Original Article

Comprehensive bioinformatics analysis of prognosis and immunotherapy in lung adenocarcinoma

Ling Peng[1], Luping Xia[2,3], Meiyu Yang[1], Yali Wen[1], Qinghua Zeng[1]

[1]Department of Respiratory and Critical Care Medicine, Jiangxi Provincial Key Laboratory of Respiratory Diseases, Jiangxi Institute of Respiratory Diseases, The First Affiliated Hospital, Jiangxi Medical College, Nanchang University, Nanchang, China; [2]China-Japan Friendship Jiangxi Hospital, National Regional Center for Respiratory Medicine, Nanchang, China; [3]Department of Respiratory and Critical Care Medicine, Shangrao People's Hospital, Shangrao, China

Contributions: (I) Conception and design: Q Zeng; (II) Administrative support: Q Zeng; (III) Provision of study materials or patients: L Peng; (IV) Collection and assembly of data: L Xia; (V) Data analysis and interpretation: M Yang, Y Wen; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Qinghua Zeng, MD, PhD. Department of Respiratory and Critical Care Medicine, Jiangxi Provincial Key Laboratory of Respiratory Diseases, Jiangxi Institute of Respiratory Diseases, The First Affiliated Hospital, Jiangxi Medical College, Nanchang University, 17 Yongwai Street, Donghu District, Nanchang 330038, China. Email: zqh196945@163.com.

Background: Research has shown that genetic mutations play an important role in the prognosis of lung adenocarcinoma (LUAD). However, the genes that influence the prognosis and immunotherapy of lung cancer patients have not yet been thoroughly studied. In this study, data from The Cancer Genome Atlas (TCGA) Program and other databases were used to identify the survival-related genes in LUAD.

Methods: First, the TCGA database was used to screen key LUAD genes. Second, the Gene Expression Profiling Interactive Analysis 2 (GEPIA2), University of ALabama at Birmingham CANcer (UALCAN), Tumor IMmune Estimation Resource (TIMER), Kaplan-Meier plotter, and cBioPortal databases, and a univariate Cox analysis combined with a random forest (RF) model were used to estimate gene expression, patient prognosis, and gene mutations, respectively. TIMER was also used to predict the immune function of the genes.

Results: A total of 2,138 up-regulated and 2,559 down-regulated differentially expressed genes (DEGs) were identified from TCGA-LUAD dataset. Next, four prognostic genes (i.e., *CENPH*, *SLC35F4*, *TESMIN*, and *TERT*) were identified as the key genes. The expression levels of all four genes were higher in LUAD tissues than those in the normal lung tissues, but only *CENPH* and *TESMIN* were correlated with poor overall survival (OS). The four genes were also found to be associated with immunoinfiltration.

Conclusions: Of the four key genes identified, *CENPH* and *TESMIN* would not only contribute to the diagnosis and prognosis of LUAD but could also serve as potential immunotherapy targets for LUAD.

Keywords: Lung adenocarcinoma (LUAD); The Cancer Genome Atlas (TCGA); survival analysis; immunotherapy

Submitted Sep 14, 2024. Accepted for publication Dec 12, 2024. Published online Dec 28, 2024.
doi: 10.21037/jtd-24-1530
View this article at: https://dx.doi.org/10.21037/jtd-24-1530

Introduction

Lung cancer is the number one cause of cancer-related death and the second most common cancer worldwide (1,2). Non-small cell lung cancer (NSCLC) is the most common type of lung cancer, accounting for about 85% of lung cancers, while lung adenocarcinoma (LUAD) is currently the most common histological subtype of NSCLC (3-5). Patients diagnosed with advanced disease have a poor prognosis. The occurrence, development,

and prognosis of tumors are not only associated with pathological classification and clinical stage, but are also closely correlated with abnormal genes (6). Continuous improvements in early prevention and the use of targeted drugs have transformed lung cancer treatments in clinical settings, prolonged the progression-free survival and overall survival (OS) of patients, and improved the prognosis of patients (7). However, the prognosis for advanced lung cancer patients remains poor. Thus, biomarkers urgently need to be identified to aid in the diagnosis, treatment, and prognosis prediction of lung cancer patients.

Messenger RNA (mRNA) is a type of single-stranded ribonucleic acid that is transcribed from a strand of DNA as a template and carries genetic information to guide protein synthesis. It connects the genetic message in DNA to the translation and expression of proteins, and plays an important role in life activities (8). In cancer treatment, mRNA can be used as a biomarker and a target in cancer therapy, and research has shown that the mRNA responsible for encoding tumor antibodies may also stimulate effective anti-tumor immunity (9). Other signaling pathways have also been shown to regulate the mRNA of many proteins (10), and many tumor cell derived exosomes also contain related mRNA that promotes cell proliferation. When mRNA is transported to receptor cells, it can produce functional proteins, promote the horizontal transmission of genetic information, and play a key role in gene expression. Presently, mRNA plays a major role in the occurrence, development, diagnosis, and treatment of tumors. Recently, mRNA tumor vaccines have gradually been applied to the treatment of cancer (11). Additionally, mRNA therapy, as an emerging gene therapy, has also begun to play a role in the treatment of various types of tumors.

