scientific reports

OPEN

Constructing a sixteen lactate-related gene risk signature for LUAD to predict the prognosis and TME by machine learning

Jinjie Wang[1,2], Deshen Shan[1,2], Hailong Tang[1], Ming Zhao[1] & Hao Wang[1✉]

Although it is the most common subtype of lung cancer in clinical practice, lung adenocarcinoma (LUAD) was proven to be associated with a poor prognosis. In recent years, lactate metabolism has been considered an important biological mechanism in lung cancer. However, the mechanisms of lactate-related genes in tumour microenvironment (TME) of LUAD are unknown. Lactate-related genes modification patterns in 701 LUAD samples from TCGA and GEO database were systematically assessed on the basis of 16 lactate-related genes (LRGs). We also identified the correlation of these clusters to the TME and immune cell infiltration in the TME. After unsupervised clustering analysis was performed, the LUAD samples were divided into three lactate-related gene phenotypes on the basis of 36 prognostic lactate-related genes (PLRGs). These molecular subtypes have different immune cell infiltration characteristics and pathway enrichment. Using LASSO, a 16-LRG risk signature was constructed. The lactate-related signature demonstrated a stable and accurate ability to predict the prognosis of LUAD in patients. Furthermore, the lactate-related signature demonstrated good predictive ability for immune infiltrating cells, tumour mutation burden, and response to immunotherapy. The 16-LRG risk signature was subsequently verified in the GSE50081 GEO cohort. This study revealed the significant clinical utility of the 16-LRG risk signature in the understanding the TME and prognosis of LUAD. The 16-LRG risk signature is conducive to understanding immune cell infiltration in the LUAD TME and contributes to the selection of more effective immunotherapy strategies.

Abbreviations

LUAD	Lung adenocarcinoma
TME	Tumor microenvironment
LRG	Lactate-related gene
PLRGs	Prognostic lactate-related genes
LASSO	Least absolute shrinkage and selection operator
TCGA	The cancer genome atlas database
GEO	Gene expression omnibus database
AUC	Under the curve
OS	Overall survival
ssGSEA	Single-sample gene-set enrichment analysis
DEGs	Differentially expressed genes
ROC	Receiver operating characteristic
HR	Hazard ratio
PCA	Principal component analysis

[1]Department of Thoracic Surgery, Affiliated Huishan Hospital of Xinling College, Nantong University, Wuxi Huishan District People's Hospital, 2 Zhanqian North Road, Luoshe Town, Huishan District, Wuxi 214187, People's Republic of China. [2]Jinjie Wang and Deshen Shan: have contributed equally to this work. ✉email: 13848216885@163.com

Lung cancer is considered to be the most common cancer worldwide and is correlated with the prevalence of tobacco use[1,2]. In the past few decades, the incidence of lung cancer in China has been increasing[3]. The proportion of non-small cell lung cancer (NSCLC) among all types of lung cancer is 85%[4], and lung adenocarcinoma (LUAD) is the most common type of NSCLC[5]. Recently, research has shown the high heterogeneity and complexity of LUAD, which is not reflected in the traditional histological classification of LUAD. These findings show that tumour classification requires reliable prognostically relevant factors to select more effective treatment strategies for patients with lung cancer. Lactic acid provides energy for lung tumour cells, more so than glucose does[6]. Therefore, the lactate pathway may become a novel LUAD treatment target.

Going back to 1930, Otto Warburg proved that cancer cells prefer glycolysis even when their oxygen content is normal; this phenomenon is known as the "Warburg effect"[7]. Lactate, the reduction product of pyruvate via lactate dehydrogenase (LDH), is the final product of glycolysis[8]. In recent years, the role of lactate in tumours has been gradually discovered; for example, lactate is the tricarboxylic acid (TCA) cycle carbon source of NSCLC and provides energy for tumour cells[6], Bergers G and Fendt SM reported that lactate directly promotes the invasion and migration of cancer cells[9] and Karin Fischer at el. reported that a high concentration of lactate in the TME blocks the output of lactate in T cells, thus disrupting their metabolism and function[10]. These studies show that lactate metabolism has the potential to become a new target for tumor therapy[11].

Many genes have been confirmed to regulate lactate metabolism[12–16]. Although research related to lactate metabolism has become a hot topic in cancer treatment research, the correlation between the prognosis of lung adenocarcinoma (LUAD) in patients and the expression of lactate-related genes remains unclear. Therefore, to elucidate the heterogeneity of LUAD, identifying the molecular characteristics of lactate-related genes may aid in the early diagnosis and prevention of LUAD and provide clinical insights for the personalized treatment of individual LUAD patients.

In this study, 701 LUAD samples from the GEO cohort and the TCGA cohort were classified into phenotypes on the basis of 36 genes related to lactate metabolism. LASSO and Cox regression analyses were used to construct a new score model to quantify the lactate-related patterns of individual patients. Furthermore, the lactate-related signature demonstrated good predictive ability for immune infiltrating cells, tumour mutation burden, and response to immunotherapy. The validity of the 16-LRG risk signature was subsequently verified in the GEO cohort GSE50081. These results indicate that lactation modification plays a nonnegligible role in shaping the characteristics of individual tumour microenvironments. Therefore, we established a scoring system to quantify the lactation modification patterns of individual patients, aiming to address early diagnosis and prognosis prediction in patients with LUAD.

Materials and methods

Acquisition of LUAD dataset

We downloaded three LUAD datasets: the TCGA-LUAD, GSE30219, and GSE37745 datasets. We removed LUAD samples with missing clinical data (such as stage and survival time), ultimately yielding 701 LUAD samples.

Acquisition of LRGs

We accessed the MSigDB website (https://www.gsea-msigdb.org) and searched for the specific subset using "lactate" as the keyword. We then downloaded the retrieved specific subset and ultimately obtained 285 LRGs.

Consensus clustering of PLRGs

The "limma" R software package was used for survival analysis, and 36 genes were identified as prognostic lactate-related genes (PLRGs) of LUAD, with $P < 0.01$. By employing the Euclidean distance metric and the partition around median (PAM) clustering algorithm, the 'Consensus Cluster Plus' R package was used for cluster analysis, and 1000 cycles were conducted to ensure the reliability of the classification results. We chose the case with the best clustering stability between k values of 2 and 9, where k = 3.

Application of GSVA and ssGSEA

GSVA was used to analyse the differences in gene set enrichment among the three lactate-related clusters. The gene set "c2.cp.kegg.v7.4.symbols.gmt" from the MSigDB database (https://www.gsea-msigdb.org) was obtained to conduct GSEA. As an extension of the GSEA method, ssGSEA was applied to calculate the degree of infiltration of 23 kinds of immune cells in each sample with the TIMER, CIBERSORT, CIBERSORT–ABS, QUANTISEQ, MCPCOUNTER, XCELL and EPIC methods. Using the 'GSEABase' and 'GSVA' R packages, differential expression-related pathways and immune cells were identified[17].

DEGs

DEGs among LUAD patients in the three LA modification patterns were identified with adjusted $P < 0.0001$ and log2|FC|>1 by Bonferroni method[18,19]. GO annotation and KEGG pathway enrichment analyses were performed on the genes identified by the 'clusterProfiler' R package[20–22].

Construction and validation of the lactate-related signature

Univariate Cox analysis was performed to screen PLRGs. LASSO was used to construct a sixteen-gene prognostic signature based on the screened LRGs. The risk score was determined with the formula: risk score = $\sum$(gene × coefficient). To verify the 16-LRG risk signature, an external dataset, GSE50081, was downloaded.

Statistical analysis

R 4.2.0 is used for data processing.

Results

Screening and genetic variation profile of the prognostic lactate-related genes in LUAD

First, a total of 285 genes were downloaded from the MSigDB, and 36 genes were identified as prognostic lactate-related genes (PLRGs) of LUAD, with $P < 0.01$ (Table 1). Prevalent CNV changes were detected in the PLRGs. Among them, SLC25A4, SLC13A5 and C1QBP had more CNV deletions. However, the CNV amplification of NDUFS6, COX6B1 and MRPL12 was greater (Fig. 1A). The somatic mutations and copy numbers of 36 lactate-related genes in LUAD were analysed. Only 75 (13.37%) of the 561 samples had lactate-related gene mutations, and the overall mutation frequency was low (Fig. 1B). The location of CNV alterations on the chromosome for PLRGs were determined (Fig. 1C). Finally, we analysed the differences in the mRNA expression of selected genes between tumour and normal samples (Fig. 1D).

