*CORRESPONDENCE
Jianhua Zhang
✉ 972659434@qq.com

†These authors have contributed equally to
this work

Integrated analysis of single-cell and Bulk RNA sequencing reveals a malignancy-related signature in lung adenocarcinoma

Mengxi Wu[1†], Zhenyu Wu[2†], Jun Yan[1†], Jie Zeng[3], Jun Kuang[1], Chenghua Zhong[1], Xiaojia Zhu[1], Yijun Mo[1], Quanwei Guo[1], Dongfang Li[1], Jianfeng Tan[1], Tao Zhang[1] and Jianhua Zhang[1*]

[1]Department of Thoracic Surgery, Shenzhen Hospital, Southern Medical University, Shenzhen, China,
[2]Department of Urology, The First People's Hospital of Foshan, Foshan, China, [3]Department of
Thoracic Surgery, Guangzhou First People's Hospital, South China University of Technology,
Guangzhou, China

Background: Lung adenocarcinoma (LUAD), the most common histotype of lung cancer, may have variable prognosis due to molecular variations. The research strived to establish a prognostic model based on malignancy-related risk score (MRRS) in LUAD.

Methods: We applied the single-cell RNA sequencing (scRNA-seq) data from Tumor Immune Single Cell Hub database to recognize malignancy-related geneset. Meanwhile, we extracted RNA-seq data from The Cancer Genome Atlas database. The GSE68465 and GSE72094 datasets from the Gene Expression Omnibus database were downloaded to validate the prognostic signature. Random survival forest analysis screened MRRS with prognostic significance. Multivariate Cox analysis was leveraged to establish the MRRS. Furthermore, the biological functions, gene mutations, and immune landscape were investigated to uncover the underlying mechanisms of the malignancy-related signature. In addition, we used qRT-PCR to explore the expression profile of MRRS-constructed genes in LUAD cells.

Results: The scRNA-seq analysis revealed the markers genes of malignant celltype. The MRRS composed of 7 malignancy-related genes was constructed for each patient, which was shown to be an independent prognostic factor. The results of the GSE68465 and GSE72094 datasets validated MRRS's prognostic value. Further analysis demonstrated that MRRS was involved in oncogenic pathways, genetic mutations, and immune functions. Moreover, the results of qRT-PCR were consistent with bioinformatics analysis.

Conclusion: Our research recognized a novel malignancy-related signature for predicting the prognosis of LUAD patients and highlighted a promising prognostic and treatment marker for LUAD patients.

KEYWORDS

lung adenocarcinoma, ScRNA-seq, malignancy, gene signature, prognosis

1 Introduction

Lung cancer is a malignant tumor with the highest mortality rate in the world (1). Lung adenocarcinoma (LUAD) is the most common histotype of lung cancer, accounting for around 40% of all lung malignancies with an increasing prevalence (2). Despite the employment of novel therapies including targeted therapy and immunotherapy, the prognosis for patients with LUAD was dismal (3). Although patients may have comparable pathology, anatomical location, and clinical staging, their survival outcomes will likely vary because of molecular differences. Therefore, it is necessary to exploit newly prognostic molecular biomarkers, which will be of assistance in improving the prognosis of overall survival (OS) as well as the therapy effectiveness for LUAD patients.

In recent years, next-generation sequencing has been widely used, but conventional NGS does not detect cellular heterogeneity (4–6). Single-cell RNA sequencing (scRNA-seq) can be used to detect the genome, transcriptome, and other multi-omics of single cells. It is a powerful approach to dissect cellular heterogeneity, which can specifically obverse the changes in the tumor microenvironment (TME) (7). It is known that tumor cells are surrounded by TME, including a variety of immune cells, stromal cells, extracellular matrix molecules, and various cytokines (8). As a key priority, tumor cells play a vital role in the occurrence and development of tumors. In this study, we extracted tumor cell subpopulations and identified tumor cell marker genes through scRNA-seq.

In the research, we applied the scRNA-seq data from Tumor Immune Single Cell Hub (TISCH) database to obtain a gene expression map from the level of single cells in LUAD. Next, we extracted transcriptome data and associated clinical information from TCGA database (TCGA-LUAD) and Gene Expression Omnibus (GEO) database. Subsequently, we used TCGA cohort as the training set while the GSE68465 and GSE72094 cohorts as the validation sets. By linking associated genes with clinical cases of LUAD, we focused on investigating the effect of the malignancy-related signature on foretelling the mortality of LUAD patients and exploring their underlying mechanisms on tumor growth and progression.

2 Materials and methods

2.1 Dataset source and preprocessing

First, we downloaded the single cell RNA sequencing (scRNA-seq) dataset (GSE117570) (9) of 4 LUAD patients from the TISCH database[1]. RNA-sequencing (RNA-seq) and the matched clinical characteristics of 460 LUAD patients were derived from the TCGA database[2]. We downloaded two datasets (439 LUAD patients for GSE68465 and 386 patients for GSE72094) of RNA expression data and complete clinical data from the GEO database[3] (10, 11). The

batch effect was adjusted through the "sva" R package. The IMvigor210 cohort (12) of bladder cancer patients treated with anti-PD-L1 therapy was obtained through the "IMvigor210CoreBiologies" R package, and the GSE91061 dataset (13) that accepted anti-PD-1 and anti-CTLA4 treatment was also attained to predict the efficiency of immunotherapy. The basic information of series was shown in Supplementary Table 1.

2.2 ScRNA-seq analysis

Quality control and normalization of scRNA-seq data were processed with the "Seurat" R package. Cell clusters were derived from the TISCH database. According to the documentation, the cell type was annotated by the description provided by the original study, the markers of malignant cells, and the "inferCNV" R package. FindAllMarkers function was used to determine and annotate gene markers for different cell clusters with thresholds of p.adjust< 0.05 and log2 [Foldchange] > 0.3. We extracted the marker genes of the malignant celltype for further study.

2.3 Generation of malignancy-related signature

To establish the malignancy-related signature, we employed TCGA cohort as the training set, while the GSE68465 and GSE72094 datasets were the validation sets. Univariate Cox analysis was performed to explore prognosis-associated genes ($p<$ 0.001). Random survival forest (RSF) analysis was then conducted using the "randomForestSRC" R package to further narrow down the prognostic gene panel. In RSF analysis, variables were ranked by minimal depth, of which a smaller value indicated greater predictiveness. Next, multivariate Cox regression analysis was used to establish the optimal malignancy-related signature based on respective coefficients (β) and gene expression levels (Exp). This formula was used to calculate each patient's malignancy-related risk score (MRRS). Subsequently, we divided the patients into high- and low-risk groups based on the median MRRS. The Kaplan-Meier approach was applied to determine the prognostic difference between the two groups. We further evaluate the correlations between the MRRS and clinical features including age, gender, clinical stage, TN stages, and smoking status. Univariate and multivariate Cox analyses were utilized to assess the prognostic significance of MRRS. Meanwhile, we collected the GSE68465 and GSE72094 cohorts to verify MRRS's predictive efficacy.

2.4 Functional enrichment analysis

To investigate the underlying mechanism regarding MRRS, differentially expressed genes were obtained between the high- and low-risk cohorts. First, we performed Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses using the "clusterProfiler" R package. GO and KEGG terms with $p<$ 0.05 were visualized by the "circlize"

---

1   http://tisch.comp-genomics.org/

2   https://portal.gdc.cancer.gov/

3   https://www.ncbi.nlm.nih.gov/geo/

R package. GSVA was employed to determine the differences between the two cohorts on the oncogenic hallmark pathways (h.all.v7.1.symbols) deposited in the MSigDB database (*p*.adjust< 0.001). Gene Set Enrichment Analysis (GSEA) was performed between the two groups for the same hallmark pathways with the "GSEA" R package (FDR< 0.25, NES > 1, and *p*.adjust< 0.05). Kaplan-Meier method was employed to determine the prognostic significance of the overlapping hallmark pathways of GSVA and GSEA.

