

Original Article

Comprehensive analysis of protein phosphatase 1 regulatory inhibitor subunit 14B, a molecule related to tumorigenesis, poor prognosis, and immune cell infiltration in lung adenocarcinoma

Yu-Xuan Zheng^{1*}, Shuo Shi^{1*}, Xiao-Hong Jiang^{2*}, Kai-Cheng Liu³, Zhao-Jie Qin⁴, Yong-Yong Wang¹, Zi-Hao Li¹, Ming-Wu Chen¹

¹Department of Cardio-Thoracic Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, People's Republic of China; ²Department of Orthopedic, Affiliated Minzu Hospital of Guangxi Medical University, Nanning, Guangxi, People's Republic of China; ³Department of Bone and Joint Surgery, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, People's Republic of China; ⁴Department of Orthopedic, The People's Hospital of Hechi, Hechi 547600, Guangxi, People's Republic of China. *Equal contributors.

Received September 17, 2022; Accepted December 28, 2022; Epub February 15, 2023; Published February 28, 2023

Abstract: Objective: To explore the relationship between Protein Phosphatase 1 Regulatory Inhibitor Subunit 14B (PPP1R14B) and the occurrence of lung adenocarcinoma (LUAD). Method: PPP1R14B expression was investigated using various databases, and its molecular functions and pathways were evaluated using Gene Set Variation Analysis (GSVA) and Gene Set Enrichment Analysis (GSEA). Then, the correlation between tumor mutations and PPP1R14B expression was analyzed. Furthermore, the regulation network and expression pathway axes of PPP1R14B were constructed. The correlation analysis between PPP1R14B and immune cell infiltration was performed using deconvolution algorithm analysis and the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm. Finally, quantitative real-time polymerase chain reaction (qRT-PCR) and immunohistochemical (IHC) staining of the clinical samples were used for expression validation. Results: PPP1R14B showed high expression in tumor tissue. PPP1R14B was associated with T and N stages and poor prognosis and was linked to the cell cycle, DNA repair, and low immune response. High PPP1R14B expression was associated with high tumor mutation rates. The upstream and downstream genes of PPP1R14B were identified, along with the construction of a protein-protein interaction network (PPI network) and the expression pathway axes of PPP1R14B. PPP1R14B expression was associated with poor immune cell infiltration and a negative correlation between PPP1R14B and mast cell and eosinophil infiltration. Conclusion: This study reveals high PPP1R14B expression in LUAD, its contribution to poor prognosis, molecular function, biological pathways, and impact on immune cell infiltration, and provides great insight into the role of PPP1R14B in LUAD tumorigenesis.

Keywords: Lung adenocarcinoma, protein phosphatase 1 regulatory inhibitor subunit 14B, tumorigenesis, immune cell infiltration

Introduction

Lung cancer is an increasingly prevalent global public health concern. In 2020, the number of patients with lung cancer reached 2,206,771, while mortality increased to 684,996 deaths worldwide. Lung cancer is the most fatal cancer among all cancer types [1]. Lung adenocarcinoma (LUAD) is a pathological type of lung cancer that originates from the mucous epithelium and glandular epithelium of the bronchi [2], and

is currently the most prevalent lung cancer type worldwide [3]. The high incidence rate and unfavorable prognosis of advanced clinical stage, as a feature of LUAD, remain. Pertinent biomolecules have been identified for the diagnosis of LUAD in recent years; however, the molecular mechanisms that underlie its occurrence and growth have yet to be elucidated.

Protein serine/threonine phosphatase 1 (PP1) is a member of the protein phosphatases fami-

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

ly. The phosphatases encoded by this gene family dephosphorylate the serine and threonine residues of proteins. Studies have shown that PP1 performs various functions in different cells by interacting with multiple regulatory subunits [4].

Protein Phosphatase 1 Regulatory Inhibitor Subunit 14B (PPP1R14B) was first described by Lagercrantz et al. and includes three introns and four exons that are located upstream of the PLCB3 gene [5]. As a PP1 inhibitor, it has been found in diverse tissue types and inhibits the functional domain of PP1, leading to multiple downstream changes [6]. Research has shown that PPP1R14B plays a crucial role in cell migration and regulation of endothelial cell retraction [7]. Furthermore, PPP1R14B was found to be highly expressed in glioma [8] and was linked to poor prognosis. In addition, PPP1R14B was found to be highly expressed in chronic lymphocytic leukemia [9] and ovarian clear cell carcinoma [10] gene expression datasets, also contributing to poor prognosis in these neoplasms. These studies imply that PPP1R14B may be involved in the occurrence of cancer and may lead to low survival rates. However, the levels of PPP1R14B expression, clinical features, functions, and mechanisms in LUAD remain unknown.

The expression and function of a gene are a consequence of the co-expression and interaction of the upstream and downstream genes. Exploration of the interaction network of downstream genes helps us to better understand the downstream molecular interactions and the primary function of PPP1R14B. Furthermore, constructing a protein-protein interaction (PPI) network of genes downstream of PPP1R14B allows a more thorough exploration of its pathway and molecular function. Competing endogenous RNAs (ceRNAs) function as a molecular expression regulatory mechanism. Increasing research has shown that ceRNA crosstalk plays an essential role in overall gene transcription regulation, thereby regulating multiple cell functions [11]. Furthermore, several studies have shown that the ceRNA network participates in tumorigenesis and tumor progression and metastasis [12]. Zhao et al. showed that LINC00466 regulates PPP1R14B expression by sponging miR-137 [8]. Since no prior studies

have focused on the ceRNA network that functions upstream of PPP1R14B, to further investigate long non-coding RNA (lncRNAs) and microRNAs (miRNAs) and how they interact with PPP1R14B, a ceRNA network targeting PPP1R14B was constructed.

Transcription factors (TFs) refer to one type of protein that functions as a regulatory molecule that recognizes and combines the regulatory element domains of genes and affects transcription [13]. The misregulation of TF leads to multiple diseases [14]; yet none of the upstream TFs of PPP1R14B have been reported in recent studies.

The tumor microenvironment (TME) consists of cancer cells and endothelial, stromal, and immune cells. Crosstalk of a cluster of immune cells contributes to clinical features and challenges cancer treatment [15]. However, no studies have reported an association between PPP1R14B with the TME in LUAD. Investigating the interaction of PPP1R14B with the TME will help us better understand the bioprocess related to PPP1R14B.

This study explored the relationship between PPP1R14B and clinicopathological features and prognosis based on multiple databases. Furthermore, gene set variation analysis (GSVA) and gene set enrichment analysis (GSEA) were used to investigate the prospective biofunction of PPP1R14B in LUAD. To determine the main biofunction of PPP1R14B, we constructed a PPI network and identified the hub genes of PPP1R14B. To explore the regulatory network of PPP1R14B, ceRNA and TFs, as upstream genes of PPP1R14B, were predicted and used to hypothesize and propose PPP1R14B signaling axes. Next, we evaluated the correlation between PPP1R14B and its axes with the TME. Finally, we performed quantitative real-time polymerase chain reaction (qRT-PCR) and immunohistochemical (IHC) staining of clinical samples to validate PPP1R14B expression at the mRNA and protein levels, respectively. In conclusion, our work revealed the role of PPP1R14B in the LUAD tumorigenesis and its molecular function, biological pathways, and impact on immune cell infiltration. This study may provide novel insights into the diagnosis, treatment, and understanding of the occurrence of lung adenocarcinoma.

Material and methods

Data acquisition

LUAD mRNA count data and fragments per kilobase of exon model per million mapped fragments (fpkm) data in The Cancer Genome Atlas (TCGA) database were acquired from the University of California, Santa Cruz Xena platform (UCSC Xena (xenabrowser.net)) which included 510 tumor tissue specimens and 58 normal tissue specimens. Fpkm data were transformed into log₂-TPM data. Clinical information (T stage, N stage, and M stage) and survival data (overall survival time and survival status) were also included. Gene expression data for GSE31210 [16] were acquired from the Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo/>) database which included 226 tumor tissue samples and 20 normal tissue samples. Pathological and survival data were also analyzed, and the data were log₂-transformed. Gene expression data for E-MEXP-231 were acquired from the Array-Express database (<https://www.ebi.ac.uk/arrayexpress/>) which included 49 tumor tissue samples and 9 normal tissue samples. TCGA-LUAD somatic copy number variant (CNV), single-nucleotide variant (SNV), and tumor mutation burden (TMB) data were acquired from UCSC Xena (UCSC Xena (xenabrowser.net)).

Pan-cancer overview and identification of differential PPP1R14B expression

To explore PPP1R14B expression in multiple types of cancer, the R package UCSCXenaShiny [17] was used to visualize differential PPP1R14B expression in tumor tissues and paracancerous normal tissues for various types of cancer from the PCAWG and TCGA databases.

To investigate PPP1R14B, the differential expression levels in tumor and normal tissues at different TNM and clinical stages were obtained from TCGA and GSE31210 and extracted for visualization.

The expression data in ArrayExpress were used to validate PPP1R14B expression at the mRNA level. Expression data in the CPTAC database were used to validate PPP1R14B expression at the protein level and were acquired and visualized using Ualcan [18]. Immunohistochemical

staining results from the Human Protein Atlas (HPA, www.proteinatlas.org) were used to validate PPP1R14B expression at the protein level.

Survival analysis of PPP1R14B

We investigated whether PPP1R14B expression affects patient's prognosis. Tumor samples from the TCGA and GSE31210 datasets were separated into high- or low-expression groups, according to PPP1R14B expression levels. Kaplan-Meier (KM) survival analyses of overall survival time (OS) in the two data sets were performed using the R package Survival, version 3.2-13 (<https://cran.r-project.org/web/packages/survival/index.html>) and visualized using the R package Survminer (<https://cran.rstudio.com/web/packages/survminer/index.html>).

Identification of differential expression genes (DEGs) and functional analysis of PPP1R14B

To identify DEGs of PPP1R14B, high and low PPP1R14B expression tumor samples in the TCGA and GSE31210 cohorts were evaluated using the R package Limma (<http://www.biocductor.org/packages/release/bioc/html/limma.html>) [19]. To explore the molecular functions of PPP1R14B, GSEA [20] and GSEA (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) were deployed in the TCGA and GSE31210 datasets, respectively, with R packages. The enriched gene sets "H: hallmark gene sets" for GSEA analysis, "C5: ontology gene sets" and "C2: curated gene sets" for GSEA were downloaded from the MSigDB database.

Genetic mutation and variants analysis

Since the GSEA and GSEA results showed that biological processes "G2M CHECKPOINT" and "DNA REPAIR" were enriched, we inferred that there might be an enrichment of genetic mutations and variant accumulation in the high PPP1R14B expression group tumors. To validate our hypothesis, the landscape of single-nucleotide variant (SNV) and copy number variation (CNV) data was visualized using the R package ComplexHeatmap [21] and a bar plot, respectively. SNV and CNV event counts for each sample were collected and compared. SNV or CNV may be involved in the process of neoantigens, which could act as potential markers for immunotherapy response [22]. To

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

inquire about the underlying relationship, the neoantigen data of LUAD were downloaded from the TCGA database (<https://tcia.at/home>) for follow-up analysis.

Construction of a PPI network of PPP1R14B downstream genes and identification of hub genes

Although we explored the potential biological function of PPP1R14B, the downstream gene interactions and primary biological function of PPP1R14B remained unknown. To identify the downstream genes of PPP1R14B, DEGs with adjusted *P*-Values < 0.05 and log fold change (logFC) \geq 1.0 in both cohorts were collected for the following analysis. To precisely predict downstream genes of PPP1R14B, upregulated DEGs in both datasets were first taken at the intersection and then analyzed using STRING database analysis [23] (www.string-db.org/). Functional or pathway enrichment analyses in the STRING database were automatically performed when the PPI network was constructed. To identify hub genes, genes involved in the four enriched functions or pathways marked in four colors were considered hub genes. Principles of function or pathway choice in each term list: false discovery rate < 0.01, with the highest strength score in one term. For all the false discovery rates > 0.01 in the one-term list, we choose the mini false discovery rate-enriched term.

Construction of a ceRNA network of PPP1R14B

To investigate molecule-regulated PPP1R14B, multiple databases were used to predict the molecules associated with PPP1R14B. To explore miRNAs that target PPP1R14B, R package MultiMiR [24] was used. According to the ceRNA theory, miRNAs that were predicted to inversely correlate with PPP1R14B (Pearson coefficient $R < -0.2$, $P < 0.05$) were considered miRNA-targeted PPP1R14B. miRNAs with differential expression levels between tumor and normal tissues that contributed to prognosis were considered to be the core miRNAs of PPP1R14B.

lncRNA-targeted miRNAs were predicted using lncBase [25] (<https://diana.e-ce.uth.gr/lncbasev3>), starbase [26] (<https://starbase.sysu.edu.cn/>) and LncACTdb [27] ([\[bigdata.net/LncACTdb/\]\(http://bigdata.net/LncACTdb/\)\). According to ceRNA theory, the lncRNAs that were predicted to inversely correlate with miRNA expression \(Pearson coefficient \$R < -0.2\$, \$P < 0.05\$ \) were considered lncRNA-targeted miRNAs. lncRNAs with differential expression levels between tumor and normal tissues that contributed to prognosis were considered core lncRNAs.](http://www.bio-</p></div><div data-bbox=)

Identification of transcription factor (TF) targets of PPP1R14B

Although we constructed a ceRNA network for PPP1R14B as an essential molecular regulator, the transcription factors related to PPP1R14B remained unknown. To further investigate the regulatory network of PPP1R14B, we investigated the upstream TFs of PPP1R14B using the online tool NetworkAnalyst [28] (<https://www.networkanalyst.ca>). According to the definition of a TF [13], predicted TFs with positive expression relationships (Pearson coefficient $R > 0.4$, $P < 0.05$) in the TCGA and GSE31210 cohorts were considered TFs that target PPP1R14B. TFs with differential expression levels between tumor and normal tissues that contributed to prognosis were considered core TFs.