Three phenotypes were identified by prognostic lactate-related genes

The correlations between PLRGs are shown in a network diagram (Fig. 2A). For further analysis, we applied a consensus clustering algorithm to classify LUAD samples into subgroups on the basis of the 36 PLRGs. When k = 2−9, k = 3 indicates the best stability; thus, LUAD patients were classified into three different lactate-related molecular clusters (LA clusters) (Fig. 2B–D). Principal component analysis (PCA) revealed three different distribution characteristics, which indicated significant consensus clustering results (Fig. 2E). Among LAcluster A (n = 180), LAcluster B (n = 291), and LAcluster C (n = 230), LUAD cases in LAcluster B had the best prognosis, and LUAD cases in LAcluster C had the worst prognosis (Fig. 2F). Different clinicopathological characteristics of individual PLRGs are shown in a heatmap (Fig. 2G).

| id	HR	HR.95L	HR.95H	*p* value	km
LDHA	1.887588	1.536741	2.318536	1.40E-09	7.37E−11
NDUFA12	1.878866	1.424798	2.477641	7.89E-06	2.97E−05
SLC3A1	1.529748	1.245935	1.878211	4.91E-05	0.064595
MRPL12	1.455573	1.197723	1.768935	0.000161	1.79E−05
ETHE1	1.396104	1.168366	1.668232	0.00024	9.74E−06
COX8A	1.581672	1.236595	2.023044	0.000261	3.16E−06
SLC25A42	0.654933	0.519879	0.82507	0.000328	0.000241
MRPL44	1.670862	1.261496	2.213071	0.000344	2.43E−05
PYGL	1.252772	1.102126	1.42401	0.000566	8.64E−06
FBP1	0.806944	0.713913	0.912099	0.000599	0.000131
DNM1L	1.484825	1.181857	1.865458	0.000686	7.89E−05
HAGH	0.637038	0.490876	0.826722	0.000697	0.000143
NDUFS6	1.386938	1.139045	1.688782	0.001131	5.74E−05
MRPL3	1.505122	1.175941	1.92645	0.001166	0.000195
LDHD	0.822562	0.729505	0.927489	0.001429	0.000778
CYP27A1	0.816377	0.720571	0.924921	0.001446	0.000334
TACO1	1.632588	1.206848	2.208518	0.001475	0.001681
MDH2	1.472153	1.15913	1.869708	0.001521	4.66E−05
DARS2	1.334788	1.115581	1.597068	0.001605	0.000303
NDUFA9	1.487514	1.160408	1.906827	0.001724	0.000532
BOLA3	1.382767	1.126374	1.697521	0.001953	8.42E−05
PNPT1	1.341981	1.105817	1.628582	0.002897	5.62E−06
MLYCD	0.60541	0.432375	0.847693	0.003477	0.002483
NDUFB3	1.501713	1.142739	1.973453	0.003531	0.00013
HPDL	1.19563	1.058395	1.35066	0.004075	0.001567
ISCU	0.638953	0.46832	0.871758	0.004717	0.006262
MRPS7	1.44729	1.116911	1.875395	0.00517	0.000408
MRPS16	1.45413	1.116423	1.89399	0.005491	0.001485
C1QBP	1.306624	1.078194	1.583451	0.006375	0.005107
SLC13A5	1.272535	1.067241	1.51732	0.007254	0.000765
ACAT1	0.758524	0.619908	0.928136	0.007269	0.000948
NDUFA8	1.462401	1.107408	1.931191	0.007382	0.005244
SLC25A4	0.75261	0.608875	0.930276	0.00858	0.00014
COX6B1	1.313202	1.071148	1.609954	0.008763	0.002856
COQ2	1.425049	1.09027	1.862626	0.009527	0.000657
USP18	1.234567	1.051933	1.448909	0.009885	0.01041

Table 1. 36 prognostic lactate-related genes (PLRGs) of LUAD.
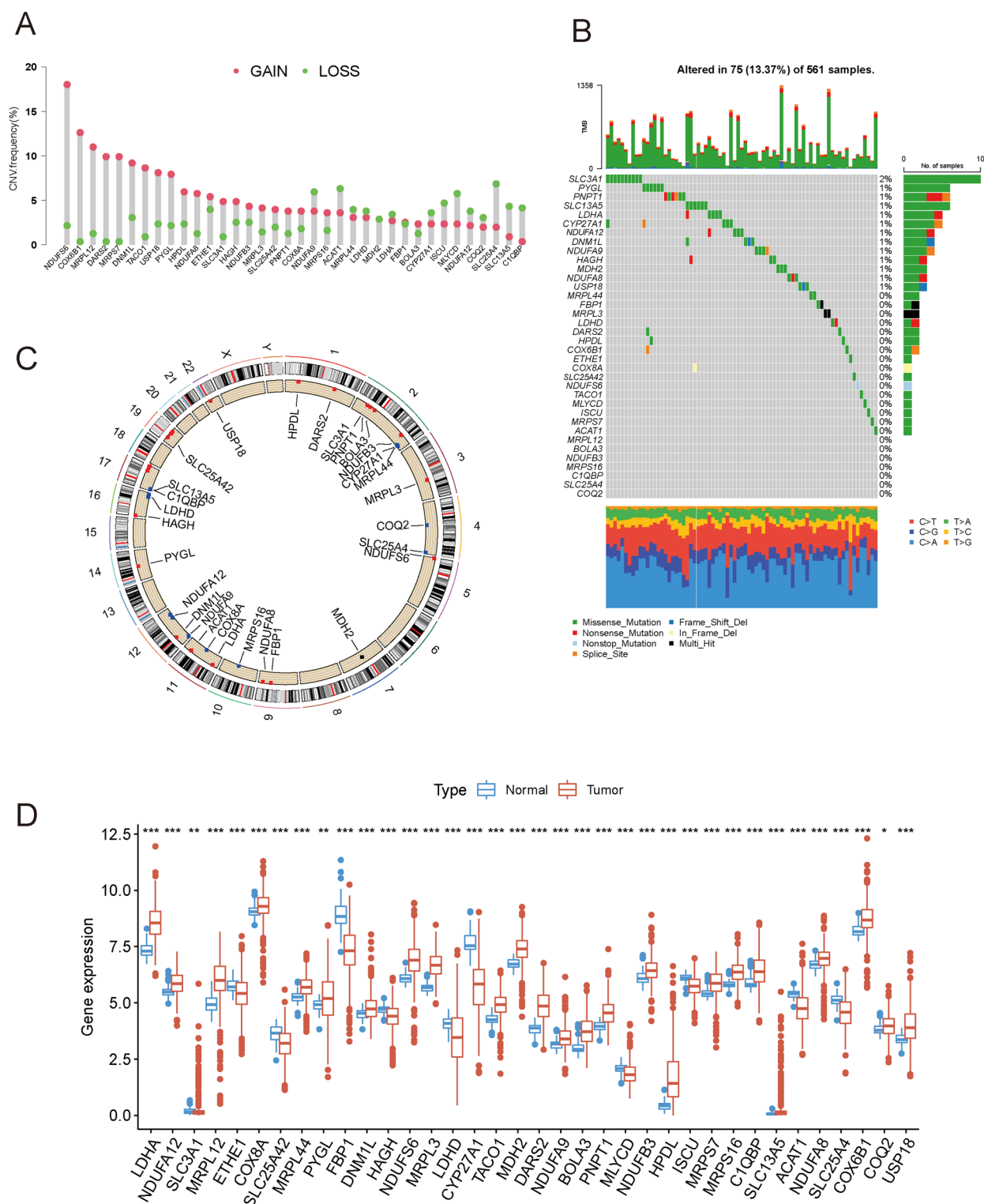
Fig. 1. Genetic variation of PLRGs. (**A**) The CNV variation frequency of PLRGs. (**B**) Mutation frequency of PLRGs. (**C**) Location of the CNV alteration of PLRGs on chromosomes. (**D**) Boxplot of PLRGs expression levels.