2.5 Somatic mutation analysis

The somatic mutations of LUAD patients were extracted from TCGA database. The "maftools" R package explored the specific somatic mutation variations in different MRRS groups. Next, we investigated the mutually co-occurring or exclusive mutations, tumor-causing genes, and enrichment of known oncogenic pathways between the two cohorts. Tumor mutation burden (TMB) reflecting total mutation counts for each LUAD patient was computed and tested for correlation with MRRS. In addition, we analyzed the predictive value of TMB and MRRS on the survival outcomes in terms of the MRRS risk cohorts.

2.6 Immune landscape analysis and treatment response prediction

We compared the high- and low-risk groups' immune cell abundance, immune function, and immune checkpoints. The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm[4] (14) was applied to predict the potential immunotherapy response based on the RNA expression profile of LUAD patients. The IMvigor210 and GSE91061 datasets were also used to determine the correlation between the MRRS and potential immunotherapy efficacy. Meanwhile, we investigated the chemotherapy response of the two groups, and the "oncoPredict" R package predicted the therapeutic effect of chemotherapeutic drugs for each patient.

2.7 RNA extraction and quantitative real-time PCR

Human LUAD cells (A549, H1299, SKLU1, and H1250) and normal bronchial epithelial cells (16HBE) were obtained from Procell Life Science and Technology (Wuhan, China). Total RNA of the cells was isolated using TRIzol reagent (BioTeke, Beijing, China). Subsequently, qRT-PCR was performed utilizing the HiScript II Q RT SuperMix for qPCR (Bioer Technology, Hangzhou, China). The 2−△△Ct method was used to calculate the relative expression of each lnRNA. The differences in the expression levels of LUAD cells and normal cells were assessed

using the unpaired t test. The primer sequences were showed in Supplementary Table 2.

2.8 Statistical analysis

We conducted all statistical analyses in this research using R software (version 4.2.2). All codes of the analysis was uploaded in GitHub[5]. The unpaired Student's t-test was used to assess the differences between continuous variables. The chi-square test was utilized to examine the relationship between categorical factors. Statistical significance for most analyses was empirically set at a two-tailed *p*< 0.05.

3 Results

3.1 ScRNA-seq analysis

Figure 1 depicted the flowchart for this research. ScRNA-seq data of 4 LUAD patients were downloaded from the GSE117570 dataset. After data processing, standardization, and data filtering, a total of 11,453 cells were obtained for subsequent analysis. After unsupervised clustering of all cells, 22 clusters were obtained (Figure 2A), which were visualized after dimensionality reduction by UMAP. According to the annotation of the TISCH database, we classified the cells into 9 cell types, including CD8 T cells, dendritic cells, fibroblasts, gland mucous cells, malignant cells, mast cells, myofibroblasts, pit mucous cells, and plasma cells (Figures 2B–D). The differentially expressed genes of all cell types were displayed in Figure 2E. Considering the key role of malignant cells in tumor tissue, we extracted the markers genes of malignant cell type for further study.

3.2 Construction of malignancy-based model

The marker genes of malignant cells were incorporated into the following study. In TCGA cohort, univariate Cox regression analysis recognized 80 prognosis-associated genes (Figure 3A). RSF analysis further identified 10 model-constructed candidates based on the minimal depth method (Figures 3B, C). Seven vital genes were ultimately chosen to form the MRRS using multivariate Cox regression, namely LDHA, CAMTA1, MYLIP, GALNT3, PERP, CD99, and ABL2. The formula was: $MRRS = \sum_{i=1}^{7}(Exp_i \ast \beta_i)$ (Table 1). Based on the median MRRS, the patients were separated into high-risk and low-risk. The OS of the high-risk group was significantly shorter than that of the low-risk group (Figure 3D). Figure 3E showed these patients' MRRS distribution, survival status, and MRRS profile. Further analyses suggested that a higher MRRS was correlated with worse clinical stage and TN stages (Figure 4A). In addition, it was observed that MRRS and clinical

---

4 http://tide.dfci.harvard.edu/

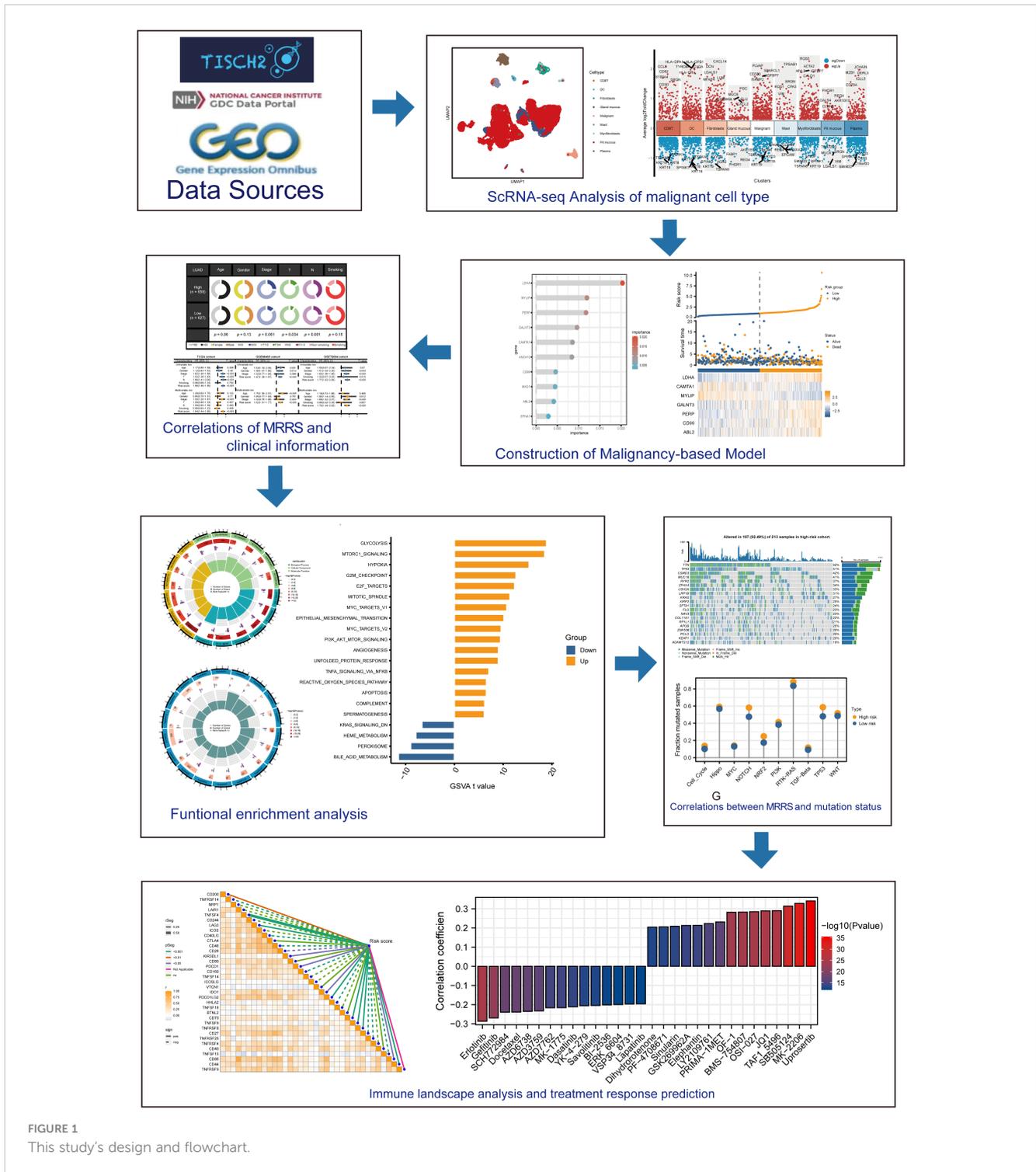5 https://github.com/No-Potato/LUAD-MRRS

FIGURE 1

This study's design and flowchart.

variables were closely associated with OS in the univariate Cox regression. Further multivariate Cox analysis indicated that MRRS was an independent prognostic factor (Figure 4B). ROC analysis also confirmed the predictive efficacy of MRRS (AUC = 0.717, Figure 4C). The outcomes of the GSE68465 and GSE72094 cohorts also validated MRRS's prognostic value (Figures 3F–I, 4). These outcomes showed that MRRS was a highly reliable prognostic indicator.

