Heliyon 10 (2024) e32341

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

# The prognostic and clinical value of genes associate with immunity and amino acid Metabolism in Lung Adenocarcinoma

Yuxin Zhang [a], Yuehui Wang [a], Ruoxuan Zhang [a], Quanwang Li [b,*]

[a] Beijing University of Chinese Medicine, No.11, North Third Ring East Road, Chaoyang District, Beijing, 100029, China
[b] Dongfang Hospital, Beijing University of Chinese Medicine, No. 6 fangxingyuan, Fengtai District, Beijing, 100078, China

ARTICLE INFO

*Keywords:*
Lung adenocarcinoma
Immunity
Amino acid metabolism
GEO
TCGA
Gene
PLK1
RRM2
TRIP13
HHMR

ABSTRACT

*Background:* Lung adenocarcinoma (LUAD) is the commonest subtype of primary lung cancer. A comprehensive analysis of the association of immunity with amino acid metabolism in LUAD is critical for understanding the disease.
*Methods:* The present study examined LUAD and noncancerous cases from the TCGA database. Differentially expressed genes (DEGs) between LUAD and noncancerous tissues were detected by analyzing processed expression profiles. We cross-referenced the up-regulated DEGs with Immune and Amino Acid Metabolism-related genes (I&AAMGs), resulting in Immune and Amino Acid Metabolism related differentially expressed genes (IAAAMRDEGs). The STRING database was employed to analyze PPI on IAAAMRDEGs, obtaining excavated hub genes, whose biological processes, molecular functions and cellular components were examined with GO/KEGG. Potential mechanisms related to LUAD were investigated by GSEA and GSVA. A prognostic model was built by LASSO-COX analysis, taking into consideration risk scores and prognostic factors to determine biomarkers affecting LUAD occurrence and prognosis.
*Results:* Totally 377 genes were detected at the intersection of upregulated DEGs and I&AAMGs. Analysis of PPI on these 377 IAAAMRDEGs yielded 17 hub genes. A LASSO regression analysis was utilized to assess the prognostic values of the 17 hub genes. Validation using the combined dataset confirmed 4 genes, e.g., polo-like kinase (PLK1), Ribonucleotide Reductase Subunit M2 (RRM2), Thyroid Hormone Receptor Interactor 13 (TRIP13), and Hyaluronan-Mediated Motility Receptor (HHMR). The model's accuracy was further assessed by ROC curve analysis and the COX model. In addition, immunohistochemical staining obtained from the HPA database, revealed enhanced PLK1 expression in LUAD samples.
*Conclusion:* LUAD pathogenesis is highly associated with immunity and amino acid metabolism. The PLK1, RRM2, TRIP13, and HMMR genes have prognostic values for LUAD. PLK1 upregulation in LUAD might be involved in tumorigenesis by modulating the cell cycle and represents a potential prognostic factor in clinic.

## 1. Introduction

Lung cancer, a prevalent malignant tumor, poses a significant threat to human health. The latest statistical report on the global cancer burden, released by the International Agency for Research on Cancer of the World Health Organization (WHO/IARC), reveals

* Corresponding author.
*E-mail address:* quanwangli@126.com (Q. Li).

https://doi.org/10.1016/j.heliyon.2024.e32341
Received 9 February 2024; Received in revised form 2 June 2024; Accepted 2 June 2024
Available online 6 June 2024

that lung cancer ranks second in terms of incidence and first in mortality [1,2]. Among the different types of lung cancer, lung adenocarcinoma (LUAD), classified as non-small cell lung cancer (NSCLC), stands as the most common subtype of primary lung cancer, accounting for 40 % of all lung cancer cases and displaying an upward trend [3]. Consequently, the identification of stable and reliable tumor markers becomes crucial to screen patients for poor prognosis and to offer more aggressive treatment options.

Amino acids play a vital role as essential nutrients for both tumor cells and immune cells. Both tumor cells and immune cells exhibit specific and distinctive amino acid requirements [4,5]. T cells, which are central to the immune system, play a pivotal role in inducing apoptosis of tumor cells and inhibiting tumor occurrence and development through various mechanisms [6,7]. The activation, differentiation, and function of T cells heavily depend on amino acid transport and metabolism. Furthermore, modern studies have shown that malignant tumor cells undergo rapid growth, characterized by exuberant tissue metabolism, abnormal cell growth, and accelerated anabolism and catabolism. Many tumors overexpress enzymes that degrade amino acids, which provide energy and metabolites for anabolic processes and also act as a mechanism for immune evasion of cancer cells [8]. Studies have shown that alanine-serine-transporter 2 (ASCT2), an essential amino acid transporter for life, can regulate the differentiation of CD4+T cell clusters, thus promoting anti-tumor immune response and achieving the goal of inhibiting tumor growth [9]. Hence, a comprehensive exploration of the interplay between immunity, amino acid metabolism, and LUAD is necessary to decipher the conditions under which specific genes promote or inhibit tumor survival.

In this paper, we analyze the expression changes of genes related to Immunity and Amino Acid Metabolism in LUAD by downloading and sorting the LUAD expression profile data and clinical information of patients from The Cancer Genome Atlas (TCGA) database. Our objective is to explore their correlation with diagnosis and prognosis and ultimately identify key genes such as PLK1, RRM2, TRIP13, and HHMR. Polo-like kinase (PLK) belongs to the serine/threonine protein kinase polo family and includes five subtypes, namely Polo-like kinase 1–5. PLK1, a known cell cycle regulator, governs cell mitosis, cytokinesis, DNA damage response, and development [10]. It is often associated with lower survival rates and frequently observed in various human cancers due to its overexpression.

## 2. Materials and methods

### 2.1. Data Acquisitions

The expression matrix of the Lung adenocarcinoma (LUAD) dataset (TCGA-LUAD) (https://portal.gdc.cancer.gov/) was downloaded from The Cancer Genome Atlas (TCGA) through the R package TCGAbiolinks (version 2.30.0) [11]. This dataset comprised 539 LUAD samples (cancer group, designated as LUAD) and 59 adjacent samples (normal group, designated as normal). All samples included in this study had matched clinical information and were standardized into Fragments Per Kilobase per Million (FPKM) data format. The corresponding clinical data were obtained from the UCSC Xena database [11](http://genome.ucsc.edu). The Tpm data and count sequencing data of the TCGA-LUAD dataset were standardized using the R-package limma (version 3.58.1) [12].

Additionally, we acquired LUAD-related datasets GSE118370 and GSE40275 from the Gene Expression Omnibus (GEO)(https://www.ncbi.nlm.nih.gov/) [13] using the R-package GEOquery (version 2.70.0) [14]. The GSE118370 dataset consisted of microarray gene-expression profiles of 6 LUAD patients and 6 matched normal tissues adjacent to cancer, derived from Homo Sapiens and the data platform GPL570 [Hg–U133 _ Plus _ 2] Affymetrix Human Genome U133 Plus 2.0 Array. On the other hand, the GSE40275 dataset comprised microarray gene expression profile data of 12 LUAD patients and 12 perfectly matched normal tissue samples adjacent to cancer, also from Homo Sapiens, with the data platform GPL15974 Human Exon 1.0 ST Array [CDF: Brain Array Version 9.0.1, HsEx 10stv2 _ Hs _ REFSEQ].

GeneCards database [15] was utilized to gather Immune-related genes (IRGs) and Amino Acid Metabolism-related genes (AAMGs) using "immune" and "amino acid metabolism" as search keywords, respectively. Additionally, the MSigDB (Molecular Signatures Database) [16] database was employed to conduct searches with keywords "Immune" and "amino acid metabolisms," resulting in the download of all relevant genomes. These datasets were merged, ultimately providing 3531 genes related to both immune and amino acid metabolisms (I&AAMGS).

### 2.2. Analysis of Differentially Expressed Genes(DEGs) related to immunity and acid metabolism between the normal and LUAD groups

To explore the potential mechanisms, related biological characteristics, and pathways of differential genes in the cancer group and normal group of LUAD, we first standardized the TCGA-LUAD, GSE118370, and GSE40275 datasets using the R-package limma. Subsequently, the TCGA-LUAD dataset was divided into the cancer group and normal group to analyze the differences in the processed expression profile data. Differentially expressed genes (DEGs) between different groups of the TCGA-LUAD dataset were identified, and the results of the differential analysis were visually represented using a volcano plot created with the ggplot2 R package. We selected the differentially expressed genes with $|\log FC| > 2$ and $P.adj < 0.05$. The obtained differentially expressed genes were then merged with I&AAMGs and visualized using a Venn diagram to identify Immune and Amino Acid Metabolism related differentially expressed genes (IAAAMRDEGs) in LUAD. The expression levels of these genes in the TCGA-LUAD dataset were visualized using the R-package pheatmap (version 1.0.12).

### 2.3. Protein-protein interaction (PPI)

Protein-protein interaction (PPI) involves the association of individual proteins through their interactions. To explore PPI

networks, we utilized the STRING database [17], which enables the search for known protein interactions and the prediction of interactions between proteins. In this study, we constructed a PPI network based on differentially expressed genes with a minimum required interaction score of 0.400. The closely related local regions within the PPI network may represent molecular complexes with specific biological functions. We employed the Markov Clustering (MCL) algorithm, a graph-based clustering method, to initially cluster the PPI network, setting the expansion parameter to 3. Subsequently, we used the Maximal Clique Centrality (MCC) algorithm to determine the importance of each node based on the maximum clique theory. Additionally, we calculated the importance of each node using the Degree algorithm, which is based on the number of nodes adjacent to each node. The Edge Percolated Component (EPC) algorithm identified the maximum connected subgraph in the network by removing edges, thereby determining the importance of each node. Moreover, the Maximum Neighborhood Component (MNC) algorithm measured the importance of each node based on the maximum number of neighborhoods to which each node belongs. Lastly, the density of maximum network neighborhood connectivity (DNNC) algorithm assessed the importance of each node by calculating the maximum neighborhood density to which each node belongs. Using these five algorithms, we assigned scores to IAAAMRDEGs related to other nodes in the PPI network. We ranked the IAAAMRDEGs based on these scores and selected the top 30 genes as hub genes (mRNAs) associated with LUAD disease.

### 2.4. Gene Ontology (GO) enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis

For large-scale function enrichment research, we conducted analysis using the Gene Ontology (GO) [18], which encompasses biological process (BP), molecular function (MF), and cellular component (CC). Additionally, we utilized the Kyoto Encyclopedia of Genes and Genomes (KEGG) [19], a comprehensive database containing information on genomes, biological pathways, diseases, and drugs. To perform GO annotation analysis on IAAAMRDEGs, we employed the R-package clusterProfiler (version 4.10.1) [20]. The selection criteria for entries were set at $P < 0.05$ and FDR value (q.value) $< 0.2$. We applied the Benjamini-Hochberg (BH) method to correct the P values.

### 2.5. Gene Set Enrichment Analysis (GSEA)

Gene Set Enrichment Analysis (GSEA) [21] evaluates the distribution trend of genes in a predefined gene set within a gene list sorted by phenotypic correlation to determine its contribution to the phenotype. In our study, we utilized the R-package clusterProfiler to perform GSEA on all genes in the cancer group and normal group of the TCGA-LUAD dataset. The GSEA enrichment analysis parameters were set as follows: 123 seeds, 1000 permutations, a minimum of 10 genes per gene set, and a maximum of 500 genes. The Benjamini-Hochberg (BH) method was used to correct the P values. We utilized the "c2.cp.v7.2.symbols" gene set from the Molecular Signatures Database (MSigDB) and considered significant enrichment at a threshold of $P < 0.05$ and FDR value (q.value) $< 0.25$.

### 2.6. Gene Set Variation Analysis (GSVA)

Gene Set Variation Analysis (GSVA) [22] is a nonparametric unsupervised analysis method used to evaluate gene set enrichment results of microarray nuclear transcriptomes. It transforms the expression matrix of genes between different samples into the expression matrix of gene sets between samples, thus determining whether different pathways are enriched among different samples. For this analysis, we obtained the "h.all.v7.4.symbols.gmt" gene set from the MSigDB database. GSVA analysis was performed on the TCGA-LUAD dataset at the gene expression level, and the differences in functional enrichment between the two groups were calculated. Significant enrichment was determined at a threshold of $P < 0.05$.

### 2.7. Least absolute shrinkage and selection operator (LASSO)

To develop the prognosis model of hub genes for LUAD, we employed 10-fold cross-validation with 123 seeds for regression using the Least Absolute Shrinkage and Selection Operator (LASSO) method [23]. We ran 1000 cycles to prevent overfitting and improve the model's generalization ability. LASSO regression is often used for constructing prognosis models. By increasing the penalty term (the absolute value of lambda $\times$ slope) on top of linear regression, the model's overfitting is reduced and its generalization ability is enhanced. The risk score was determined through the LASSO regression prognosis model, and the grouping and survival outcomes of each cancer group sample were visualized on the risk factor map.

### 2.8. Cox regression analysis

To investigate the clinical prognostic value of the screened hub genes in LUAD, we included the expression of IAAAMRDEGs from the TCGA-LUAD dataset in a multivariate Cox regression analysis. The significance threshold for the P value was set at 0.1. Based on the results of the multivariate Cox regression analysis, we constructed a forest plot to visualize the outcomes. Furthermore, a nomogram was established to predict the 1-year, 3-year, and 5-year survival of LUAD patients. A nomogram is a graphical representation that depicts the functional relationship between multiple independent variables through intersecting line segments in a plane rectangular coordinate system. It scores and characterizes the variables in the multivariate regression model using a specific scale and calculates the total score to predict the probability of events. The accuracy and resolution of the nomogram were evaluated using calibration curves. Calibration curves evaluate the prediction performance of the model by plotting the fit of actual probabilities against the probabilities predicted by the model under different circumstances. The "rms" R package was employed to construct the nomograms

and calibration curves. Additionally, the decision curve analysis (DCA) was used to assess the prediction efficacy of 1-year, 3-year, and 5-year survival outcomes in LUAD patients. DCA is a straightforward method for evaluating clinical prediction models, diagnostic tests, and molecular markers. The R-package ggDCA (version 1.1) [24] was utilized to draw the DCA curves.