Next, we performed GSVA enrichment analysis. Compared with LAcluster A, LAcluster B was more enriched in immune activation, including the JAK–STAT and Fc epsilon RI signalling pathways[23,24]. Moreover, the enriched pathways in LAcluster A were closely related to cellular mechanisms, and LAcluster C was predominantly associated with tumour metabolism, such as the P53 signalling pathway and purine metabolism (Fig. 2H). Finally, using ssGSEA, the enrichment grades of 23 kinds of immune cells in 3 LAclusters were calculated. As
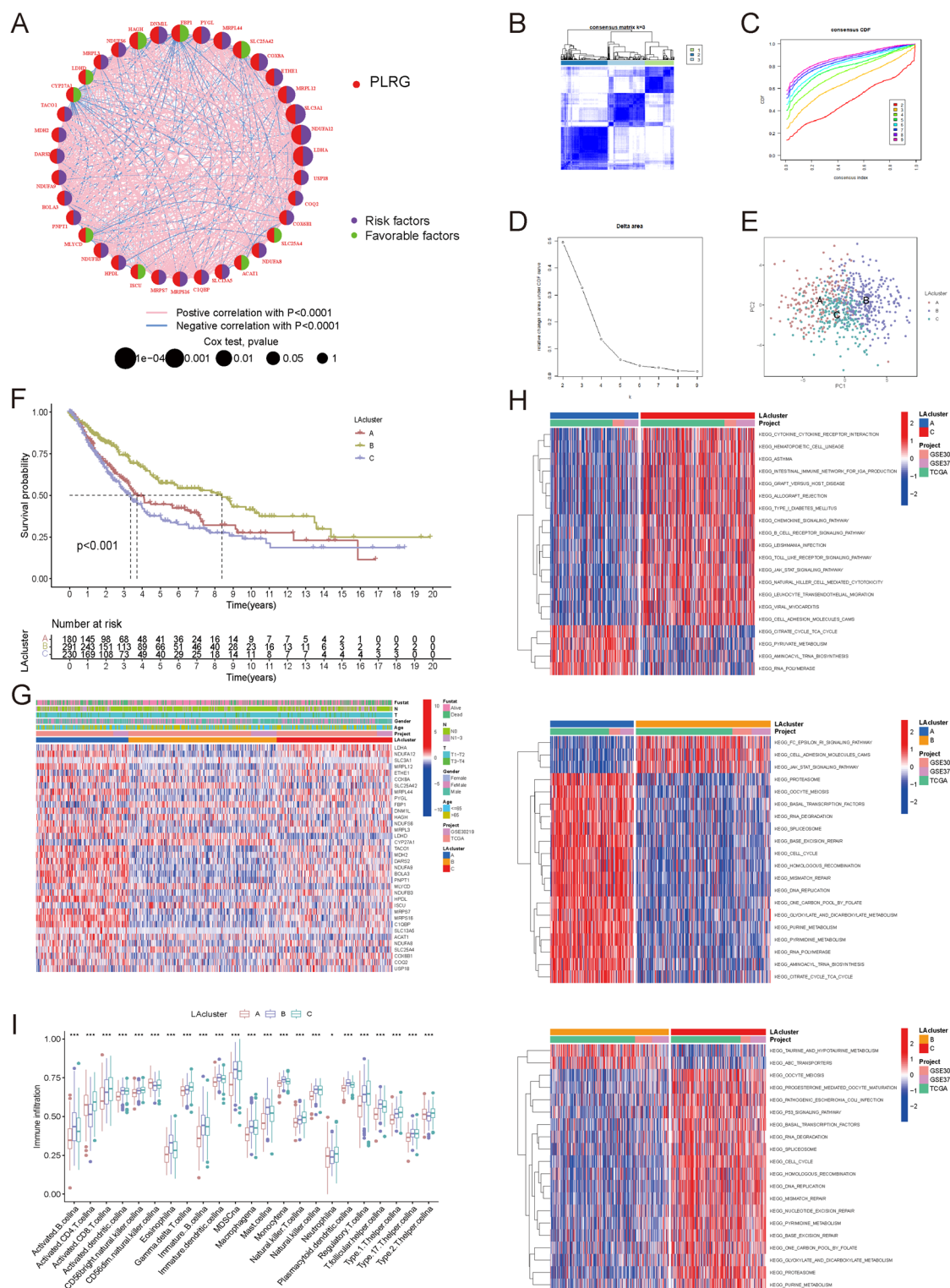
Fig. 2. The construction of LAclusters. (**A**) Correlations and prognosis of the 36 PLRGs in patients with LUAD. (**B**) Unsupervised clustering of the 36 PLRGs and optimal consensus matrices for k = 3. (**C**) Consensus clustering CDF for k = 2–9 (**D**) The CDF curve of consensus clustering for k = 2–9. (**E**) PCA of three LAclusters. (**F**) The overall survival of three LAclusters by KM curves. (**G**) Heatmaps of the distribution of 36 PLRGs in the three LAclusters. (**H**) GSEA analysis in the three LAclusters. (**I**) Degree of immune cells infiltration of three LAclusters.

shown in Fig. 2I, 13 kinds of immune cells had increased degrees of immune infiltration in LAcluster B, which again proved that LAcluster B was associated with immune activation.

To further investigate the latent biological function of the lactate-related subtypes, we analysed the differences in LA-related gene expression among the three subtypes and identified 1208 overlapping LA phenotype-related DEGs (Supplementary Fig. 1A). Supplementary Table 1 shows the list of DEGs. GO and KEGG analyses were subsequently performed, and the significantly enriched biological processes are shown in Supplementary Fig. 1B–E.

Establishment of a sixteen-gene prognostic signature

Cox regression analysis revealed 36 prognosis-related genes (Fig. 3A). Next, LASSO Cox regression was used to identify key genes, and the coefficient of genes was computed (Fig. 3B, C). Sixteen genes, including six protective genes (SLC25A42, HAGH, LDHD, MLYCD, ISCU and ACAT1) and ten risk genes (LDHA, NDUFA12, SLC3A1, ETHE1, MRPL44, PYGL, MDH2, DARS2, HPDL and NDUFA8), were selected to construct the prognostic signature. The coefficients of the sixteen genes are listed in Table 2. The risk score of each sample was determined by the formula: risk score = $\sum$(gene × coefficient). The samples were subsequently divided into a high-risk group (HRG) and a low-risk group (LRG) according to the median risk score. The distribution of risk scores and sample survival status are shown in Fig. 3D. The heatmaps show the expression of 16 genes in the two risk groups (Fig. 3E). In addition, survival analysis revealed that compared with LRG samples, HRG samples from LUAD patients were associated with poorer prognosis and worse survival (Fig. 3F). Cox univariate (Fig. 3G) and multivariate regression (Fig. 3H) analyses revealed that the 16-LRG risk signature was related to the overall survival rate of LUAD patients and could predict the prognosis of LUAD in patients.

Assessment of the correlation between the prognostic signature and clinical features

A ROC curve was used to evaluate the predictive value of the prognostic models compared with that of common cancer pathological features. The AUC of the 16-LRG risk signature was 0.692 (Fig. 4A). The areas under the curve were 0.692, 0.684 and 0.697 for the 1-year, 3-year and 5-year curves, respectively (Fig. 4B). A nomogram was constructed to assess the intrinsic value of the 16-LRG risk signature. (Fig. 4C). Calibration plots for 1-year, 3-year and 5-year OS were constructed to visualize the accuracy of the nomograms (Fig. 4D). We constructed a heatmap of the different clinicopathological characteristics associated with individual genes (Fig. 4E). The attribute changes in individual LUAD samples were visualized by an alluvial diagram (Fig. 4F).

We further analysed the OS of patients with different clinicopathological features in the high- and low-risk score groups. The samples were stratified by patient age into an older group (> 65) and a younger group (≤ 65), and the high-risk score group had worse survival than the low-risk score group did regardless of age (Supplementary Fig. 2A, B). As shown in Supplementary Fig. 2C and D, the outcomes of both sexes could be predicted with the 16-LRG risk signature. LUAD patients in the high-risk score group had worse OS in all stages (N0, N1–3, T1–T2 and T3–T4) (Supplementary Fig. 2E–H). To summarize, stratified analysis revealed that the 16-LRG risk signature had prognostic value regardless of age, sex, T stage and N stage.

GSEA, GO and KEGG analysis

The GSEA results revealed that tumours in the HRG were more enriched in tumour-promoting pathways and cellular mechanisms, such as the p53 signalling pathway, purine metabolism, the cell cycle, homologous recombination and DNA replication. In addition, the metabolic pathways associated with LRG, such as taurine and hypo taurine metabolism and primary bile acid biosynthesis, were more enriched (Fig. 5A). DEGs were defined as those whose log2|FC| was > 1 and whose FDR was < 0.05 (Fig. 5B). GO and KEGG enrichment analyses were then performed on the DEGs (Fig. 5C–F). Overall, the results revealed that the 16-LRG risk signature was associated with tumour and biological metabolism.

Immune microenvironment analysis of 16-LRG risk signature

Firstly, using different methods, we explored the immune microenvironment of two risk groups (Fig. 6A). The link between the prognostic signature and TME-infiltrating immune cells was analysed (Fig. 6B). Among the 23 kinds of common immune cells, 14 differed significantly between the HRG and LRG (Fig. 6C). Interestingly, immune cells with tumour suppressive effects, such as activated B cells, eosinophils, T-follicular helper cells and plasmacytoid dendritic cells, were more highly infiltrated in tumours in the LRG[25–27]. Through immune function analysis, APC coinhibition, CCR, cytolytic activity, MHC class-I, T-cell coinhibition and parainflammation were shown to be activated in the HRG. HLA and type-II IFN responses were more strongly activated in the LRG (Fig. 6D). To determine whether patients in the HRG had better immunotherapy responses, the differences in immune target expression between the two risk groups were analysed (Fig. 6E). Effective immunotherapeutic targets such as PDCD1 (PD-1), CD274 (PD-L1) and IDO1 may provide better therapeutic effects for HRG patients.