3.3 Functional enrichment analysis

To investigate the underlying mechanism regarding MRRS, we performed GO and KEGG analyses and the results revealed that MRRS was related to receptor ligand activity, signaling receptor activator activity, extracellular matrix organization, myeloid leukocyte migration, cell cycle, chromosomal region, and mitotic nuclear division (Figures 5A, B). The above GO and KEGG items
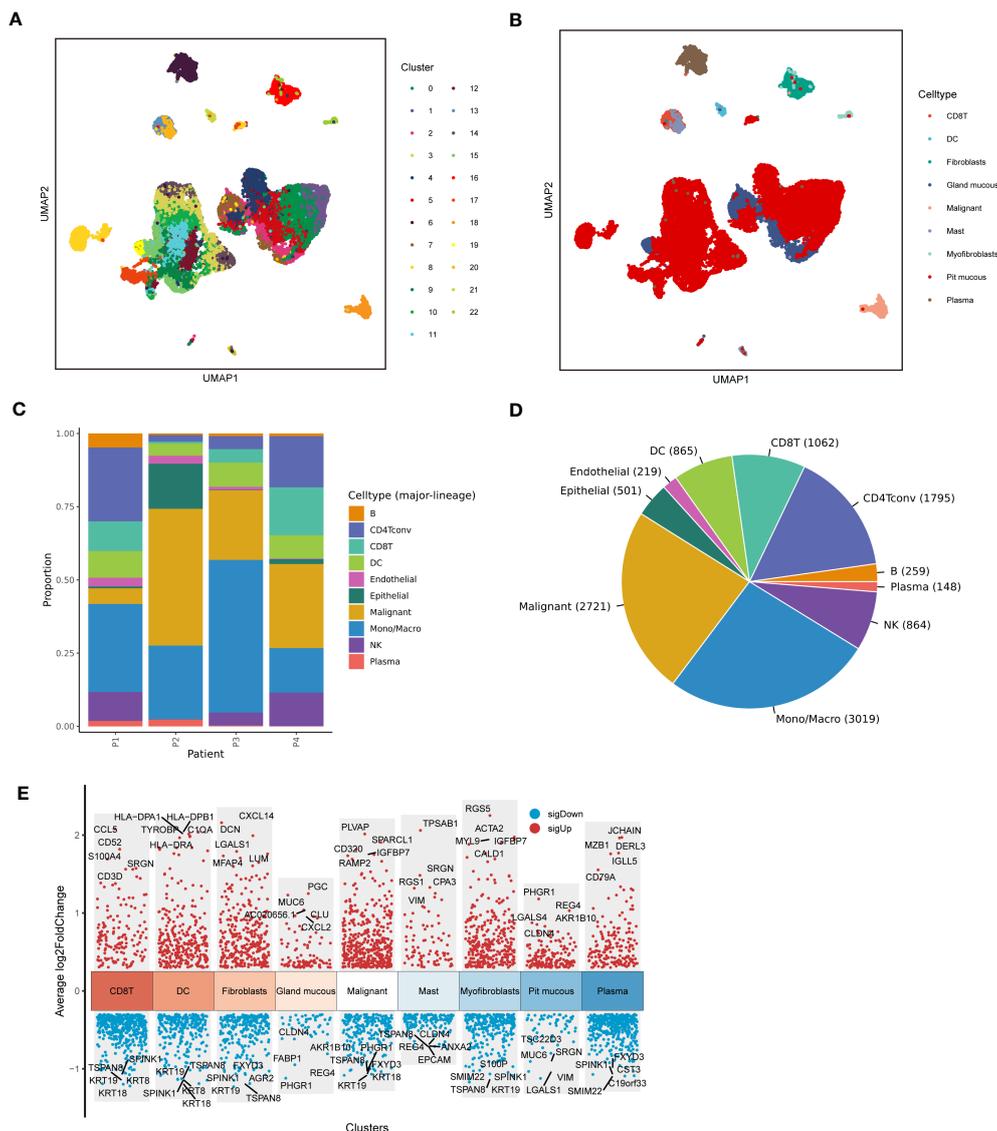
FIGURE 2

Sc-RNA seq analysis of LUAD. **(A)** The UMAP clustering map showed different clusters. **(B)** The UMAP clustering map showed different celltypes. **(C, D)** The proportion and construction of different celltypes. **(E)** Marker genes of different celltypes.

suggested that MRRS may be involved in oncogenic pathways, tumor mutations, and immune functions. Subsequently, 50 oncogenic hallmark pathways were included in GSVA, whose outcomes indicated that 17 hallmark pathways were significantly increased in high-risk patients while 4 pathways were decreased in low-risk patients (Figure 5C). GSEA confirmed that 20 were significantly upregulated in the high-risk cohort, while 2 were upregulated in the low-risk cohort (Figure 5D). The pathways obtained by intersection were analyzed by the Kaplan-Meier method, and different OS probabilities were observed for several well-known oncogenic pathways such as E2F_TARGETS, HYPOXIA, G2M_CHECKPOINT, and MYC_TARGETS_V1 (Figure 5E). In summary, MRRS participated in multiple biological functions, especially oncogenic pathways in LUAD.