### 2.9. Statistical analyses

The specific technical roadmap of our study was depicted in the figure (Fig. 1). All data processing and analyses were conducted using R software (Version 4.1.2). Continuous variables were presented as mean ± standard deviation. The Wilcoxon rank-sum test was used for comparing continuous variables between two groups, while the statistical significance of normally distributed variables was estimated using the independent Student t-test. The Kruskal-Wallis test was applied to compare three or more groups. For categorical variables, the Chi-square test or Fisher's exact test was used to analyze the statistical significance between two groups. LASSO regression analysis was performed using the glmnet R-package [25], and receiver operating characteristic (ROC) curves were generated using the pROC R-package [26]. Unless otherwise specified, all results were calculated using Spearman correlation analysis, and P values were two-tailed, with statistical significance set at $P < 0.05$.

## 3. Results

### 3.1. Analysis of Differentially Expressed Genes(DEGs) related to Immunity and Amino Acid Metabolism between the normal and LUAD groups

From the TCGA-LUAD dataset, a total of 18,436 genes were identified, with 1,666 genes highly expressed and 564 genes lowly expressed (Fig. 2A). We further cross-referenced the up-regulated DEGs with I&AAMGs, resulting in 377 genes that were related to both immune and amino acid metabolism (IAAAMRDEGs) (Table 1). The expression patterns of these 377 IAAAMRDEGs in the TCGA-LUAD dataset were depicted in a Venn diagram (Fig. 2B) and represented as a heat diagram (Fig. 2C).

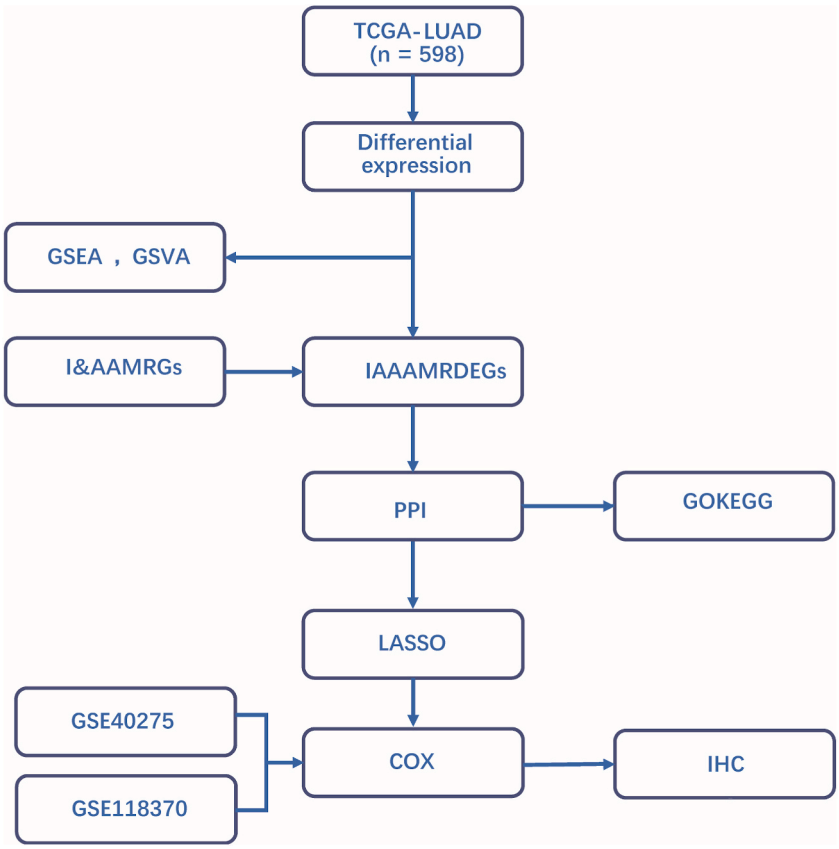



**Fig. 1.** Technical roadmap. TCGA, the cancer genome atlas. LUAD, Lung adenocarcinoma. AAMGs, Amino Acid Metabolism genes. IAAAMRDEGs, Immune and Amino Acid Metabolism related differentially expressed genes. GO, Gene Ontology. KEGG, Kyoto Encyclopedia of Genes and Genomes. GSEA, Gene Set Enrichment Analysis. GSVA, Gene Set Variation Analysis. PPI, Protein-protein interaction. LASSO, least absolute shrinkage and selection operator. ROC curve, receiver operating characteristic curve.

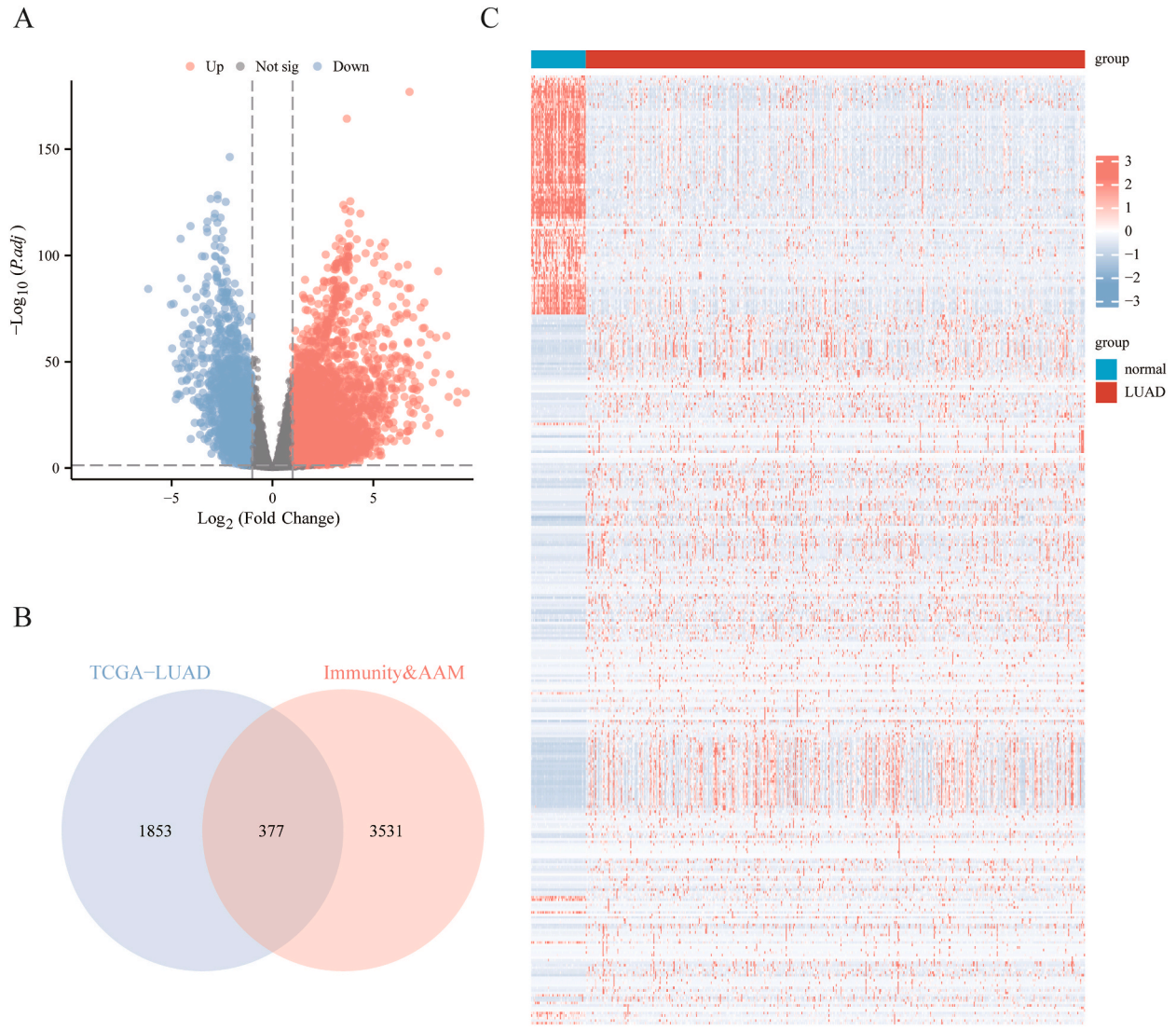



**Fig. 2.** Differentially expressed genes analysis. A: Volcano Plot for Differential Expression Analysis between the Cancer Group (Subgroup: Tumor) and Normal Group (Subgroup: Normal) of the TCGA-LUAD. The X-axis is log FC, and the larger the absolute value is, the larger the corrected P value is, indicating the larger the multiple of the difference is. The Y-axis is the corrected P value, and the larger the logarithm of log10 is, indicating the more significant the difference is. B: Venn Diagram Representation of the Intersection of Differentially Expressed Genes and IAAAAMRDEGs from the TCGA-LUAD Dataset. C: Heat Map Showing the Expression of IAAAMRDEGs on the TCGA-LUAD Dataset. Red represents high expression, blue represents low expression. LUAD, Lung Adenocarcinoma, a type of lung cancer. AAM, Amino Acid Metabolism. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

### *3.2. PPI interaction network analysis of immune-amino acid metabolism related genes*

To analyze the interactions between proteins encoded by 337 IAAAMRDEGs and predict the PPI network, we constructed a network using the STRING database. Then we used MCL algorithm for clustering and identify 31 closely related IAAAMRDEGs in PPI network. We utilized five algorithms, namely MCC, DMNC, MNC, EPC, and Degree, to mine scores related to other PPI network nodes for the IAAAMRDEGs. The top 30 node genes scored by each algorithm were identified as hub genes associated with LUAD disease. Taking the scoring result of MCC as an example, we used the cytoscape software to visually display the interaction relationship (Fig. 3A). We can know that altogether 17 hub genes included in the top30 scored of the five algorithms (Fig. 3B–Table S1). These hub genes are RRM2, UBE2C, CDT1, PLK1, MAD21.1, KIF11, KIF23, TRIP13, CDK1, CDC20, BIRC5, CCNA2, CDC25C, CCNB2, AURKA, HMMR and CENPF. The visualization of the interaction relationship they saw is shown in the figure (Fig. 3C). The specific gene scoring levels are shown in Tables S2–S6. These hub genes have shown a high degree of correlation and importance in a variety of independent bioinformatics approaches, indicating that they may play a central role in LUAD disease. And we found that these 17 hub genes showed a strong interaction relationship. It helps us to identify potential therapeutic targets or the focus of disease mechanism research more easily.

**Table 1**
List of gene symbol of IAAAMRDEGs.

| Gene Symbol | | | | | | | |
|---|---|---|---|---|---|---|---|
| ZIC2 | CXCR1 | KIF1A | SIGLEC11 | CHRNB4 | HOXC11 | PROM2 | TAL1 |
| ABCA3 | CYP27B1 | KIF23 | SIM2 | CHRNG | HOXC13 | PSAT1 | TESC |
| ACADL | CYP2F1 | KIF5A | SIX6 | CILP | HOXC9 | PTGES | TFAP2A |
| ACVRL1 | CYP4B1 | KIRREL2 | SLC13A5 | CIT | HOXD13 | PTPRZ1 | TFF1 |
| ADAM8 | CYP4F3 | KLF4 | SLC1A1 | CLEC14A | HPDL | PYCR1 | TFF2 |
| ADAMTS8 | DAO | KLHDC8A | SLC1A7 | CLPS | HRH3 | RAB26 | TH |
| ADCYAP1R1 | DCAF4L2 | KLK12 | SLC22A11 | CMTM2 | HTR3A | RAMP2 | THBD |
| ADH1B | DEFA4 | KLK3 | SLC25A10 | COCH | IGF2BP3 | RAMP3 | TLX1 |
| ADRB2 | DKK4 | KLK6 | SLC28A1 | COL10A1 | IGFBP1 | RECQL4 | TLX2 |
| AFP | DLC1 | KLK8 | SLC28A2 | COL1A1 | IGFBP3 | REG4 | TMEM171 |
| AGT | DLX5 | KRT14 | SLC2A1 | COL5A1 | IGLON5 | RETN | TMEM179 |
| AHNAK2 | DPEP1 | KRT15 | SLC2A2 | COL5A2 | IL1RL1 | RNF186 | TMEM59L |
| AIM2 | E2F2 | KRT17 | SLC2A5 | COL9A1 | IL31RA | RRM2 | TMEM63C |
| AK4 | E2F7 | KRT3 | SLC30A10 | COX4I2 | IL6 | RUFY4 | TMEM74 |
| AKR1C1 | E2F8 | KRT34 | SLC30A2 | CP | ITPKA | SCARA5 | TMEM88 |
| AKR1C2 | EDNRB | KRT4 | SLC39A5 | CPS1 | IVL | SCN8A | TNNC1 |
| AKR1C3 | EEF1A2 | KRT79 | SLC4A1 | CPT1B | KCNA10 | SCUBE1 | TNNI3 |
| AKR1D1 | EFNA2 | KRT83 | SLC6A15 | CPXM1 | KCNA5 | SEMA5A | TNNT1 |
| ALB | EMX1 | LAMP3 | SLC6A17 | CRABP1 | KCNE4 | SERPINA5 | TPPP3 |
| ALDH3B2 | EN2 | LCN2 | SLC6A18 | CTAG2 | KCNK9 | SERPINA9 | TRIM72 |
| ALOX15 | EPHA8 | LILRA5 | SLC6A3 | CXCL13 | KCTD4 | SFTPC | TRIP13 |
| ANGPTL7 | F2 | LIPN | SLC7A11 | CXCL14 | KIF11 | SFTPD | TROAP |
| ANKRD22 | F5 | LPO | SLC7A5 | BIRC5 | GAD1 | MUC2 | TSPAN8 |
| ANKRD33 | F7 | LRRC15 | SLC7A9 | BLM | GALR3 | MUC4 | TTR |
| ANKRD34B | FABP4 | LRRC2 | SLC9A3 | C11orf91 | GAST | MUC5AC | TUBB1 |
| ANLN | FABP7 | MAD2L1 | SNX22 | C1QL1 | GCG | MUC5B | ZBTB16 |
| AOC3 | FAM180B | MAG | SOX21 | CA10 | GCGR | MUC6 | ZG16B |
| APOBEC3B | FBXO32 | MAGEA6 | SPN | CA4 | GCLC | MYBPH | |
| ARHGEF15 | FGFR4 | MARCO | SPRR1B | CA9 | GDF2 | MYCN | |
| ARNTL2 | FGR | MB | SST | CACNA1I | GFAP | NDRG4 | |
| ARTN | FHL1 | MC5R | SSTR1 | CACNA2D2 | GH2 | NGB | |
| ASPA | FOSB | MELTF | STRA6 | CALML3 | GIMAP6 | NKX2-2 | |
| ATP4A | FOXD1 | MKRN3 | STX11 | CALML5 | GJB4 | NKX2-5 | |
| AURKA | FPR2 | MMP1 | STYK1 | CALML6 | GLDC | NKX3-2 | |
| B3GNT6 | FRMD3 | MMP11 | SUSD2 | CARD14 | GNGT1 | NKX6-1 | |
| BARHL2 | FTCD | MMP13 | SYNGR3 | CAV1 | GPA33 | NLRC4 | |
| BCL2L10 | FUT6 | MNX1 | SYNGR4 | CAV3 | GPD1 | NMUR1 | |
| BDNF | FXYD1 | MSI1 | SYT2 | CBX2 | GPR35 | NPTX1 | |
| BHMT | FZD9 | MSR1 | SYT5 | CCDC54 | GPT | OTC | |
| BHMT2 | G6PC2 | MUC16 | SYT7 | CCL24 | GPT2 | OVOL1 | |
| CD93 | HAL | PDK4 | TUBB3 | CCNA2 | GPX3 | PADI4 | |
| CDA | HAO1 | PEBP4 | TUBB4A | CCNB2 | GREB1 | PAH | |
| CDC20 | HBB | PGLYRP3 | UBE2C | CCNE1 | GREM1 | PANX2 | |
| CDC25C | HBEGF | PHOX2A | UNC93A | CD300LG | GRIN1 | PAX7 | |
| CDCA7 | HECW1 | PIP5K1B | VGF | CD36 | GRK5 | PCDH8 | |
| CDH3 | HGD | PKHD1L1 | VSIG4 | CD52 | GSTM5 | PCK1 | |
| CDK1 | HGFAC | PKMYT1 | VWF | CD79A | H2AC18 | PCSK1 | |
| CDKN2A | HMGA1 | PLA2G4A | WNK2 | CDT1 | HMMR | PNMA5 | |
| CDO1 | HMGB3 | PLK1 | XDH | CEACAM8 | HNF1A | PRAM1 | |
| CGA | HOXA11 | PRKG2 | ZIC4 | CENPF | HNF4A | PRKCG | |