Tumor mutation burden of the 16-LRG risk signature

Because the TMB is associated with the efficacy of immunotherapy, cancer-related gene mutation data were utilized to assess the TMB levels in the two risk groups. First, the top 15 genes with the highest mutation frequency are shown in Fig. 7A and B by waterfall plots. We defaulted to plotting the TMB data transformed by log10 and calculated the TMB scores for both groups. Compared with the LRG, the HRG had higher TMB (Fig. 7C). High TMB indicates a poor prognosis in various cancers. In our study, the OS of the high-risk group was poor, and the results were consistent with those of other studies. As shown in Fig. 7D, low TMB is associated with a worse prognosis in patients with LUAD. Interestingly, low-risk patients with different TMBs had better
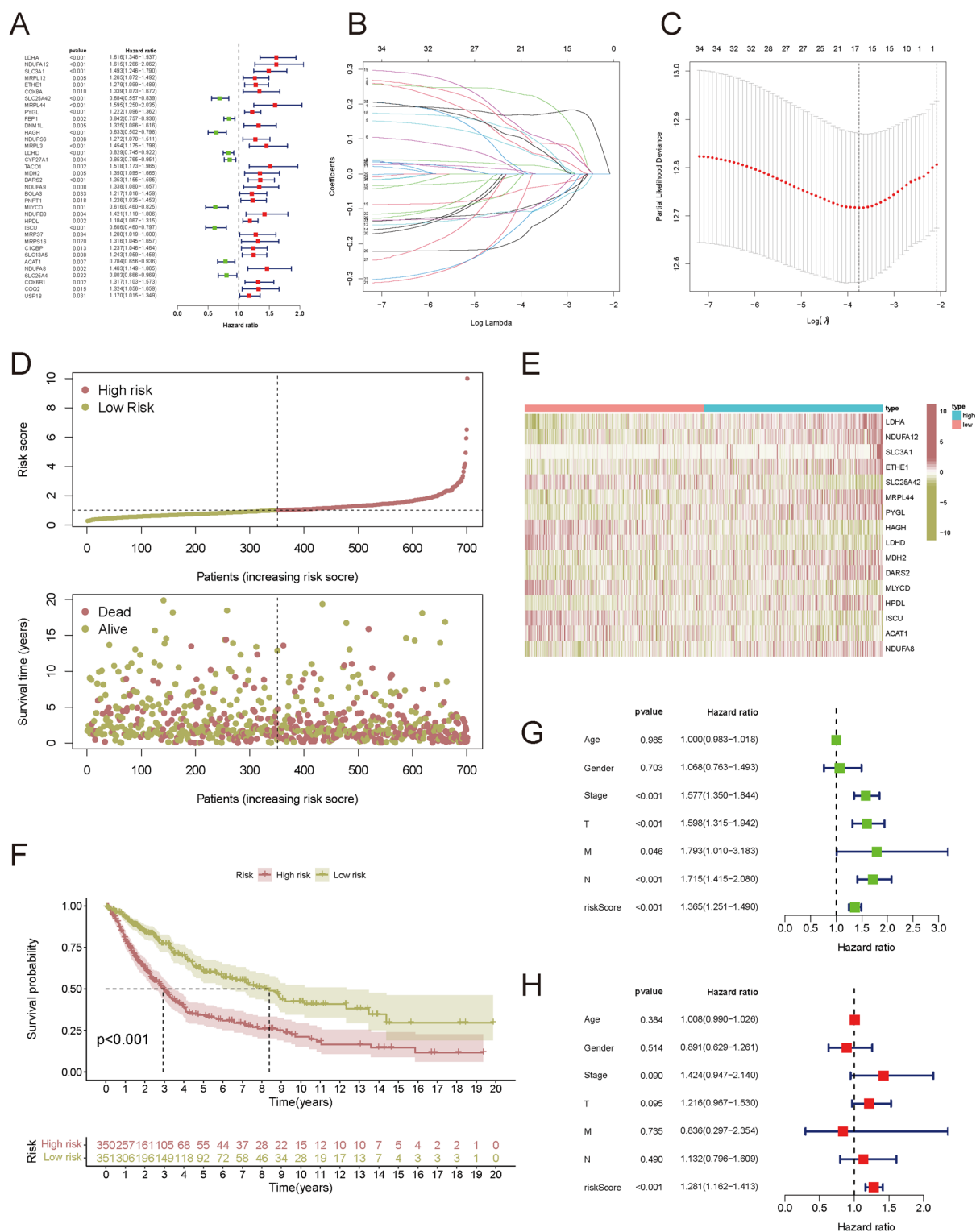
Fig. 3. Establishing the 16-LRG risk signature and analysis of independent prognostic factor. (**A**) Forest plots showing 36 prognostic LRGs based on Cox univariate regression analysis. (**B**, **C**) Cvfit and lambda curves of LASSO regression by minimum criteria. (**D**) Distribution of the risk scores and survival status in high- and low-risk groups. (**E**) Distribution of the 16 genes in high- and low-risk groups. (**F**) The overall survival of two risk groups by KM curves. Univariate Cox regression analysis (**G**) and multivariate Cox regression analysis (**H**) of clinical features regarding the 16-LRG risk signature.

id	coef
LDHA	0.158335
NDUFA12	0.083557
SLC3A1	0.264693
ETHE1	0.091899
SLC25A42	− 0.12164
MRPL44	0.084094
PYGL	0.039661
HAGH	− 0.15262
LDHD	− 0.05697
MDH2	0.183567
DARS2	0.144324
MLYCD	− 0.16598
HPDL	0.015465
ISCU	− 0.20216
ACAT1	− 0.16381
NDUFA8	0.1566

Table 2. The coefficients of the 16 key genes determined by Lasso Cox regression.

OS than high-risk patients with different TMBs did (Fig. 7E). In summary, compared with the TMB, the 16-LRG risk signature was more reliable for predicting the prognosis of LUAD.

Verification of the 16-LRG risk signature with an independent LUAD dataset

The GEO dataset GSE50081 was downloaded for independent external validation. The risk scores of the GSE50081 patients were calculated by the same formula. By the median risk score, the GSE50081 patients were separated into the HRG and the LRG. The distributions of patient risk scores and survival status are shown in Fig. 8A. As expected, the K‒M survival curve revealed that patients in the HRG had worse overall survival (Fig. 8B). The expression of 16 LRGs in GSE50081 patients was drawn as a heatmap (Fig. 8C). As shown by the ROC curve, the AUC of this risk score model was 0.721, which was greater than that of the other common clinical features (Fig. 8D). Time-dependent ROC curves were constructed, and the AUCs at 1, 2, and 3 years were 0.721, 0.623, and 0.657, respectively (Fig. 8E). Finally, we conducted expression and survival analyses on the genes used to establish the model (Supplementary Fig. 3A,B).The Results confirmed the reliability of these genes.

Discussion

LUAD, a highly heterogeneous tumour with a low overall survival rate, is the primary subtype of NSCLC[28]. Although we have made some progress in the chemotherapy and immunotherapy of LUAD, the overall treatment results are still not satisfactory. Owing to the complexity of the TME and strong heterogeneity of LUAD, many patients exhibit drug resistance and insensitivity to treatment[29]. Therefore, identifying a new LUAD classification to contribute to better patient treatment strategies is urgent. In recent years, increasing evidence has confirmed the role of lactic acid in promoting processes such as migration, immune escape and proliferation in cancer[30]. High lactate levels have the connection with increased metastasis, tumour recurrence and poor result[11]. These findings suggest that lactic acid can become a tumour treatment target. The role of lactic acid in lung cancer is receiving increasing attention. Glucose uptake and increased glycolytic activity are metabolic characteristics of lung cancer cells[31]. Previous studies have proven that human lung tumours may use lactate as a carbon source[32]. In terms of lung cancer immunity, lactate from tumours inactivates immune cell proliferation and effector function by acidifying the tumour microenvironment and upregulating PD-L1 expression[10,33]. According to the recent literature, increased extracellular lactate levels through the Notch1/TAZ axis inhibit cytotoxic T-cell activity, leading to the invasive characteristics of lung cancer cells[34]. Therefore, research on the mechanism through which lactate promotes immunotherapy efficacy in LUAD is important. However, the relationship between LUAD and the comprehensive role of lactate-related genes is not well defined. Therefore, in this study, we explored risk signatures for predicting prognosis in LUAD patients and the immune status of their TME. Ultimately, a 16-LRG risk signature was constructed to predict the prognosis and effect of immunotherapy of LUAD in patients.

In this study, on the basis of the expression of the 36 PLRGs, LUAD patients were divided into three subtypes. Using GSVA enrichment analysis, we further analysed the reasons for the differences among the three subtypes, and it was obvious from the findings that the enriched pathways associated with immune activation in LAcluster B and LAcluster A were significantly enriched in cellular mechanisms and that LAcluster C was predominantly related to tumour mechanisms. Finally, the correlation between the three lactate-related subtypes and TME cell infiltration was assessed.

