Ammann, R. W., Lankisch, P. G., Andersen, J. R., et al. (1993). Pancreatitis and the risk of pancreatic cancer. International pancreatitis study group. *N. Engl. J. Med.* 328 (20), 1433–1437. doi:10.1056/NEJM199305203282001

Mao, F., Duan, H., Allamyradov, A., Xin, Z., Du, Y., Wang, X., et al. (2021). Expression and prognostic analyses of SCAMPs in pancreatic adenocarcinoma. *Aging (Albany NY)* 13 (3), 4096–4114. doi:10.18632/aging.202377

Masiero, M., Simoes, F. C., Han, H. D., Snell, C., Peterkin, T., Bridges, E., et al. (2013). A core human primary tumor angiogenesis signature identifies the endothelial orphan receptor ELTD1 as a key regulator of angiogenesis. *Cancer Cell.* 24 (2), 229–241. doi:10.1016/j.ccr.2013.06.004

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

Pathria, G., Lee, J. S., Hasnis, E., Tandoc, K., Scott, D. A., Verma, S., et al. (2019). Translational reprogramming marks adaptation to asparagine restriction in cancer. *Nat. Cell. Biol.* 21 (12), 1590–1603. doi:10.1038/s41556-019-0415-1

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007

Runa, F., Hamalian, S., Meade, K., Shisgal, P., Gray, P. C., and Kelber, J. A. (2017). Tumor microenvironment heterogeneity: Challenges and opportunities. *Curr. Mol. Biol. Rep.* 3 (4), 218–229. doi:10.1007/s40610-017-0073-7

Rydyznski Moderbacher, C., Kim, C., Mateus, J., Plested, J., Zhu, M., Cloney-Clark, S., et al. (2022). NVX-CoV2373 vaccination induces functional SARS-CoV-2-specific CD4+ and CD8+ T cell responses. *J. Clin. Invest.* 132 (19), e160898. doi:10.1172/JCI160898

Shen, S., Gui, T., and Ma, C. (2017). Identification of molecular biomarkers for pancreatic cancer with mRMR shortest path method. *Oncotarget* 8 (25), 41432–41439. doi:10.18632/oncotarget.18186

Singhi, A. D., Koay, E. J., Chari, S. T., and Maitra, A. (2019). Early detection of pancreatic cancer: Opportunities and challenges. *Gastroenterology* 156 (7), 2024–2040. doi:10.1053/j.gastro.2019.01.259

Son, J., Lyssiotis, C. A., Ying, H., Wang, X., Hua, S., Ligorio, M., et al. (2013). Glutamine supports pancreatic cancer growth through a KRAS-regulated metabolic pathway. *Nature* 496 (7443), 101–105. doi:10.1038/nature12040

Sosa, V., Moline, T., Somoza, R., Paciucci, R., Kondoh, H., and Lleonart, M. E. (2013). Oxidative stress and cancer: An overview. *Ageing Res. Rev.* 12 (1), 376–390. doi:10.1016/j.arr.2012.10.004

Springfeld, C., Jager, D., Buchler, M. W., Strobel, O., Hackert, T., Palmer, D. H., et al. (2019). Chemotherapy for pancreatic cancer. *Presse Med.* 48 (32), e159–e174. doi:10.1016/j.lpm.2019.02.025

Strickler, J. H., Hanks, B. A., and Khasraw, M. (2021). Tumor mutational burden as a predictor of immunotherapy response: Is more always better? *Clin. Cancer Res.* 27 (5), 1236–1241. doi:10.1158/1078-0432.CCR-20-3054

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830.e14. doi:10.1016/j.immuni.2018.03.023

Tibshirani, R. (1997). The lasso method for variable selection in the Cox model. *Stat. Med.* 16 (4), 385–395. doi:10.1002/(sici)1097-0258(19970228)16:4<385:aid-sim380>3.0.co;2-3

Wang, A., Dai, H., Gong, Y., Zhang, C., Shu, J., Luo, Y., et al. (2019). ANLN-induced EZH2 upregulation promotes pancreatic cancer progression by mediating miR-218-5p/LASP1 signaling axis. *J. Exp. Clin. Cancer Res.* 38 (1), 347. doi:10.1186/s13046-019-1340-7

Weng, W., Zhang, Z., Huang, W., Xu, X., Wu, B., Ye, T., et al. (2020). Identification of a competing endogenous RNA network associated with prognosis of pancreatic adenocarcinoma. *Cancer Cell. Int.* 20, 231. doi:10.1186/s12935-020-01243-6

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

Wu, M., Li, X., Zhang, T., Liu, Z., and Zhao, Y. (2019). Identification of a nine-gene signature and establishment of a prognostic nomogram predicting overall survival of pancreatic cancer. *Front. Oncol.* 9, 996. doi:10.3389/fonc.2019.00996

Wu, Z., Huang, X., Cai, M., Huang, P., and Guan, Z. (2022). Novel necroptosis-related gene signature for predicting the prognosis of pancreatic adenocarcinoma. *Aging (Albany NY)* 14 (2), 869–891. doi:10.18632/aging.203846

Yan, C., Niu, Y., Li, F., Zhao, W., and Ma, L. (2022). System analysis based on the pyroptosis-related genes identifies GSDMC as a novel therapy target for pancreatic adenocarcinoma. *J. Transl. Med.* 20 (1), 455. doi:10.1186/s12967-022-03632-z

Yang, J., Wei, X., Hu, F., Dong, W., and Sun, L. (2022). Development and validation of a novel 3-gene prognostic model for pancreatic adenocarcinoma based on ferroptosis-related genes. *Cancer Cell. Int.* 22 (1), 21. doi:10.1186/s12935-021-02431-8

Yang, W., Soares, J., Greninger, P., Edelman, E. J., Lightfoot, H., Forbes, S., et al. (2013). Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* 41, D955–D961. doi:10.1093/nar/gks1111

Yodsurang, V., Tanikawa, C., Miyamoto, T., Lo, P. H. Y., Hirata, M., and Matsuda, K. (2017). Identification of a novel p53 target, COL17A1, that inhibits breast cancer cell migration and invasion. *Oncotarget* 8 (34), 55790–55803. doi:10.18632/oncotarget.18433

Yu, Y., Yan, Y., Niu, F., Wang, Y., Chen, X., Su, G., et al. (2021). Ferroptosis: A cell death connecting oxidative stress, inflammation and cardiovascular diseases. *Cell. Death Discov.* 7 (1), 193. doi:10.1038/s41420-021-00579-w

Zhang, Y., Li, F., Liu, L., Jiang, H., Hu, H., Du, X., et al. (2019). Salinomycin triggers endoplasmic reticulum stress through ATP2A3 upregulation in PC-3 cells. *BMC Cancer* 19 (1), 381. doi:10.1186/s12885-019-5590-8

Frontiers in Pharmacology

Explores the interactions between chemicals and living beings

The most cited journal in its field, which advances access to pharmacological discoveries to prevent and treat human disease.

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