IAAAMRDEGs - Immune and Amino Acid Metabolism related differentially expressed genes.

### *3.3. GO/KEGG enrichment analysis of genes related to immune-amino acid metabolism*

To analyze the biological functions and signaling pathways of the 17 hub genes associated with immune-amino acid metabolism in LUAD, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of the 17 hub genes. After converting gene IDs, we analyzed the hub genes through GO and KEGG (Table 2). The GO annotations of the 17 hub genes included Cellular Component (CC), Biological Process (BP), and Molecular Function (MF), providing insights into the functional enrichment of DEGs. The GO analysis results showed that the 17 hub genes were primarily enriched in biological process (BP) such as nuclear division(GO:0000280), mitotic nuclear division (GO: 0140014) and mitotic cell cycle phase transition (GO: 0044772) in LUAD. Spindle (GO:0005819), spindle pole (GO:0000922), mitotic spindle (GO:0072686), and other cellular components (CC); microtubule binding (GO:0008017), tubulin binding(GO:0015631), histone kinase activity (GO: 0035173), and other molecular function (MF) (Fig. 4A and B). KEGG analysis revealed associations between the 17 hub genes and signaling pathways. Specifically, the 17 hub genes were mainly linked to progesterone-mediated oocyte maturation (hsa04914), cell cycle (hsa04110), and oocyte meiosis (hsa04114) (Fig. 4A and B). The results of GOKEGG enrichment analysis were visualized using a chord diagram, combining logFC with

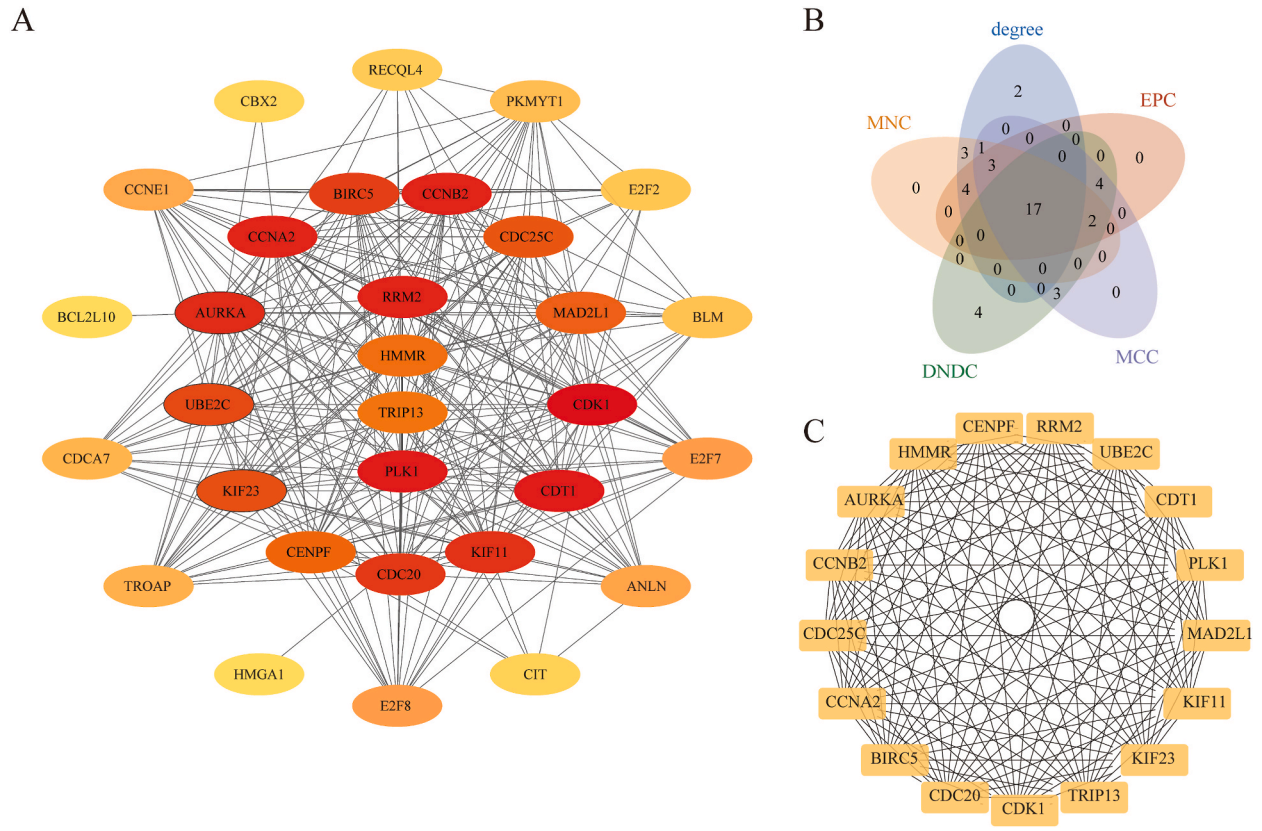



**Fig. 3.** PPI network. A. The IAAAMRDEGs were subjected to a Protein-Protein Interaction (PPI) network analysis using the MCL algorithm in the STRING database. The resulting PPI network was visualized using the MCC algorithm to display the node scores. The node colors ranged from yellow to red, indicating scores from low to high. B. The intersection of IAAAMRDEGs was analyzed using five scoring algorithms: MCC, EPC, MNC, DNNC, and Degree. The top 30 genes with the highest scores from each algorithm were identified as the hub genes. C. The PPI network for the hub genes was constructed. PPI: Protein-Protein Interaction. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

**Table 2**
GOKEGG enrichment analysis results of IAAAMRDEGs.

| Ontology | ID | Description | GeneRatio | BgRatio | pvalue | p.adjust |
|---|---|---|---|---|---|---|
| BP | GO:0044772 | mitotic cell cycle phase transition | 12/16 | 440/18800 | 3.89e-17 | 7.9e-15 |
| BP | GO:0140014 | mitotic nuclear division | 11/16 | 293/18800 | 4.45e-17 | 7.9e-15 |
| BP | GO:0000280 | nuclear division | 12/16 | 446/18800 | 4.59e-17 | 7.9e-15 |
| BP | GO:0048285 | organelle fission | 12/16 | 493/18800 | 1.53e-16 | 1.98e-14 |
| BP | GO:1901990 | regulation of mitotic cell cycle phase transition | 10/16 | 321/18800 | 1.34e-14 | 1.39e-12 |
| CC | GO:0005819 | spindle | 10/17 | 402/19594 | 2.02e-13 | 1.11e-11 |
| CC | GO:0072686 | mitotic spindle | 6/17 | 160/19594 | 3.1e-09 | 7.9e-08 |
| CC | GO:0000922 | spindle pole | 6/17 | 169/19594 | 4.31e-09 | 7.9e-08 |
| CC | GO:0005876 | spindle microtubule | 5/17 | 83/19594 | 7.18e-09 | 9.87e-08 |
| CC | GO:0030496 | midbody | 6/17 | 203/19594 | 1.29e-08 | 1.42e-07 |
| MF | GO:0008017 | microtubule binding | 5/17 | 272/18410 | 3.63e-06 | 0.0002 |
| MF | GO:0015631 | tubulin binding | 5/17 | 376/18410 | 1.75e-05 | 0.0005 |
| MF | GO:0035173 | histone kinase activity | 2/17 | 16/18410 | 9.56e-05 | 0.0019 |
| MF | GO:0008022 | protein C-terminus binding | 3/17 | 179/18410 | 0.0006 | 0.0082 |
| MF | GO:0016538 | cyclin-dependent protein serine/threonine kinase regulator activity | 2/17 | 50/18410 | 0.0010 | 0.0113 |
| KEGG | hsa04914 | Progesterone-mediated oocyte maturation | 7/13 | 102/8164 | 6.23e-11 | 1.87e-09 |
| KEGG | hsa04110 | Cell cycle | 7/13 | 126/8164 | 2.8e-10 | 3.69e-09 |
| KEGG | hsa04114 | Oocyte meiosis | 7/13 | 131/8164 | 3.69e-10 | 3.69e-09 |
| KEGG | hsa04115 | p53 signaling pathway | 3/13 | 73/8164 | 0.0002 | 0.0014 |
| KEGG | hsa05166 | Human T-cell leukemia virus 1 infection | 4/13 | 222/8164 | 0.0003 | 0.0019 |

GO - Gene Ontology; KEGG - Kyoto Encyclopedia of Genes and Genomes; IAAAMRDEGs - Immune and Amino Acid Metabolism related differentially expressed genes; BP - biological process; CC - cellular component; MF - molecular function.

circle diagrams (Fig. 4C and D). The circle diagram highlighted that the mitotic cell cycle phase transition (GO: 0044772) showed significantly up-regulated expression, and progesterone-mediated oocyte maturation (HSA04914) was the KEGG pathway with significantly up-regulated expression.