Next, we identified 16 LRGs among the 36 PLRGs to construct a risk signature by LASSO and Cox regression analyses. LDHA catalyses the conversion of pyruvate to lactate, fuelling histone lactylation, which promotes gene expression linked to cancer progression and immune evasion[35]. As a component of mitochondrial complex
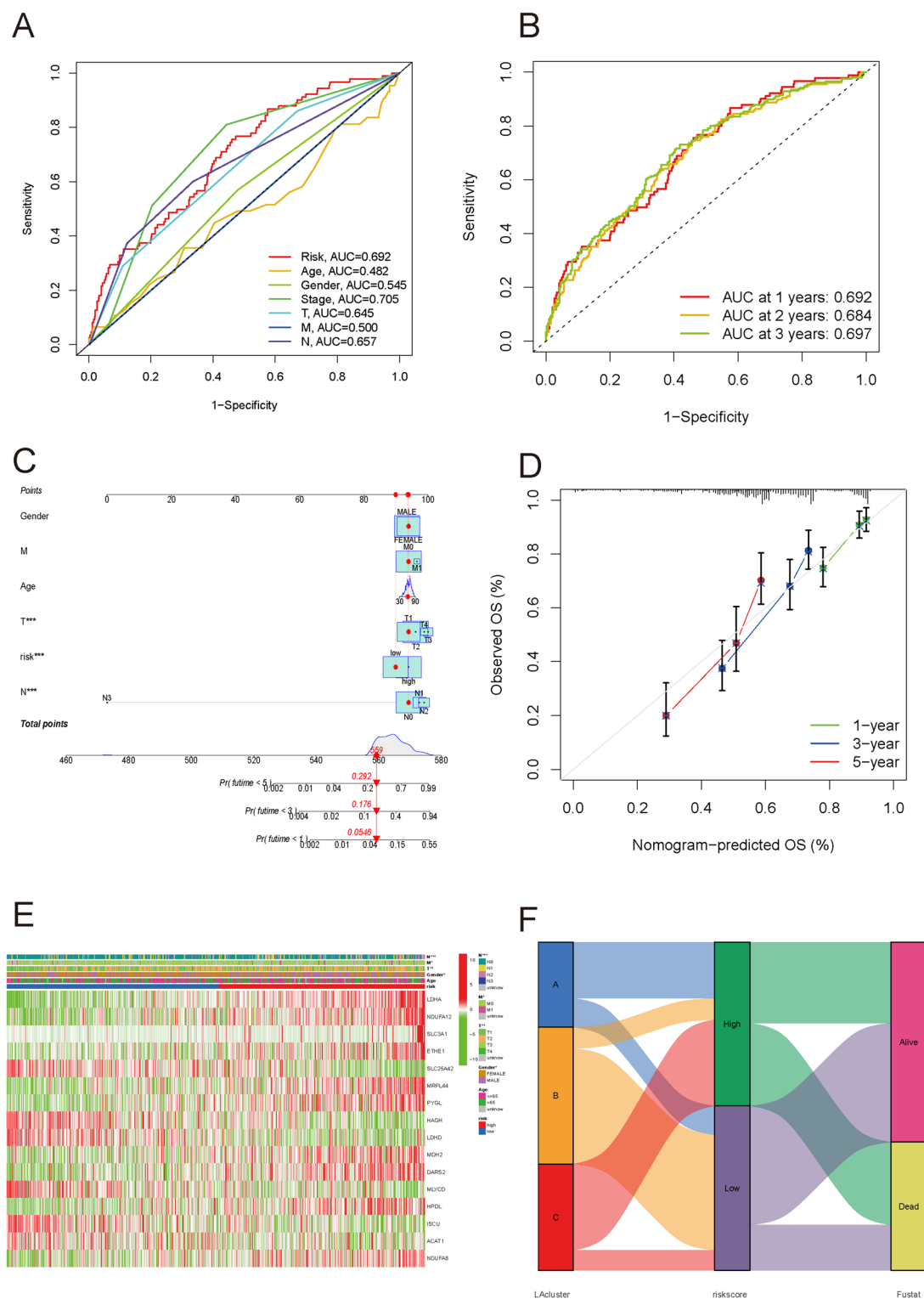
Fig. 4. Correlation between clinical information and the 16-LRG risk signature. (**A**) ROC curves of clinical features and risk scores. (**B**) ROC curves of the 16-LRG risk signature forecasting 1-, 2-, and 3-year OS. (**C**) Nomogram of the 16-LRG risk signature, age, gender, and TNM stage. (**D**) Calibration curves for internal validation of the nomogram according to 1-, 3-, and 5-year OS rate of LUAD samples. (**E**) The distribution of different clinicopathological characteristics in the 16 LRGs. (**F**) The alluvial diagram of the transformations of LAcluster, 16-LRG risk signature (riskscore), and survival status (Fustat).
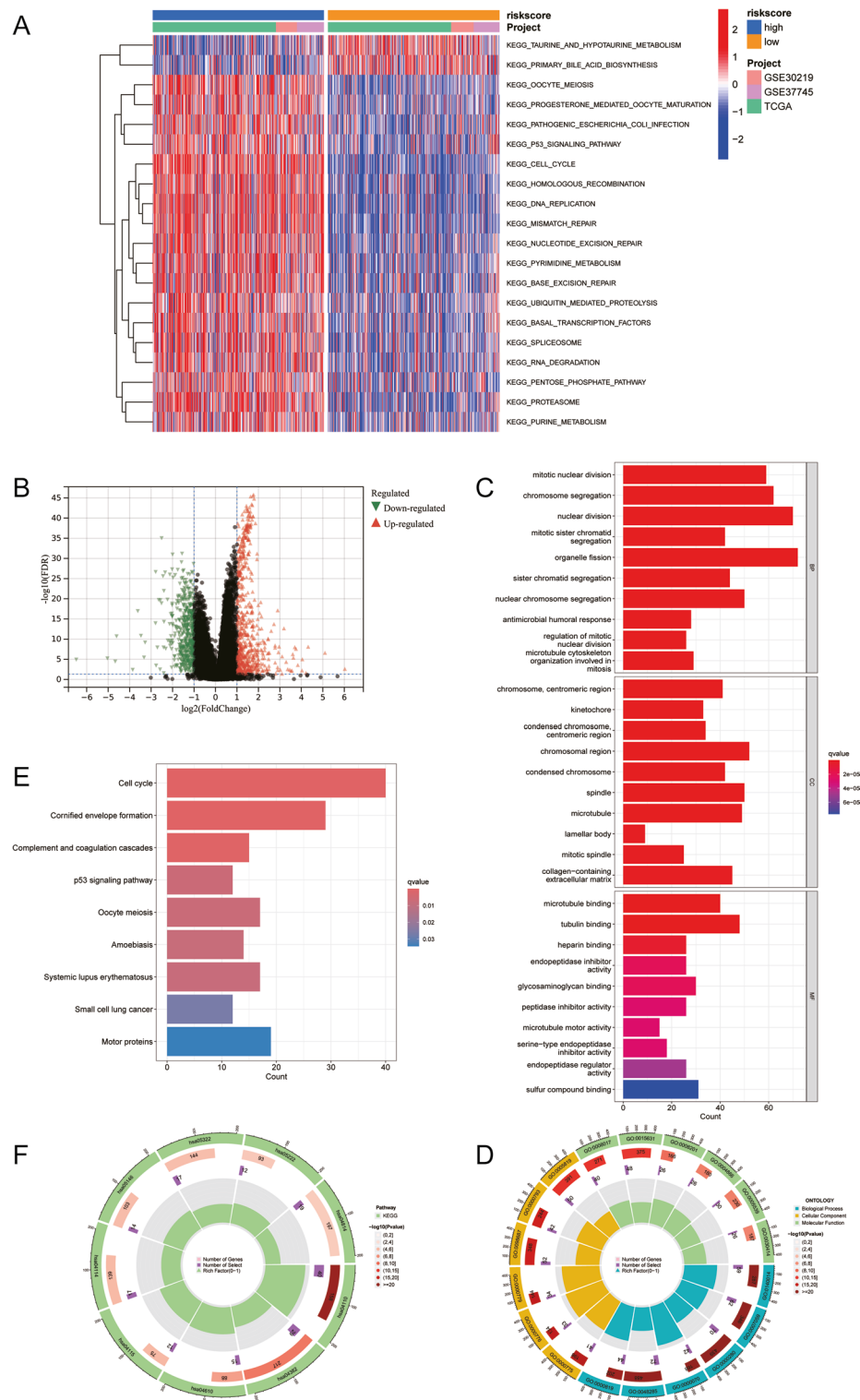
Fig. 5. GSEA, GO and KEGG analysis. (**A**) GSEA analysis of the high- and low-risk groups. (**B**) Volcano plot of DEGs between high- and low-risk groups. (**C**, **D**) GO analysis by DEGs. (**E**, **F**) KEGG analysis indicating pathways were significant enriched.