Immune function analysis of PPP1R14B

Since the GSEA and GSVA results revealed “IL6 JAK STAT3 SIGNALING”, “INFLAMMATORY RESPONSE”, and “ACTIVATION OF IMMUNE RESPONSE” as being linked to PPP1R14B expression, we speculated that the expression of PPP1R14B is involved in shaping the TME. To investigate our hypothesis, we performed a series of analyses. To explore immune infiltration in tumors, we performed a single-sample gene set enrichment (ssGSEA) algorithm [29] in the R package “gsva” [30], Xcell [31], Cibersort (CBS) [32], MCPcounter (MCP) [33], and TIMER [34] in R package “ImmuneDeconv” [35] to estimate immune infiltration in the TCGA dataset. The estimated algorithm [36] revealed an abundance of immune cells in tumors. To investigate whether PPP1R14B expression was involved in the adaptive immunity or functions of T cells, the Tumor Immune Dysfunction and Exclusion algorithm (TIDE) [36] (tide.dfci.harvard.edu/) was used. A correlation analysis between the estimated immune cell infiltration results and PPP1R14B axial genes was then used to uncover the potential relationship. The Pearson correlation coeffi-

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

cient $|R| \geq 0.2$, p -value < 0.05 revealed the correlation between genes and immune cell infiltration. Only genes in axes that were related to one immune cell were considered in the relationship between the PPP1R14B axes and the immune cell. Since the GSEA results showed that the biological process “major histocompatibility complex (MHC) class II protein complex” was enriched, we inferred that PPP1R14B expression may be involved in the transcription of MHC molecules. To validate our hypothesis, we compared the expression of MHC molecules. To validate the results in the TCGA cohort, we repeated the same process as described above in the GSE31210 cohort.

Sample collection

This study was approved by the Ethics Committee Review Board of the First Affiliated Hospital of the Guangxi Medical University. All tissue collections were notified to the patient and subject to consent.

Thirty-one patients who underwent lung resection surgery at The First Affiliated Hospital of Guangxi Medical University in 2022 were enrolled in this study. Each patient was diagnosed with lung adenocarcinoma by a board-certified clinical pathologist. Samples collected from each patient-included tumor tissue and paired adjacent normal lung tissue, at a 5 cm distance.

RNA extraction and qRT-PCR

To investigate differences in PPP1R14B mRNA expression between adenocarcinoma tissue and adjacent normal lung tissue, qRT-PCR was performed to evaluate 31 pairs of samples. Tissues were treated with TRIzol reagent (15596-026, Ambion, USA) to extract the total RNA. The HiScript II Q Select RT SuperMix reagent (R223-01, Vazyme Biotech Co., Ltd., China) was used for reverse transcription. The process was performed at 50°C for 15 min and 85°C for 5 s, followed by a rest for 10 min at 4°C to obtain complementary DNA (cDNA). qRT-PCR was performed using cDNA, SYBR Green Master Mix (Q111-02, Vazyme) and primers (Sangon, China). The primer sequences are listed in **Table 2**. The cycling program was conducted using a PCR amplifier (EDC-810, Eastwin Scientific Equipment, China) for 40 cycles (degeneration at 95°C for 15 s, anneal-

ing at 60°C for 60 s, and melting curve analysis at 60°C for 60 s). Each sample was run independently in triplicate. Gene expression was normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH), and the relative gene expression of mRNAs was calculated using the $2^{-\Delta\Delta CT}$ method.

Immunohistochemical staining

To investigate the differences in PPP1R14B protein expression between adenocarcinoma tissue and adjacent normal lung tissue, we performed immunohistochemical staining. Frozen tissue samples were sent to the pathology department and cut into sections at the First Affiliated Hospital of Guangxi Medical University. The sections were immersed in 3% hydrogen peroxide for 15 min at room temperature to block endogenous peroxidase. Goat serum was used to block the sections at room temperature for 30 min. After blocking, sections were incubated with antibodies against PPP1R14B overnight at 4°C. Following incubation, the sections were washed with PBS and incubated with the corresponding secondary antibodies (K5007; Dako) at 37°C for 30 min. Subsequently, the sections were washed with PBS and stained with diaminobenzidine chromogenic solution (Beyotime, China). All sections were counterstained with hematoxylin. After staining, the sections were dehydrated with ethanol and dimethylbenzene and sealed. Fluorescence images were obtained using a standard electronic microscope.

Statistical analysis

Statistical analyses were performed using R, version 4.1.2. Two-group differential expression analysis was performed using the t-test. The Wilcoxon rank-sum test was used for comparisons that did not satisfy the t-test. A paired t-test was used to analyze the qRT-PCR results. The log-rank test was used to perform KM survival analysis. The Kruskal-Wallis test was used for multiple group comparisons. Pearson's correlation analysis was used to evaluate the relationship between these two items. The threshold of $P < 0.05$ indicates statistical significance of the correlation, and $P < 0.05$ was set as the significant difference threshold ($P > 0.05$, NS; $P < 0.05$, *; $P < 0.01$, **; $P < 0.001$, ***; $P < 0.0001$, ****).

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

Table 1. Data summary

Characteristic	Tumor cases	Normal lung cases
TCGA cohort		
n	510	58
T stage, n (%)		
T1	167 (32.7%)	
T2	276 (54.1%)	
T3	45 (8.8%)	
T4	19 (3.7%)	
Tx	3 (5.9%)	
N stage, n (%)		
N0	327 (64.1%)	
N+	171 (33.5%)	
Nx	12 (2.4%)	
M stage, n (%)		
M0	343 (67.3%)	
M1	25 (4.9%)	
Mx	142 (27.8%)	
GSE31210 cohort		
n	226	20
Clinical stage, n (%)		
I	168 (74.3%)	
II	58 (11.4%)	
E-MEXP-231 cohort		
n	49	9

Table 2. Primer sequences

Gene	Sequence (5' to 3')
PPP1R14B	Forward: CAAGGGAAGGTCACCGTCAAGTATG
	Reverse: GCTCATCCACGTCAATCTCCAGTTC
GAPDH	Forward: CAGGAGGCATTGCTGATGAT
	Reverse: GAAGGCTGGGGCTCATT

Results

PPP1R14B is upregulated in LUAD

The summary and clinical information of each cohort are shown in **Table 1**. The pan-cancer analysis revealed that PPP1R14B was upregulated in tumor samples compared to paracancerous normal tissues in most types of cancer in the PCAWG (**Figure 1A**) and TCGA databases (**Figure 1B**). The TCGA cohort results showed that PPP1R14B mRNA expression was higher in tumor tissues than in normal lung tissues (**Figure 1C**). Furthermore, PPP1R14B was differentially expressed at various T stages, with a

tendency toward higher expression in samples of higher clinical T stage (**Figure 1D**). In addition, PPP1R14B expression was higher in node metastasis (N+) samples (**Figure 1E**) and showed no significant difference between M stages (**Figure 1F**). Moreover, the qRT-PCR results (**Figure 1G**, **Table S1**) and immunohistochemical staining (**Figure 1H** and **1I**) clearly showed that PPP1R14B expression was elevated in tumor tissue compared to adjacent normal lung tissue in clinical patient samples. The results of the validated GSE31210 cohort showed that PPP1R14B mRNA expression was higher in tumor tissues than that in normal lung tissues (**Figure S1A**) and was more highly expressed in samples of clinical stage II than of stage I (**Figure S1B**). The results of the validated E-MEXP-231 cohort in ArrayExpress (**Figure S1C**), CPTAC (**Figure S1D**), and HPA (**Figure S1E** and **S1F**) databases revealed that PPP1R14B was upregulated in cancer tissue compared to normal lung tissue at the protein level, as revealed in the TCGA and GSE31210 results.

PPP1R14B contributes to poor prognosis

The KM plots indicated a significant difference between the high and low PPP1R14B expression groups in the TCGA (**Figure 2A**) and GSE31210 cohorts (**Figure 2B**). Thus, PPP1R14B expression contributes to a poor prognosis in LUAD.

Acquisition of PPP1R14B DEGs and PPP1R14B function analysis

R package limma was used to evaluate differences between the high and low PPP1R14B expression groups in TCGA and GSE31210 cohorts. With GSEA enrichment analysis, the results of the TCGA datasets revealed that the hallmark pathways of “MITOTIC SPINDLE”, “G2M CHECKPOINT”, “E2F TARGETS”, “DNA REPAIR”, and “MYC TARGETS” etc. were enriched in the high PPP1R14B expression group and significantly differed compared to the two low PPP1R14B expression groups. In contrast, the hallmark pathway “IL2 STAT5 SIGNALING” was significantly enriched in the low PPP1R14B expression group (**Figure 3A**). Similar enrichment results were observed in

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

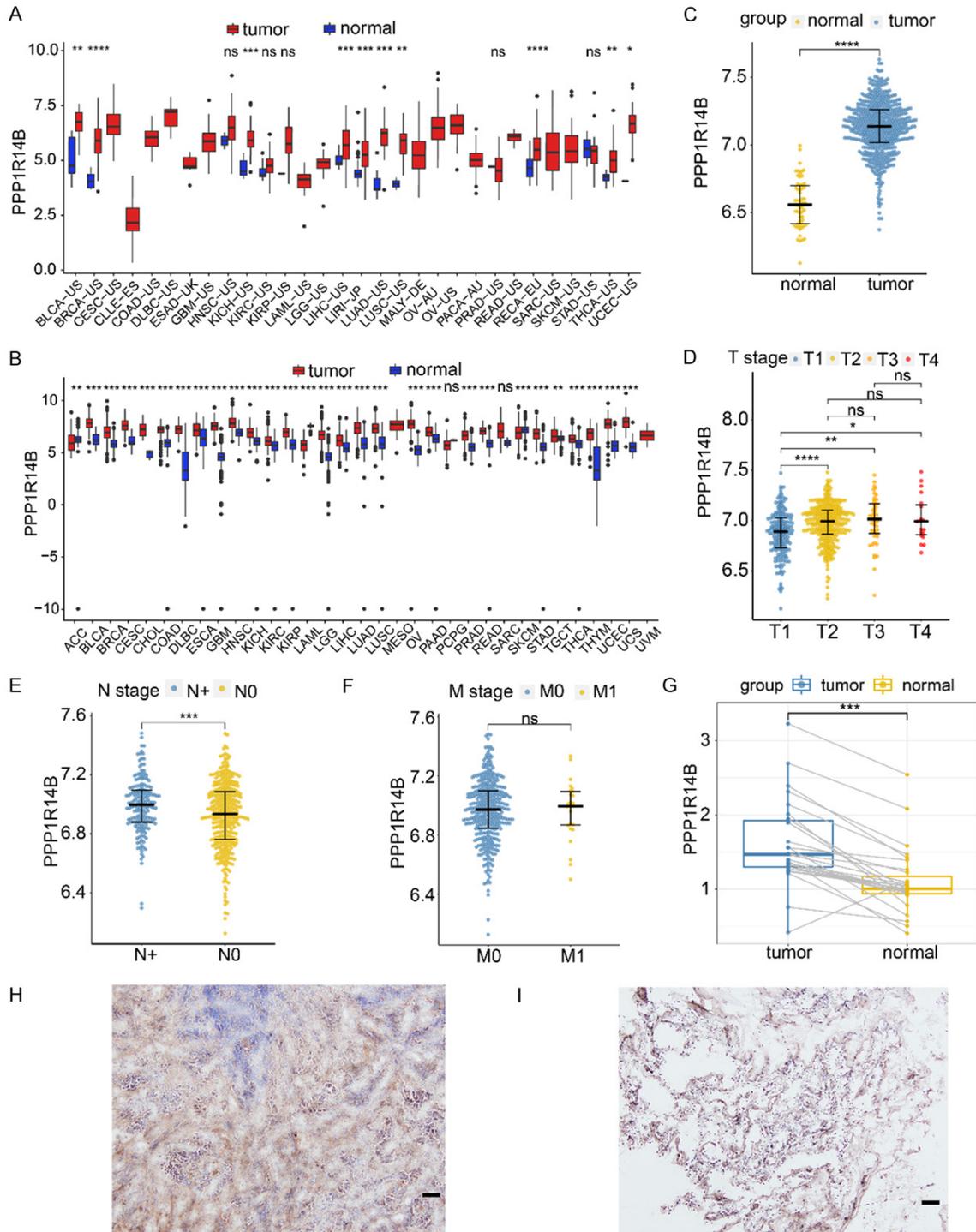


Figure 1. Protein Phosphatase 1 Regulatory Inhibitor Subunit 14B (PPP1R14B) was upregulated in lung adenocarcinoma (LUAD) and linked to multiple clinical characteristics. (A) Differential PPP1R14B expression in the tumor and normal groups from multiple cancers in the TCGA database, visualized by R package Xena Shiny. (B) Differential PPP1R14B expression in the tumor and normal groups from multiple cancers in the WGCNA database, visualized by R package Xena Shiny. (C) PPP1R14B expression was higher in tumor samples than that in normal lung samples in the TCGA cohort. T: tumor, N: normal. (D) PPP1R14B expression was linked to clinical T stage in the TCGA cohort. (E) PPP1R14B expression was linked to clinical N stage in the TCGA cohort. N+: N1, N2, N3 stage. (F) PPP1R14B expression did not significantly differ in clinical M stage samples in the TCGA cohort. (G) Relative PPP1R14B mRNA expression detected by qRT-PCR in adenocarcinoma and adjacent normal lung tissue from each patient. (H) Representative immunohistochemical staining of adenocarcinoma tissue and (I) adjacent normal lung tissue. ns, $P > 0.05$; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$. Scale bar: 100 μm .