### 3.4. GSEA of immune-amino acid metabolism related genes in high and low risk groups

To assess the influence of gene expression levels on the disparity between the LUAD group and normal group, we performed GSEA on all genes expressed in the TCGA-LUAD dataset. The criteria for significant enrichment were set at $P < 0.05$ and FDR value (q.value)

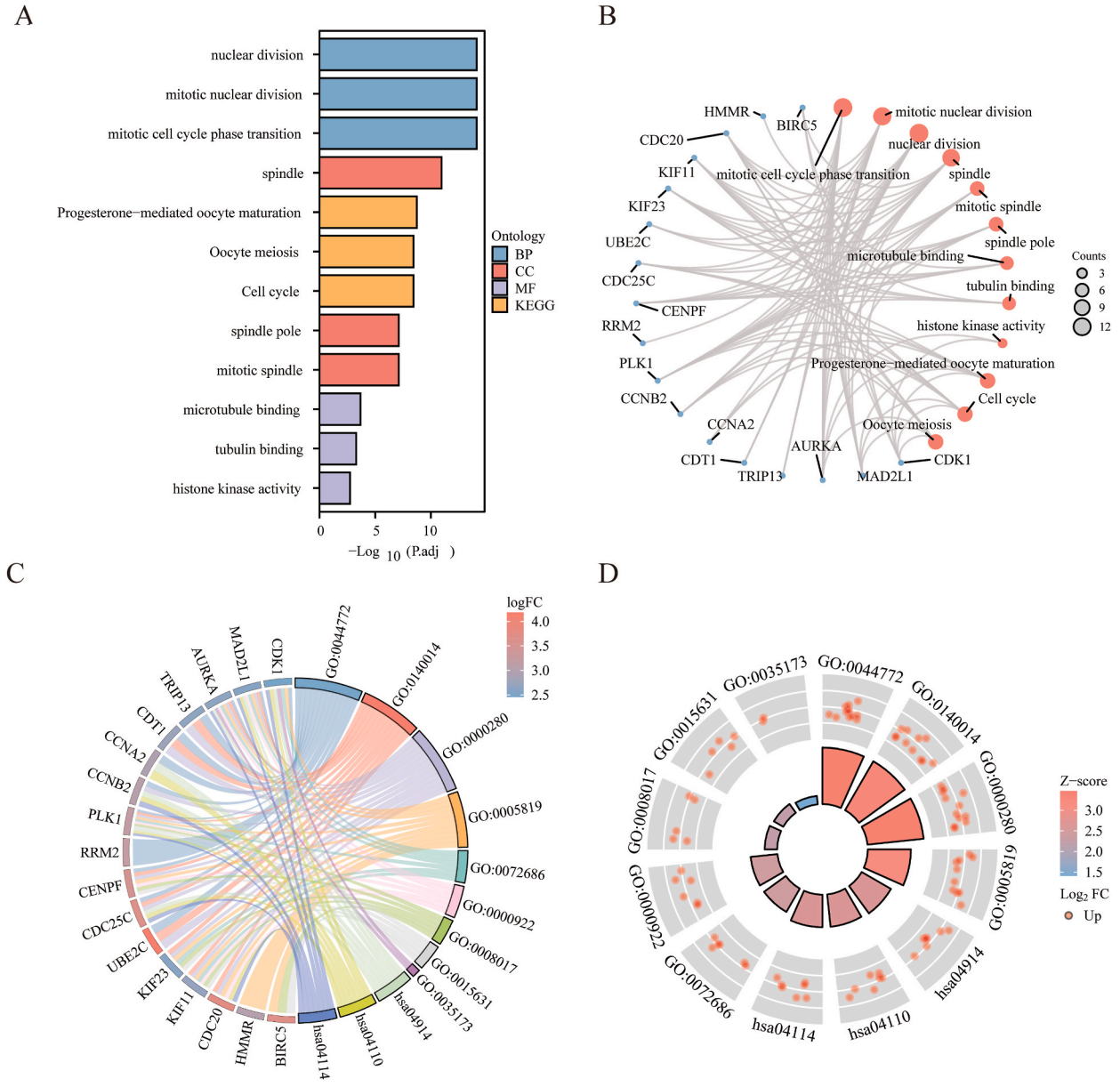



**Fig. 4.** Heat map of IAAAMRDEGS and GOKEGG enrichment analysis. A. Histogram presentation of results from the GO/KEGG enrichment analysis of HubGenes. B. Divergent network diagram presentation of results from the GO/KEGG enrichment analysis of hubgenes. In this diagram, red dots represent specific pathways, and blue circles represent specific genes. C. Chords from HubGenes combined with those from logFC's GO/KEGG enrichment analysis. D. Results from the combined logFC GO/KEGG enrichment analysis of Hub genes are presented in circles. In the network diagram (B), the red dots represent up-regulated genes (logFC >0), and the hubgenes in the diagram are all up-regulated genes. The screening criteria for GO/KEGG enrichment items are $P < 0.05$ and FDR value (q.value) < 0.25.GO: Gene Ontology, BP: Biological Process, CC: Cellular Component, MF: Molecular Function, KEGG: Kyoto Encyclopedia of Genes and Genomes. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

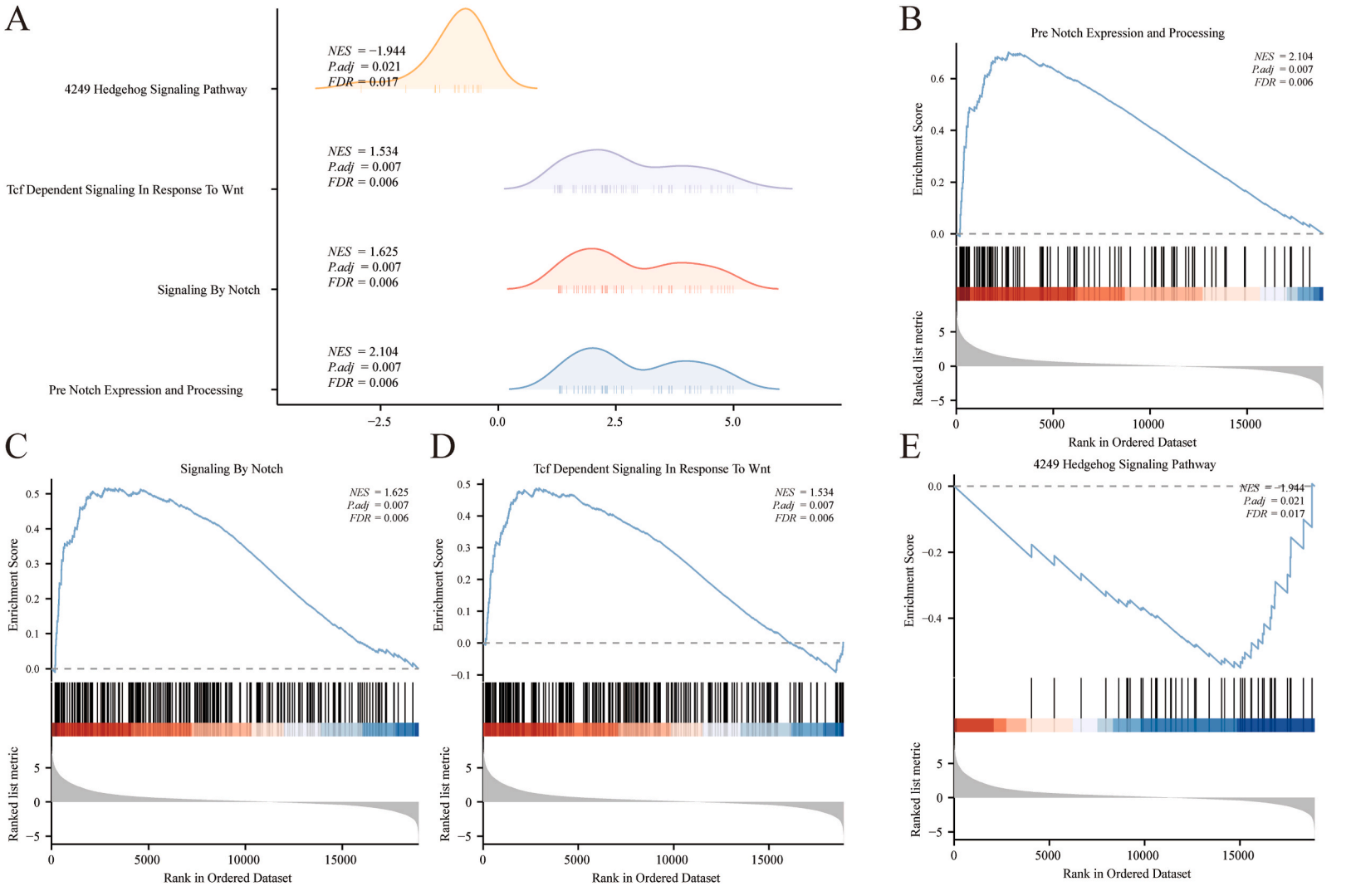



**Fig. 5.** GSEA. A. The GSEA of dataset TCGA-LUAD revealed four main biological characteristics. These are as follows: B. Pre Notch Expression and Processing (Fig. 5B) C. Signaling by NOTCH (Fig. 5C) D. TCF Dependent Signaling in Response to Wnt (Fig. 5D) E. 4249 Hedgehog Signaling Pathway (Fig. 5E) The screening criteria for significant enrichment in GSEA analysis were set at $P < 0.05$ and a FDR value (q.value) $< 0.25$. GSEA: Gene Set Enrichment Analysis.

< 0.25. The results revealed significant enrichment in several biological processes, cell components, and molecular functions. Specifically, the genes in the TCGA-LUAD dataset exhibited significantly enriched in pre-notch expression and processing (Fig. 5B), signaling by NOTCH (Fig. 5C), TCF dependent signaling in response to Wnt (Fig. 5D), 4249 hedgehog signaling pathway (Fig. 5E), and other pathways (Table 3). Mountain and path maps were generated to visualize these enrichments in the TCGA-LUAD dataset (Fig. 5A and 5B–E).

### 3.5. GSVA of immune-amino acid metabolism related genes in high and low risk groups

To explore the discrepancies in hallmark gene sets between the LUAD group (group: Tumor) and the Normal group (group: normal), we analyzed the expression of all genes in the TCGA-LUAD dataset using GSVA. The GSVA analysis showed significant differences in 46 hallmark gene sets between the LUAD group and the normal group ($P < 0.05$, Fig. 6A, Table 4). Among these, we selected 45 hallmark gene sets that demonstrated significant differences ($P < 0.001$) or played important roles in tumor occurrence and development. A group comparison chart was created (Fig. 6B) to visually display the results. The genes in the TCGA-LUAD dataset were notably enriched in hallmark pathways such as adipogenesis, allograft rejection, androgen response, angiogenesis, and apical junction, among others.

### 3.6. Screening prognostic related genes by lasso regression analysis

To assess the prognostic value of the 17 Hub genes in the TCGA-LUAD dataset, we performed LASSO regression analysis to construct a prognostic model (Fig. 7A and B), which included 4 Hub genes. Excluding normal samples, the cancer group was divided into a high-risk group (grouping: High Risk) and a low-risk group (grouping: Low Risk) based on the median of the Risk Score obtained from the LASSO model. The high-risk grouping results of the LASSO regression were visualized and presented in a risk factor map (Fig. 7C). To validate the prognosis model of LASSO, we conducted a statistical analysis of the clinical information of LUAD patients obtained from the TCGA-LUAD dataset (Table 5). We then merged the GSE118370 and GSE40275 datasets to create a Combined Dataset for verification. Using the variable coefficients from the LASSO model, we constructed a Risk Score for the cancer group samples in the combined dataset. Similar to the TCGA-LUAD dataset, the cancer group in the combined dataset was categorized into a high-risk group (grouping: High Risk) and a low-risk group (grouping: Low Risk). The expression levels of the four Hub genes in both datasets were grouped according to high and low risks, and a comparison chart was generated (Fig. 7D and E) to show the statistical significance of the expression level differences ($P < 0.05$ is considered statistically significant). From the chart, we can observe consistent results for four genes (PLK1, RRM2, TRIP13, and HHMR) in the merged dataset compared to the TCGA-LUAD dataset.

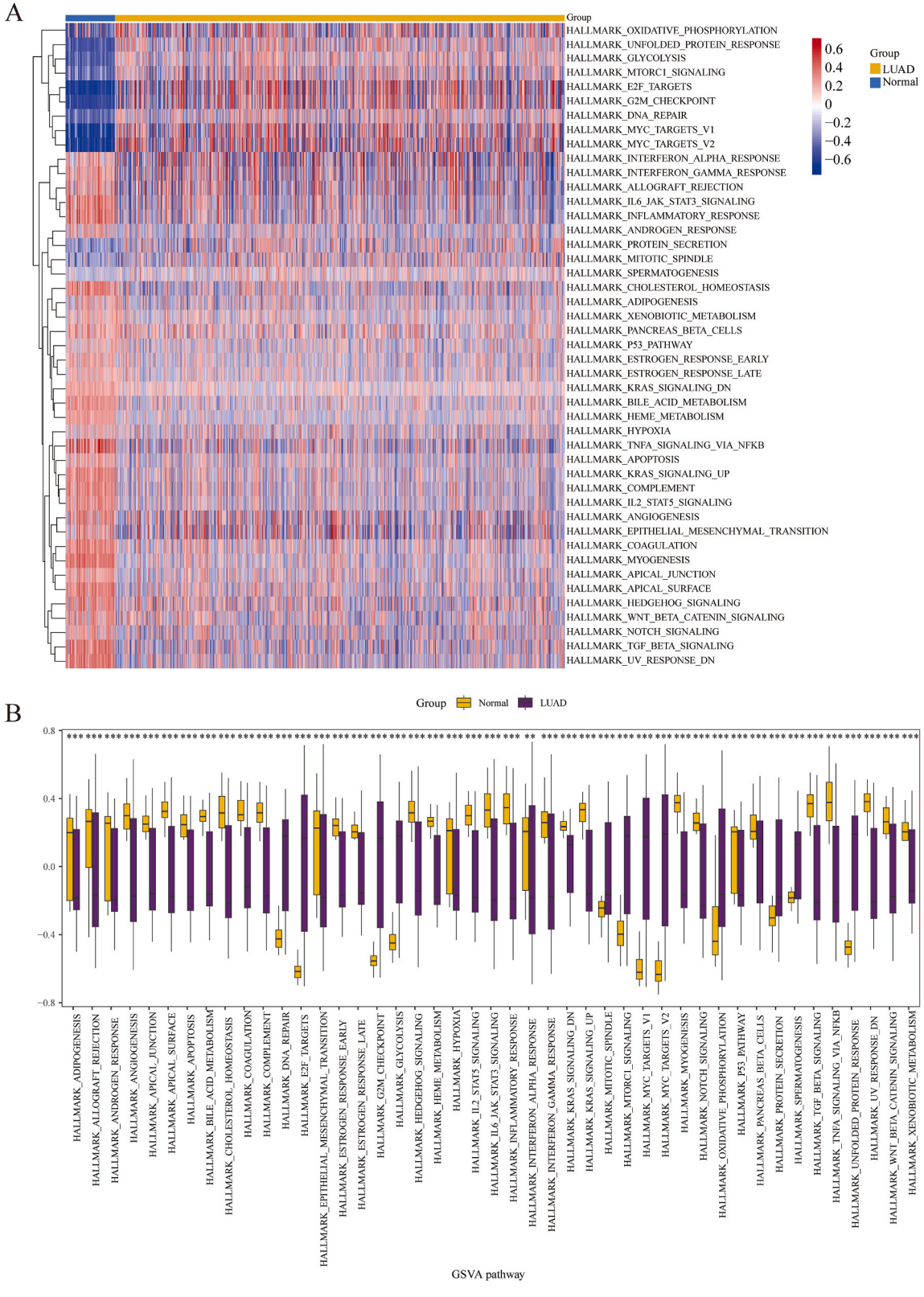
### 3.7. Correlation between the expression of key genes and the occurrence of LUAD