3.4 Somatic mutation analysis

Gene mutations landscape between the high- and low-risk groups were shown in waterfall plots (Figures 6A, B). The genes with the highest mutation frequencies in the high-risk cohort were TTN, TP53, CSMD3, MUC16, and RYR2, while those in the low-risk cohort were TP53, TTN, MUC16, LRP1B, and CSMD3. Furthermore, the co-occurring or exclusive mutations across the top 25 mutated genes between the two cohorts were also exhibited, with no significant differences observed (Figure 6C). The mutation enrichment of known oncogenic pathways showed no significant difference between the high- and low-risk teams (Figure 6D). Further analysis also confirmed the positive correlation between TMB and MRRS, and higher TMB had a
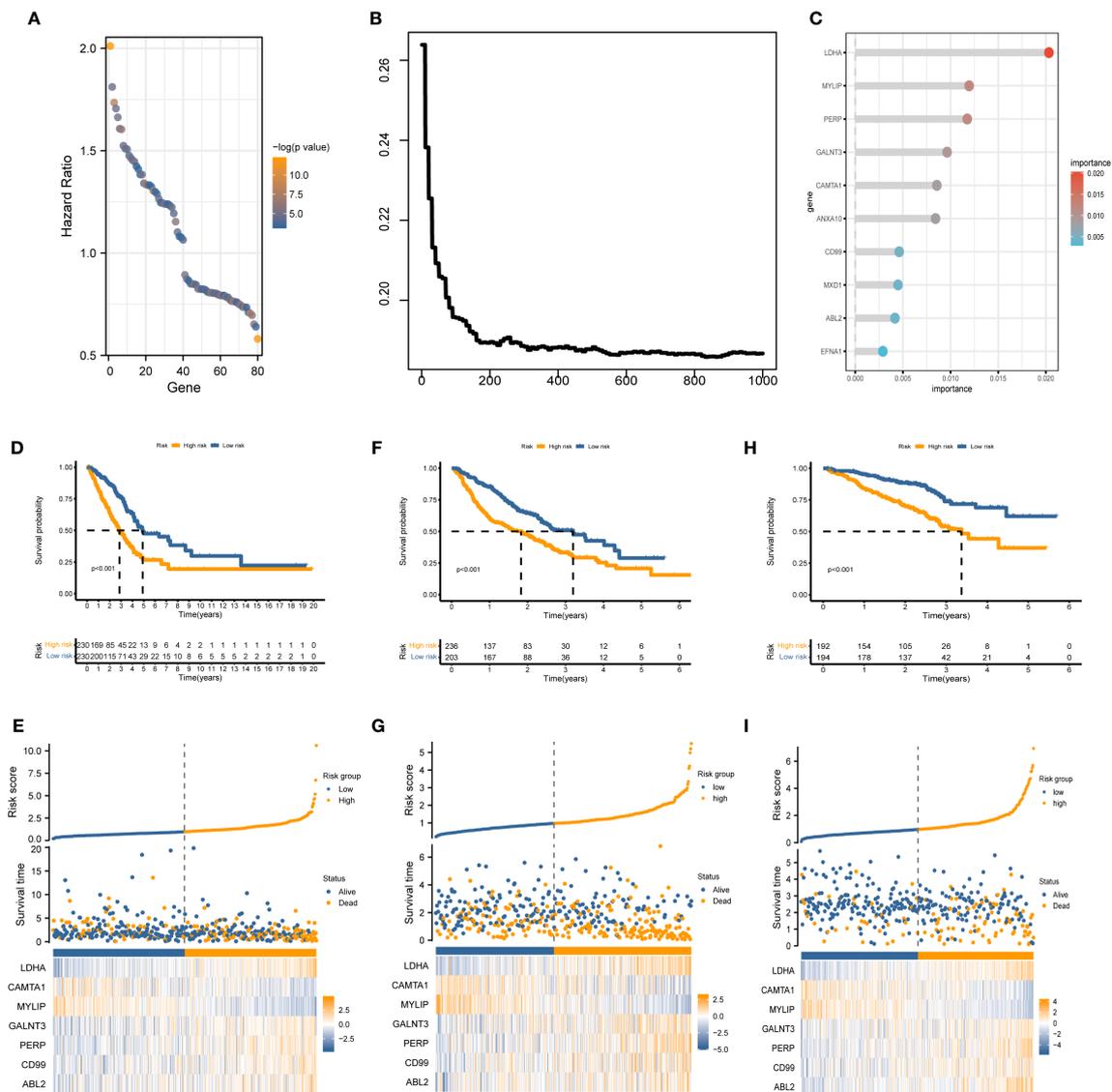
FIGURE 3
The establishment of MRRS and verification of its prognostic efficiency. **(A)** Univariate Cox regression analysis recognized 80 prognosis-associated genes. **(B)** Correlations between error rate and classification trees. **(C)** The relative importance of prognosis-associated genes. **(D)** The Kaplan-Meier method unveiled a significantly worse OS of the high-risk cohort compared to the low-risk cohort. **(E)** The illustrations of all patient's survival conditions, risk variations, and MRRS distributions. **(F−I)** The outcomes of the GSE68465 and GSE72094 cohorts also validated MRRS's prognostic value.

TABLE 1 The prognostic significance of the 7-genes signature.

MRRS-related gene	Coef
LDHA	0.444007897
CAMTA1	-0.159218411
MYLIP	-0.337064301
GALNT3	0.128249103
PERP	0.111381641
CD99	0.15017571
ABL2	0.234266034

better OS (Figures 6E, F). Survival analysis suggested that low TMB and high MRRS had the worst prognoses (Figure 6G). In conclusion, comprehensive analyses disclosed the mutation variations between high- and low-risk cohorts, and multiple remarkable genes and pathways showed significant mutation abnormalities between the cohorts.

3.5 Immune landscape analysis and treatment response prediction

Immune landscape analysis revealed higher abundances of T cells, B cells, and NK cells in the low-risk group compared to the
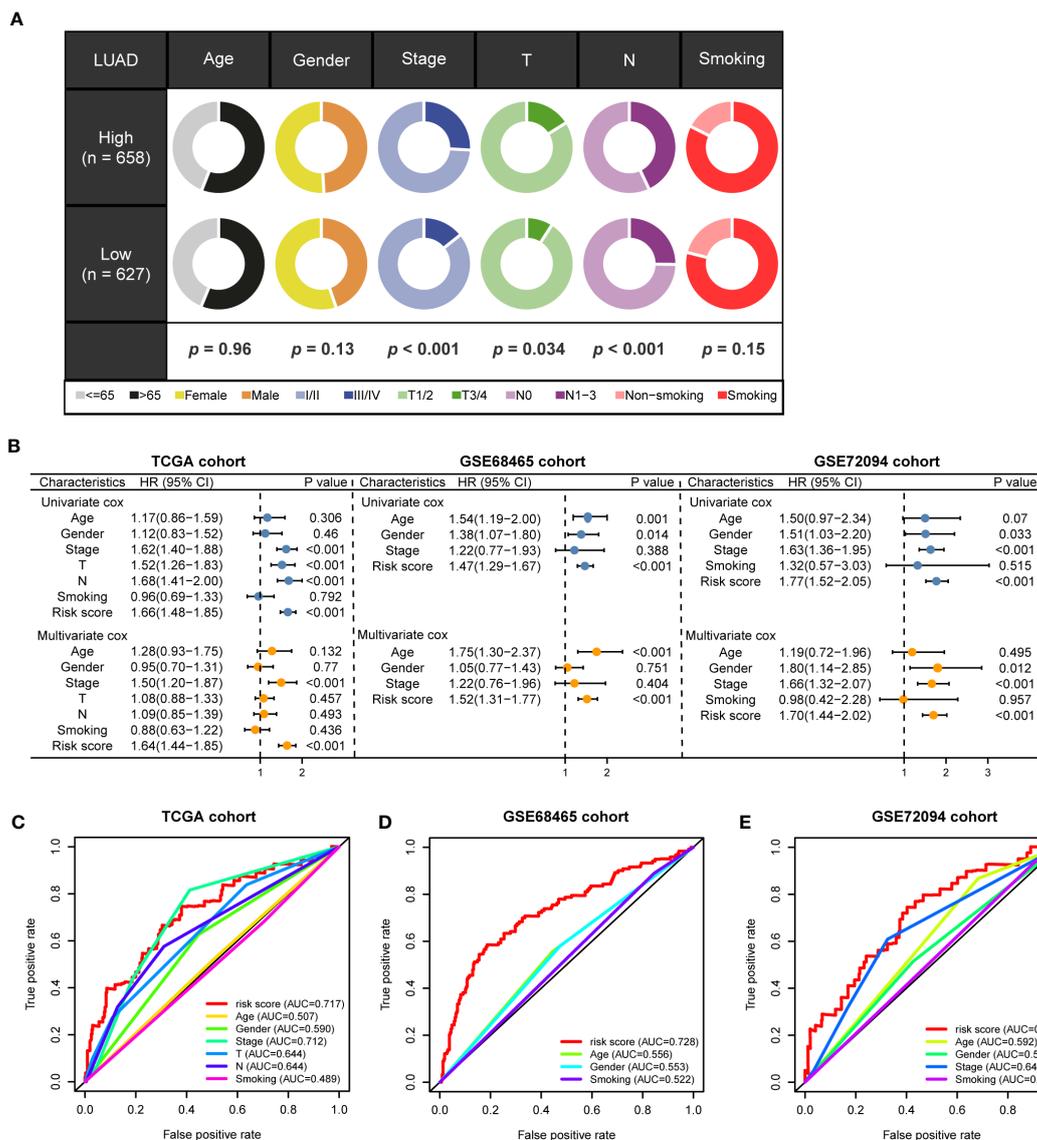
FIGURE 4
The prognostic value of MRRS and clinical variables. **(A)** Relationship between MRRS and clinical features. **(B)** Univariate and multivariate Cox regression analyses of MRRS and clinical features in TCGA, GSE68465, and GSE91061 cohorts. **(C–E)** The ROC method revealed the prognostic significance of MRRS in TCGA, GSE68465, and GSE72094 cohorts, respectively.

high-risk group (Figure 7A). Most anti-tumor immune functions such as HLA function and T cell co-stimulation were relatively decreased in the high-risk team (Figure 7B). Besides, we also found higher expressions of immunosuppressive receptors and inhibitory ligands (PDCD1, PDCD1LG2, and LAG3) in high-risk patients (Figure 7C). Meanwhile, the TIDE algorithm identified no significant difference in immunotherapy response between the high- and low-risk groups (Figure 7D). The predicted outcomes of the IMvigor210 cohort and GSE91061 cohort also supported the above conclusion (Figure 7E). Considering the poor response to immunotherapy in LUAD, the chemotherapy response of LUAD patients with different MRRSs was assessed by the "oncoPredict" R package. Our findings indicated that high-risk patients had significantly lower IC50 values in several chemotherapy molecules

including Erlotinib, Gefitinib, SCH772984, and Docetaxel (Figure 7F). Overall, the immune landscape analyses demonstrated that MRRS was associated with different immune responses, and chemotherapy may be more effective than immunotherapy for the high-risk patients.