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

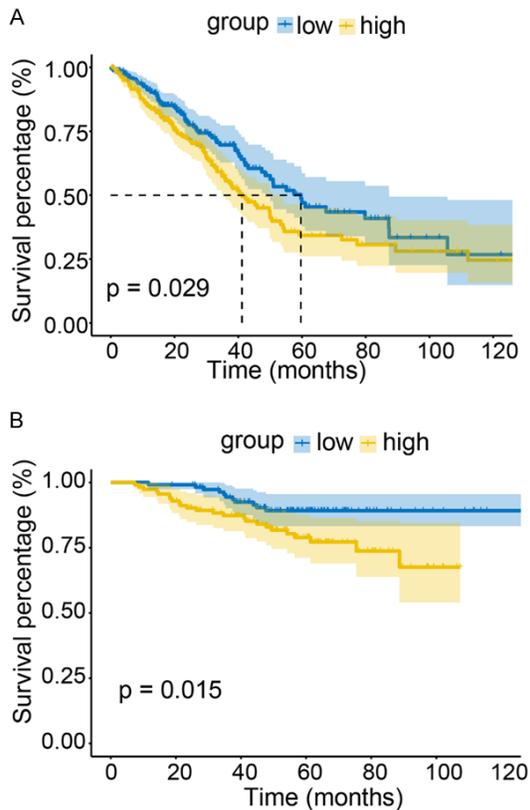


Figure 2. Analysis of the prognostic value of PPP1R14B in lung adenocarcinoma. A. Kaplan-Meier (KM) plot of PPP1R14B in the TCGA dataset. The comparison was between high and low expression (497 cases). B. KM survival analysis of PPP1R14B in the GSE31210 dataset (224 patients).

the validated GSE31210 cohort (**Figure 3B**). In the GSEA enrichment analysis, the results of TCGA datasets showed that the biological functions or pathways of “MITOTIC NUCLEAR DIVISION”, “REGULATION OF DNA REPAIR”, “NUCLEAR CHROMOSOME”, “DAMAGED DNA BINDING”, “P53 SIGNALING PATHWAY”, “MISMATCH REPAIR” etc. were enriched in the high PPP1R14B expression group. In contrast, “ACTIVATION OF IMMUNE RESPONSE”, “MHC CLASS II PROTEIN COMPLEX”, and “CELL ADHESION MOLECULES CAMS” were enriched in the low PPP1R14B expression group. Similar enrichment results were observed in the validated GSE31210 dataset (**Figure 3C**). The results indicated that PPP1R14B expression is involved in cell division, DNA repair, and immune response inhibition.

Genetic mutation and copy number variation analysis

The GSEA and GSEA results revealed that PPP1R14B is involved in DNA repair-related bio-

logical functions, including “BASE EXCISION REPAIR”, “HOMOLOGOUS RECOMBINATION”, and “MISMATCH REPAIR”. We hypothesize that the somatic mutation in the high PPP1R14B expression group accumulated to a greater extent than in the low PPP1R14B expression group. To validate our hypothesis, somatic mutation data were analyzed and compared between high and low PPP1R14B expression groups. The waterfall plot shows an overview of the top-ranked mutation-number genes between the two groups (**Figure 4A**). The bar plot shows an outline of the CNV distribution in all samples (**Figure 4B**). The results of the tumor mutation burden (TMB, **Figure 4C**), CNV event count (**Figure 4D**), and neoantigens (**Figure 4E**) were all significantly higher in the high PPP1R14B expression group than those in the low expression group.

PPI network construction of PPP1R14B downstream genes and the location of hub genes

DEGs with adjusted *P*-Values < 0.05 and logFC ≥ 1.0 in both datasets were collected for analysis (**Table S2**). Ninety-three upregulated DEGs were obtained in the TCGA cohort (**Figure 5A**), and 149 upregulated DEGs were obtained in the GSE31210 cohort (**Figure 5B**). Twenty-seven intersecting genes in both datasets were selected for PPI network construction (**Figure 5C**). Among these, 12 genes were extracted, and a PPI network was constructed using the STRING database (**Figure 5D**). Five genes (*CDC20*, *CDC25C*, *UBE2C*, *CENPA*, and *NUF2*) were enriched in four biological functions or pathways, which included “cell cycle” and “nuclear deviation” (**Table 3**), and finally, these were considered as hub genes. In the TCGA cohort, hub genes were upregulated in tumor samples (**Figure S2A**) and were associated with T stage (**Figure S2B**) and N stage (**Figure S2C**). No significant differences were observed in the M stage, and the results are not shown. In the GSE31210 cohort, hub genes were upregulated in tumor samples (**Figure S2D**) and were associated with clinical stage (**Figure S2E**). All hub genes contributed to poor prognosis in the TCGA (**Figure S2F**) and GSE31210 (**Figure S2G**) cohorts.

ceRNA network construction

With R package MultiMiR (for prediction results see **Table S3**), according to the ceRNA theory, three miRNAs (hsa-let-7b-5p, hsa-miR-219a-5p, hsa-miR-330-5p) that targeted PPP1R14B

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

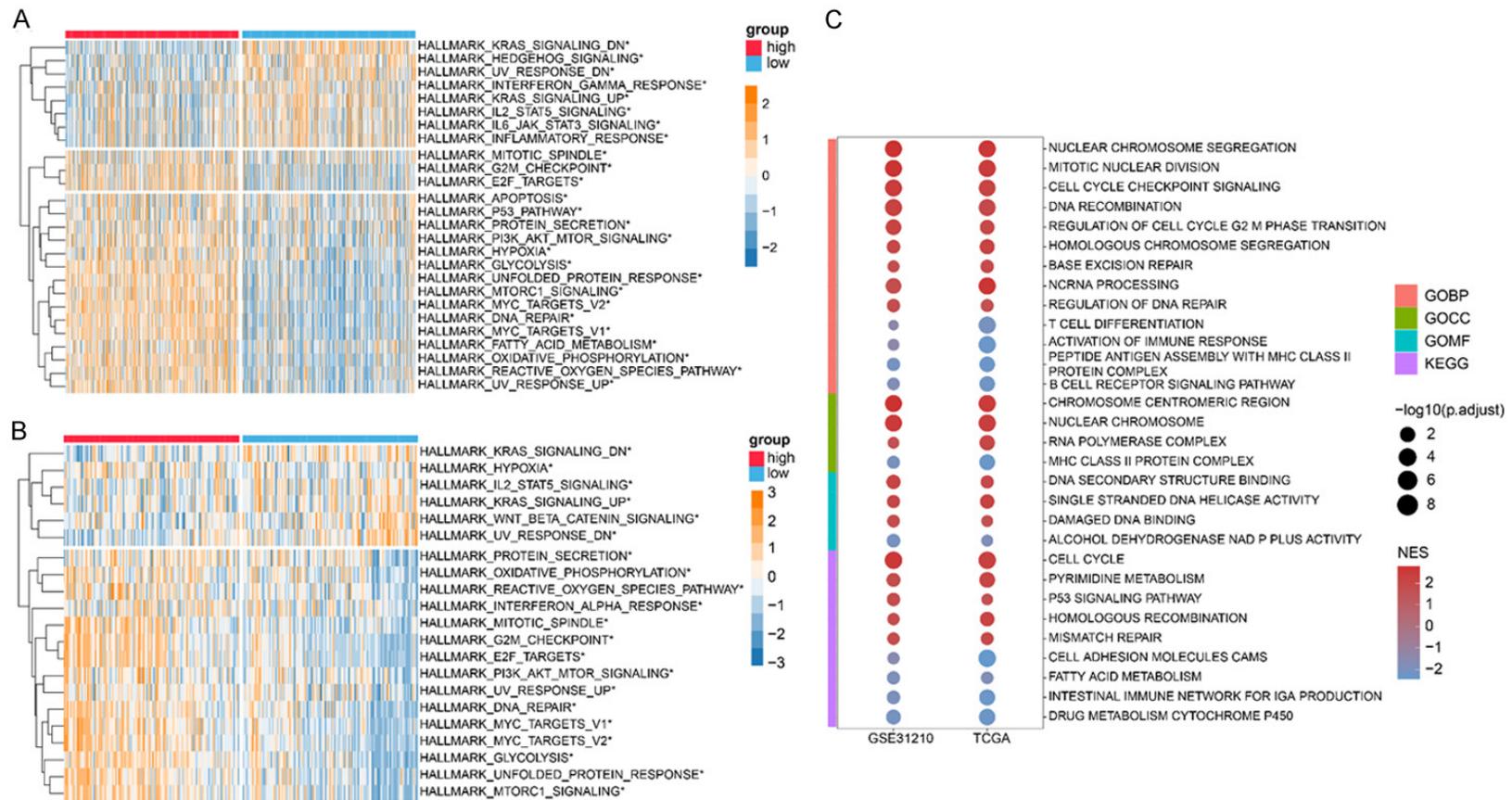


Figure 3. Biological function analysis of PPP1R14B. The heatmap shows the compared enrichment of hallmark pathways by Gene Set Variation Analysis (GSVA) in high and low PPP1R14B expression groups in the TCGA (A) and GSE31210 (B) cohorts. * Indicates significant difference between the two groups. The dot plot shows part of the upregulated biological functions or pathways that were identified by gene set enrichment analysis (GSEA) in high PPP1R14B expression groups in the TCGA and GSE31210 cohorts (C).

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

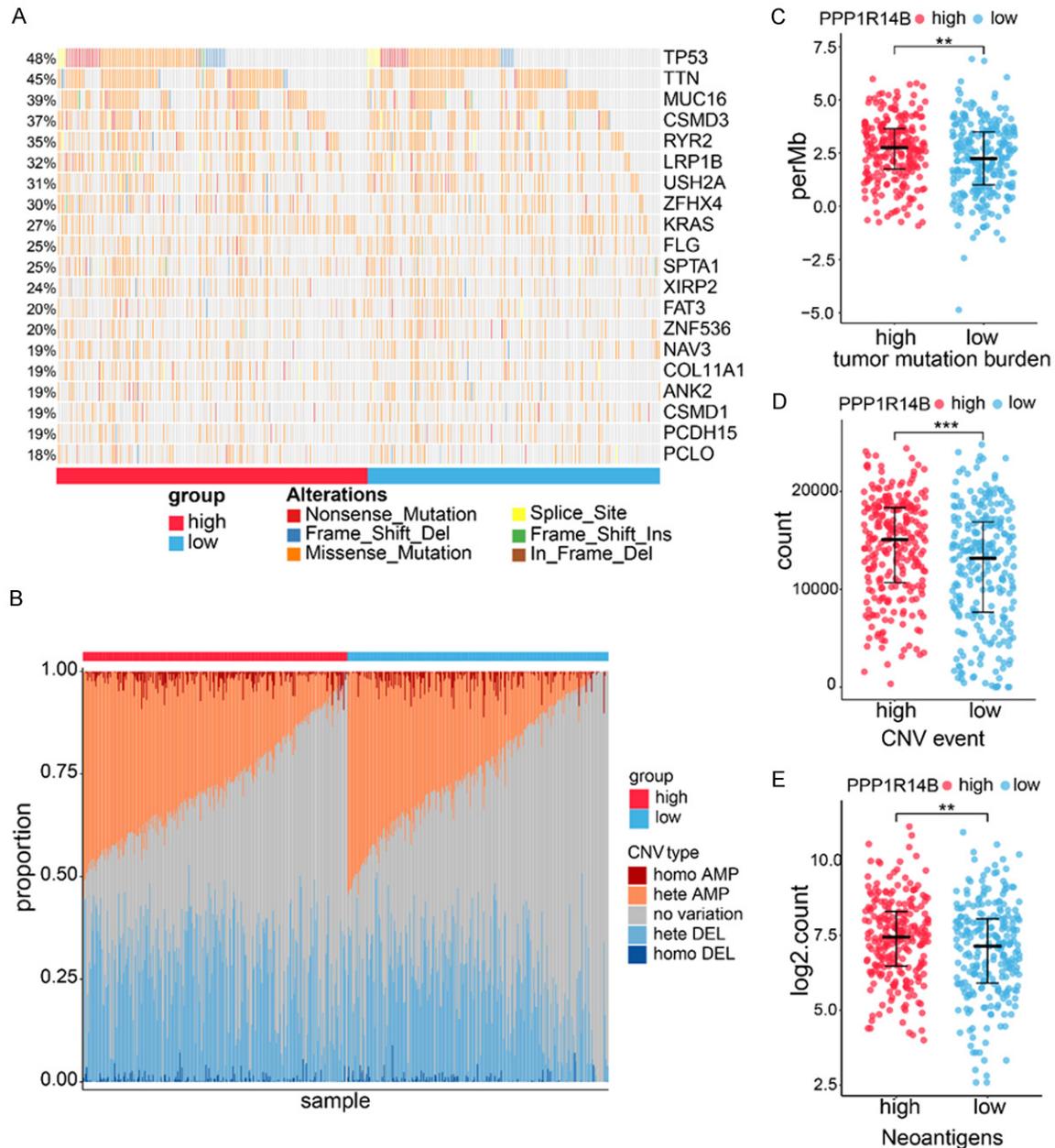


Figure 4. Somatic mutation analysis. (A) The waterfall plot shows the distribution of the 20 top-rank single nucleotide site variation genes in the high and low PPP1R14B expression groups. (B) The bar plot shows the copy number variation (CNV) type distribution in all samples. Comparisons of tumor mutation burden (TMB) (C), CNV events count (D), and neoantigens (E) between high and low PPP1R14B expression groups. **, $P < 0.01$; ***, $P < 0.001$. Homo: homozygosity; hete: heterozygosity; amp: amplification; del: deletion.