In this study, LUAD gene expression profile data were downloaded from The Cancer Genome Atlas (TCGA) Program, a public database, and a univariate Cox analysis combined with a random forest (RF) model was used to identify the genes associated with prognosis in LUAD. An R bioinformation tool and online software were then used to perform functional analysis and prediction analysis of these genes. Finally, we successfully identified four genes critical to the prognosis of LUAD, and found that their abnormal expression was associated with poor prognosis in LUAD. We present this article in accordance with the STREGA reporting checklist (available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-1530/rc).

Highlight box

Key findings

- We identified genes associated with lung adenocarcinoma (LUAD) prognosis and immunotherapy.

What is known, and what is new?

- Messenger RNA (mRNA) can be used as a biomarker and target for cancer therapy, and the mRNA responsible for encoding tumor antibodies may also stimulate effective anti-tumor immunity.
- LUAD genes expression profile data were downloaded from public databases, and four genes critical to the prognosis of LUAD were successfully identified. The abnormal expression of these genies is associated with a poor prognosis in LUAD.

What is the implication, and what should change now?

- Through bioinformatics analysis, we have discovered molecular markers that can serve as prognostic markers for the diagnosis and treatment of lung cancer, adding some assistance to overcome the difficulties in lung cancer diagnosis and treatment.

Methods

RNA-sequencing transcriptome data of samples

The RNA-sequencing expression profiles of LUAD were downloaded from TCGA (https://portal.gdc.cancer.gov/, up to March 29, 2023), a publicly available, open-access database. The dataset comprised 600 samples, of which 541 LUAD tumor tissue samples and 59 normal lung tissues samples were included. The clinical data of LUAD patients were also obtained from TCGA, and 484 patients with complete survival data were included in our survival analysis. The transcripts per kilobase million and count values were acquired from TCGA-LUAD database. The data workflow type was STAR-Count files. Only the protein coding genes were selected for further study. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Identification of DEGs

The LUAD tissues were compared to the normal tissue samples using R package limma (version 3.56.2), and the differentially expressed genes (DEGs) were identified. The DEGs with a P value <0.05 and $|\log_2$ fold change (FC)$|$ >1 were considered statistically significant. Finally, the DEGs were visualized in volcano plots, heatmaps, and principal component analysis (PCA) plots using R package ggplot2 (version 3.3.0) and tinyarray (version 2.3.0).

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses

To determine the function of the DEGs in LUAD, GO and KEGG pathway analyses were conducted using R package clusterProfiler and DOSE. The GO analysis was conducted to examine the biological processes (BPs), cellular components (CCs), and molecular functions (MFs) of the DEGs. A P value <0.01 was considered statistically significant.

Identification of the diagnostic prognosis-related genes

The randomForestSRC and randomSurvivalForest package in R (version 4.3.0) were used to construct a RF model and select the biomarkers that contribute to prognosis in LUAD. A P value <0.01 was set as the cut-off value for the univariable Cox analysis, and a P value <0.3 was set as the cut-off value for the RF analysis. The prognosis-related genes were intersected with the DEGs to select the hub prognosis-related DEGs.

Survival analysis of DEGs

Kaplan-Meier plotter (http://kmplot.com/analysis/index.php?p=background) online software was used to predict the survival of the LUAD patients. To verify the results, the "survival" and "survminer" packages in R (version 4.3.0) were used to visualize the Kaplan-Meier survival analysis of the data of the 515 LUAD patients from TCGA database, and the patients were divided into two groups according to the median value of the risk-score model for the survival analysis.

Expression analysis of the hub prognosis-related DEGs

The expression of the hub DEGs was verified using the online website University of ALabama at Birmingham CANcer (UALCAN; https://ualcan.path.uab.edu/), Tumor IMmune Estimation Resource (TIMER; https://cistrome.shinyapps.io/timer/), and Gene Expression Profiling Interactive Analysis 2 (GEPIA2; http://gepia2.cancer-pku.cn/#index) databases. A P value <0.05 was set as the threshold.

Genomic alteration analysis of the hub prognosis-related DEGs

The cBio Cancer Genomics Portal (12) (c-BioPortal; https://www.cbioportal.org/) database was used to conduct the genomic alteration analysis of the LUAD hub prognosis-related DEGs.

Immune cell infiltration analysis of the hub prognosis-related DEGs

The TIMER (13) (https://cistrome.shinyapps.io/timer/) database was used to explore immune cell infiltration in LUAD.

Statistical analysis

The differences between the groups were examined using the Student's *t*-test. The data were presented as the mean ± standard deviation (SD). The statistical analysis was performed using GraphPad Prism (version 9.0). A P value <0.05 (two-tailed) was considered statistically significant.