3.6 Expression level of MRRS-constructed genes in LUAD cells

In addition, we validated the expression levels of MRRS-constructed genes (LDHA, CAMTA1, MYLIP, GALNT3, PERP, CD99, and ABL2) between the LUAD cells (A549, H1299, SKLU1, and H1250) and normal bronchial epithelial cells (16HBE) using
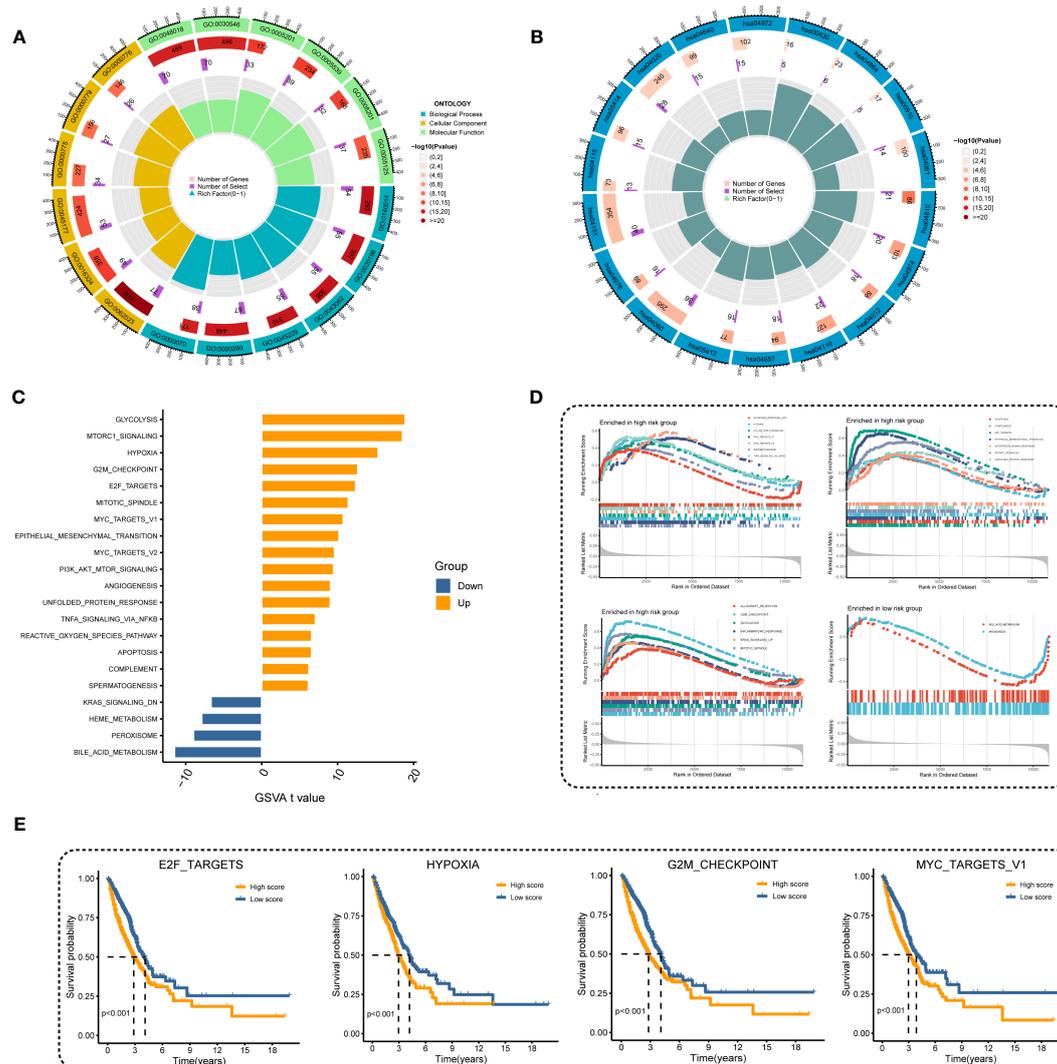
FIGURE 5
Investigation of underlying mechanism regarding MRRS. **(A)** GO enrichment analysis of MRRS. **(B)** KEGG pathway analysis of MRRS. **(C)** Determination of oncogenic hallmark pathways in terms of the MRRS risk cohorts utilizing GSVA. **(D)** The GSEA outcomes for the hallmark pathways between the high- and low-risk patients. **(E)** Kaplan-Meier curve uncovered the OS in overlapping hallmark pathways between GSVA and GSEA.

qRT-PCR. As shown in Figure 8, the expression of LDHA, GALNT3, PERP, and ABL2 was significantly upregulated in LUAD cell lines, while MYLIP was significantly upregulated in 16HBE cells. However, the expression of CAMTA1 and CD99 showed no significant difference between the LUAD cells and normal bronchial epithelial cells. Overall, the results of qRT-PCR were consistent with bioinformatics analysis.

4 Discussion

Patients with lung cancer, with the highest fatality rate in the world, had a dismal prognosis, with 5-year survival rates of less than 20% (3). Patients with advanced LUAD now have new hope thanks to the development and application of immunotherapy and targeted therapies medications in recent years (15). However, due

to molecular variances, patients with the same pathological type and clinical stage may have varying prognoses. Therefore, to predict the prognosis of LUAD patients, further molecular indicators must be investigated. Increasing bioinformatic articles get published recent years and achieved excellent efficacy (6, 16–19). The scRNA-seq has good application potential in disease research since it can obtain gene expression maps at the level of a single cell and identify heterogeneous tissue samples in groups (20–22). In our study, we amalgamated data from TCGA, GSE68465, and GSE72094 cohorts to acquire MRRS. ROC analysis verified the advantageous predictive efficacy of MRRS with an AUC of 0.717. Multiple datasets were utilized to investigate the biological functions of MRRS, which augmented the reliability of the results.