were obtained (Figure 6A). Using lncRNA databases (for prediction results see Table S4), five lncRNAs targeting hsa-let-7b-5p, one lncRNA targeting hsa-miR-219a-5p, and five lncRNAs targeting hsa-miR-330-5p were obtained (Figure 6B). A ceRNA network was constructed as shown (Figure 6C). Notably,

hsa-let-7b-5p was differentially expressed in tumor and normal lung tissues (Figure 6D), linked to prognosis (Figure 6E), and identified as a core miRNA. Furthermore, TMPO-AS1 was differentially expressed in tumor and normal lung tissues (Figure 6F), linked to prognosis (Figure 6G), and identified as a core lncRNA.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

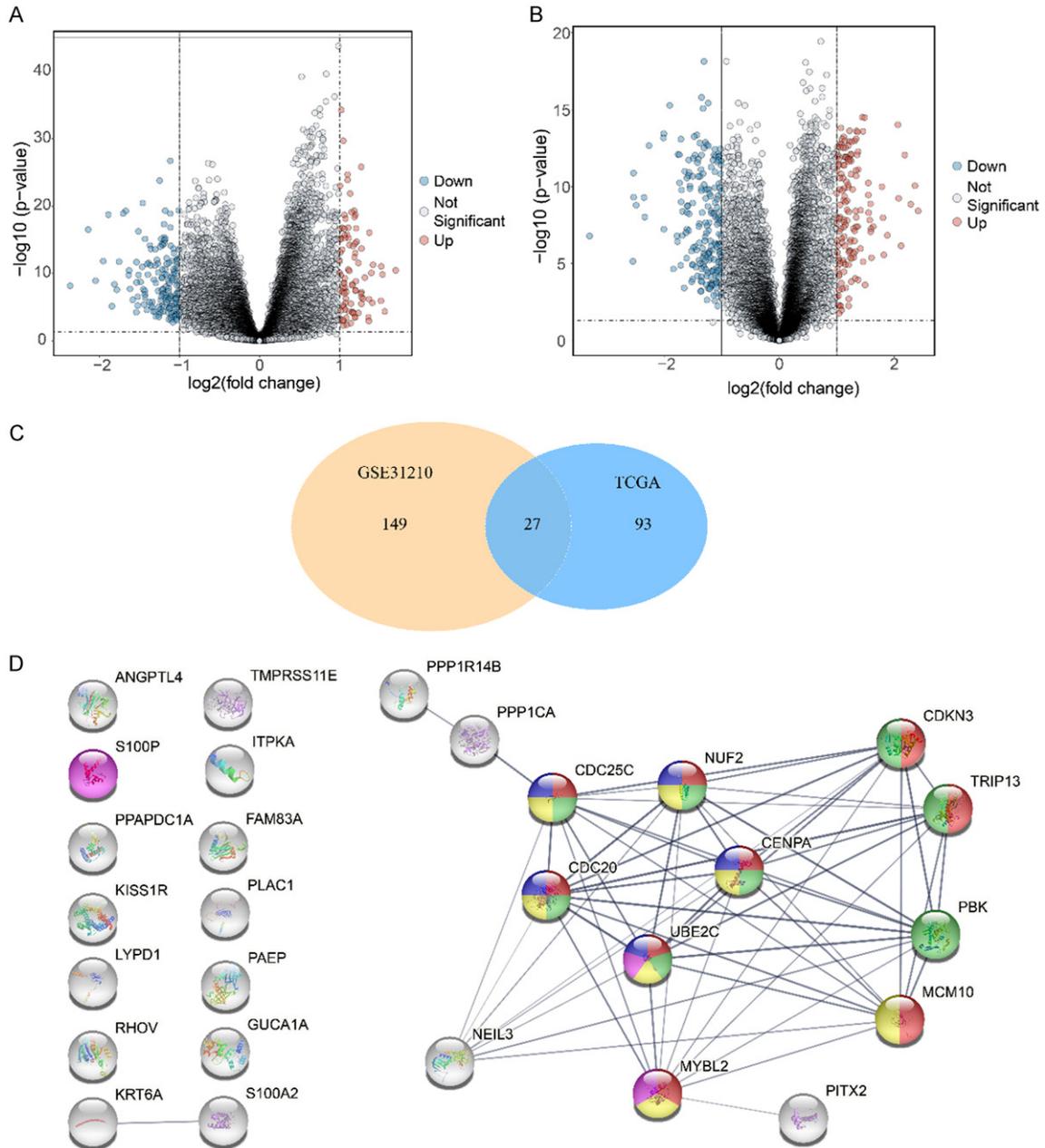


Figure 5. Construction of a protein-protein interaction (PPI) network of downstream genes and a search of hub genes. The volcano plot shows the differential expression genes (DEGs) between high and low PPP1R14B expression groups in the TCGA (A) and GSE31210 (B) cohorts. Each dot represents one gene. DEGs with p.adjust values < 0.05 and log fold change (logFC) < 1 or p.adjust values < 0.05 and logFC > 1 were respectively highlight in blue and red. (C) The intersection of significantly upregulated DEGs (p.adjust value < 0.05 and logFC > 1) in TCGA and GSE31210 cohorts. (D) PPI network of the intersection genes. Colors indicate the different biological functions or pathway genes that were involved. Five hub genes were identified.

Hub genes and TFs contribute to tumorigenesis and affect prognosis

Five TFs targeting PPP1R14B were obtained (for prediction results see [Table S5](#)), with co-expression related to PPP1R14B in the TCGA

cohort (**Figure 7A**) and the GSE31210 cohort (**Figure 7B**). FOXM1 was differentially expressed in tumor and normal lung tissues in the TCGA (**Figure 7C**) and GSE31210 (**Figure 7D**) cohorts, linked to prognosis (**Figure 7E** and **7F**), and thus considered a core TF. Finally, we iden-

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

Table 3. Annotated legend of the enriched items from the STRING database

Category	Term	Description	Count in network	Strength	FDR	Node
GO Process	GO:0007049	Cell cycle	10 of 695	1.03	0.00014	
STRING clusters	CL:4849	Mitotic Spindle Checkpoint, and mitotic nuclear division	8 of 148	1.61	8.48E-08	
Reactome	HSA-69278	Cell Cycle, Mitotic	7 of 518	1.01	0.0044	
UniProt Keywords	KW-0498	Mitosis	5 of 270	1.14	0.0183	
WikiPathways	WP2361	Gastric cancer network	3 of 28	1.91	0.0062	

Strength: description of the significance of the enrichment effect. FDR: false discovery rate.

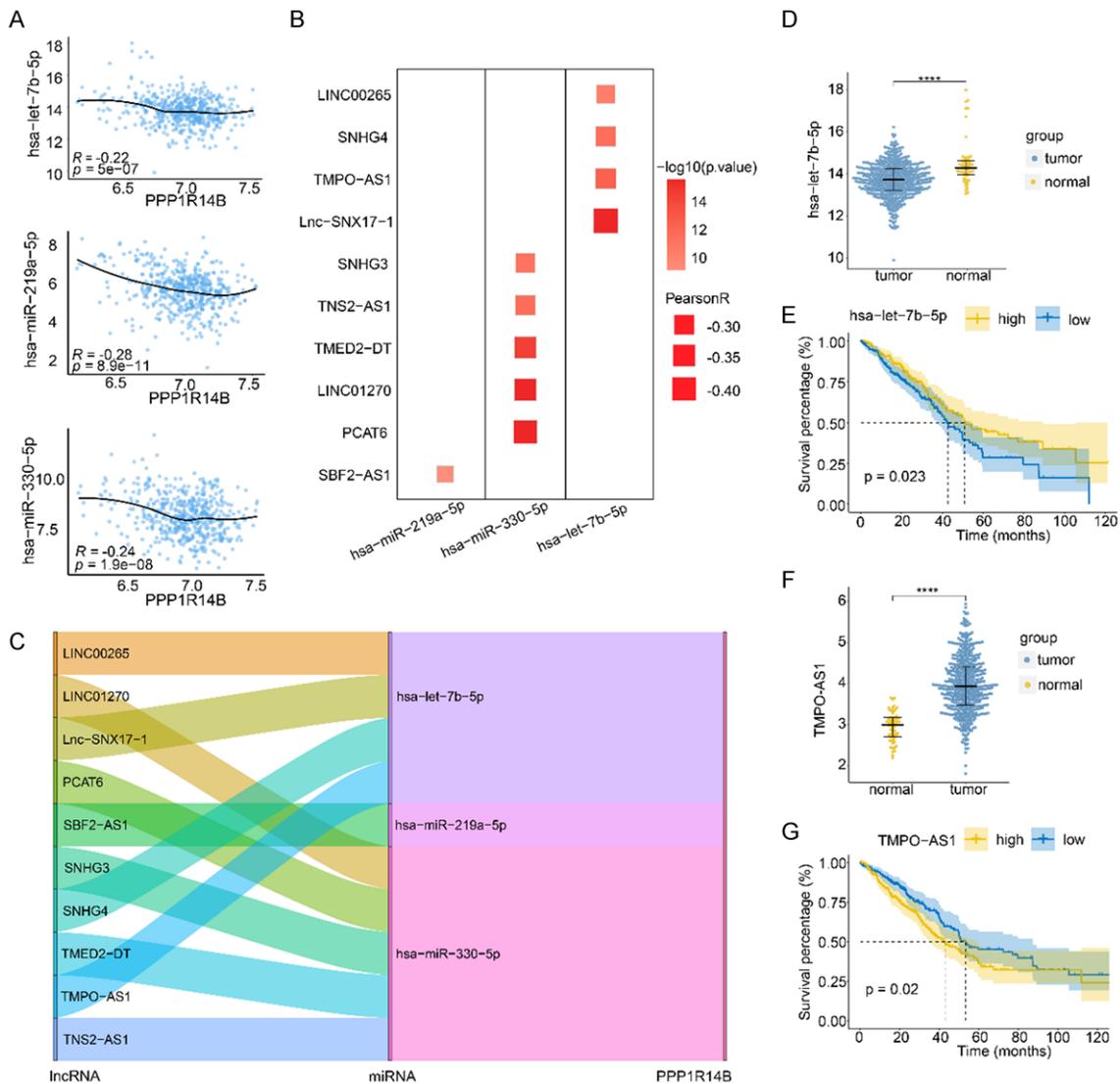


Figure 6. Construction of a competing endogenous RNAs (ceRNA) network target for PPP1R14B. A. The three predicted microRNAs (miRNAs) that target PPP1R14B with negative expression correlation. B. The five predicted long non-coding RNAs (lncRNAs) that target hsa-let-7b-5p, one predicted lncRNAs that target hsa-miR-219a-5p, and five predicted lncRNAs that target hsa-miR-330-5p, with negative expression correlation with the targeted miRNA. C.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

CeRNA network of PPP1R14B. D. Core miRNA hsa-let-7b-5p differential expression between tumor and normal lung tissue. E. Survival analysis of core miRNA hsa-let-7b-5p. F. Core lncRNA TMPO-AS1 differential expression between tumor and normal lung tissue. G. Survival analysis of core lncRNA TMPO-AS1. ****, $P < 0.0001$. R: Pearson correlation coefficient.