Results

Identification of DEGs in TCGA-LUAD

TCGA-LUAD data were analyzed using R package limma, tinyarray, and clusterProfiler. We conducted a PCA plot analysis and found that the tumor and normal samples were well distributed in the two clusters of the gene signature (*Figure 1A*). We analyzed the expression of 541 LUAD samples and 59 normal samples from the TCGA database, and we found that 2,138 genes were up-regulated and 2,559 genes were down-regulated (P<0.05, log|FC| >1). The volcano plots and heatmap of these genes are shown in *Figure 1B,1C*.

Enrichment analysis of DEGs

The DEG data of TCGA-LUAD patients were subjected to GO and KEGG enrichment analyses using R package clusterProfiler. The results of the GO analysis showed that the top three BPs were extracellular structure organization, external encapsulating structure organization, and axoneme assembly (*Figure 2A*), the top three CCs were the collagen-containing extracellular matrix, ion channel complex, and apical part of cell (*Figure 2B*), and the top three MFs were passive transmembrane transporter activity, channel activity, and gated channel activity (*Figure 2C*). Additionally, the KEGG pathway analysis showed that the top three significantly enriched pathways were the neuroactive

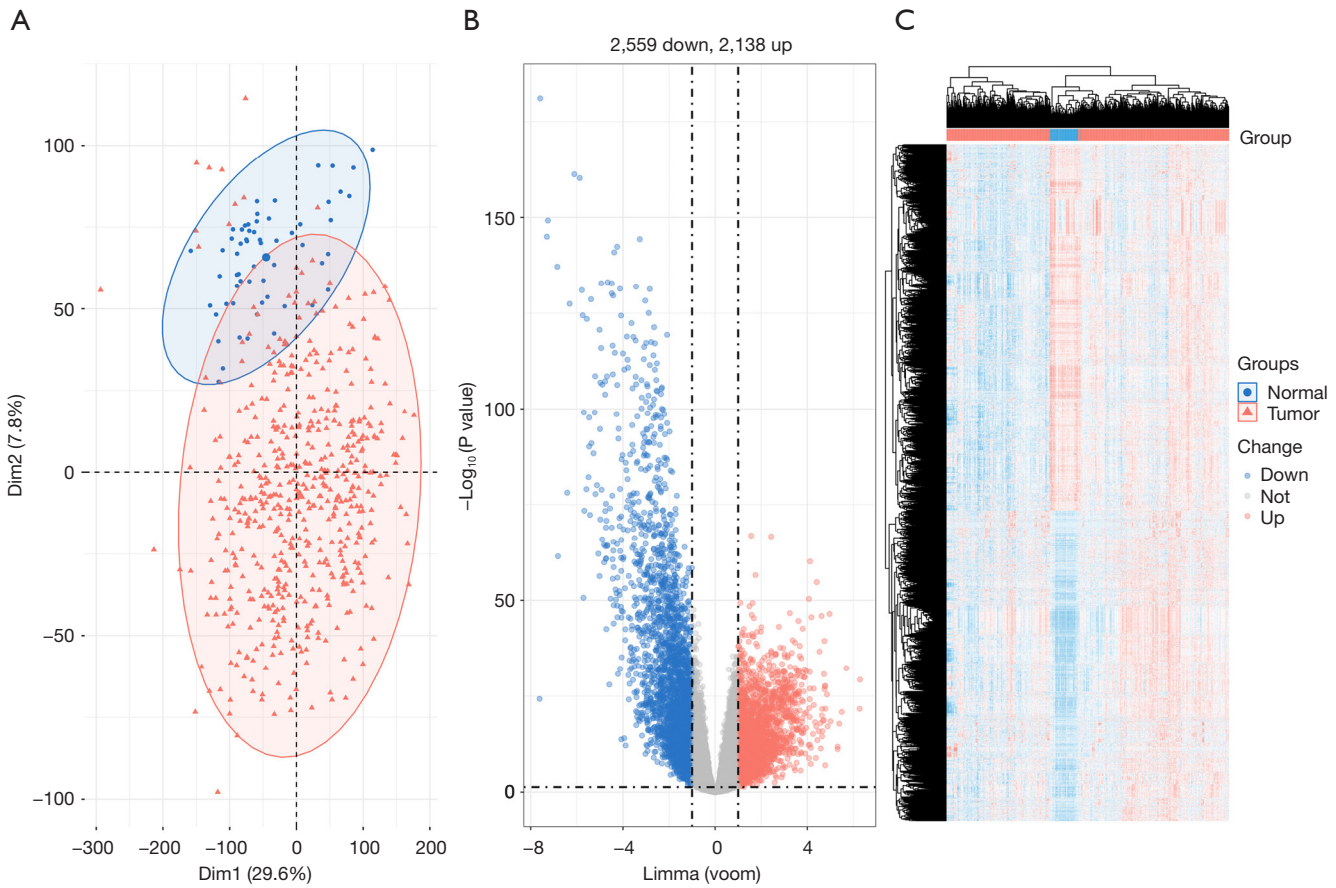


Figure 1 DEGs in TCGA-LUAD. (A) PCA plots of LUAD and normal samples. Red means up, blue means down. (B) Volcano plots of DEGs in TCGA-LUAD. (C) Heatmaps of TCGA-LUAD. DEGs, differentially expressed genes; TCGA, The Cancer Genome Atlas; LUAD, lung adenocarcinoma; PCA, principal component analysis.

ligand-receptor interaction pathway, the systemic lupus erythematosus pathway, and the neutrophil extracellular trap formation pathway (*Figure 2D*).