Furthermore, GO and KEGG analyses demonstrated that MRRS may be correlated with chromosomal region, mitotic
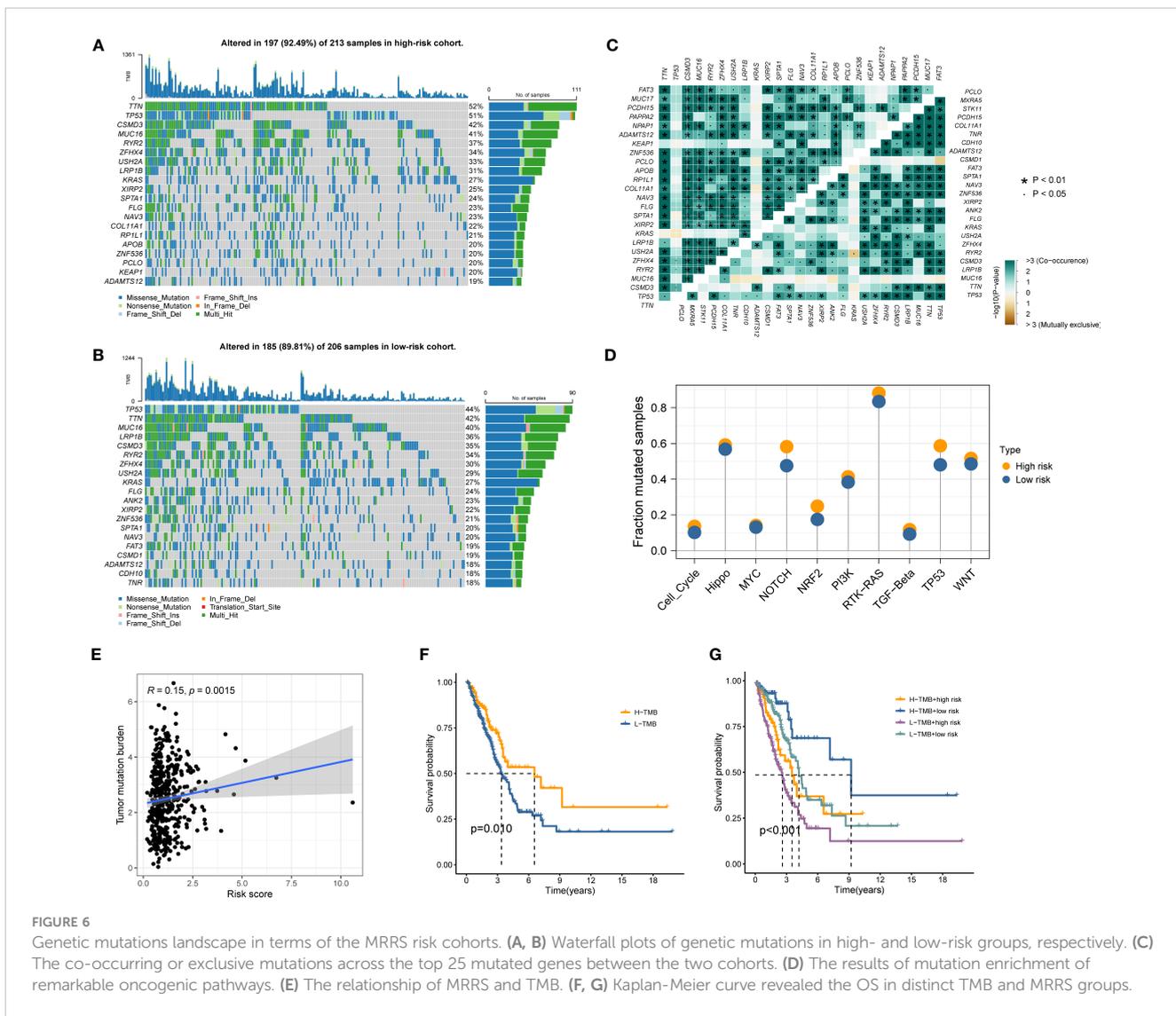
FIGURE 6
Genetic mutations landscape in terms of the MRRS risk cohorts. **(A, B)** Waterfall plots of genetic mutations in high- and low-risk groups, respectively. **(C)** The co-occurring or exclusive mutations across the top 25 mutated genes between the two cohorts. **(D)** The results of mutation enrichment of remarkable oncogenic pathways. **(E)** The relationship of MRRS and TMB. **(F, G)** Kaplan-Meier curve revealed the OS in distinct TMB and MRRS groups.

nuclear division, extracellular matrix organization, myeloid leukocyte migration, cell cycle, and signaling receptor activator activity. The findings of GSVA and GSEA implied that MRRS may modify the tumor's biological behavior by participating in multiple oncogenic hallmark pathways. Signal pathways including E2F TARGETS, HYPOXIA, G2M CHECKPOINT, and MYC TARGETS V1 were found to be strongly correlated with patient prognosis by Kaplan-Meier analysis. It has been reported that upregulation of the E2F signaling promoted the proliferation and progression of LUAD (23). The MYC pathway's aberrant activation can contribute to the progression of LUAD and its metastasis (24, 25). As a result, by regulating certain oncogenic pathways, MRRS may have an impact on a patient's prognosis.

The genetic mutations between the high- and low-risk groups were further investigated. TTN, TP53, CSMD3, MUC16, and RYR2 were the top five genes for mutation frequency in the high-risk team, while TP53, TTN, MUC16, LRP1B, and CSMD3 were the top

five genes in the low-risk team. LRP1B was reported as a tumor suppressor gene in non-small-cell lung cancer (26). A recent study showed that patients with cancer who had LRP1B mutations had improved clinical outcomes when receiving immune checkpoint inhibitors (27). Additionally, we discovered that MRRS and TMB had a positive correlation and that their combination can more accurately predict the prognosis of the patient. TMB has emerged as a valuable biomarker for evaluating the efficacy of immunotherapy in LUAD because it represents a mutagenesis process triggered by intracellular and environmental factors (28, 29). Our research identified mutation variations in the MRRS risk cohorts. MRRS may take part in a variety of aberrant mutations involving oncologic genes and pathways to modulate the growth and progression of LUAD.

Immune landscape analysis revealed that the high-risk group had substantially more immune cells and immune functions. Consequently, an additional investigation discovered that the
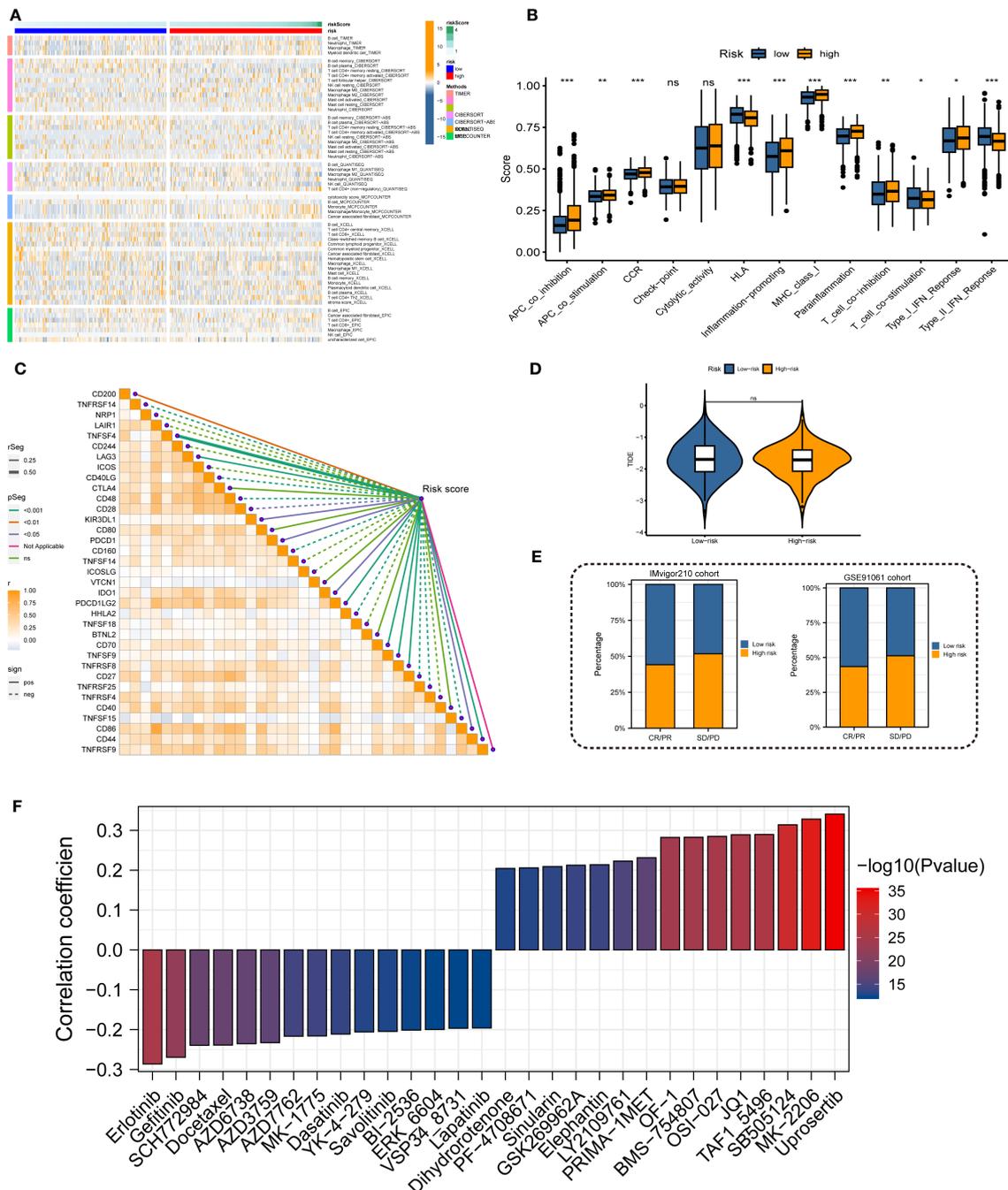
FIGURE 7
Immune landscape and treatment response prediction. **(A)** Estimation of immune cell infiltration in high- and low-risk teams. **(B)** Explorations of immunological responses in terms of the MRRS risk groups. **(C)** Correlations between MRRS and immune checkpoints. **(D)** TIDE algorithm identified the difference in immunotherapy response between high- and low-risk groups. **(E)** The prediction of immunotherapy response using IMvigor210 and GSE72094 cohorts. **(F)** The prediction of chemotherapy response of LUAD patients with different MRRSs. *$p< 0.05$, **$p< 0.01$, ***$p< 0.001$, ns, not significant.