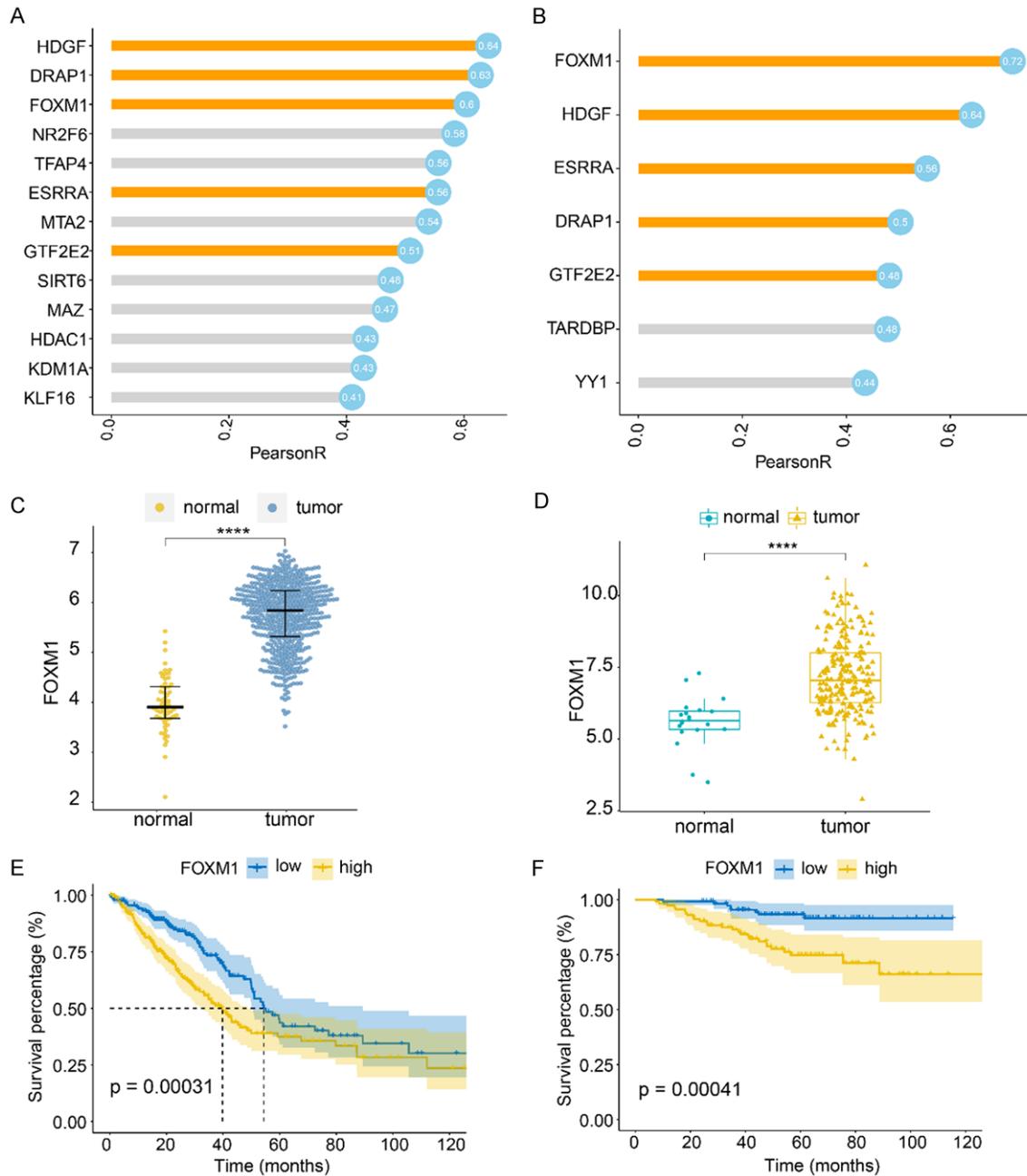


Figure 7. Identification of transcription factors (TFs) that target PPP1R14B. A. Predicted TFs and their correlation with expression in the TCGA data set. The number represents the Pearson's R correlation coefficient (all p values < 0.01). The orange bar indicates that the predicted TFs that correlated with PPP1R14B in both the TCGA and GSE31210 data sets. B. Predicted transcription factors expression correlation with PPP1R14B in the GSE31210 data set (all p value < 0.01). C, D. Core TF differential expression between the tumor and normal lung tissue in the TCGA and GSE31210 datasets. E, F. Survival analysis of core TFs in the TCGA and GSE31210 datasets. ****, $P < 0.0001$. Pearson R: Pearson correlation coefficient. ****, $P < 0.0001$.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

tified two potential axes of PPP1R14B: TMPO-AS1/hsa-let-7b-5p/PPP1R14B/hub genes and FOXM1/PPP1R14B/hub genes.

PPP1R14B expression links to immune infiltration

The landscape of immune cells estimated by ssGSEA showed high immune cell infiltration in the high-expression groups. Most of the immune cell infiltrations showed significant differences between the two groups. The results of the immune score evaluated by the Estimate algorithm showed a significant difference between the high and low PPP1R14B expression groups. The significantly different results of the Dysfunction score and the Exclusion score, evaluated by the TIDE algorithm in the two groups, revealed that high PPP1R14B expression excluded immune cell infiltration and the dysfunction of T cells in a portion of low PPP1R14B expression samples (**Figure 8A**). Overall, the results indicated that high PPP1R14B expression contributed to the high abundance of immune cell infiltration. The heatmap shows the correlation between multiple immune cell infiltrations and genes in the PPP1R14B axes in the TCGA cohort. The results indicated an inverse correlation between PPP1R14B axis gene expression and resting mast cells (estimated by Cibersort), eosinophils, mast cells (estimated by ssGSEA), myeloid dendritic cells, and endothelial cells (evaluated by MCPcounter) (**Figure 8B; Tables 4, S6**). The results also showed that PPP1R14B expression inversely correlated with B cells ($R = -0.293$), and CD4+ T cells ($R = -0.348$) (estimated by TIMER). Among these, low mast cells (**Figure S3A**) and eosinophil (**Figure S3B**) infiltration were linked to a poor prognosis in LUAD. The expression of MHC-I molecules was significantly higher in the group with high PPP1R14B expression. In contrast, the expression of some MHC-II molecules was significantly higher in the low PPP1R14B group (**Figure S3C**). Similar results were observed in the GSE31210 cohort: high abundance of immune cell infiltration tendency, high immune score in the low PPP1R14B expression group, and high exclusion score in the high PPP1R14B expression group (**Figure S4A**). The heatmap shows an inverse correlation between eosinophils, mast cells, and endothelial cells in the GSE31210 cohort (**Figure S4B**). The comparison of MHC molecules

between the high and low PPP1R14B expression groups showed that MHC-II molecules were downregulated in the high PPP1R14B expression group (**Figure S4C**).

Discussion

In our study, we found that PPP1R14B was overexpressed in LUAD and the expression levels significantly differed in clinical features, T stage, N stage, and clinical stage. To validate the difference in PPP1R14B expression between tumor and normal samples, additional datasets in ArrayExpress, CPTAC and HPA were utilized. To examine PPP1R14B expression and its ability to serve as a clinical biomarker, we collected patient samples for qRT-PCR and immunohistochemical staining. The results revealed that PPP1R14B had significantly higher expression in adenocarcinoma tissues than that in adjacent normal lung tissues in patient samples, as expected. These results indicated that PPP1R14B participates in growth and aggression in LUAD.

Moreover, recent studies have shown that PPP1R14B is associated with poor clinical outcomes in various cancers. Hence, we analyzed the association between PPP1R14B expression and prognosis and found that PPP1R14B was associated with poor clinical outcomes in LUAD.

This study focused on the molecular functions of PPP1R14B in LUAD; therefore, we investigated PPP1R14B using GSVA and GSEA. Similar to reported PPP1R14B functions in other types of cancers [37], we found that PPP1R14B promoted the cell cycle and cell division. However, unlike those studies, we found that PPP1R14B may function via the E2F, MYC, MTORC1, and P53 pathways. The activation of these pathways can affect the cell cycle and stimulate cell division [38-41]. This strongly suggests that PPP1R14B expression contributes to tumorigenesis and is linked to the TNM stage, clinical stage, and poor prognosis. In addition, the results imply that high PPP1R14B expression in LUAD may be sensitive to the above drug targets. Furthermore, the results showed that “IL2 STAT5 SIGNALING” pathways and “ACTIVATION OF IMMUNE RESPONSE” biological functions significantly differed between the high and low PPP1R14B expression groups. A previous study

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

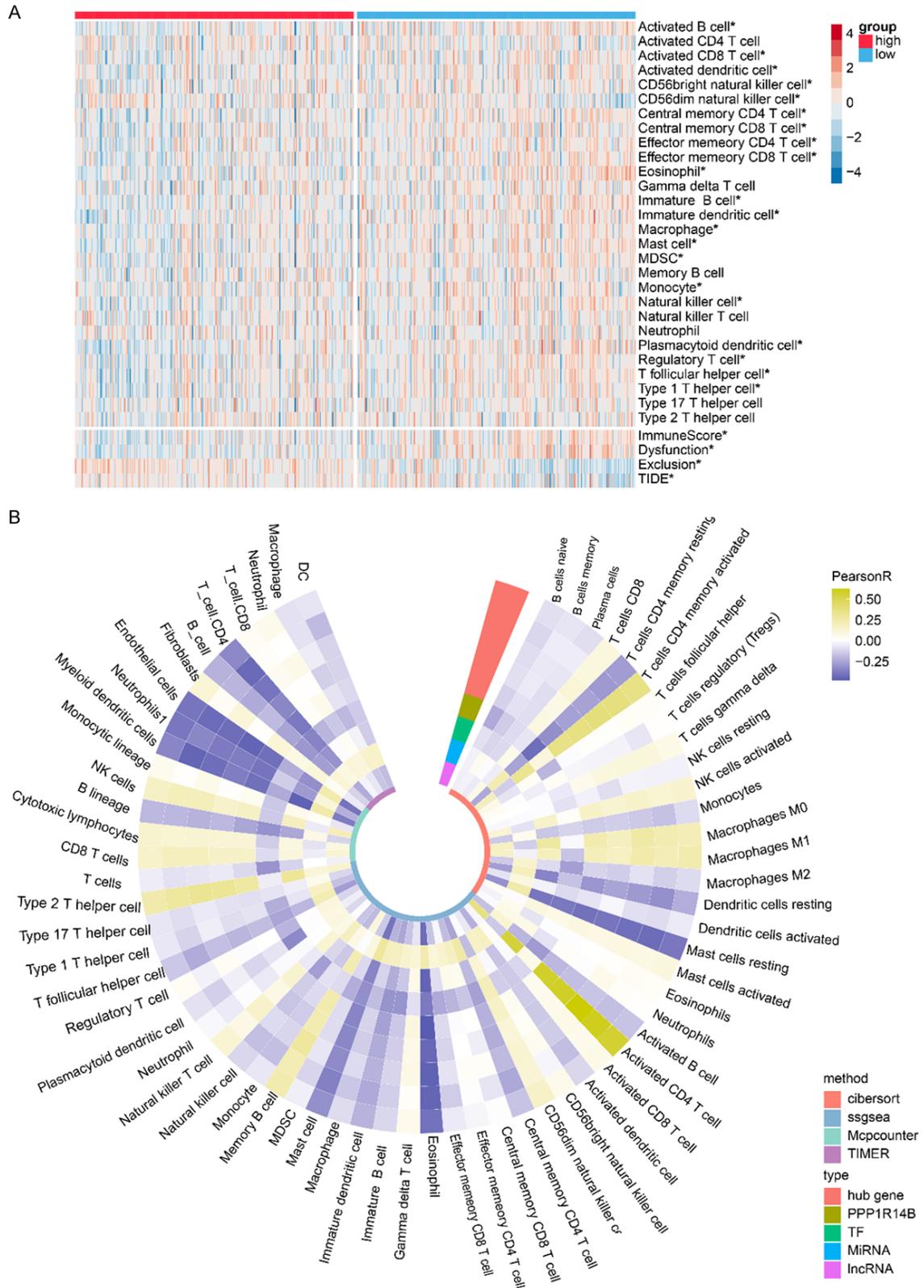


Figure 8. Immune-related analysis. A. Landscape of 28 immune cells estimated by ssGSEA, with immune scores evaluated by the Estimate algorithm, the Dysfunction and Exclusion scores evaluated by the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm, and their comparison in two groups in TCGA cohort. * Indicates a significant difference between the two groups. B. Correlation between PPP1R14B axial gene expression and immune cell infiltration in the TCGA cohort. Pearson R: Pearson correlation coefficient.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

Table 4. Correlation between immune cell infiltration and PPP1R14B axial gene expression

Gene	Mast cells resting	Eosinophil	Mast cell	Myeloid dendritic cells	Endothelial cells
Pearson Correlation Coefficient					
TMPO-AS1	-0.312	-0.428	-0.405	-0.343	-0.389
hsa-let-7b-5p	0.249	0.234	0.286	0.250	0.263
FOXM1	-0.399	-0.456	-0.354	-0.423	-0.291
PPP1R14B	-0.393	-0.388	-0.322	-0.171	-0.251
CDC20	-0.414	-0.483	-0.353	-0.345	-0.392
CDC25C	-0.401	-0.442	-0.347	-0.322	-0.434
CENPA	-0.421	-0.463	-0.333	-0.346	-0.379
NUF2	-0.431	-0.452	-0.367	-0.368	-0.390
UBE2C	-0.394	-0.413	-0.319	-0.294	-0.377
Pearson Correlation P-value					
TMPO-AS1	1.98E-21	1.45E-25	2.77E-20	1.64E-19	2.83E-22
hsa-let-7b-5p	4.71E-15	2.49E-14	5.59E-14	5.71E-13	3.89E-15
FOXM1	8.40E-21	9.94E-23	9.62E-15	7.18E-18	6.68E-16
PPP1R14B	1.03E-21	1.37E-24	4.64E-17	5.89E-13	5.71E-15
CDC20	1.04E-20	1.16E-23	3.81E-15	5.05E-16	7.09E-19
CDC25C	1.71E-20	2.50E-23	9.46E-16	1.19E-15	1.18E-20
CENPA	6.80E-20	4.84E-22	2.03E-14	7.07E-16	3.31E-18
NUF2	2.99E-22	3.27E-24	7.29E-17	7.55E-18	2.94E-20
UBE2C	8.45E-20	2.71E-22	1.25E-14	1.03E-14	1.68E-19

[38] showed that the stat5 signal pathway contributes to sustaining normal immune cell functions. Therefore, we conclude that PPP1R14B is involved in immune cell infiltration and immune response in the TME. Finally, we found that DNA repair-related pathways or functions, such as “G2M CHECKPOINT”, “DAMAGED DNA BINDING”, etc., were enriched in the high PPP1R14B expression group. We speculate that during PPP1R14B function, somatic mutations accumulate and promote tumor growth, as shown in previous studies [42]. To investigate our hypothesis, we analyzed somatic mutation data from LUAD-TCGA. As expected, the significant difference in TMB and copy number variation events that was observed in the high PPP1R14B expression group illustrated that more somatic mutations accumulate as PPP1R14B levels increase. High neoantigen levels in the high PPP1R14B expression group implied the accumulation of somatic mutations and the normal antigen presentation process.