Prognosis-related DEGs

Based on our findings, we suspected that the identified DEGs were associated with the development of LUAD. Next, to evaluate whether the DEGs could serve as diagnostic markers for LUAD, we conducted a univariate Cox analysis with a RF model of the preprocessed data. The results of the univariate analysis (*Figure 3A-3C*) showed that the four key genes, including telomerase reverse transcriptase (*TERT*), centromere protein H (*CENPH*), testis expressed metallothionein-like protein (*TESMIN*), which is also known as metallothionein-like 5 protein (*MTL5*), and solute carrier family 35 member F4 (*SLC35F4*), were significantly associated with the prognosis of LUAD patients. These four genes (i.e., *CENPH*, *TERT*, *SLC35F4*, and *TESMIN*) were up-regulated in the LUAD samples (*Table 1*).

The four prognosis-related DEGs were significantly up-regulated in LUAD tissues

The GEPIA2 database showed that the mRNA levels of the four genes were significantly higher in the LUAD tissues than those of the corresponding normal tissues (*Figure 4A*). The TIMER database results showed that the expression of the four genes was significantly higher in the LUAD tissues than that of the adjacent normal tissues (*Figure 4B*). The expression of the four genes in the LUAD and normal tissues was also examined using the UALCAN database, and we found that the expression levels of the four genes were