To investigate the correlation between the expression of the four key genes (PLK1, RRM2, TRIP13, and HHMR) and the occurrence of LUAD, we plotted ROC curves for these genes in the TCGA-LUAD dataset, using clinical status (Tumor vs Normal) as the outcome variable (Fig. 8A–D). The AUC values for PLK1, RRM2, TRIP13, and HHMR were obtained from the ROC curves. Additionally, we examined the clinical correlations of these four Hub genes in subgroups based on clinical T-stage, N-stage, M-stage, and overall survival (OS) (Fig. 8E-T). The results indicated that the expression of PLK1 was associated with the differences between T1 and T2 in clinical T-stage ($P < 0.0001$), N0 and N1 in N stage ($P < 0.05$), M0 and M1 in clinical M stage ($P < 0.01$), and the difference between survival and death in OS events ($P < 0.001$). The expression of RRM2 was associated with the differences between T1 and T2 in clinical T-stage ($P < 0.001$), N0 and N1 in N stage ($P < 0.001$), M0 and M1 in clinical M stage ($P < 0.05$), and the difference between survival and death in OS events ($P < 0.001$). The expression of TRIP13 was associated with the differences between T1 and T2 in clinical T-stage ($P < 0.001$), N0 and N1 in N stage ($P < 0.001$), and the difference between survival and death in OS events ($P < 0.001$). The expression of HHMR

**Table 3**
GSEA analysis of dataset TCGA-LUAD.

| ID | setSize | NES | pvalue | p.adjust | qvalue |
|---|---|---|---|---|---|
| REACTOME_CELL_CYCLE_CHECKPOINTS | 237 | 3.043798 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_MITOTIC_G1_PHASE_AND_G1_S_TRANSITION | 142 | 3.002095 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_DNA_REPLICATION | 137 | 2.956086 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_CELL_CYCLE_MITOTIC | 458 | 2.947152 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_G2_M_CHECKPOINTS | 134 | 2.90127 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_SYNTHESIS_OF_DNA | 110 | 2.900659 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_MITOTIC_METAPHASE_AND_ANAPHASE | 201 | 2.887159 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| WP_RETINOBLASTOMA_GENE_IN_CANCER | 84 | 2.814346 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_S_PHASE | 145 | 2.798761 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_MITOTIC_SPINDLE_CHECKPOINT | 92 | 2.77242 | 1.00E-10 | 7.58E-09 | 6.02E-09 |
| REACTOME_PRE_NOTCH_EXPRESSION_AND_PROCESSING | 106 | 2.103701 | 0.000212 | 0.00666 | 0.005539 |
| REACTOME_SIGNALING_BY_NOTCH | 233 | 1.625008 | 0.000202 | 0.00666 | 0.005539 |
| REACTOME_TCF_DEPENDENT_SIGNALING_IN_RESPONSE_TO_WNT | 231 | 1.534075 | 0.000203 | 0.00666 | 0.005539 |
| WP_4249_HEDGEHOG_SIGNALING_PATHWAY | 43 | −1.94397 | 0.001339 | 0.020584 | 0.017119 |

GSEA - Gene Set Enrichment Analysis; TCGA - the cancer genome atlas; LUAD - Lung adenocarcinoma.




*(caption on next page)*

**Fig. 6.** GSVA. A. The heat map represents the functional scores obtained from the GSVA analysis of dataset TCGA-LUAD. B. The subgroup comparison diagram of lung adenocarcinoma carcinoma group and normal group shows the enrichment pathways with significant differences in GSVA analysis of dataset TCGA-LUAD. The symbol * represents $P < 0.05$, indicating statistical significance. The symbol ***** represents $P < 0.001$, indicating very high statistical significance. GSVA, Gene Set Variation Analysis.

**Table 4**
GSVA analysis of dataset TCGA-LUAD.

| ID | logFC | AveExpr | t | P.Value | adj.P.Val | B |
|---|---|---|---|---|---|---|
| HALLMARK_E2F_TARGETS | 0.64533307 | 0.00065017 | 5.65274133 | 5.72E-08 | 2.48E-06 | 7.89501104 |
| HALLMARK_G2M_CHECKPOINT | 0.57097633 | 0.00140977 | 5.54154778 | 9.92E-08 | 2.48E-06 | 7.36695794 |
| HALLMARK_BILE_ACID_METABOLISM | −0.3840152 | −0.0001741 | −5.2577695 | 3.91E-07 | 6.51E-06 | 6.0538299 |
| HALLMARK_MYC_TARGETS_V1 | 0.5306426 | 0.00233643 | 4.80019694 | 3.20E-06 | 4.01E-05 | 4.04608105 |
| HALLMARK_DNA_REPAIR | 0.37306835 | −0.0170058 | 4.6003791 | 7.70E-06 | 7.70E-05 | 3.21390964 |
| HALLMARK_MYC_TARGETS_V2 | 0.53444311 | −0.0029183 | 4.34714325 | 2.25E-05 | 0.00018724 | 2.20002089 |
| HALLMARK_KRAS_SIGNALING_DN | −0.3238048 | −0.0046782 | −4.21351 | 3.88E-05 | 0.00027745 | 1.68386536 |
| HALLMARK_MITOTIC_SPINDLE | 0.30215716 | −0.0128773 | 3.90774626 | 0.00012953 | 0.00069603 | 0.55346012 |
| HALLMARK_MTORC1_SIGNALING | 0.33671904 | −0.007894 | 3.89486062 | 0.00013607 | 0.00069603 | 0.50739877 |
| HALLMARK_UNFOLDED_PROTEIN_RESPONSE | 0.3149531 | 0.00276347 | 3.88888542 | 0.00013921 | 0.00069603 | 0.48608353 |
| HALLMARK_INTERFERON_ALPHA_RESPONSE | 0.39771069 | 0.00346612 | 3.40079985 | 0.00081919 | 0.0037236 | −1.1592777 |
| HALLMARK_FATTY_ACID_METABOLISM | −0.267686 | 0.00117008 | −3.3611827 | 0.00093845 | 0.00391021 | −1.284354 |
| HALLMARK_ESTROGEN_RESPONSE_EARLY | −0.2497546 | −0.0005093 | −3.3374465 | 0.00101747 | 0.00391336 | −1.3586703 |
| HALLMARK_UV_RESPONSE_UP | 0.23054538 | −0.0038078 | 3.27899909 | 0.00123929 | 0.00442604 | −1.539669 |
| HALLMARK_XENOBIOTIC_METABOLISM | −0.2296009 | 0.01419187 | −3.2284017 | 0.00146686 | 0.00470819 | −1.6940551 |
| HALLMARK_UV_RESPONSE_DN | −0.2591284 | 0.00401489 | −3.2203194 | 0.00150662 | 0.00470819 | −1.7185177 |
| HALLMARK_SPERMATOGENESIS | 0.23384455 | 0.00096414 | 3.13243602 | 0.00200847 | 0.00560016 | −1.9809602 |
| HALLMARK_ANDROGEN_RESPONSE | −0.2358593 | 0.00089583 | −3.1312714 | 0.00201606 | 0.00560016 | −1.9843942 |
| HALLMARK_PI3K_AKT_MTOR_SIGNALING | 0.22336863 | −0.0058916 | 3.09357816 | 0.00227632 | 0.00576899 | −2.0949163 |
| HALLMARK_HEME_METABOLISM | −0.2063673 | −0.0026332 | −3.0893204 | 0.0023076 | 0.00576899 | −2.1073249 |
| HALLMARK_MYOGENESIS | −0.2614267 | −0.0021476 | −3.0611753 | 0.0025245 | 0.00601072 | −2.1889604 |
| HALLMARK_ADIPOGENESIS | −0.2281892 | −0.0074251 | −3.0378809 | 0.00271807 | 0.00617744 | −2.2560149 |
| HALLMARK_NOTCH_SIGNALING | 0.24004192 | −0.0155674 | 2.95403874 | 0.00353347 | 0.00768145 | −2.4935117 |
| HALLMARK_APOPTOSIS | 0.20767484 | 0.00230802 | 2.8108193 | 0.00546024 | 0.01137551 | −2.8851594 |
| HALLMARK_INTERFERON_GAMMA_RESPONSE | 0.28771787 | 0.00750707 | 2.77041929 | 0.00615517 | 0.01231034 | −2.9924071 |
| HALLMARK_PEROXISOME | −0.1953483 | −0.0063351 | −2.6362898 | 0.00907583 | 0.01745352 | −3.3381835 |
| HALLMARK_TGF_BETA_SIGNALING | 0.21603451 | 0.00212297 | 2.45542399 | 0.01497239 | 0.02680552 | −3.7791481 |
| HALLMARK_GLYCOLYSIS | 0.18346378 | −0.003386 | 2.45446576 | 0.01501109 | 0.02680552 | −3.7814063 |
| HALLMARK_ESTROGEN_RESPONSE_LATE | −0.1744409 | 0.0007029 | −2.35924 | 0.01932872 | 0.03332538 | −4.0016882 |
| HALLMARK_WNT_BETA_CATENIN_SIGNALING | 0.19436398 | −0.0015724 | 2.19545877 | 0.0293438 | 0.04890634 | −4.3612954 |
| HALLMARK_ALLOGRAFT_REJECTION | 0.21471797 | 0.00492428 | 2.10027824 | 0.037026 | 0.05971935 | −4.5589911 |

GSVA - Gene Set Variation Analysis; TCGA - the cancer genome atlas; LUAD - Lung adenocarcinoma.

was associated with the differences between T1 and T2 in clinical T stage ($P < 0.001$), N0 and N1 in N stage ($P < 0.001$), M0 and M1 in clinical M stage ($P < 0.05$), and the difference between survival and death in OS events ($P < 0.001$).

### *3.8. Multivariate Cox regression analysis of clinical correlation between expression of key genes and prognosis*

To further validate the prognostic model established by LASSO regression, we conducted multivariate Cox regression analysis to analyze the correlation between the expression of the four key genes (PLK1, RRM2, TRIP13, HHMR) and prognosis. The multivariate Cox regression model was constructed (Table 6), and a forest plot was generated to display the results (Fig. 9A). Then we performed nomogram analysis to assess the predictive power of the multivariate Cox regression model and created a nomogram (Fig. 9B). Additionally, we calibrated the nomogram for 1-year (Fig. 9C), 3-year (Fig. 9D), and 5-year (Fig. 9E) survival predictions and drew the calibration curves (Fig. 9C–E). From the calibration curves, it can be observed that the blue line representing three years is closest to the gray ideal line, indicating that the prediction effect of the model in the third year is superior to that in the first year and fifth year. Subsequently, we used decision curve analysis (DCA) to evaluate the clinical effectiveness of the constructed LASSO-Cox regression prognosis model for one year (Fig. 9F), three years (Fig. 9G), and five years (Fig. 9H). The results are shown in Fig. 9E–H, where the X-axis range of the blue line representing the model is higher than that of the all-positive red line and all-negative gray line, suggesting that the model performs better in five years but less so in one year and three years.

### *3.9. PLK1 has a higher expression level in LUAD tumor tissues*

Using the key gene PLK1 from the multivariate Cox model for further analysis, we analyzed the expression of PLK1 in LUAD tumor tissue and normal lung tissue by immunohistochemistry in the HPA database, using the HPA053229 antibody. The results revealed that the expression level of PLK1 in LUAD tumor tissue (Fig. 10B) was higher compared to normal lung tissue (Fig. 10A).

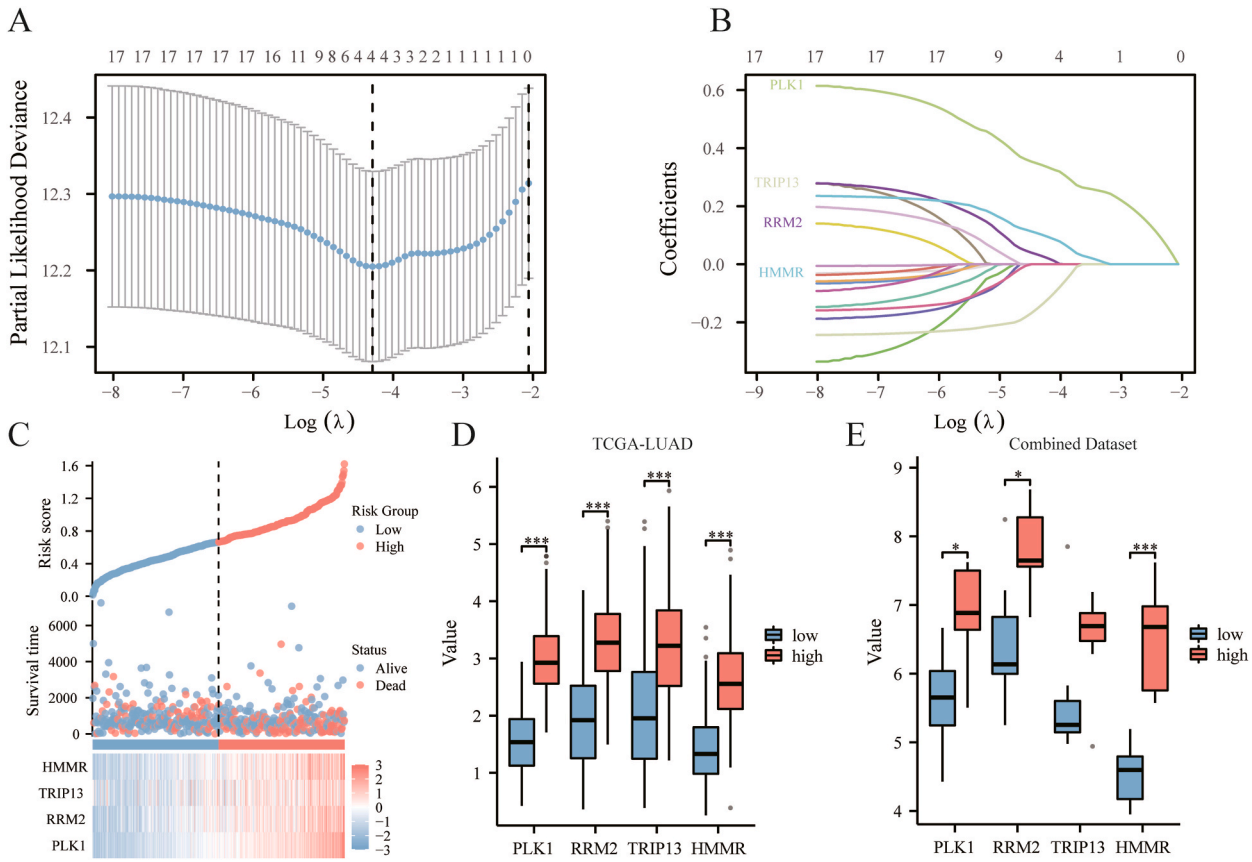



**Fig. 7.** Construction of HubGenes prognostic model and differential gene analysis between high-risk and low-risk groups of LASSO. A. The diagram represents the LASSO regression prognostic model for HubGenes. B. The variable trace diagram (B) shows the likelihood deviation value of the LASSO regression on the vertical axis and the log(λ) value on the x-axis. The x-axis represents the situation after taking the lambda coefficient of the penalty term in the LASSO regression as log, and the numbers above the x-axis indicate the number of variables with non-zero coefficients for each lambda. C. The risk factor diagram (C) of the LASSO regression diagnosis model is presented as a scatter plot. The blue dots represent surviving patients, and the red dots represent deceased patients. D. Hub genes are presented as a group comparison for high and low risk groups in dataset TCGA-LUAD (D). E. Hub genes are presented as a group comparison for high and low risk groups in the Consolidated dataset (E). TCGA, the cancer genome atlas. LUAD, Lung adenocarcinoma. LASSO, least absolute shrinkage and selection operator. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