A PPI network was constructed to explore the genes downstream of PPP1R14B. The PPI network revealed that PPP1R14B interacts with PPP1CA, a protein phosphatase 1 (PP1) sub-

unit, which activates downstream gene expression. Downstream hub genes *NUF2*, *CENPA*, *UBE2C*, *CDC20*, and *CDC25* were identified as being associated with PPP1R14B and contribute to cell proliferation. Further DEG and survival analyses suggested that PPP1R14B, together with its downstream hub genes, promotes tumor progression and leads to poor clinical prognosis. Similar to previous studies, five hub genes were upregulated in diverse cancers and were linked to cell proliferation and poor prognosis. Notably, Xie et al. found that *NUF2* was highly expressed in hepatocellular carcinoma (HCC) [43], and Wang et al. found that *CENPA* was upregulated in clear cell renal cell carcinoma [44]. Jin et al. found that *UBE2C* was highly expressed in head and neck squamous cell carcinomas [45], and Yang et al. found that *CDC20* was upregulated in HCC [46]. Xun et al. also found that *CDC25C* was highly expressed in HCC [47]. These results strongly suggest that PPP1R14B contributes to cell growth and tumorigenesis.

Numerous studies have shown that ceRNAs play a crucial role in the occurrence and progression of LUAD. As a miRNA sponge, lncRNA

reduces the amount of miRNA in cells and weakens the inhibitory effect of miRNA on silencing downstream target genes. To determine the ceRNA network targets of PPP1R14B, multiple databases were used for this study. Three PPP1R14B upstream miRNAs, hsa-let-7b-5p, hsa-miR-219a-5p, and hsa-miR-330-5p, were obtained. Subsequently, lncRNAs upstream of the three miRNAs were predicted, 11 lncRNAs were obtained, a ceRNA network was constructed, and the core miRNA hsa-let-7b-5p was identified. As previous studies showed, Pietro Di Fazio et al. found that hsa-let-7b-5p was downregulated in lung tumors [48]. Furthermore, Xi found that hsa-let-7b-5p inhibits the cell cycle in glioma cells [49]. The core lncRNA TMPO-AS1 was then identified, and recent studies have shown that TMPO-AS1 is upregulated in multiple cancers, such as bladder cancer [46] and nasopharyngeal carcinoma [47] and is associated with the cell cycle and poor prognosis. Further experiments are required to confirm the influence of these lncRNAs in other cancers.

TFs play a crucial role in molecular regulatory mechanisms. To explore the TFs that target PPP1R14B, we used NetworkAnalyst to predict TFs. Five TFs were identified as candidate TFs that target PPP1R14B. Subsequently, FOXM1 was identified as the core TF, which was upregulated and linked to poor prognosis in the TCGA and GSE31210 datasets. Previous studies have shown that FOXM1 is upregulated in various cancers, such as bladder cancer [50] and laryngeal squamous cell carcinoma [51], and promotes cell cycle progression. However, further experiments are required to validate the relationship between FOXM1 and PPP1R14B.

The LUAD microenvironment landscape helps to explain cancer aggressiveness and phenotypes. Numerous studies have shown that TME is essential for tumor occurrence and progression. To reveal the underlying relationship between PPP1R14B and the TME, we evaluated and compared the immune characteristics of the two groups and further analyzed the relationship between tumor immune infiltration and PPP1R14B axial genes. The immune cell landscape illustrated the low abundance of immune cell infiltration and the low immune score in the high PPP1R14B expression group. This, along with the GSVA results, suggest that

PPP1R14B suppressed normal immune functions or excluded immune cells from entering the tumor microenvironment by down-regulating the IL2-STAT5 signaling pathway. This was in agreement with the GSEA results, as expected. The correlation heatmap confirmed this hypothesis, and the results showed that PPP1R14B expression was linked to low infiltration of resting mast cells, eosinophils, mast cells, myeloid dendritic cells, and endothelial cells.

As Atsushi Kato's study expounded, the normal functions of eosinophils, mast cells, epithelial cells, and other immune cells and immune functions are maintained by activating the IL2-STAT5 pathway [52], which explains our results. Among the above immune cells, the low infiltration of eosinophils and mast cells contributes to poor prognosis. Previous studies reported that eosinophils in the TME showed anti-tumorigenic characteristics and benefited patients from prolonged survival [53]. However, a few studies have reported that eosinophils in the TME play a pro-tumorigenic role [54]. This may suggest that eosinophils functions are determined by the cancer type or potential subtypes of eosinophils that have yet to be explored.

Additionally, prior research has reported that mast cell infiltration in LUAD contributes to good clinical outcomes and may provide a sensitive marker for immunotherapy response [55]. Moreover, the GSEA results showed that PPP1R14B downregulated the assembly of the MHC-II protein complex. The comparison of MHC molecules illustrated that MHC-I molecules had high expression, while MHC-II molecules exhibited low expression in the high PPP1R14B expression group. Notably, MHC molecules play a crucial role in tumorigenesis. MHC-I molecules activate cytotoxic CD8+ T cells by processing and presenting antigen peptides, whereas MHC-II molecules function in tumorigenesis by activating CD4+ T cells [56, 57]. The significant difference between neoantigens and MHC molecules indicated that PPP1R14B plays a role in TME immunity by disrupting the normal functions of MHC-II molecules. This is also implied by the inverse correlation between PPP1R14B expression and CD4+ T cells ($R = -0.3484$, TIMER) and B cells ($R = -0.29324$, TIMER). In conclusion, these results suggest that PPP1R14B and its axial genes play a crucial role in the TME and explain

how PPP1R14B contributes to poor clinical outcomes in LUAD.

This study has a few limitations. First, samples were only from a subset of patients with LUAD, and future studies should include more patients with LUAD to determine the generalizability of our findings. Second, this study was based only on LUAD; thus, our PPP1R14B findings may be inapplicable to other types of cancers, and future studies should include more types of cancers. Third, previous studies have shown that somatic mutations alter gene expression and thus change cancer characteristics and phenotypes [58]. Our study only analyzed the differences in somatic mutations between the high and low PPP1R14B expression groups. Future studies should focus on differences in SNV or CNV genes and the functional changes that are affected by PPP1R14B. Finally, our experiment only focuses on PPP1R14B expression in patient samples. To better study PPP1R14B, further functional experiments on corresponding cell lines remain warranted.

In conclusion, PPP1R14B expression is upregulated in LUAD and is associated with clinical features and poor prognosis. We found that PPP1R14B enhanced tumorigenesis by promoting cell cycle progression. We constructed a PPP1R14B axis based on its upstream and downstream genes. We also found that PPP1R14B expression inhibits immune cell infiltration. We provide significant evidence that supports the hypothesis that high PPP1R14B expression levels contribute to tumorigenesis and poor clinical outcomes in LUAD.

Acknowledgements

We would like to thank all of the members who participated in this study, those who shared R language codes on Internet, and Editage (www.editage.com) for English language editing.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Ming-Wu Chen, Department of Cardio-Thoracic Surgery, The First Affiliated Hospital of Guangxi Medical University, Shuang-Yong Road, 6#, Nanning 530021, Guangxi, People's Republic of China. Tel: +86-13507888835; E-mail: chen535@126.com

References

- [1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71: 209-249.
- [2] Noguchi M, Morikawa A, Kawasaki M, Matsuno Y, Yamada T, Hirohashi S, Kondo H and Shimamoto Y. Small adenocarcinoma of the lung. Histologic characteristics and prognosis. *Cancer* 1995; 75: 2844-2852.
- [3] Travis WD, Brambilla E, Nicholson AG, Yatabe Y, Austin JHM, Beasley MB, Chirieac LR, Dacic S, Duhig E, Flieder DB, Geisinger K, Hirsch FR, Ishikawa Y, Kerr KM, Noguchi M, Pelosi G, Powell CA, Tsao MS and Wistuba I; WHO Panel. The 2015 World Health Organization classification of lung tumors: impact of genetic, clinical and radiologic advances since the 2004 classification. *J Thorac Oncol* 2015; 10: 1243-1260.
- [4] Cohen PT. Novel protein serine/threonine phosphatases: variety is the spice of life. *Trends Biochem Sci* 1997; 22: 245-251.
- [5] Lagercrantz J, Carson E, Larsson C, Nordenskjold M and Weber G. Isolation and characterization of a novel gene close to the human phosphoinositide-specific phospholipase C beta 3 gene on chromosomal region 11q13. *Genomics* 1996; 31: 380-384.
- [6] Eto M, Karginov A and Brautigan DL. A novel phosphoprotein inhibitor of protein type-1 phosphatase holoenzymes. *Biochemistry* 1999; 38: 16952-16957.
- [7] Tountas NA and Brautigan DL. Migration and retraction of endothelial and epithelial cells require PHI-1, a specific protein-phosphatase-1 inhibitor protein. *J Cell Sci* 2004; 117: 5905-5912.
- [8] Zhao M, Shao Y, Xu J, Zhang B, Li C and Gong J. LINC00466 impacts cell proliferation, metastasis and sensitivity to temozolomide of glioma by sponging miR-137 to regulate PPP1R14B expression. *Onco Targets Ther* 2021; 14: 1147-1159.
- [9] Mosquera Orgueira A, Antelo Rodriguez B, Diaz Arias JA, Diaz Varela N and Bello Lopez JL. A three-gene expression signature identifies a cluster of patients with short survival in chronic lymphocytic leukemia. *J Oncol* 2019; 2019: 9453539.
- [10] Worley MJ Jr, Liu S, Hua Y, Kwok JS, Samuel A, Hou L, Shoni M, Lu S, Sandberg EM, Keryan A, Wu D, Ng SK, Kuo WP, Parra-Herran CE, Tsui SK, Welch W, Crum C, Berkowitz RS and Ng SW. Molecular changes in endometriosis-associated ovarian clear cell carcinoma. *Eur J Cancer* 2015; 51: 1831-1842.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