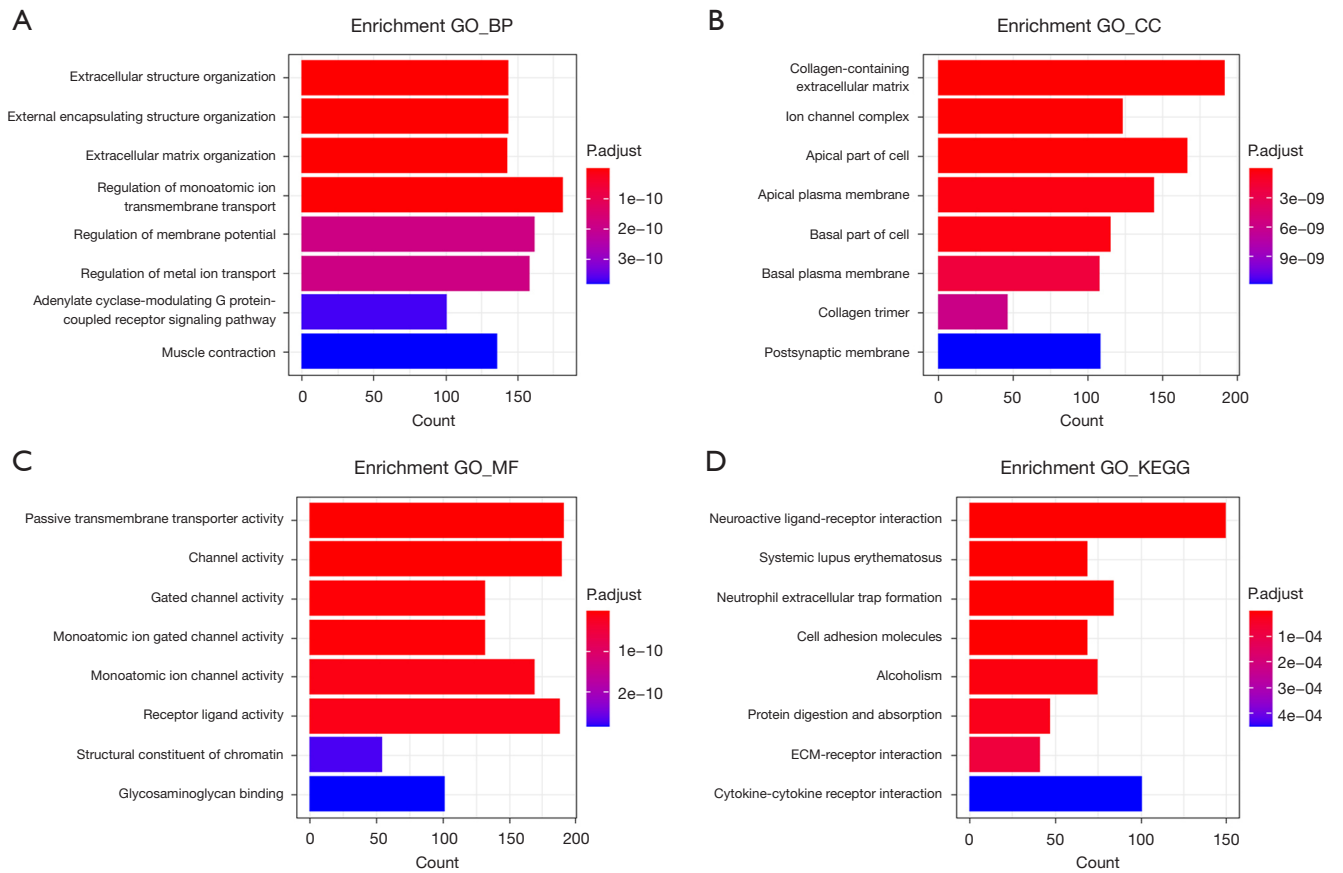


Figure 2 Enrichment analysis of the DEGs in the LUAD samples. (A-C) GO enrichment analysis. (D) KEGG pathway enrichment analysis histograms. GO, Gene Ontology; BP, biological process; CC, cellular component; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genomes; ECM, extracellular matrix; DEGs, differentially expressed genes; LUAD, lung adenocarcinoma.

higher in the LUAD tissues than those in the normal tissues (*Figure 4C*). These results suggested that the four genes were overexpressed in the cancer tissues, which is consistent with the TCGA-LUAD results.

The four prognosis-related DEGs were correlated with poor OS in LUAD

The four key genes were significantly associated with the prognosis of LUAD patients (*Figure 5*A). We then further examined the prognostic significance of the four DEGs in LUAD patients. The results of the Kaplan-Meier OS analysis showed that the high expression of all four DEGs (i.e., *CENPH*, *SLC35F4*, *TERT*, and *TESMIN*) was associated with poor OS in LUAD patients (*Figure 5B-5E*); however, the difference was only statistically significant for *CENPH* and *TESMIN*, which exerted persecute effects for patients with LUAD (*Figure 5B,5E*, P<0.001). The results of TCGA-LUAD data survival analysis, showed that the high expression of four DEGs (i.e., *CENPH*, *SLC35F4*, *TERT*, and *TESMIN*) was associated with worse OS in the LUAD patients (*Figure 5F-5I*); however, *CENPH* and *TESMIN* exerted persecute effects for patients with LUAD (*Figure 5F,5I*, P<0.05). The data for *SLC35F4* could not be analyzed, as there were only two samples in the high expression group (*Figure 5G*). TCGA-LUAD survival results were consistent with the Kaplan-Meier survival plot results. However, *CENPH*, *TERT* and *TESMIN* may have more research value in LUAD treatment.

Genomic alteration analysis of the four key prognosis-related DEGs in LUAD

A genomic alteration analysis of the four key prognosis-

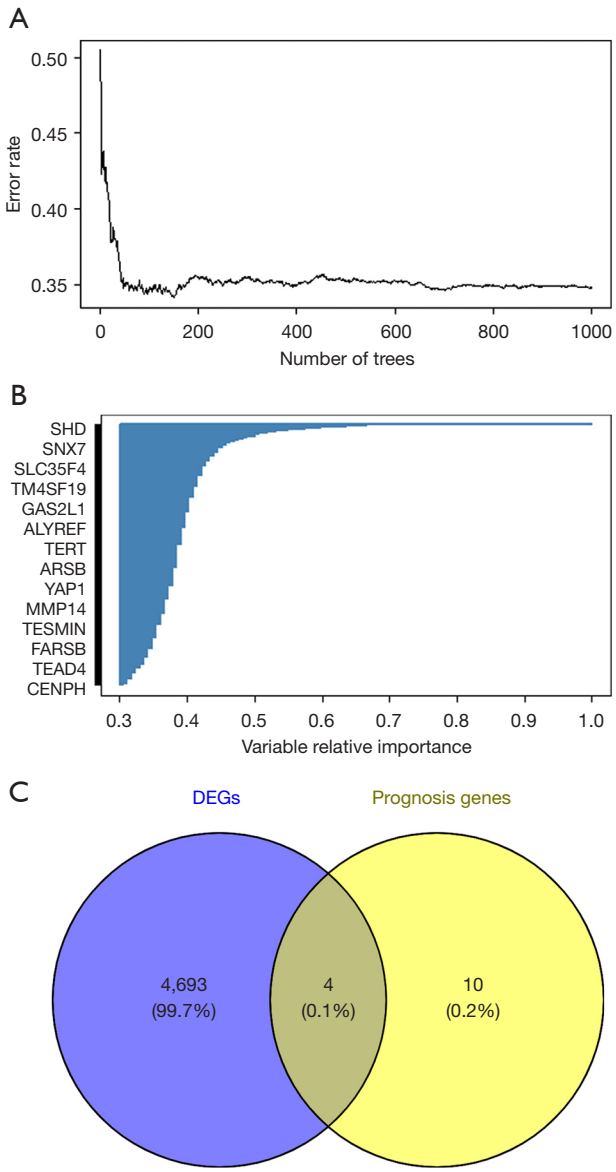


Figure 3 The prognosis-related DEGs in LUAD. (A) Error graph of the RF models (B) The prognosis-related genes in LUAD based on the univariate Cox analysis. (C) Venn diagram of the DEGs and prognosis genes. DEGs, differentially expressed genes; LUAD, lung adenocarcinoma; RF, random forest.