## 4. Discussion

LUAD represents a common form of lung cancer, accounting for approximately 40 % of all lung cancer cases [27]. Despite advancements in early detection and treatment options, the survival rate for LUAD remains low, underscoring the need for further research into the molecular mechanisms underpinning this disease. A key area of interest is the potential contribution of alterations in amino acid metabolism and immune signaling pathways, particularly involving the protein PLK1, to LUAD pathogenesis. Exploring differentially expressed genes associated with immune-amino acid metabolism in LUAD could enhance early-stage prediction, facilitate the evaluation of LUAD patient conditions, and aid in determining their prognosis, thus offering valuable insights for subsequent research and clinical practice.

In this study, we initially identified up-regulated DEGs between LUAD samples and normal lung tissue, specifically screening for genes associated with immune-amino acid metabolism, and analyzed their biological processes, molecular functions and cellular components with GO/KEGG. Subsequently, we elucidated the potential mechanisms of action and relevant biological characteristics and pathways in both the LUAD and normal groups through the utilization of GSEA and GSVA analyses. Following this comprehensive exploration, we employed LASSO-COX analysis, risk scoring, and prognostic modeling to identify PLK1, RRM2, TRIP13, and HHMR as genes significantly associated with LUAD and immune-amino acid metabolism. Additionally, we performed immunohistochemical analysis of PLK1 expression in LUAD tumor tissue and normal lung tissue using the HPA database, revealing higher expression levels of PLK1 in LUAD compared to normal tissues. This finding aligns with previous studies that have also reported elevated PLK1 protein expression in various cancer types, such as osteosarcoma [28], colon cancer [29,30], and lung squamous cell carcinoma [31].Our study exhibits a certain level of innovation. Firstly, we not only investigated gene expression differences but also integrated immune response and amino acid metabolism. This comprehensive approach offers a deeper insight into the intricate biological features of

**Table 5**
Patient Characteristics of LUAD patients in the TCGA datasets.

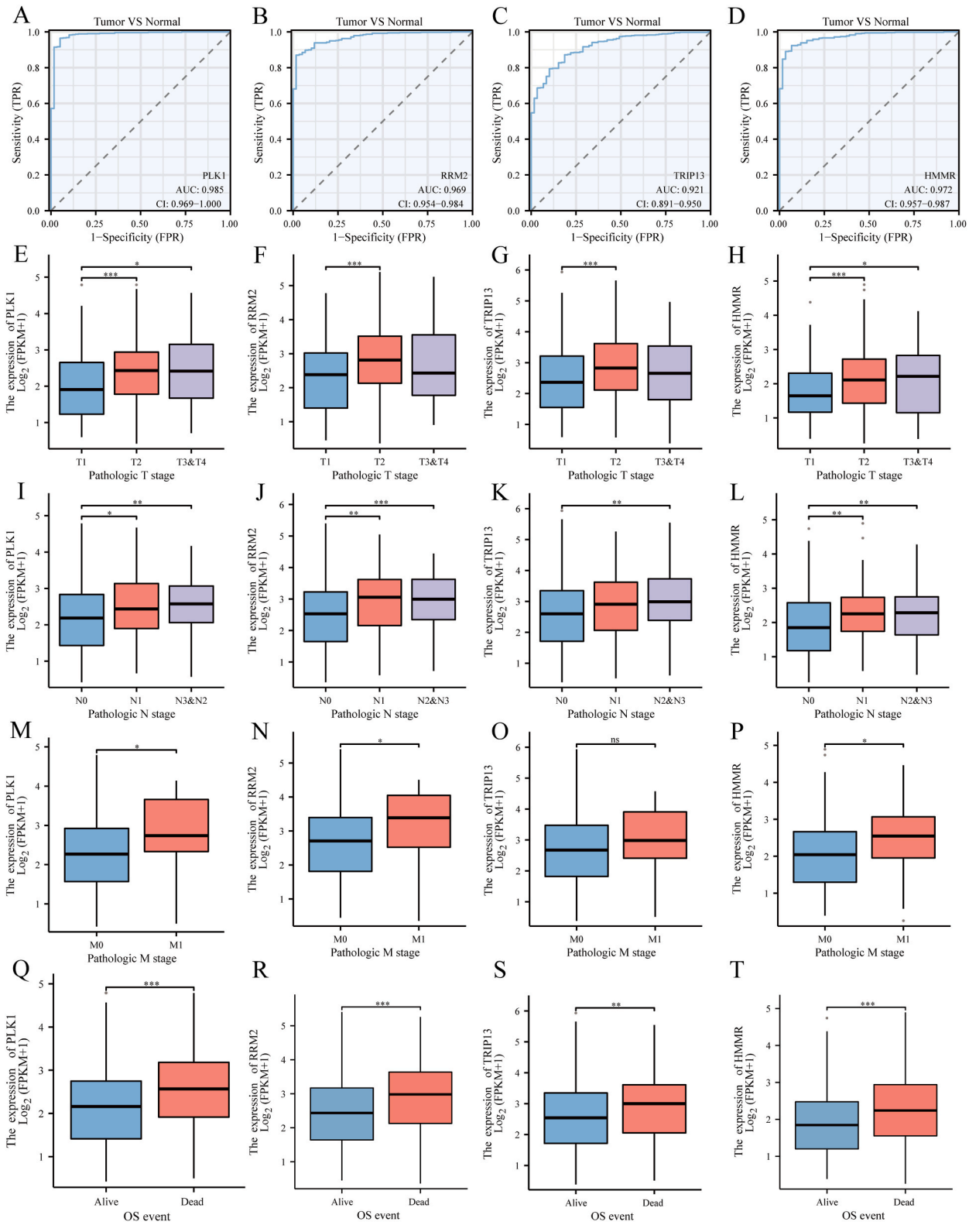
| Characteristics | overall |
|---|---|
| Pathologic T stage, n (%) | |
| T1 | 176 (32.8 %) |
| T2 | 292 (54.5 %) |
| T3 | 49 (9.1 %) |
| T4 | 19 (3.5 %) |
| Pathologic N stage, n (%) | |
| N0 | 350 (66.9 %) |
| N1 | 97 (18.5 %) |
| N2 | 74 (14.1 %) |
| N3 | 2 (0.4 %) |
| Pathologic M stage, n (%) | |
| M0 | 365 (93.6 %) |
| M1 | 25 (6.4 %) |
| Pathologic stage, n (%) | |
| Stage I | 296 (55.7 %) |
| Stage II | 125 (23.5 %) |
| Stage III | 84 (15.8 %) |
| Stage IV | 26 (4.9 %) |
| Gender, n (%) | |
| Female | 289 (53.6 %) |
| Male | 250 (46.4 %) |
| Age, n (%) | |
| $\leq 65$ | 257 (49.4 %) |
| $> 65$ | 263 (50.6 %) |
| OS event, n (%) | |
| Alive | 347 (64.4 %) |
| Dead | 192 (35.6 %) |
| OS event, n (%) | |
| Yes | 226 (41.9 %) |
| No | 313 (58.1 %) |

LUAD - Lung adenocarcinoma; TCGA - the cancer genome atlas; OS - overall survival; IQR - interquartile range.

LUAD. Secondly, we constructed a predictive model through LASSO-COX analysis and obtained four hub genes (PLK1, RRM2, TRIP13, HHMR), that might significantly contribute to the immunophenotype of LUAD, our target disease. There are a limited number of studies that link these genes with the disease phenotype at present, and they have not been sufficiently investigated in previous research. Our study has filled this gap by providing new options for clinical prognostic biomarkers.

Through five different algorithms, 17 hub genes were identified within PPI network as hub genes related to immunity and amino acid metabolism in LUAD. These algorithms measure the centrality of nodes from different angles, including interaction frequency, connectivity of neighbor nodes, etc. These genes are identified as hub genes due to their central position in the network and high degree of interaction. Their existence and interaction pattern may provide potential targets for the research and treatment of LUAD in the future. The functions and interaction of these protein and their specific mechanism of action in LUAD can be further verified by experiments in the future.

The results of GO/KEGG analysis revealed that the mitotic cell cycle phase transition was a significantly up-regulated biological process in LUAD, which is consistent with previous studies. Cell division processes play a vital role in cancer development. Error-free chromosomal segregation and cell proliferation during mitosis are central events in the life cycle. Proto-onco genes and tumor suppressor genes are directly or indirectly involved in the regulation of cell cycle, or it may be the result of genetic damage of genes encoding cyclins [32]. Additionally, the development of LUAD is also associated with the inactivation of tumor suppressor genes like p53 and Rb, as well as mutations in oncogenes such as EGFR and KRAS, all of which are closely related to cell cycle regulation [33,34]. Amino acid metabolism affects not only tumor cells but also the function of immune cells, and the normal regulation of cell cycle is of great significance to tumor cells and immune cells. Our research shows that some cell cycle regulatory genes represented by PLK1 also play a role in immune response and amino acid metabolism. Further understanding of these interactions is helpful to improve the existing chemotherapy and targeted drug therapy strategies and improve the therapeutic effect of LUAD. As we all know, steroid hormone receptors, including progesterone receptors, play an important role in the development of hormone-targeted tissue cancers, such as breast cancer and endometrial cancer [35]. While there is currently no direct evidence linking LUAD to progesterone-mediated oocyte maturation, some studies have suggested that progesterone might play a role in the occurrence and progression of LUAD. Researchers have found that the combination of estrogen and progesterone in NSCLC cells synergistically promotes the expression of vascular endothelial growth factor (VEGF) by increasing the proliferation of endothelial cells from adjacent vessels [36]. Moreover, other studies have shown that the high expression of the progesterone receptor in patients with LUAD is not related with malignancy and prognosis of LUAD [37]. Although there may be some connections between progesterone-mediated oocyte maturation and LUAD




*(caption on next page)*

**Fig. 8.** ROC curve and clinical correlation analysis. A-D. The ROC curve results for genes PLK1(A), RRM2(B), TRIP13(C), and HMMR(D) are displayed with Tumor and Normal as outcome variables. E-T. Subgroup comparison was conducted to analyze the clinical relevance of genes PLK1, RRM2, TRIP13, and HHMR in clinical T-staging, N-staging, M-staging, and OS events, respectively. In the figures, the symbol * indicates statistical significance with $P < 0.05$, while the symbol * * * indicates high statistical significance with $P < 0.001$. TPR, true positive rate. FPR, false positive rate. ROC, receiver operating characteristic curve. OS, overall survival.

**Table 6**
Cox regression to identify hub genes and clinical features associated with OS.

| Characteristics | Total(N) | Univariate analysis | | Multivariate analysis | |
|---|---|---|---|---|---|
| | | HR(95 % CI) | P value | HR(95 % CI) | P value |
| PLK1 | 530 | | <0.001 | | |
| Low | 267 | Reference | | Reference | |
| High | 263 | 1.831 (1.366–2.454) | <0.001 | 1.575 (1.019–2.436) | 0.041 |
| RRM2 | 530 | | 0.002 | | |
| Low | 266 | Reference | | Reference | |
| High | 264 | 1.587 (1.184–2.127) | 0.002 | 0.965 (0.590–1.580) | 0.888 |
| TRIP13 | 530 | | 0.003 | | |
| Low | 266 | Reference | | Reference | |
| High | 264 | 1.546 (1.156–2.069) | 0.003 | 1.054 (0.690–1.609) | 0.807 |
| HMMR | 530 | | <0.001 | | |
| Low | 267 | Reference | | Reference | |
| High | 263 | 1.678 (1.252–2.250) | <0.001 | 1.243 (0.807–1.916) | 0.323 |

OS - overall survival.

and the possibility of crossing signaling pathways for cell cycle regulation and tumor growth, these connections may be relatively indirect.