I, NDUFA12 dysfunction can increase lactate production, indirectly influencing the lactylation landscape that supports tumour growth[36]. SLC3A1 is part of a cysteine/glutamate transporter, influencing redox balance and potentially affecting the metabolic shift to lactate generation in cancer cells[37]. ETHE1 mutations cause sulfide toxicity that inhibits mitochondrial respiration, potentially diverting metabolism towards aerobic glycolysis and lactate production in tumours[38]. As a mitochondrial ribosomal protein, MRPL44 is essential for oxidative
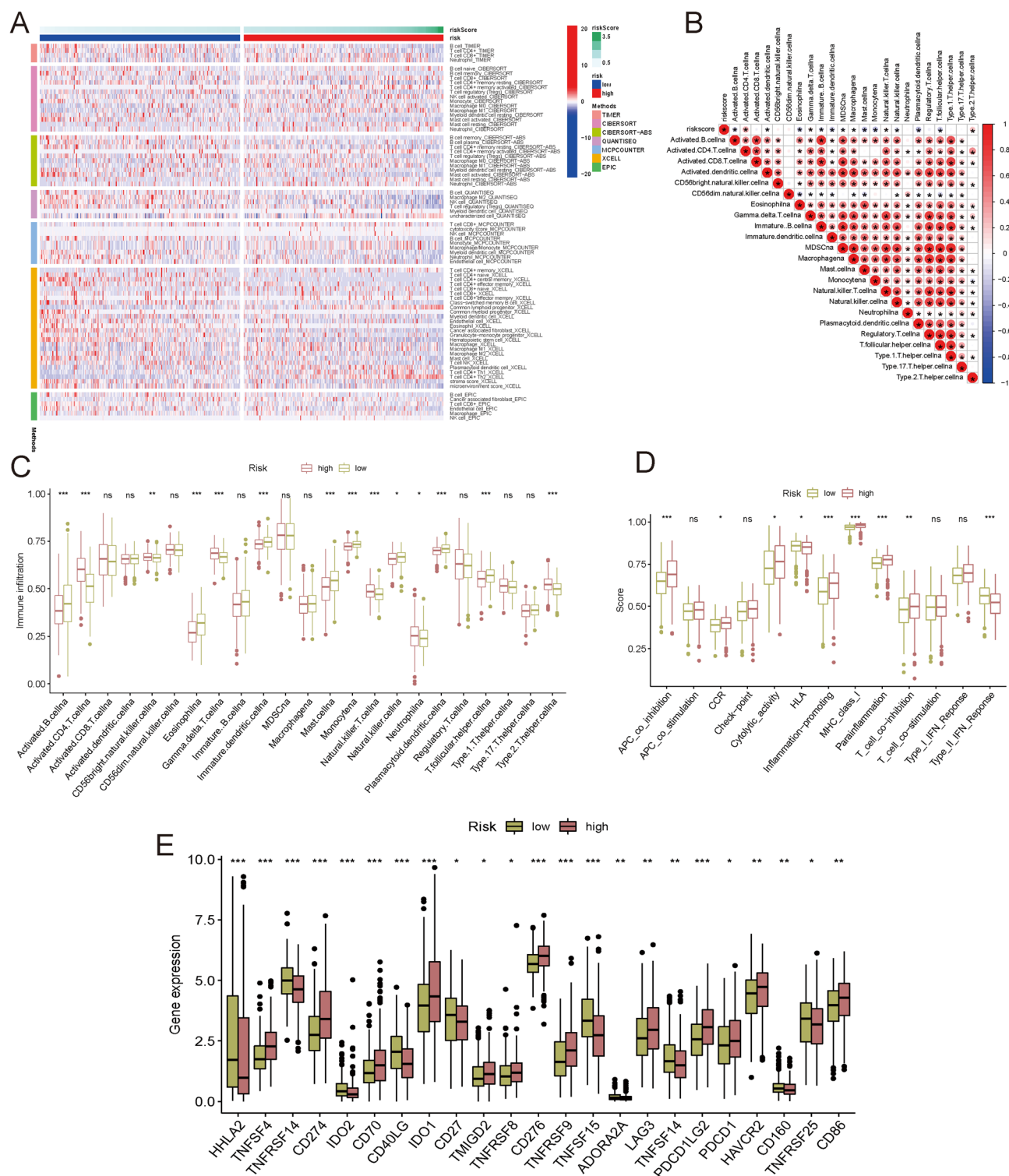
Fig. 6. Immune-related analysis of high-risk group and low-risk group. (**A**) Heatmap of immune microenvironment in high- and low-risk group. (**B**) Connection of TME-infiltrating cells and the 16-LRG risk signature. (**C**) Boxplot of immune cells infiltration. (**D**) Immune function analysis. (**E**) Immune checkpoints analyses between LRG and HRG.

phosphorylation; its deficiency can promote a glycolytic switch and increase the availability of lactate for lactylation[39]. PYGL encodes glycogen phosphorylase, which breaks down glycogen to provide glucose for glycolysis, thereby supplying the pyruvate and lactate necessary for lactylation in cancer[40]. HAGH degrades toxic glycolytic byproducts such as methylglyoxal, protecting cancer cells from damage and ensuring the continued glycolytic flux that generates a lactylation substrate[41]. LDHD oxidizes lactate back to pyruvate, potentially
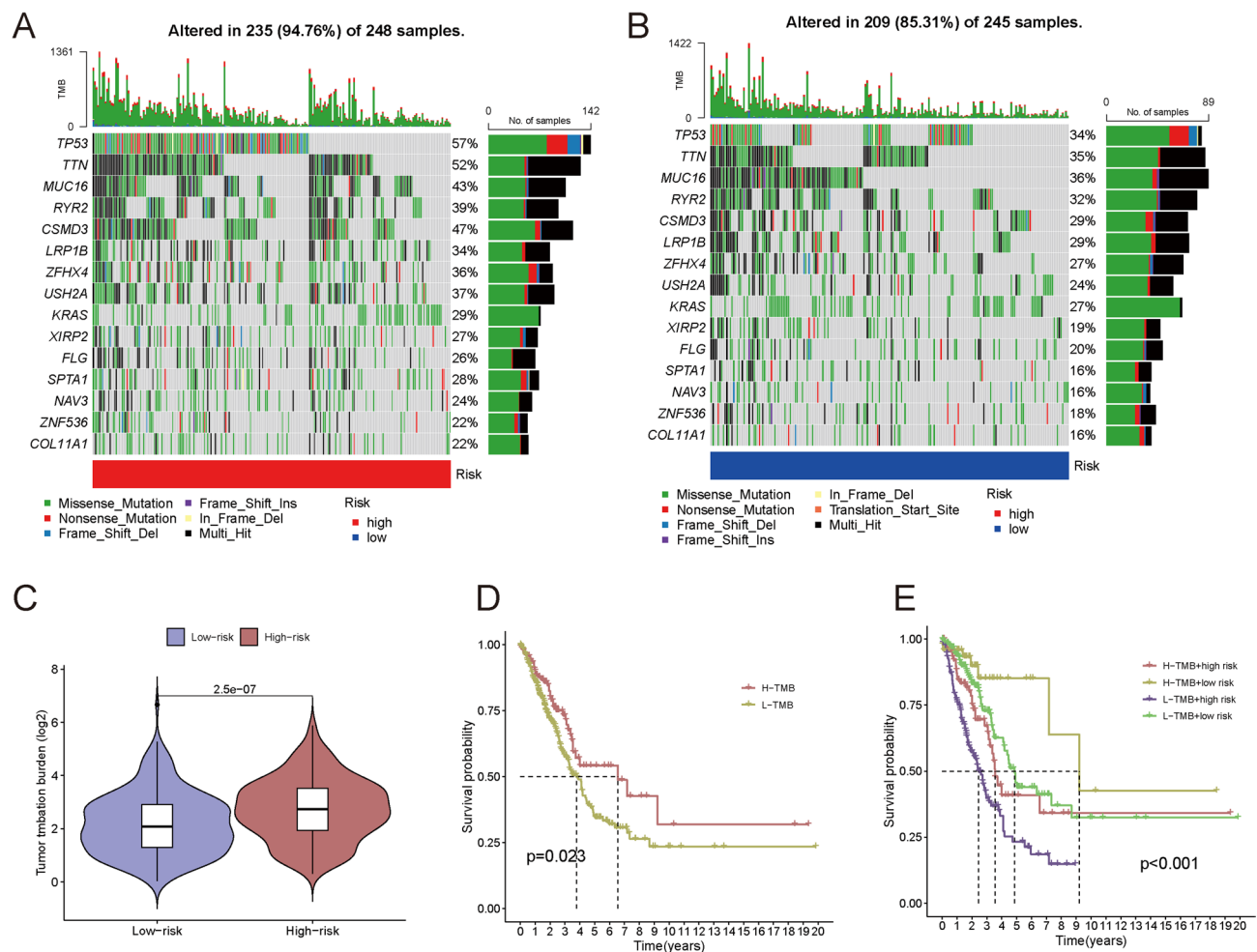
Fig. 7. Tumor mutation-related analyses. (**A, B**) Waterfall plots indicating TMB genes in HRG (**A**) and LRG (**B**). (**C**) TMB between the two risk groups. (**D**) Overall survival of high-TMB and low-TMB groups. (**E**) Overall survival of four groups.

regulating intracellular lactate levels and thereby modulating the pool available for histone lactylation in cancer cells[42]. MDH2 is a key TCA cycle enzyme that maintains metabolic homeostasis; its dysregulation can disrupt mitochondrial metabolism and promote a lactogenic environment in tumours[43]. MLYCD regulates fatty acid metabolism, and its loss can lead to metabolic rewiring, which enhances the dependence on glycolysis and lactate generation in cancer[44]. ISCU is essential for iron–sulfur cluster biogenesis; ISCU deficiency cripples mitochondrial respiration, forcing cancer cells to rely on glycolysis and lactate production[45]. ACAT1 catalyses a key step in ketone body and pyruvate metabolism, and its activity can influence the acetyl-CoA/lactyl-CoA balance, potentially by competing with or influencing lactylation[46]. As another core subunit of mitochondrial complex I, loss of NDUFA8 impairs oxidative phosphorylation, driving the glycolytic shift that increases lactate levels and lactylation in cancer[47].