related DEGs in LUAD was conducted using the cBioPortal website. The results indicated that genomic alterations of *CENPH*, *SLC35F4*, *TERT*, and *TESMIN* occurred in 2%, 1.8%, 13%, and 2.5% of LUAD patients, respectively (*Figure 6A*). The genetic alteration type and frequency of the four key prognostic DEGs showed differences in various LUAD databases (*Figure 6B*). Together, these

Table 1 The expression of the prognosis-related DEGs in LUAD

Symbol	LogFC	P value	Adj. P value	Change
TERT	4.41	1.80E−55	8.40E−54	Up
CENPH	1.35	1.39E−17	8.89E−17	Up
TESMIN (*MTL5*)	2.08	1.22E−12	5.21E−12	Up
SLC35F4	1.02	3.14E−06	7.28E−06	Up

DEGs, differentially expressed genes; LUAD, lung adenocarcinoma; FC, fold change.

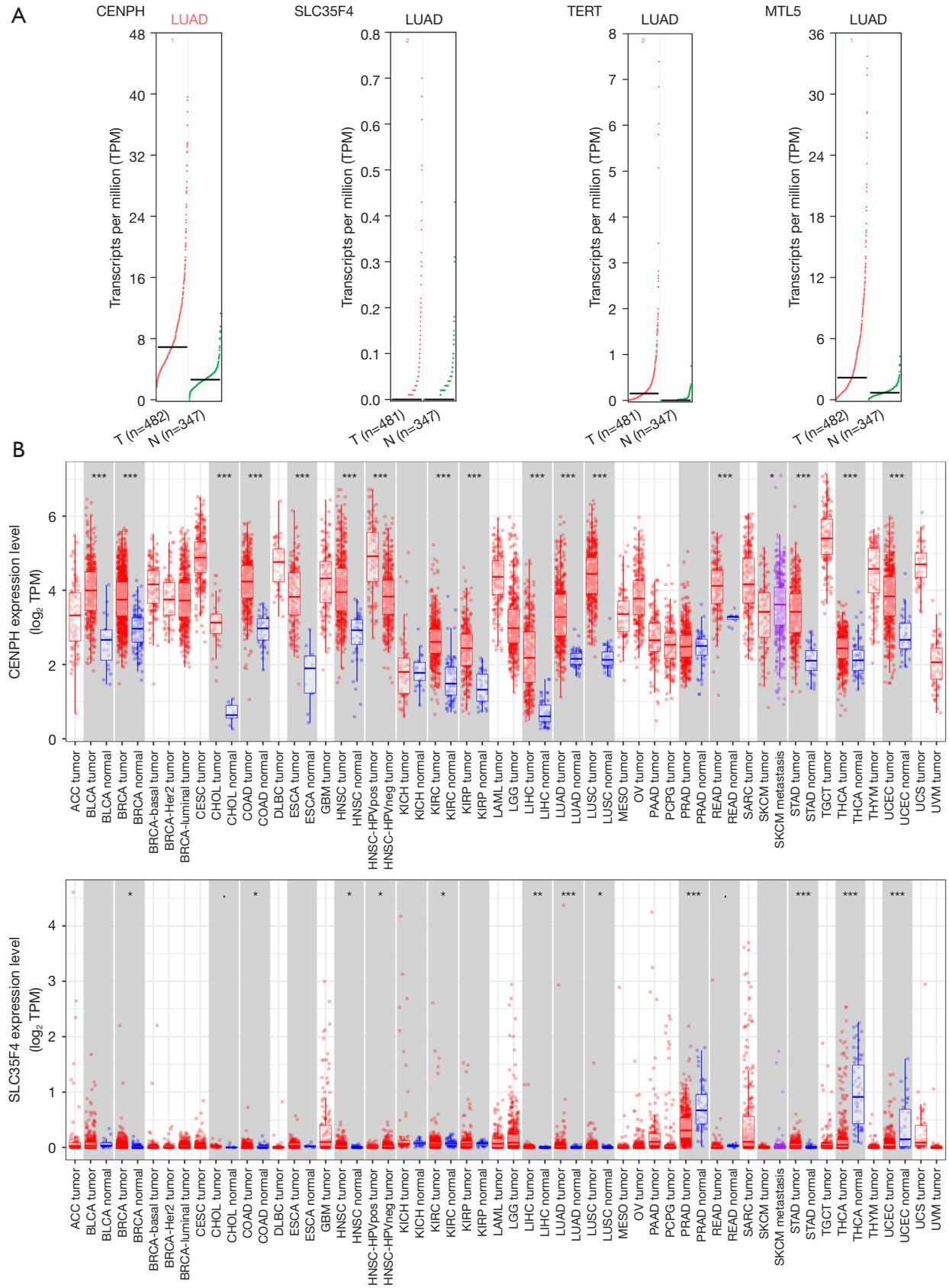
results showed that the genomic alterations of the four key prognosis-related DEGs might play an essential role in cancer onset and progression.