In our study, GSEA analysis revealed significantly enriched pathways in LUAD, including pre-notch expression and processing, signaling by NOTCH, TCF dependent signaling in response to WNT, and 4249 hedgehog signaling pathway. Notch, Wnt, and Hedgehog are all developmental signaling pathways. Previous research has reported that developmental signaling pathways, including Notch, WNT, and Hedgehog, are frequently altered in cancer stem cells (CSCs) to sustain the survival of CSCs and interact with other oncogenic signaling pathways such as MAPK, NF-κB, PI3K, and EGFR [38]. Although CSCs constitute only a small subset, accounting for less than 1 % of the tumor microenvironment, they possess the characteristics of self-renewal and contribute to tumor initiation, metastasis, spread, and resistance.

The GSVA enrichment analysis of the TCGA-LUAD dataset revealed significant enrichment in several hallmark pathways, including adipogenesis, allograft rejection, androgen response, angiogenesis, and apical junction pathways. As LUAD is a malignant tumor, its development is influenced by the tumor microenvironment, which includes factors such as fat cells, immune cells, and blood vessels. Adipogenesis, the process of fat formation in adipose tissue, has been shown to be related to the occurrence and progression of lung cancer [39]. Angiogenesis is the formation of new blood vessels. It plays a critical role in the growth and metastasis of lung cancer and other malignancies. The density of blood vessels in normal lung tissue is lower than that in LUAD, suggesting that angiogenesis may accelerate the growth and spread of LUAD. Factors such as VEGF, which promote angiogenesis, have been implicated in promoting the growth and spread of lung cancer [40]. As for androgen, modern research showed that androgen receptor exist in normal human lung, non-small cell, and small cell lung cancer tissue [41]. In both male and female patients with lung cancer, androgen levels and their response may influence tumor growth and progression [42]. Androgen response in hepatocarcinoma has attracted the attention of researchers, who have found that androgens can promote the growth and metastasis of hepatocellular carcinoma and are associated with the prognosis of the disease [43]. The apical junction is a crucial structure that interconnects cells and maintains the integrity and stability of tissues. Abnormal expression of proteins involved in cell junctions in LUAD may lead to structural damage and increased proliferation of tumor cells. Understanding the regulation of epithelial cell junctions may have implications for preventing the metastasis of LUAD and improving therapeutic strategies.

Our analysis confirms the significant overexpression of PLK1 in LUAD in TCGA data, suggesting its potential as a diagnostic marker. Specifically, in LUAD, the expression of PLK1 serves as a promising diagnostic marker, with an impressive AUC exceeding 0.9. Moreover, our findings indicate a correlation between PLK1 and the clinical stage of LUAD, particularly the T stage, providing further support for the notion that PLK1 expression might be associated with the degree of malignancy in LUAD. PLK1, a serine/threonine kinase, plays a critical role in cell cycle progression and mitosis. Its frequent overexpression in cancer cells is associated with enhanced cell proliferation, cell cycle arrest, and resistance to apoptosis [44]. Consistent with previous research, our study aligns with the evidence linking high PLK1 expression to poor prognosis in LUAD [45,46]. Recent studies suggest that PLK1 may also play a role in immune signaling pathways in cancer cells. For instance, it has been shown that PLK1-mediated phosphorylation of vimentin accelerates the transfer of cytotoxic T cells and immune escape in LUAD [47].

Moving on to ribonucleotide reductase subunit M2 (RRM2), this enzyme holds crucial importance in DNA synthesis and repair and has been associated with various cancers, including LUAD. In lung adenocarcinoma cells, RRM2 exhibits upregulation, and its expression levels have been linked to poor prognosis [48]. One potential mechanism through which RRM2 contributes to the development and progression of LUADa involves its impact on immune system function. Multiple studies suggest that RRM2 has been

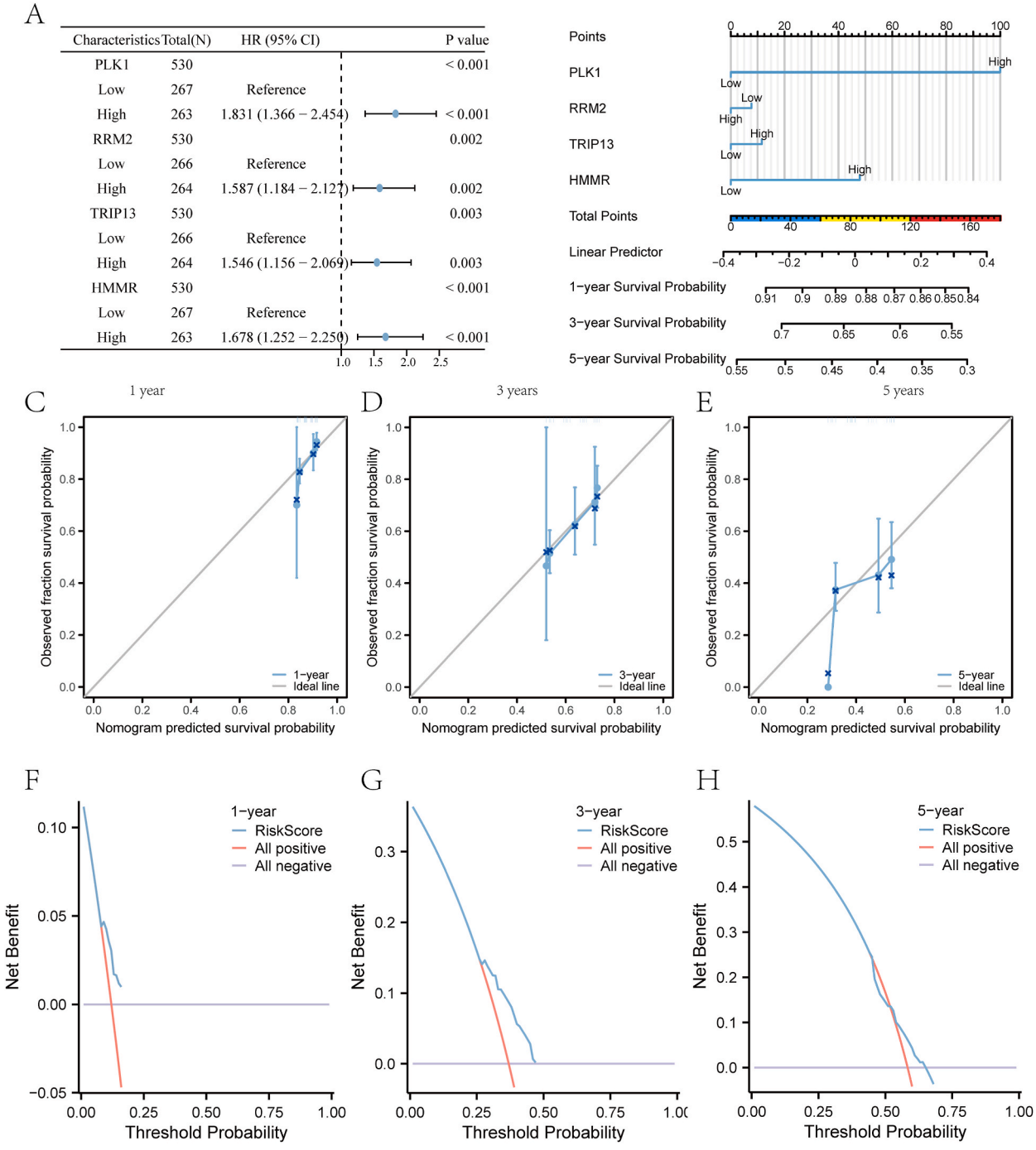



**Fig. 9.** Constructing Cox model. A-B. The forest plot (A) and nomogram (B) illustrate the results of multivariate Cox regression analysis for key genes (PLK1, RRM2, TRIP13, HHMR). C-E. The calibration curves are shown for 1-year (C), 3-year (D), and 5-year (E) nomogram analysis of the multivariate Cox regression model. F–H. The decision curve analysis (DCA) plots are presented for 1-year (F), 3-year (G), and 5-year (H) prognosis of the LASSO-Cox regression model. The X-axis in the DCA graph represents the probability threshold or Threshold Probability, and the Y-axis represents the net gain. DCA, decision curve analysis. LASSO, least absolute shrinkage and selection operator.

found to affect lung cancer progression and tumor immune cell infiltration. For example, RRM2 inhibition has been found to be effective in promoting M1 macrophage polarization and inhibiting M2 macrophage polarization in vitro and in vivo [49]. Additionally, in bladder cancer, RRM2 is positively correlated with immune checkpoints and cytotoxic T lymphocytes [50].

Thyroid Hormone Receptor Interactor 13 (TRIP13) encodes an AAA + ATPase that plays crucial roles in various cellular processes,

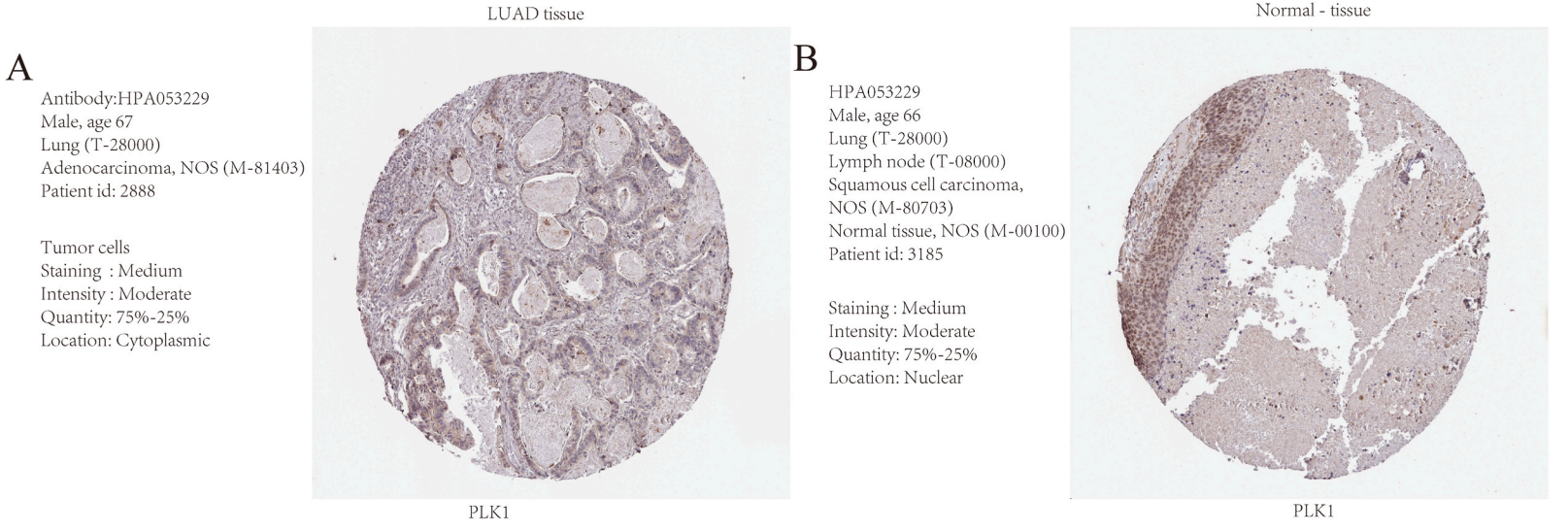



**Fig. 10.** Immunohistochemical analysis. A-B. Immunohistochemical analysis of the gene PLK1 in LUAD tissue (A) and normal tissue (B). The data were obtained from the HPA database. LUAD,Lung adenocarcinoma. HPA, human protein Atlas.

including meiotic spindle assembly and DNA repair. Recent investigations have revealed frequent amplification and overexpression of TRIP13 in diverse cancer types, including LUAD. Wei Li et al. (2016) [51] observed upregulation of TRIP13 in LUAD tissues compared to adjacent normal tissues, and its expression was correlated with poor prognosis in LUAD patients. Furthermore, Mechanistically, TRIP13 was shown to regulate the expression of multiple genes involved in cell cycle progression and apoptosis, the knockdown of TRIP13 can arrest lung cancer cells in G2/M phase and regulate the expression levels of genes related to cell cycle checkpoints [52].

Hyaluronan Mediated Motility Receptor (HMMR) is a gene encoding a protein that plays a pivotal role in tumor immune evasion and amino acid metabolism. The expression level of the HMMR gene in LUAD is closely associated with tumor occurrence and prognosis. Studies have found that high HMMR gene expression is linked to poor prognosis in LUAD patients and increased tumor recurrence [53]. Numerous studies have confirmed that HMMR is overexpressed in non-small cell lung cancer [54]. Abnormal tumor microenvironment is closely related to the tumors occurrence, and immune cells are a vital component of tumor microenvironment. Studies have identified a connection between high HMMR gene expression and tumor immune in liver cancer. Specifically, HMMR induces a macrophage-related immune response, likely by activating subgroup M2, and its high expression predicts increased infiltration levels of B cells, $CD8^+$ T cells, $CD4^+$ T cells, macrophages, neutrophils, and dendritic cells [55].

The findings of this study provide valuable insights and inspiration for the future clinical workflow and treatment of LUAD. By further studying the mechanism of the four key genes PLK1, RRM2, TRIP13 and HMMR in LUAD development, we can offer researchers more targeted treatment strategies and more choices for individualized treatment of patients with LUAD. These findings will enhance clinicians' understanding of LUAD's pathobiological characteristics enabling more accurate prognosis evaluation and treatment selection for patients. Therefore, integrating these findings into clinical practice holds the potential to improve treatment outcomes and survival rates for LUAD patients, laying the groundwork for future advancements in clinical workflows and treatment modalities.

While this study has advanced our understanding of the interplay between immunity, amino acid metabolism, and LUAD, it does have some limitations. One significant constraint is the absence of in vitro and in vivo validation for our findings. Additionally, due to the design constraints of this study, there might be other important signaling pathways related to LUAD that were not explored. Further investigation is necessary to uncover these potential pathways.