- [11] Tay Y, Rinn J and Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature* 2014; 505: 344-352.
- [12] Karreth FA and Pandolfi PP. ceRNA cross-talk in cancer: when ce-bling rivalries go awry. *Cancer Discov* 2013; 3: 1113-1121.
- [13] Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR and Weirauch MT. The human transcription factors. *Cell* 2018; 172: 650-665.
- [14] Lee TI and Young RA. Transcriptional regulation and its misregulation in disease. *Cell* 2013; 152: 1237-1251.
- [15] Vitale I, Manic G, Coussens LM, Kroemer G and Galluzzi L. Macrophages and metabolism in the tumor microenvironment. *Cell Metab* 2019; 30: 36-50.
- [16] Okayama H, Kohno T, Ishii Y, Shimada Y, Shirai-shi K, Iwakawa R, Furuta K, Tsuta K, Shibata T, Yamamoto S, Watanabe S, Sakamoto H, Kumamoto K, Takenoshita S, Gotoh N, Mizuno H, Sarai A, Kawano S, Yamaguchi R, Miyano S and Yokota J. Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res* 2012; 72: 100-111.
- [17] Wang S, Xiong Y, Zhao L, Gu K, Li Y, Zhao F, Li J, Wang M, Wang H, Tao Z, Wu T, Zheng Y, Li X and Liu XS. UCSCXenaShiny: an R/CRAN package for interactive analysis of UCSC xena data. *Bioinformatics* 2021; 38: 527-9.
- [18] Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK and Varambally S. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia* 2017; 19: 649-658.
- [19] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47.
- [20] Hanzelmann S, Castelo R and Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013; 14: 7.
- [21] Gu Z, Eils R and Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016; 32: 2847-2849.
- [22] Robbins PF, Lu YC, El-Gamil M, Li YF, Gross C, Gartner J, Lin JC, Teer JK, Cliften P, Tycksen E, Samuels Y and Rosenberg SA. Mining exomic sequencing data to identify mutated antigens recognized by adoptively transferred tumor-reactive T cells. *Nat Med* 2013; 19: 747-752.
- [23] Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, Doncheva NT, Legeay M, Fang T, Bork P, Jensen LJ and von Mering C. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res* 2021; 49: D605-D612.
- [24] Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, Mahaffey S, Rossi S, Calin GA, Bemis L and Theodorescu D. The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res* 2014; 42: e133.
- [25] Karagkouni D, Paraskevopoulou MD, Tastsoglou S, Skoufos G, Karavangeli A, Pierros V, Zacharopoulou E and Hatzigeorgiou AG. DIANA-LncBase v3: indexing experimentally supported miRNA targets on non-coding transcripts. *Nucleic Acids Res* 2020; 48: D101-D110.
- [26] Li JH, Liu S, Zhou H, Qu LH and Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res* 2014; 42: D92-97.
- [27] Wang P, Li X, Gao Y, Guo Q, Wang Y, Fang Y, Ma X, Zhi H, Zhou D, Shen W, Liu W, Wang L, Zhang Y, Ning S and Li X. LncACTdb 2.0: an updated database of experimentally supported ceRNA interactions curated from low- and high-throughput experiments. *Nucleic Acids Res* 2019; 47: D121-D127.
- [28] Zhou G, Soufan O, Ewald J, Hancock REW, Basu N and Xia J. NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res* 2019; 47: W234-W241.
- [29] Rooney MS, Shukla SA, Wu CJ, Getz G and Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 2015; 160: 48-61.
- [30] Hanzelmann S, Castelo R and Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics* 2013; 14: 7.
- [31] Aran D, Hu Z and Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* 2017; 18: 220.
- [32] Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M and Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015; 12: 453-457.
- [33] Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautes-Fridman C, Fridman WH and de Reynies A. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol* 2016; 17: 218.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

- [34] Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, Jiang P, Shen H, Aster JC, Rodig S, Signoretti S, Liu JS and Liu XS. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. *Genome Biol* 2016; 17: 174.
- [35] Sturm G, Finotello F and List M. Immunedeconv: an R package for unified access to computational methods for estimating immune cell fractions from bulk RNA-sequencing data. *Methods Mol Biol* 2020; 2120: 223-232.
- [36] Fu J, Li K, Zhang W, Wan C, Zhang J, Jiang P and Liu XS. Large-scale public data reuse to model immunotherapy response and resistance. *Genome Med* 2020; 12: 21.
- [37] Deng M, Peng L, Li J, Liu X, Xia X and Li G. PPP1R14B is a prognostic and immunological biomarker in pan-cancer. *Front Genet* 2021; 12: 763561.
- [38] Baluapuri A, Wolf E and Eilers M. Target gene-independent functions of MYC oncoproteins. *Nat Rev Mol Cell Biol* 2020; 21: 255-267.
- [39] Kent LN and Leone G. The broken cycle: E2F dysfunction in cancer. *Nat Rev Cancer* 2019; 19: 326-338.
- [40] Szwed A, Kim E and Jacinto E. Regulation and metabolic functions of mTORC1 and mTORC2. *Physiol Rev* 2021; 101: 1371-1426.
- [41] Joerger AC and Fersht AR. The p53 pathway: origins, inactivation in cancer, and emerging therapeutic approaches. *Annu Rev Biochem* 2016; 85: 375-404.
- [42] Glover TW, Wilson TE and Arlt MF. Fragile sites in cancer: more than meets the eye. *Nat Rev Cancer* 2017; 17: 489-501.
- [43] Xie X, Jiang S and Li X. Nuf2 is a prognostic-related biomarker and correlated with immune infiltrates in hepatocellular carcinoma. *Front Oncol* 2021; 11: 621373.
- [44] Wang Q, Xu J, Xiong Z, Xu T, Liu J, Liu Y, Chen J, Shi J, Shou Y, Yue C, Liu D, Liang H, Yang H, Yang X and Zhang X. CENPA promotes clear cell renal cell carcinoma progression and metastasis via Wnt/beta-catenin signaling pathway. *J Transl Med* 2021; 19: 417.
- [45] Jin Z, Zhao X, Cui L, Xu X, Zhao Y, Younai F, Mes-sadi D and Hu S. UBE2C promotes the progression of head and neck squamous cell carcinoma. *Biochem Biophys Res Commun* 2020; 523: 389-397.
- [46] Yang G, Wang G, Xiong Y, Sun J, Li W, Tang T and Li J. CDC20 promotes the progression of hepatocellular carcinoma by regulating epithelial-mesenchymal transition. *Mol Med Rep* 2021; 24: 483.
- [47] Xun R, Lu H and Wang X. Identification of CDC25C as a potential biomarker in hepatocellular carcinoma using bioinformatics analysis. *Technol Cancer Res Treat* 2020; 19: 1533033820967474.
- [48] Di Fazio P, Maass M, Roth S, Meyer C, Grups J, Rexin P, Bartsch DK and Kirschbaum A. Expression of hsa-let-7b-5p, hsa-let-7f-5p, and hsa-miR-222-3p and their putative targets HMGA2 and CDKN1B in typical and atypical carcinoid tumors of the lung. *Tumour Biol* 2017; 39: 1010428317728417.
- [49] Xi X, Chu Y, Liu N, Wang Q, Yin Z, Lu Y and Chen Y. Joint bioinformatics analysis of underlying potential functions of hsa-let-7b-5p and core genes in human glioma. *J Transl Med* 2019; 17: 129.
- [50] Yi L, Wang H, Li W, Ye K, Xiong W, Yu H and Jin X. The FOXM1/RNF26/p57 axis regulates the cell cycle to promote the aggressiveness of bladder cancer. *Cell Death Dis* 2021; 12: 944.
- [51] Wu T, Sun Y, Sun Z, Li S, Wang W, Yu B and Wang G. Hsa_circ_0042823 accelerates cancer progression via miR-877-5p/FOXM1 axis in laryngeal squamous cell carcinoma. *Ann Med* 2021; 53: 960-970.
- [52] Kato A. Group 2 innate lymphoid cells in airway diseases. *Chest* 2019; 156: 141-149.
- [53] Davis BP and Rothenberg ME. Eosinophils and cancer. *Cancer Immunol Res* 2014; 2: 1-8.
- [54] Grisaru-Tal S, Itan M, Klion AD and Munitz A. A new dawn for eosinophils in the tumour micro-environment. *Nat Rev Cancer* 2020; 20: 594-607.
- [55] Bao X, Shi R, Zhao T and Wang Y. Mast cell-based molecular subtypes and signature associated with clinical outcome in early-stage lung adenocarcinoma. *Mol Oncol* 2020; 14: 917-932.
- [56] Dersh D, Holly J and Yewdell JW. A few good peptides: MHC class I-based cancer immunosurveillance and immunoevasion. *Nat Rev Immunol* 2021; 21: 116-128.
- [57] Sledzinska A, Vila de Mucha M, Bergerhoff K, Hotblack A, Demane DF, Ghorani E, Akarca AU, Marzolini MAV, Solomon I, Vargas FA, Pule M, Ono M, Seddon B, Kassiotis G, Ariyan CE, Korn T, Marafioti T, Lord GM, Stauss H, Jenner RG, Peggs KS and Quezada SA. Regulatory T cells restrain interleukin-2- and blimp-1-dependent acquisition of cytotoxic function by CD4(+) T cells. *Immunity* 2020; 52: 151-166.e6.
- [58] Mace A, Kutalik Z and Valsesia A. Copy number variation. *Methods Mol Biol* 2018; 1793: 231-258.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

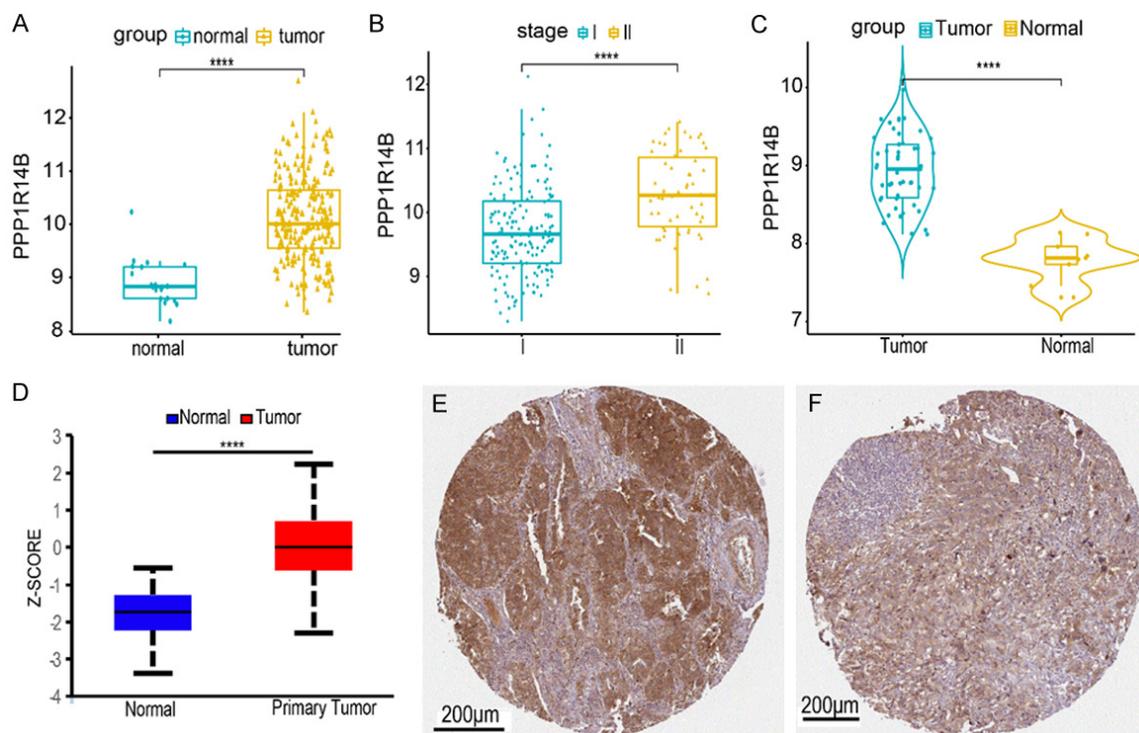


Figure S1. Validation of PPP1R14B expression. Comparison between tumor and normal tissue in the GSE31210 cohort (A). Comparison between clinical stage I and stage II in the GSE31210 cohort (B). Comparison between tumor and normal tissue in the E-MEXP-231 cohort (C). At the protein expression level, a comparison between tumor and normal tissue in the CPTAC dataset (D). (E, F) Immunohistochemical staining of PPP1R14B from the Human Protein Atlas database. ****, $P < 0.0001$.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