The 16-LRG risk signature classified patients into LRG and HRG groups on the basis of the median risk score, and patients in the HRG group had worse survival than those in the LRG group did. Moreover, univariate and multivariate Cox regression analyses suggested that the 16-LRG risk signature was an independent prognostic marker for LUAD. ROC curves and nomograms have been widely applied to predict the prognosis of cancer in patients[48,49]. ROC analysis was used to visualize time-associated results in patients with LUAD. Moreover, a nomogram was constructed by combining the risk score with clinical factors.

Different immune cells play different roles in the tumour microenvironment. For example, the interaction and activation between B cells and macrophages are considered key humoral immune effects in inhibiting tumour progression[50]. It has been shown that the extensive infiltration of eosinophils in tumours is related to the long-term survival of patients[25]. Neutrophils were found to have key functions in carcinogenesis[51,52]. Studies have confirmed that neutrophils antagonize metastasis in lung cancer[53]. The infiltration of T-follicular helper cells indicates that these cells are involved in the antitumour immunity of NSCLC[26]. In this study, more activated B cells, eosinophils, immature dendritic cells, mast cells, monocytes, natural killer cells, plasmacytoid dendritic cells, and T follicular helper cells were present in the TME in the HRG than in the LRG, and most of these cells had positive effects on tumour immunity, which explains why LRG patients have better OS. These
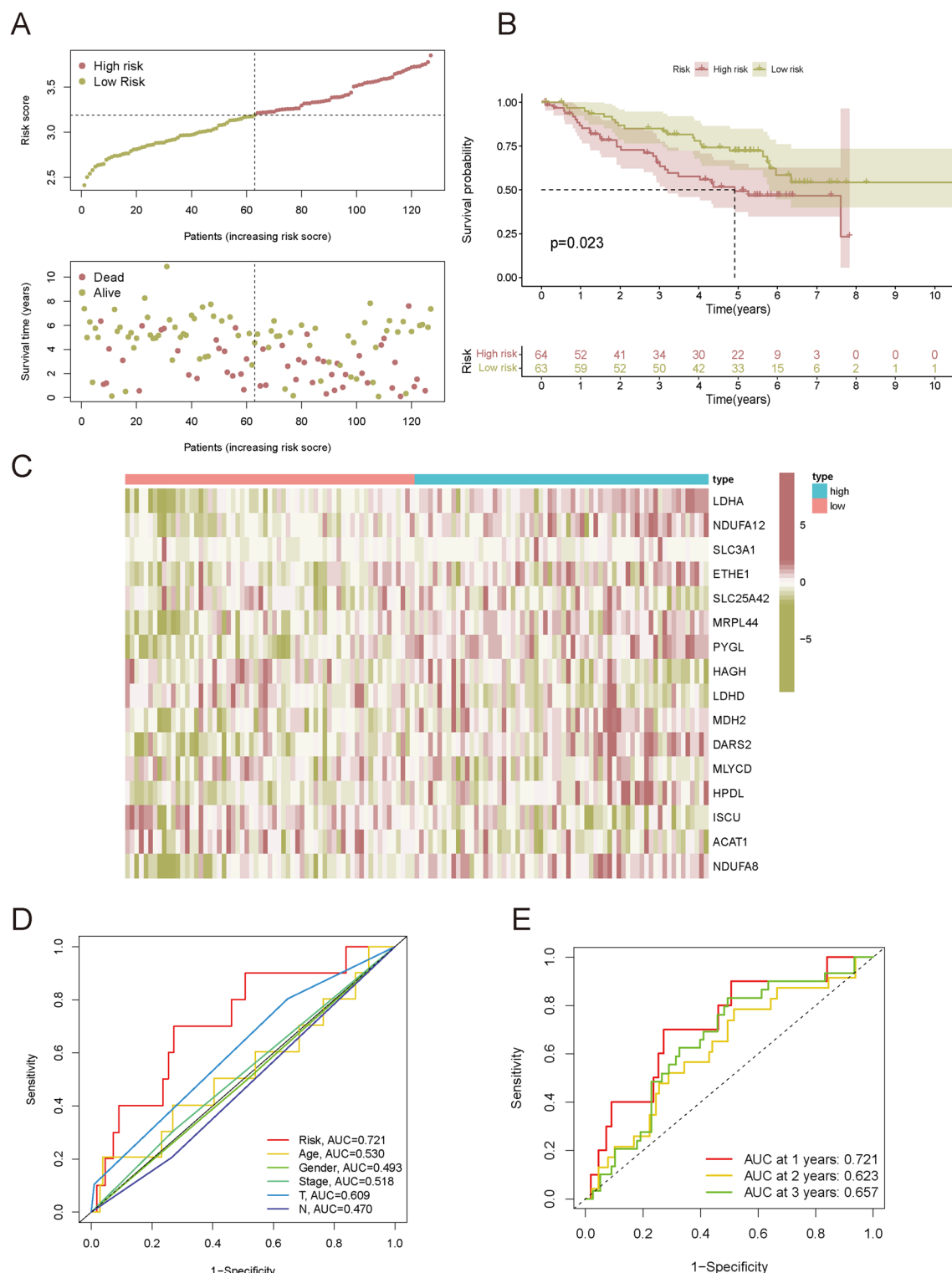
Fig. 8. Verifying of the 16-LRG risk signature by independent LUAD dataset. (**A**) The risk-score distribution and survival status. (**B**) Survival curves of two risk groups. (**C**) Expression of the 16 genes in two risk groups. (**D**) ROC curves of clinical features and risk scores. (**E**) ROC curves of the 16-LRG risk signature.

findings may help us to understand the relationships among lactate levels, TME cell infiltration and LUAD. Many studies have confirmed the link between lactate levels and immunosuppression. Lactate helps to establish an immunosuppressive environment conducive to cancer cell growth and immune escape[54].

This study has several limitations. We did not conduct sufficient in-depth research on the 16 lactate-related genes (LRGs) selected, and their detailed mechanisms of action still need further exploration. Future research

should focus on elucidating the detailed mechanisms of these genes in tumour progression and immune response and further verifying their effectiveness and mechanisms as tumour biomarkers and therapeutic targets through more in vitro and in vivo experiments. Moreover, this prognostic model requires more clinical patient information to verify its feasibility for clinical application.

Conclusion

Three lactate-related molecular subtypes with different prognoses were found in LUAD. Sixteen LRGs were identified, and relevant models were constructed. Both the ROC curve and the nomogram confirmed the prognostic value of the 16-LRG risk signature, which can serve as an independent prognostic factor for LUAD in patients. Considering the heterogeneity of LUAD, the 16-LRG risk signature enhances our characterization of TME cell infiltration and suggests more effective immunotherapy strategies.

Data availability

The datasets generated during and/or analysed during the current study are available from GEO [https://www.ncbi.nlm.nih.gov/geo/] (https://www.ncbi.nlm.nih.gov/geo) and TCGA official website (https://portal.gdc.cancer.gov/repository).