## 5. Conclusion

LUAD is a complex disease characterized by dysregulation in various biological processes, including amino acid metabolism and immune signaling pathways. The prognosis model was constructed by LASSO, and the dataset TCGA-LUAD was verified with the combined datasets (GSE118370, GSE40275). The results showed that there were four identical genes, namely, PLK1, RRM2, TRIP13 and HHMR. We drew the ROC curves of four hub genes, and the AUCs of PLK1, RRM2, TRIP13 and HHMR were all greater than 0.9, indicating that the genes related to immunity and amino acid metabolism selected by us were accurate in predicting the presence of LUAD. In addition, multivariate Cox regression analysis revealed that the expression levels of four hub genes were correlated with the prognosis of LUAD. PLK1, RRM2, TRIP13 and HMMR as the pivotal protein, may serve as a crucial link between these different processes by regulating both amino acid metabolism and immune cell function. Further research is essential to comprehensively understand how PLK1 contributes to LUAD development, and to develop targeted therapies that can exploit these pathways for potential therapeutic advantages.

## Ethical approval

Review and/or approval by an ethics committee was not needed for this study because this study does not contain any studies with human participants or animals performed by any of the authors.

## Data availability statement

Data are available in a public, open access repository. All data relevant to the study are included in the article or uploaded as supplemental information. The datasets for this study can be found in the GEO (https://www.ncbi.nlm.nih.gov/) and TCGA(https://portal.gdc.cancer.gov/). The GSE118370 dataset derived from Homo Sapiens and the data platform GPL570 [HG-U133 _Plus _ 2] Affymetrix Human Genome U133 Plus 2.0 Array. The GSE40275 dataset derived from Homo Sapiens, with the data platform GPL15974 Human Exon 1.0 ST Array [CDF: Brainarray Version 9.0.1, HsEx10stv2 _ Hs _ REFSEQ]. The corresponding clinical data were obtained from the UCSC Xena database [11](http://genome.ucsc.edu).

## CRediT authorship contribution statement

**Yuxin Zhang:** Writing – original draft, Methodology, Conceptualization. **Yuehui Wang:** Formal analysis, Data curation. **Ruoxuan Zhang:** Writing – review & editing, Writing – original draft. **Quanwang Li:** Project administration, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e32341.

## References


[1] LungCancer2020:Epidemiology,Etiology,andPrevention, Clin. Chest Med. 41 (1) (2020) 1–24. Mar.
[2] A.M. Noone, N. Howlader, M. Krapcho, et al., SEER Cancer Statistics Review, 1975–2015, National Cancer Institute, Bethesda, MD, April 2018. Available online: https://seer.cancer.gov/csr/1975_2015/. based on November 2017 SEER data submission, posted to the SEER website.
[3] B.D. Hutchinson, G.S. Shroff, M.T. Truong, et al., Spectrum of lung adenocarcinoma, Semin. Ultrasound CT MR 40 (3) (2019) 255–264.
[4] A. Luengo, D.Y. Gui, M.G. Vander Heiden, Targeting metabolism for cancer therapy, Cell Chem. Biol. 24 (9) (2017) 1161–1180.
[5] B. Faubert, A. Solmonson, R.J. Deberardinis, Metabolic reprogramming and cancer progression, Science 368 (6487) (2020) eaaw5473.
[6] M. Nakaya, Y. Xiao, X. Zhou, et al., Inflammatory T cell responses rely on amino acid transporter ASCT2 facilitation of glutamine uptake and m TORC1 kinase activation, Immunity 40 (5) (2014) 692–705.
[7] P.J. Siska, J.C. Rathmell, T cell metabolic fitness in antitumor immunity, Trends Immunol. 36 (4) (2015) 257–264.
[8] E. Ananieva, Targeting amino acid metabolism in cancer growth and anti-tumor immune response, World J. Biol. Chem. 6 (4) (2015) 281–289, https://doi.org/10.4331/wjbc.v6.i4.281.
[9] M. Nakaya, Y. Xiao, X. Zhou, et al., Inflammatory T cell responses rely on amino acid transporter ASCT2 facilitation of glutamine uptake and mTORC1 kinase activation, Immunity 40 (5) (2014) 692–705, https://doi.org/10.1016/j.immuni.2014.04.007.
[10] S. Iliaki, R. Beyaert, I.S. Afonina, Polo-like kinase 1(PLK1) signaling in cancer and beyond, Biochem. Pharmacol. 193 (2021) 114747.
[11] M.J. Goldman, et al., Visualizing and interpreting cancer genomics data via the Xena platform, Nat. Biotechnol. 38 (6) (2020) 675–678.
[12] M.E. Ritchie, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic acids research 43 (7) (2015).
[13] T. Barrett, et al., NCBI GEO: archive for functional genomics data sets–update, Nucleic Acids Res. 41 (Database issue) (2013) D991–D995.
[14] S. Davis, P.S. Meltzer, GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor, Bioinformatics 23 (14) (2007) 1846–1847.
[15] G. Stelzer, et al., The GeneCards suite: from gene data mining to disease genome sequence analyses, Curr Protoc Bioinformatics 54 (2016) 1.30.1–1.30.33.
[16] A. Liberzon, et al., The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell Syst 1 (6) (2015) 417–425.
[17] D. Szklarczyk, et al., STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, Nucleic acids research 47 (D1) (2019) D607–D613.
[18] C. Gene Ontology, Gene Ontology Consortium: going forward, Nucleic Acids Res. 43 (Database issue) (2015) D1049–D1056.
[19] M. Kanehisa, S.J.N.a.r. Goto, KEGG: kyoto encyclopedia of genes and genomes 28 (1) (2000) 27–30.
[20] G. Yu, et al., clusterProfiler: an R package for comparing biological themes among gene clusters, OMICS 16 (5) (2012) 284–287.
[21] A. Subramanian, et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles, Proc Natl Acad Sci U S A 102 (43) (2005) 15545–15550.
[22] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7.
[23] R.J.S.i.m. Tibshirani, The lasso method for variable selection in the Cox model 16 (4) (1997) 385–395.
[24] T. Tataranni, C. Piccoli, Dichloroacetate (DCA) and cancer: an Overview towards clinical Applications, Oxid. Med. Cell. Longev. 2019 (2019) 8201079.
[25] S. Engebretsen, J. Bohlin, Statistical predictions with glmnet, Clin Epigenetics 11 (1) (2019) 123.
[26] J.N. Mandrekar, Receiver operating characteristic curve in diagnostic test assessment, J. Thorac. Oncol. 5 (9) (2010) 1315–1316.
[27] X. Sun, Y. Mao, X. Zang, J. Zhang, G. Tang, F. Guo, A. Li, Optimization of comprehensive therapy for patients with mixed-type lung adenocarcinoma: a multicenter retrospective study, Transl. Lung Cancer Res. 8 (1) (2019) 31–39.
[28] Y.S. Chou, C.C. Yen, W.M. Chen, et al., Cytotoxic mechanism of PLK1 inhibitor GSK461364 against osteosarcoma:Mitotic arrest,apoptosis,cellular senescence, and synergistic effect with paclitaxel, Int. J. Oncol. 48 (3) (2016) 1187–1194.
[29] N. Dasgupta, B.K. Thakur, A. Ta, et al., Polo-like kinase 1 expression is suppressed by CCAAT/enhancer-binding protein alpha to mediate colon carcinoma cell differentiation and apoptosis, Biochim. Biophys. Acta Gen. Subj. 1861 (7) (2017) 1777–1787.
[30] M.J. Fernández-Aceñero, D. Cortés, T. Gómez del Pulgar, et al., PLK-1 expression is associated with histopathological response to neoadjuvant therapy of hepatic metastasis of colorectal carcinoma, Pathol. Oncol. Res. 22 (2) (2016) 377–383.
[31] H. Li, H. Wang, Z. Sun, et al., The clinical and prognostic value of polo-like kinase 1 in lung squamous cell carcinoma patients:immunohistochemical analysis, Biosci. Rep. (2017) [Epub ahead of print].
[32] T. Otto, P. Sicinski, Cell cycle proteins as promising targets in cancer therapy, Nat. Rev. Cancer 17 (2) (2017) 93–115.
[33] B. Vincenzi, G. Schiavon, M. Silletta, et al., Cell cycle alterations and lung cancer, Histol. Histopathol. 21 (4) (2006) 423–435.
[34] Cancer Genome Atlas Research Network, Comprehensive molecular profiling of lung adenocarcinoma [published correction appears in Nature 514 (7521) (2014 Oct 9) 262. Rogers, K [corrected to Rodgers, K]] [published correction appears in Nature. 2018 Jul;559(7715):E12]. Nature.
[35] G.G. Kuiper, E. Enmark, M. Pelto-Huikko, S. Nilsson, J.A. Gustafsson, Cloning of a novel receptor expressed in rat prostate and ovary, Proc Natl Acad Sci U S A. 93 (12) (1996) 5925–5930.
[36] D.C. Marquez-Garban, V. Mah, M. Alavi, et al., Progesterone and estrogen receptor expression and activity in human non-small cell lung cancer, Steroids 76 (9) (2011) 910–920.
[37] Q. He, M. Zhang, J. Zhang, et al., Correlation between epidermal growth factor receptor mutations and nuclear expression of female hormone receptors in non-small cell lung cancer: a meta-analysis, J. Thorac. Dis. 7 (9) (2015) 1588–1594.
[38] J.A. Clara, C. Monge, Y. Yang, N. Takebe, Targeting signalling pathways and the immune microenvironment of cancer stem cells - a clinical update, Nat. Rev. Clin. Oncol. 17 (4) (2020) 204–232.
[39] Hu Chan, Fan Yi, Xu Yuan, Zhijian Hu, Yiming Zeng, Research progress of lipid metabolism in the field of occurrence, development, diagnosis and treatment of lung cancer, Journal of Shanghai Jiaotong University (Medical Edition) [J] 42 (12) (2022) 1766–1771.
[40] R.S. Apte, D.S. Chen, N. Ferrara, VEGF in signaling and disease:beyond discovery and development, Cell 176 (6) (2019) 1248–1264.
[41] U. Kaiser, J. Hofmann, M. Schilli, et al., Steroid-hormone receptors in cell lines and tumor biopsies of human lung cancer, Int. J. Cancer 67 (3) (1996) 357–364.
[42] D. Rades, C. Setter, O. Dahl, S.E. Schild, F. Noack, The prognostic impact of tumor cell expression of estrogen receptor-α, progesterone receptor, and androgen receptor in patients irradiated for nonsmall cell lung cancer, Cancer 118 (1) (2012) 157–163.
[43] W.L. Ma, L.B. Jeng, H.C. Lai, P.Y. Liao, C. Chang, Androgen receptor enhances cell adhesion and decreases cell migration via modulating β1-integrin-AKT signaling in hepatocellular carcinoma cells, Cancer Lett. 351 (1) (2014) 64–71.
[44] K. Strebhardt, Multifaceted polo-like kinases: drug targets and antitargets for cancer therapy, Nat. Rev. Drug Discov. 9 (8) (2010 Aug) 643–660.
[45] Z.X. Wang, D. Xue, Z.L. Liu, et al., Overexpression of polo-like kinase 1 and its clinical significance in human non-small cell lung cancer, Int. J. Biochem. Cell Biol. 44 (1) (2012) 200–210.
[46] H. Li, H. Wang, Z. Sun, Q. Guo, H. Shi, Y. Jia, The clinical and prognostic value of polo-like kinase 1 in lung squamous cell carcinoma patients: immunohistochemical analysis, Biosci. Rep. 37 (4) (2017) BSR20170852.
[47] H.R. Jang, S.B. Shin, C.H. Kim, J.Y. Won, R. Xu, D.E. Kim, H. Yim, Correction: PLK1/vimentin signaling facilitates immune escape by recruiting Smad2/3 to PD-L1 promoter in metastatic lung adenocarcinoma, Cell Death Differ. 29 (10) (2022 Oct) 2106.

[48] C. Ma, H. Luo, J. Cao, C. Gao, X. Fa, G. Wang, Independent prognostic implications of RRM2 in lung adenocarcinoma, J. Cancer 11 (23) (2020) 7009–7022. Published 2020 Oct 17.
[49] B. Tang, W. Xu, Y. Wang, et al., Identification of critical ferroptosis regulators in lung adenocarcinoma that RRM2 facilitates tumor immune infiltration by inhibiting ferroptotic death, Clin Immunol 232 (2021) 108872.
[50] Z. Zhou, Q. Song, Y. Yang, L. Wang, Z. Wu, Comprehensive Landscape of RRM2 with immune infiltration in Pan-cancer, Cancers 14 (12) (2022) 2938. Published 2022 Jun 14.
[51] W. Li, G. Zhang, X. Li, et al., Thyroid hormone receptor interactor 13 (TRIP13) overexpression associated with tumor progression and poor prognosis in lung adenocarcinoma, Biochem. Biophys. Res. Commun. 499 (3) (2018) 416–424.
[52] Q. Zhang, Y. Dong, S. Hao, Y. Tong, Q. Luo, P. Aerxiding, The oncogenic role of TRIP13 in regulating proliferation, invasion, and cell cycle checkpoint in NSCLC cells, Int. J. Clin. Exp. Pathol. 12 (2019) 3357–3366.
[53] X. Ma, M. Xie, Z. Xue, J. Yao, Y. Wang, X. Xue, J. Wang, HMMR associates with immune infiltrates and acts as a prognostic biomaker in lung adenocarcinoma, Comput. Biol. Med. 151 (Pt A) (2022 Dec) 106213.
[54] R. He, S. Zuo, A robust 8-gene prognostic signature for early-stage non-small cell lung cancer, Front. Oncol. 9 (2019) 693.
[55] X. Lei, M. Zhang, B. Guan, Q. Chen, Z. Dong, C. Wang, Identification of hub genes associated with prognosis, diagnosis, immune infiltration and therapeutic drug in liver cancer by integrated analysis, Hum Genomics 15 (1) (2021) 39. Published 2021 Jun 29.