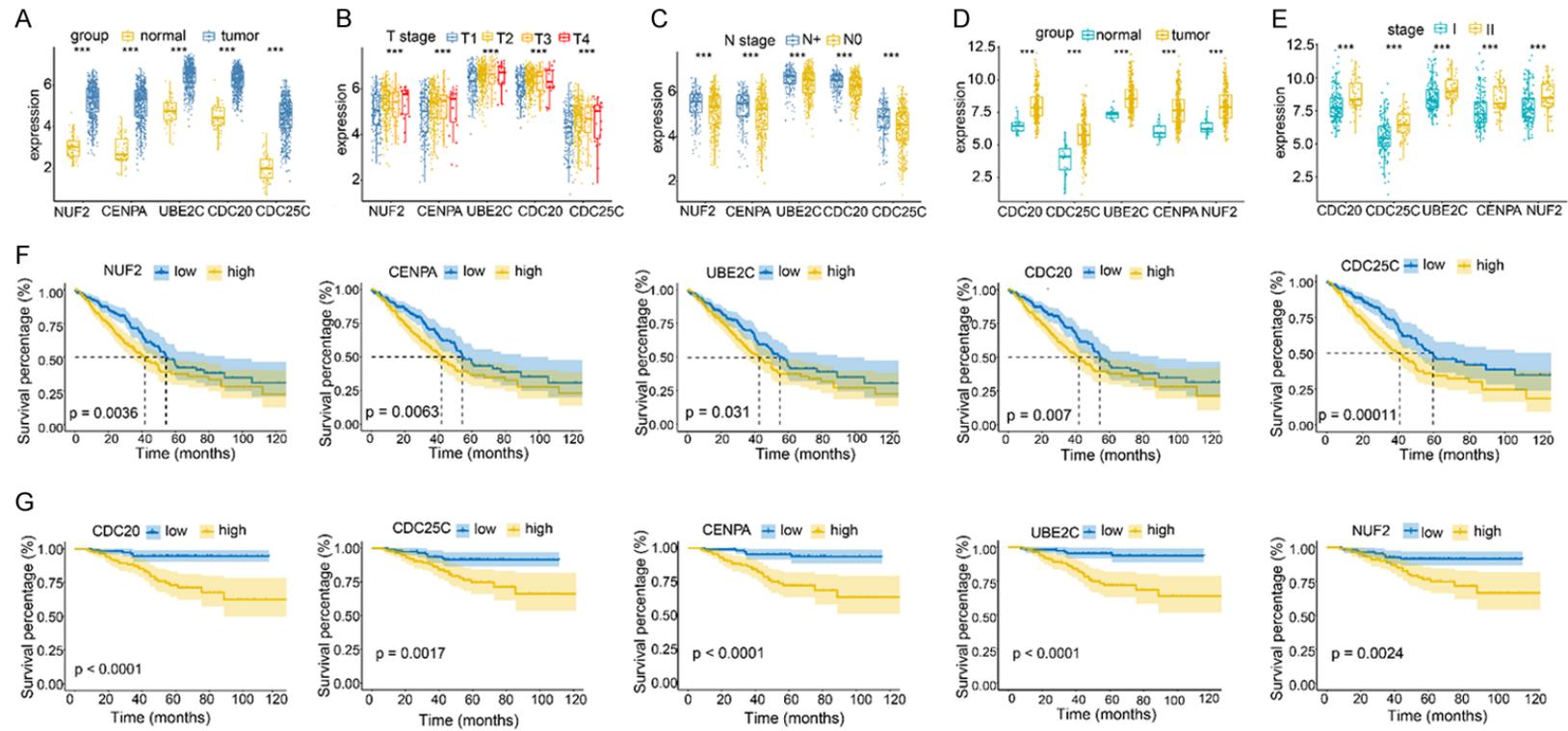


Figure S2. Hub gene differential expression analysis and survival analysis. A. Hub gene differential expression in the tumor and normal lung groups in the TCGA data set. B. Hub gene differential expression by clinical T stage in the TCGA dataset. C. Hub gene differential expression in the clinical N stage in the TCGA dataset. N+: N1, N2, N3 stage. D. PPP1R14B expression differences in the tumor and normal lung groups in the GSE31210 cohort. E. PPP1R14B expression differences by clinical stage in the GSE31210 cohort. F. KM survival analysis of hub genes in the TCGA data set. The comparison was between high and low expression. G. KM survival analysis of hub genes in the GSE31210 dataset. ***, $P < 0.00$.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

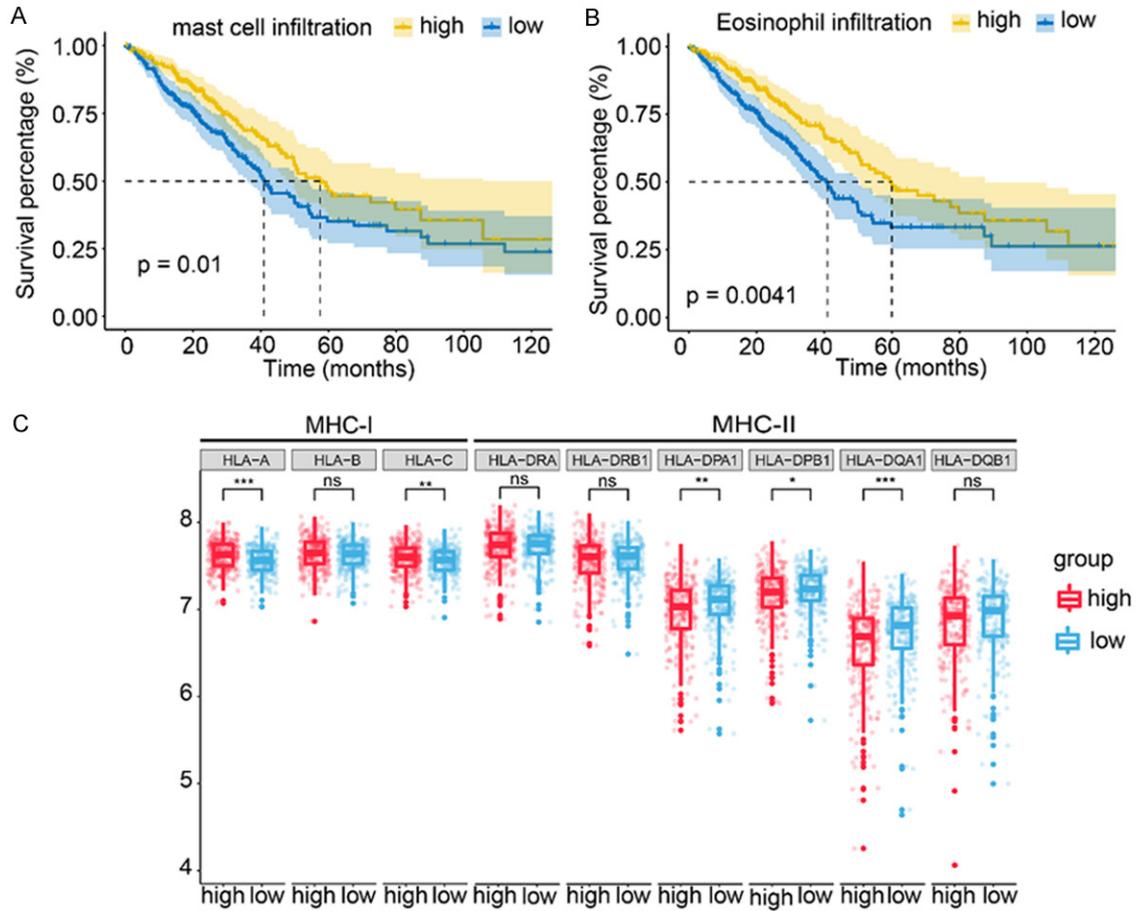
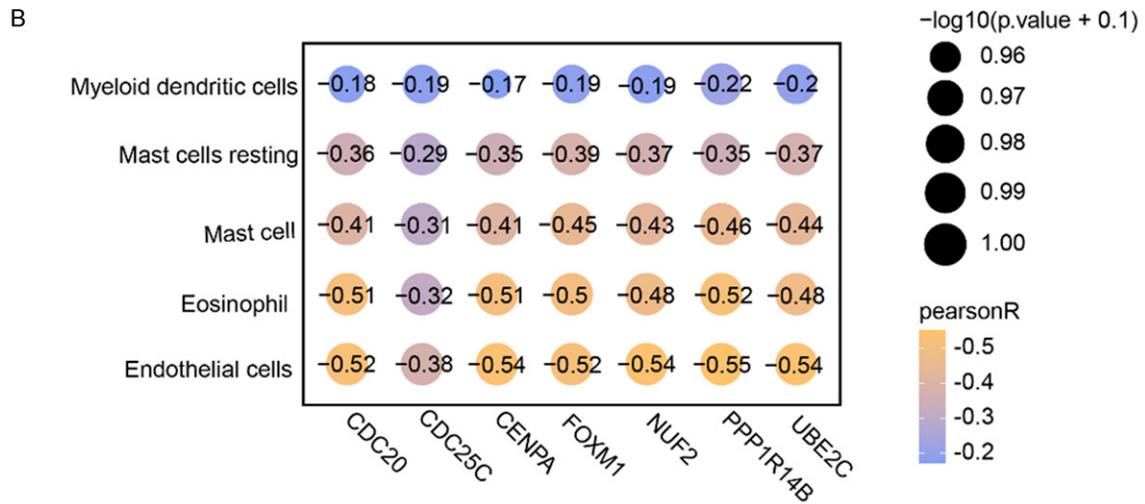
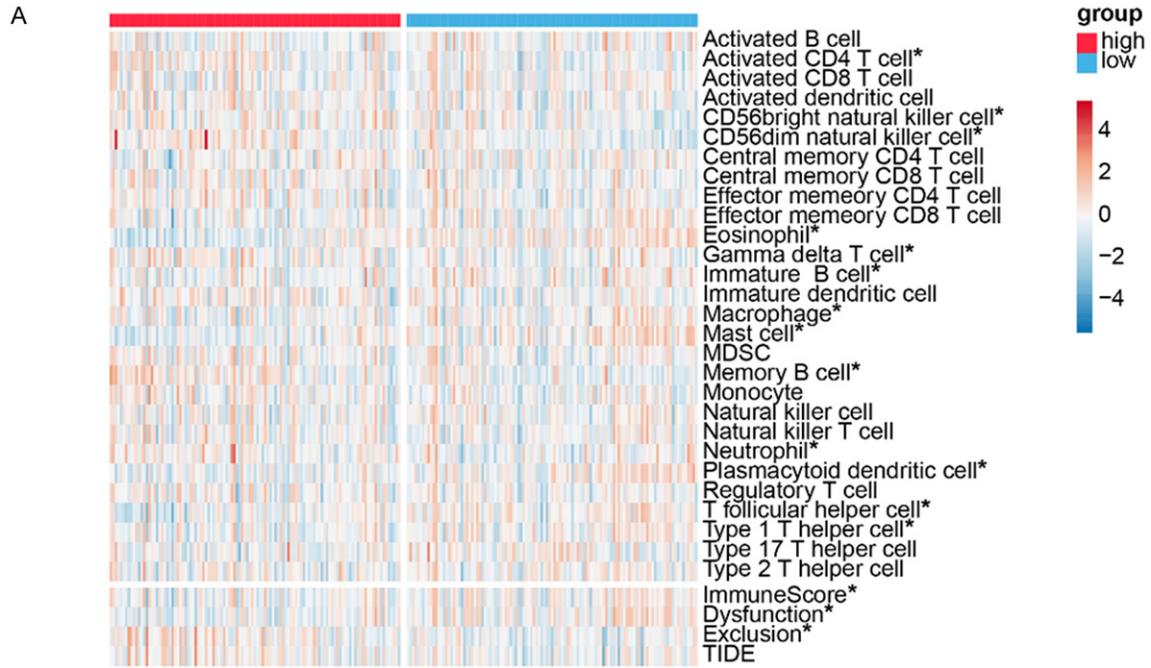


Figure S3. Survival-related analysis and major histocompatibility complex (MHC) molecule expression analysis. A. KM plot for mast cell infiltration. B. KM plot for eosinophil infiltration. C. Comparison between tumor and normal tissue of MHC molecule expression. Ns, P > 0.05; *, P < 0.05; **, P < 0.01; ***, P < 0.001.

Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis



Protein phosphatase 1 regulatory inhibitor subunit 14B in LUAD tumorigenesis

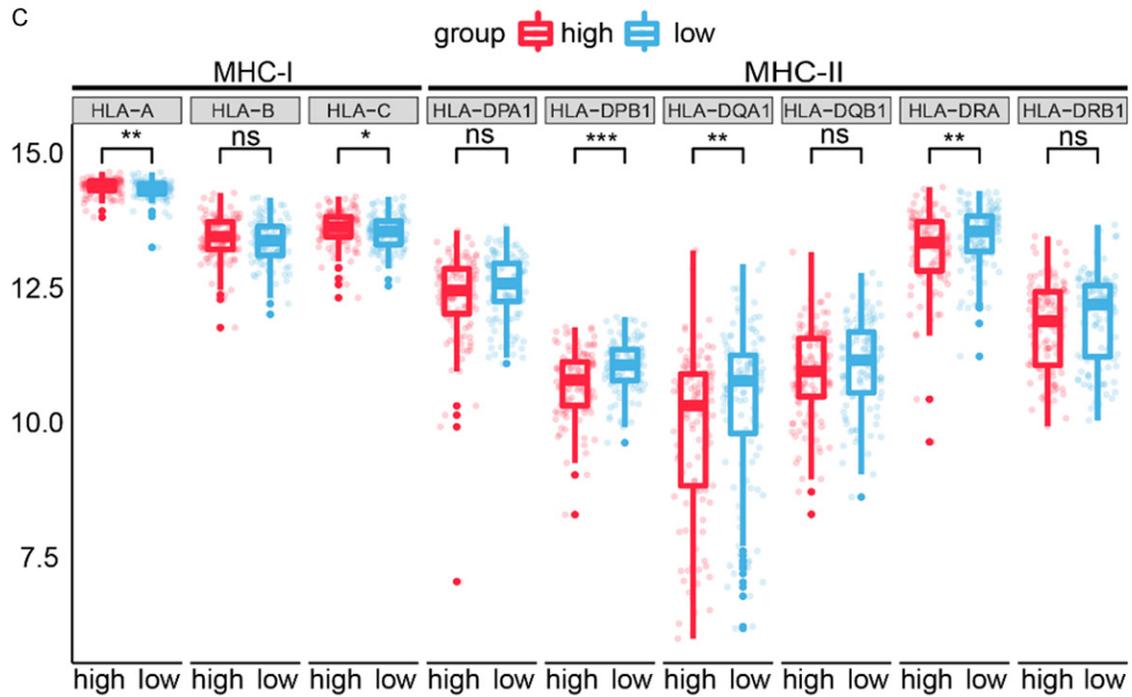


Figure S4. Immune analysis in the GSE31210 cohort for validation. A. Landscape of 28 immune cells, immune score, Dysfunction score, and Exclusion score, and their comparison between two groups in the GSE31210 cohort. * Indicates a statistically significant difference between the two groups. B. Correlation between PPP1R14B axial gene expression and immune cell infiltration in the GSE31210 cohort. C. Comparison between tumor and normal tissue of MHC molecule expression. ns, $P > 0.05$; *, $P < 0.05$; **, $P < 0.01$, ***, $P < 0.001$.