high-risk patients had a high expression of immunosuppressive receptors and inhibitory ligands (PDCD1, PDCD1LG2, and LAG3). Immune checkpoint inhibitors, as we all know, have recently significantly improved the prognosis for patients with lung cancer (30), but they are only beneficial in a small subset of patients with

LUAD. The majority of LUAD patients were still receiving targeted therapy and chemotherapy as their basic treatments. The results revealed that the high-risk group responded better to erlotinib, gefitinib, SCH772984, and docetaxel. Overall, the immune landscape analysis showed that MRRS was involved in many
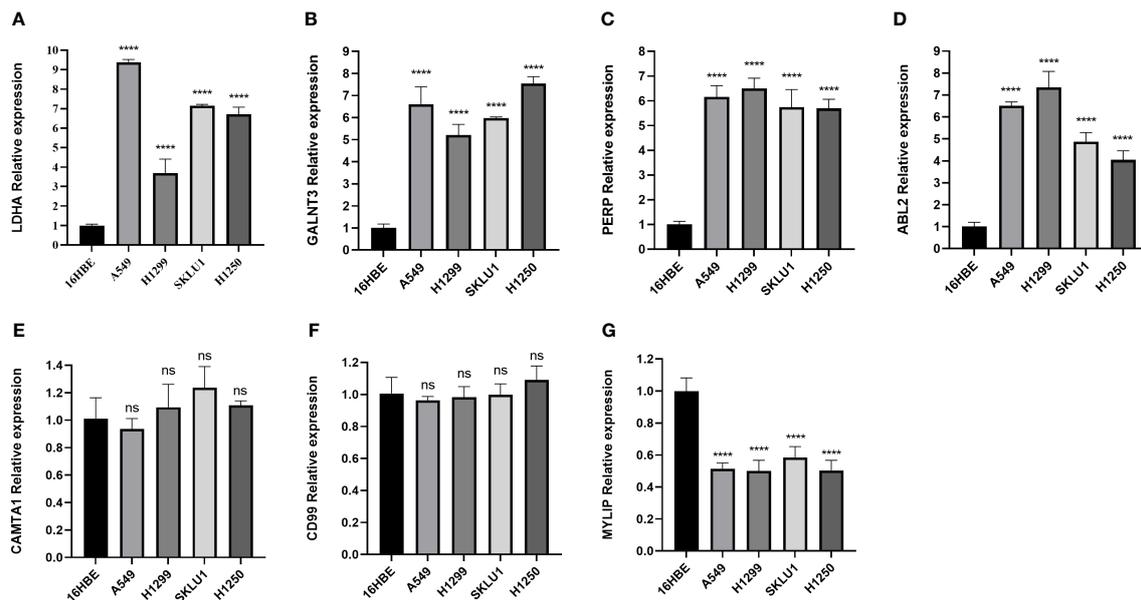
FIGURE 8
The qRT-PCR result for the MRRS-constructed genes. **(A)** LDHA. **(B)** GALNT3. **(C)** PERP, **(D)** ABL2, **(E)** CAMTA1, **(F)** CD99, **(G)** MYLIP.

immune responses and that, for high-risk patients, chemotherapy may be more efficient than immunotherapy.

In our research, the LDHA, CAMTA1, MYLIP, GALNT3, PERP, CD99, and ABL2 genes are vital ingredients of the MRRS. Among them, LDHA showed a substantial correlation with the prognosis of LUAD, which is consistent with earlier researches (31, 32). According to a recent study, phosphorylation and activation of LDHA can promote tumor invasion and metastasis (33). By regulation of LncRNA SGMS1-AS1 and miR-106a-5p38, MYLIP, a potential tumor suppressor gene in LUAD, may prevent the proliferation, invasion, and migration of LUAD cells (34). GALNT3 was developed to prevent lung cancer by preventing self-renewal and the development of a favorable tumor microenvironment (35). Previous studies showed that the knockdown of ABL2 dramatically reduced the brain metastases of LUAD cells (36, 37). According to the above findings, possible treatments targeting these essential genes may exist in the future.

Our study still had significant shortcomings, though. Since this research was retrospective, additional treatment and relapse records as well as prospective clinical investigations are needed to verify our findings. These essential genes will require additional experimental validation, and additional *in vivo* or *in vitro* experiments are needed to investigate the specific mechanisms of the genes.

5 Conclusion

In conclusion, we developed a novel malignancy-related signature, termed MRRS, which demonstrated a strong predictive capacity for OS in LUAD patients. Our findings shed light on the underlying mechanisms of LUAD progression and suggest that MRRS may serve as a promising prognostic and therapeutic marker 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

JHZ designed and started the research. MW, ZW, and JY analyzed the data and drafted the manuscript. JZ, JK, CZ, and XZ revised the manuscript. YM, QG, DL, JT, and TZ collected the data. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

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/fonc.2023.1198746/full#supplementary-material

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2021) 71(3):209–49. doi: 10.3322/caac.21660

2. Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. *Transl Lung Cancer Res* (2016) 5(3):288–300. doi: 10.21037/tlcr.2016.06.07

3. Hirsch FR, Scagliotti GV, Mulshine JL, Kwon R, Curran WJJr., Wu YL, et al. Lung cancer: current therapies and new targeted treatments. *Lancet* (2017) 389(10066):299–311. doi: 10.1016/s0140-6736(16)30958-8

4. Liang J, Cai W, Sun Z. Single-cell sequencing technologies: current and future. *J Genet Genomics* (2014) 41(10):513–28. doi: 10.1016/j.jgg.2014.09.005

5. Lai C, Wu Z, Li Z, Yu H, Li K, Tang Z, et al. A robust signature of immune-related long non-coding RNA to predict the prognosis of bladder cancer. *Cancer Med* (2021) 10(18):6534–45. doi: 10.1002/cam4.4167

6. Wu Z, Wang Y, Yan M, Liang Q, Li B, Hou G, et al. Comprehensive analysis of the endoplasmic reticulum stress-related long non-coding RNA in bladder cancer. *Front Oncol* (2022) 12:951631. doi: 10.3389/fonc.2022.951631

7. Xue Z, Huang K, Cai C, Cai L, Jiang CY, Feng Y, et al. Genetic programs in human and mouse early embryos revealed by single-cell RNA sequencing. *Nature* (2013) 500(7464):593–7. doi: 10.1038/nature12364