References

1. Bray, F. et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **68**(6), 394–424 (2018).
2. Cai, Y., Sheng, Z., Dong, Z. & Wang, J. EGFR inhibitor CL-387785 suppresses the progression of lung adenocarcinoma. *Curr. Mol. Pharmacol.* **16**(2), 211–216 (2023).
3. Gao, S. et al. Lung cancer in People's Republic of China. *J. Thorac. Oncol.* **15**(10), 1567–1576 (2020).
4. Blandin Knight, S. et al. Progress and prospects of early detection in lung cancer. *Open Biol.* **7**(9), 170070 (2017).
5. Imielinski, M. et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **150**(6), 1107–1120 (2012).
6. Faubert, B. et al. Lactate metabolism in human lung tumors. *Cell* **171**(2), 358-371e359 (2017).
7. Warburg, O. On the facultative anaerobiosis of cancer cells and its use in chemotherapy. *Munch Med. Wochenschr.* **103**, 2504–2506 (1961).
8. Askar, N., Jarrar, Y., Gharaibeh, M. & Alqudah, M. Upregulation of beta 1 and arachidonic acid metabolizing enzymes in the mouse hearts and kidneys after sub chronic administration of rofecoxib. *Curr. Mol. Pharmacol.* **16**(3), 381–392 (2023).
9. Bergers, G. & Fendt, S. M. The metabolism of cancer cells during metastasis. *Nat. Rev. Cancer* **21**(3), 162–180 (2021).
10. Fischer, K. et al. Inhibitory effect of tumor cell-derived lactic acid on human T cells. *Blood* **109**(9), 3812–3819 (2007).
11. Doherty, J. R. & Cleveland, J. L. Targeting lactate metabolism for cancer therapeutics. *J. Clin. Invest.* **123**(9), 3685–3692 (2013).
12. Wang, K. et al. Di-methylation of CD147-K234 promotes the progression of NSCLC by enhancing lactate export. *Cell Metab.* **33**(1), 160-173e166 (2021).
13. Halestrap, A. P. The SLC16 gene family - structure, role and regulation in health and disease. *Mol. Aspects Med.* **34**(2–3), 337–349 (2013).
14. Sha, L. et al. Shikonin inhibits the Warburg effect, cell proliferation, invasion and migration by downregulating PFKFB2 expression in lung cancer. *Mol. Med. Rep.* **24**(2), 560 (2021).
15. Zhao, Y. et al. HCAR1/MCT1 Regulates tumor ferroptosis through the lactate-mediated AMPK-SCD1 activity and its therapeutic implications. *Cell Rep.* **33**(10), 108487 (2020).
16. Updegraff, B. L. et al. Transmembrane protease TMPRSS11B promotes lung cancer growth by enhancing lactate export and glycolytic metabolism. *Cell Rep* **25**(8), 2223-2233e2226 (2018).
17. Abdollahi, E. et al. Immunomodulatory therapeutic effects of curcumin on M1/M2 macrophage polarization in inflammatory diseases. *Curr. Mol. Pharmacol.* **16**(1), 2–14 (2023).
18. Zhang, B. et al. m(6)A regulator-mediated methylation modification patterns and tumor microenvironment infiltration characterization in gastric cancer. *Mol Cancer* **19**(1), 53 (2020).
19. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic. Acids Res.* **43**(7), e47 (2015).
20. Sijie, W. et al. iTRAQ-based proteomic analysis unveils NCAM1 as a novel regulator in doxorubicin-induced cardiotoxicity and DT-010-exerted cardioprotection. *Curr. Pharm. Anal.* **20**(9), 966–977 (2024).
21. Kanehisa, M., Furumichi, M., Sato, Y., Matsuura, Y. & Ishiguro-Watanabe, M. KEGG: Biological systems database as a model of the real world. *Nucleic Acids Res.* **53**(D1), D672–D677 (2024).
22. Kanehisa, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**(11), 1947–1951 (2019).
23. O'Shea, J. J. & Plenge, R. JAK and STAT signaling molecules in immunoregulation and immune-mediated disease. *Immunity* **36**(4), 542–550 (2012).
24. Nigro, E. A. et al. Antitumor IgE adjuvanticity: Key role of Fc epsilon RI. *J. Immunol.* **183**(7), 4530–4536 (2009).
25. Davis, B. P. & Rothenberg, M. E. Eosinophils and cancer. *Cancer Immunol. Res.* **2**(1), 1–8 (2014).
26. Ma, Q. Y. et al. Function of follicular helper T cell is impaired and correlates with survival time in non-small cell lung cancer. *Int. Immunopharmacol.* **41**, 1–7 (2016).
27. Mitchell, D., Chintala, S. & Dey, M. Plasmacytoid dendritic cell in immunity and cancer. *J. Neuroimmunol.* **322**, 63–73 (2018).
28. Yang, D., Liu, Y., Bai, C., Wang, X. & Powell, C. A. Epidemiology of lung cancer and lung cancer screening programs in China and the United States. *Cancer Lett.* **468**, 82–87 (2020).
29. Liu, W. J., Du, Y., Wen, R., Yang, M. & Xu, J. Drug resistance to targeted therapeutic strategies in non-small cell lung cancer. *Pharmacol. Ther.* **206**, 107438 (2020).
30. Hirschhaeuser, F., Sattler, U. G. & Mueller-Klieser, W. Lactate: a metabolic key player in cancer. *Cancer Res.* **71**(22), 6921–6925 (2011).
31. Vanhove, K. et al. The metabolic landscape of lung cancer: New insights in a disturbed glucose metabolism. *Front Oncol.* **9**, 1215 (2019).
32. Hensley, C. T. et al. Metabolic heterogeneity in human lung tumors. *Cell* **164**(4), 681–694 (2016).
33. Feng, J. et al. Tumor cell-derived lactate induces TAZ-dependent upregulation of PD-L1 through GPR81 in human lung cancer cells. *Oncogene* **36**(42), 5829–5839 (2017).

34. Xie, M., Fu, X. G. & Jiang, K. Notch1/TAZ axis promotes aerobic glycolysis and immune escape in lung cancer. *Cell Death Dis.* **12**(9), 832 (2021).
35. Sharma, D., Singh, M. & Rani, R. Role of LDH in tumor glycolysis: Regulation of LDHA by small molecules for cancer therapeutics. *Semin. Cancer Biol.* **87**, 184–195 (2022).
36. Park, S. I. et al. NDUFA12 as a functional target of the anticancer compound ertredin in human hepatoma cells as revealed by label-free chemical proteomics. *J. Proteome Res.* **23**(1), 130–141 (2024).
37. Jiang, Y. et al. Cysteine transporter SLC3A1 promotes breast cancer tumorigenesis. *Theranostics* **7**(4), 1036–1046 (2017).
38. She, X. et al. ETHE1 dampens colorectal cancer angiogenesis by promoting TC45 Dephosphorylation of STAT3 to inhibit VEGF-A expression. *Cell Death Dis.* **15**(8), 631 (2024).
39. Lee, J. et al. A metabolic phenotype based on mitochondrial ribosomal protein expression as a predictor of lymph node metastasis in papillary thyroid carcinoma. *Medicine* **94**(2), e380 (2015).
40. Ji, Q. et al. PYGL-mediated glucose metabolism reprogramming promotes EMT phenotype and metastasis of pancreatic cancer. *Int. J. Biol. Sci.* **19**(6), 1894–1909 (2023).
41. Toor, S. M., Wani, S. & Albagha, O. M. E. Comprehensive transcriptomic profiling of murine osteoclast differentiation reveals novel differentially expressed genes and LncRNAs. *Front. Genet.* **12**, 781272 (2021).
42. Lv, M. et al. CDK7-YAP-LDHD axis promotes D-lactate elimination and ferroptosis defense to support cancer stem cell-like properties. *Signal Transduct. Target. Ther.* **8**(1), 302 (2023).
43. Pei, X. et al. Palmitoylation of MDH2 by ZDHHC18 activates mitochondrial respiration and accelerates ovarian cancer growth. *Sci. China Life Sci.* **65**(10), 2017–2030 (2022).
44. Zhou, L. et al. Fatty acid oxidation mediated by Malonyl-CoA decarboxylase represses renal cell carcinoma progression. *Can. Res.* **83**(23), 3920–3939 (2023).
45. Luo, J. et al. ISCU-p53 axis orchestrates macrophage polarization to dictate immunotherapy response in esophageal squamous cell carcinoma. *Cell Death Dis.* **16**(1), 462 (2025).
46. Sun, T. & Xiao, X. Targeting ACAT1 in cancer: from threat to treatment. *Front. Oncol.* **14**, 1395192 (2024).
47. Xiang, H. et al. NDUFA8 promotes cell viability and inhibits ferroptosis and cisplatin sensitivity by stabilizing Fe-S clusters in cervical cancer. *Naunyn-Schmiedeberg's Arch. Pharmacol.* https://doi.org/10.1007/s00210-025-04237-1 (2025).
48. Mandrekar, J. N. Receiver operating characteristic curve in diagnostic test assessment. *J. Thorac. Oncol.* **5**(9), 1315–1316 (2010).
49. Rose, P. G. et al. Nomograms predicting progression-free survival, overall survival, and pelvic recurrence in locally advanced cervical cancer developed from an analysis of identifiable prognostic factors in patients from NRG oncology/gynecologic oncology group randomized trials of chemoradiotherapy. *J. Clin. Oncol.* **33**(19), 2136–2142 (2015).
50. Tokunaga, R. et al. B cell and B cell-related pathways for novel cancer treatments. *Cancer Treat. Rev.* **73**, 10–19 (2019).
51. Ocana, A., Nieto-Jimenez, C., Pandiella, A. & Templeton, A. J. Neutrophils in cancer: Prognostic role and therapeutic strategies. *Mol. Cancer* **16**(1), 137 (2017).
52. Coffelt, S. B., Wellenstein, M. D. & de Visser, K. E. Neutrophils in cancer: neutral no more. *Nat Rev Cancer* **16**(7), 431–446 (2016).
53. Granot, Z. et al. Tumor entrained neutrophils inhibit seeding in the premetastatic lung. *Cancer Cell* **20**(3), 300–314 (2011).
54. Ngwa, V. M., Edwards, D. N., Philip, M. & Chen, J. Microenvironmental Metabolism Regulates Antitumor Immunity. *Cancer Res* **79**(16), 4003–4008 (2019).

Acknowledgements

Author contributions

All authors contributed to the study conception and design. Material preparation and data collection were performed by Jinjie Wang and Hailong Tang. The analysis were performed by Deshen Shan. The first draft of the manuscript was written by Jinjie Wang and Deshen Shan.The review and editing were performed by Ming Zhao and Hao Wang. All authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Ethical approval

This study did not require ethical approval.

Consent to participate

This study did not require ethical approval.

Consent for publication

Not applicable.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-24770-6.

Correspondence and requests for materials should be addressed to H.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.