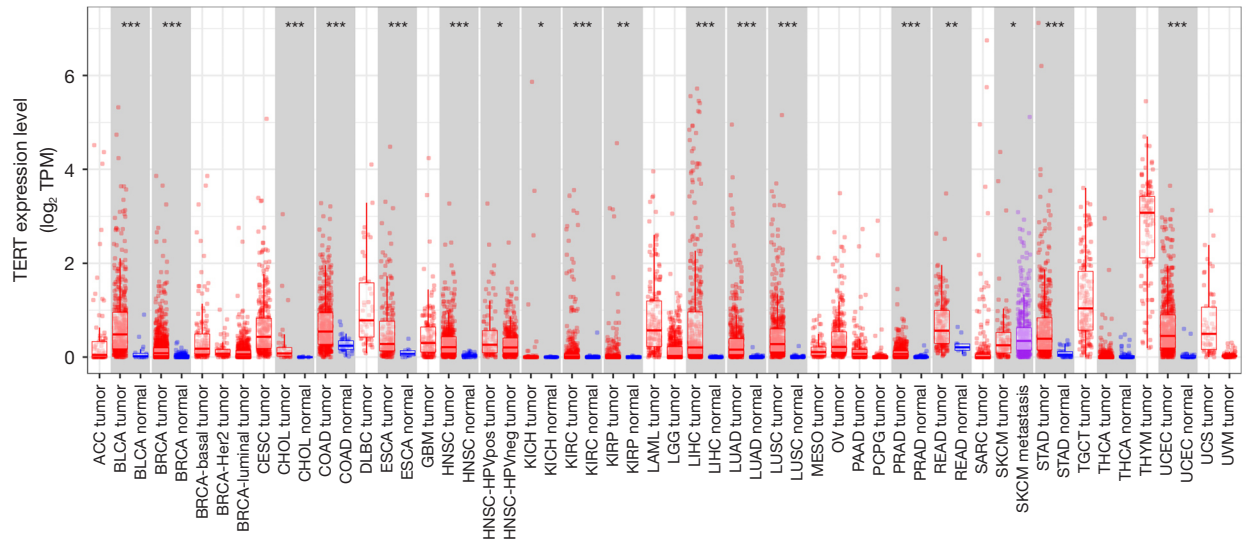
Immune cell infiltration analysis

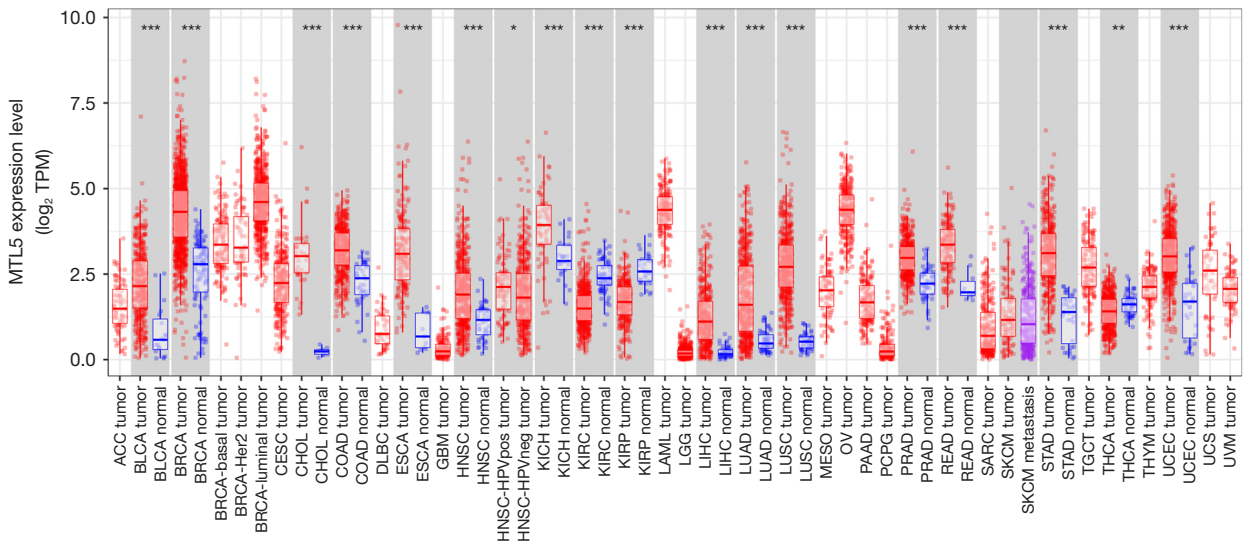
To further explore whether somatic copy number alteration (SCNA) and the degree of the cancer's immune invasion were associated with prognosis, the TIMER database was used to assess the correlation between the expression of the prognosis-related genes and immune infiltration. A negative correlation was found between *CENPH* expression and the infiltration of B cells (Cor =–0.123, P=0.007), and macrophage cells (Cor =–0.154, P<0.001; *Figure 7A*). While a positive correlation was found between *SLC35F4* expression and the infiltration of $CD4^+$ T cells (Cor =0.115, P=0.01; *Figure 7B*). A negative correlation was also found between *TERT* expression and the infiltration of $CD8^+$ T cells (Cor =–0.118, P=0.009), macrophage cells (Cor =–0.196, P<0.001), neutrophil cells (Cor =–0.112, P=0.01), and dendritic cells (Cor =–0.174, P<0.001; *Figure 7C*). Additionally, a negative correlation was found between *TESMIN (MTL5)* expression and the infiltration of B cells (Cor =–0.173, P<0.001; *Figure 7D*). These results showed that the four genes in the risk-score formula were closely associated with the immune infiltration process in LUAD, especially *CENPH*, *TERT*, and *TESMIN*, which may be the reason why their expression affected prognosis.