8. Hanahan D, Coussens LM. Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* (2012) 21(3):309–22. doi: 10.1016/j.ccr.2012.02.022

9. Wang F, Zhang Y, Hao Y, Li X, Qi Y, Xin M, et al. Characterizing the metabolic and immune landscape of non-small cell lung cancer reveals prognostic biomarkers through omics data integration. *Front Cell Dev Biol* (2021) 9:702112. doi: 10.3389/fcell.2021.702112

10. Shedden K, Taylor JM, Enkemann SA, Tsao MS, Yeatman TJ, Gerald WL, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med* (2008) 14(8):822–7. doi: 10.1038/nm.1790

11. Schabath MB, Welsh EA, Fulp WJ, Chen L, Teer JK, Thompson ZJ, et al. Differential association of STK11 and TP53 with KRAS mutation-associated gene expression, proliferation and immune surveillance in lung adenocarcinoma. *Oncogene* (2016) 35(24):3209–16. doi: 10.1038/onc.2015.375

12. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* (2018) 554(7693):544–8. doi: 10.1038/nature25501

13. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell* (2017) 171(4):934–949.e16. doi: 10.1016/j.cell.2017.09.028

14. Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat Med* (2018) 24(10):1550–8. doi: 10.1038/s41591-018-0136-1

15. Zhang X, Zeng L, Li Y, Xu Q, Yang H, Lizaso A, et al. Anlotinib combined with PD-1 blockade for the treatment of lung cancer: a real-world retrospective study in China. *Cancer Immunol Immunother* (2021) 70(9):2517–28. doi: 10.1007/s00262-021-02869-9

16. Xie J, Zhang J, Tian W, Zou Y, Tang Y, Zheng S, et al. The pan-cancer multi-omics landscape of FOXO family relevant to clinical outcome and drug resistance. *Int J Mol Sci* (2022) 23(24). doi: 10.3390/ijms232415647

17. Zeng J, Wu Z, Luo M, Xu X, Bai W, Xie G, et al. Development and validation of an endoplasmic reticulum stress long non-coding RNA signature for the prognosis and immune landscape prediction of patients with lung adenocarcinoma. *Front Genet* (2023) 14:1024444. doi: 10.3389/fgene.2023.1024444

18. Zheng S, Liang JY, Tang Y, Xie J, Zou Y, Yang A, et al. Dissecting the role of cancer-associated fibroblast-derived biglycan as a potential therapeutic target in immunotherapy resistance: a tumor bulk and single-cell transcriptomic study. *Clin Transl Med* (2023) 13(2):e1189. doi: 10.1002/ctm2.1189

19. Wang Y, Hou K, Jin Y, Bao B, Tang S, Qi J, et al. Lung adenocarcinoma-specific three-integrin signature contributes to poor outcomes by metastasis and immune escape pathways. *J Transl Int Med* (2021) 9(4):249–63. doi: 10.2478/jtim-2021-0046

20. Liu Y, Xun Z, Ma K, Liang S, Li X, Zhou S, et al. Identification of a tumour immune barrier in the HCC microenvironment that determines the efficacy of immunotherapy. *J Hepatol* (2023) 78(4):770–82. doi: 10.1016/j.jhep.2023.01.011

21. Wang Y, Fan JL, Melms JC, Amin AD, Georgis Y, Barrera I, et al. Multimodal single-cell and whole-genome sequencing of small, frozen clinical specimens. *Nat Genet* (2023) 55(1):19–25. doi: 10.1038/s41588-022-01268-9

22. Zhang J, Lu T, Lu S, Ma S, Han D, Zhang K, et al. Single-cell analysis of multiple cancer types reveals differences in endothelial cells between tumors and normal tissues. *Comput Struct Biotechnol J* (2023) 21:665–76. doi: 10.1016/j.csbj.2022.12.049

23. Schaal C, Chellappan S. Nicotine-mediated regulation of nicotinic acetylcholine receptors in non-small cell lung adenocarcinoma by E2F1 and STAT1 transcription factors. *PloS One* (2016) 11(5):e0156451. doi: 10.1371/journal.pone.0156451

24. Wei C, Dong X, Lu H, Tong F, Chen L, Zhang R, et al. LPCAT1 promotes brain metastasis of lung adenocarcinoma by up-regulating PI3K/AKT/MYC pathway. *J Exp Clin Cancer Res* (2019) 38(1):95. doi: 10.1186/s13046-019-1092-4

25. Zhou Y, Zhao Y, Ma W, Zhang L, Jiang Y, Dong W. USF1-CHCHD4 axis promotes lung adenocarcinoma progression partially via activating the MYC pathway. *Discovery Oncol* (2022) 13(1):136. doi: 10.1007/s12672-022-00600-3

26. Liu CX, Musco S, Lisitsina NM, Yaklichkin SY, Lisitsyn NA. Genomic organization of a new candidate tumor suppressor gene, LRP1B. *Genomics* (2000) 69(2):271–4. doi: 10.1006/geno.2000.6331

27. Brown LC, Tucker MD, Sedhom R, Schwartz EB, Zhu J, Kao C, et al. LRP1B mutations are associated with favorable outcomes to immune checkpoint inhibitors across multiple cancer types. *J Immunother Cancer* (2021) 9(3). doi: 10.1136/jitc-2020-001792

28. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *N Engl J Med* (2018) 378(22):2093–104. doi: 10.1056/NEJMoa1801946

29. Jardim DL, Goodman A, De Melo Gagliato D, Kurzrock R. The challenges of tumor mutational burden as an immunotherapy biomarker. *Cancer Cell* (2021) 39(2):154–73. doi: 10.1016/j.ccell.2020.10.001

30. Aldarouish M, Wang C. Trends and advances in tumor immunology and lung cancer immunotherapy. *J Exp Clin Cancer Res* (2016) 35(1):157. doi: 10.1186/s13046-016-0439-3

31. Yu C, Hou L, Cui H, Zhang L, Tan X, Leng X, et al. LDHA upregulation independently predicts poor survival in lung adenocarcinoma, but not in lung squamous cell carcinoma. *Future Oncol* (2018) 14(24):2483–92. doi: 10.2217/fon-2018-0177

32. Wang Z, Embaye KS, Yang Q, Qin L, Zhang C, Liu L, et al. Establishment and validation of a prognostic signature for lung adenocarcinoma based on metabolism-related genes. *Cancer Cell Int* (2021) 21(1):219. doi: 10.1186/s12935-021-01915-x

33. Jin L, Chun J, Pan C, Alesi GN, Li D, Magliocca KR, et al. Phosphorylation-mediated activation of LDHA promotes cancer cell invasion and tumour metastasis. *Oncogene* (2017) 36(27):3797–806. doi: 10.1038/onc.2017.6

34. Liu T, Yang C, Wang W, Liu C. LncRNA SGMS1-AS1 regulates lung adenocarcinoma cell proliferation, migration, invasion, and EMT progression via

miR-106a-5p/MYLI9 axis. *Thorac Cancer* (2021) 12(14):2104–12. doi: 10.1111/1759-7714.14043

35. Park MS, Yang AY, Lee JE, Kim SK, Roe JS, Park MS, et al. GALNT3 suppresses lung cancer by inhibiting myeloid-derived suppressor cell infiltration and angiogenesis in a TNFR and c-MET pathway-dependent manner. *Cancer Lett* (2021) 521:294–307. doi: 10.1016/j.canlet.2021.08.015

36. Hoj JP, Mayro B, Pendergast AM. A TAZ-AXL-ABL2 feed-forward signaling axis promotes lung adenocarcinoma brain metastasis. *Cell Rep* (2019) 29(11):3421–3434.e8. doi: 10.1016/j.celrep.2019.11.018

37. Hoj JP, Mayro B, Pendergast AM. The ABL2 kinase regulates an HSF1-dependent transcriptional program required for lung adenocarcinoma brain metastasis. *Proc Natl Acad Sci U.S.A.* (2020) 117(52):33486–95. doi: 10.1073/pnas.2007991117