Discussion

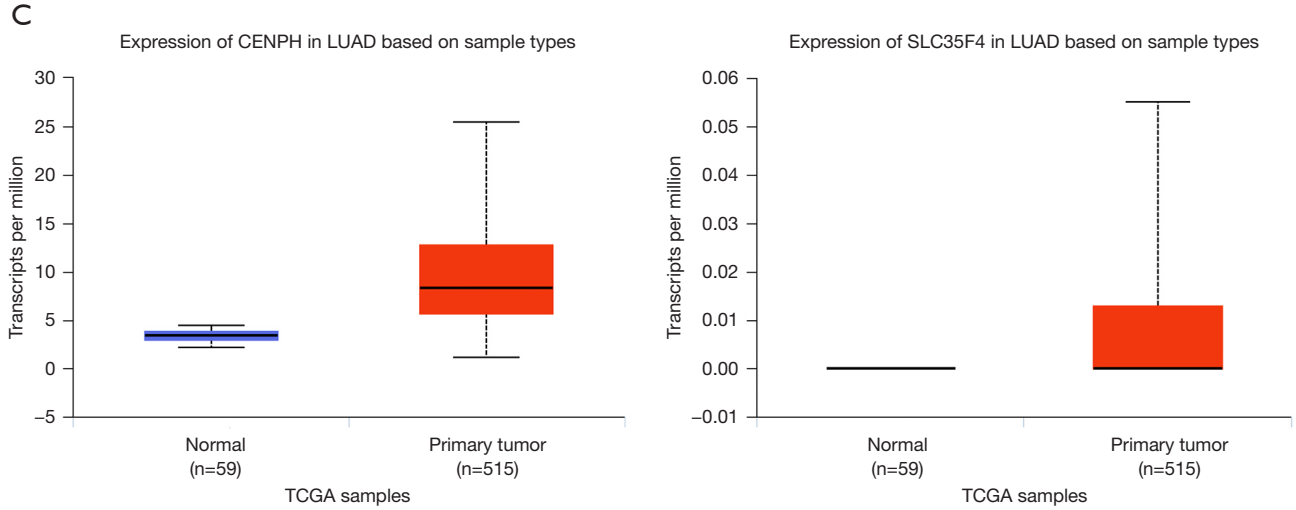
Many related studies have been conducted on LUAD, and a number of targeted immune drugs have been developed (14); however, many patients become resistant to drugs (15), and the mortality rate of LUAD remains very high. Therefore, research needs to be conducted to improve the therapeutic effects of treatments and the prognosis of


A
CENPH
LUAD
SLC35F4
LUAD
TERT
LUAD
MTL5
LUAD
Transcripts per million (TPM)
T (n=482)
N (n=347)
T (n=481)
N (n=347)
T (n=481)
N (n=347)
T (n=482)
N (n=347)
B
CENPH expression level (log2 TPM)
SLC35F4 expression level (log2 TPM)
ACC tumor
BLCA tumor
BLCA normal
BRCA tumor
BRCA normal
BRCA-basal tumor
BRCA-Her2 tumor
BRCA-luminal tumor
CESC tumor
CHOL tumor
CHOL normal
COAD tumor
COAD normal
DLBC tumor
ESCA tumor
ESCA normal
GBM tumor
HNSC tumor
HNSC normal
HNSC-HPVpos tumor
HNSC-HPVneg tumor
KICH tumor
KICH normal
KIRC tumor
KIRC normal
KIRP tumor
KIRP normal
LAML tumor
LGG tumor
LIHC tumor
LIHC normal
LUAD tumor
LUAD normal
LUSC tumor
LUSC normal
MESO tumor
OV tumor
PAAD tumor
PCPG tumor
PRAD tumor
PRAD normal
READ tumor
READ normal
SARC tumor
SKCM tumor
SKCM metastasis
STAD tumor
STAD normal
TGCT tumor
THCA tumor
THCA normal
THYM tumor
UCEC tumor
UCEC normal
UCS tumor
UVM tumor





C



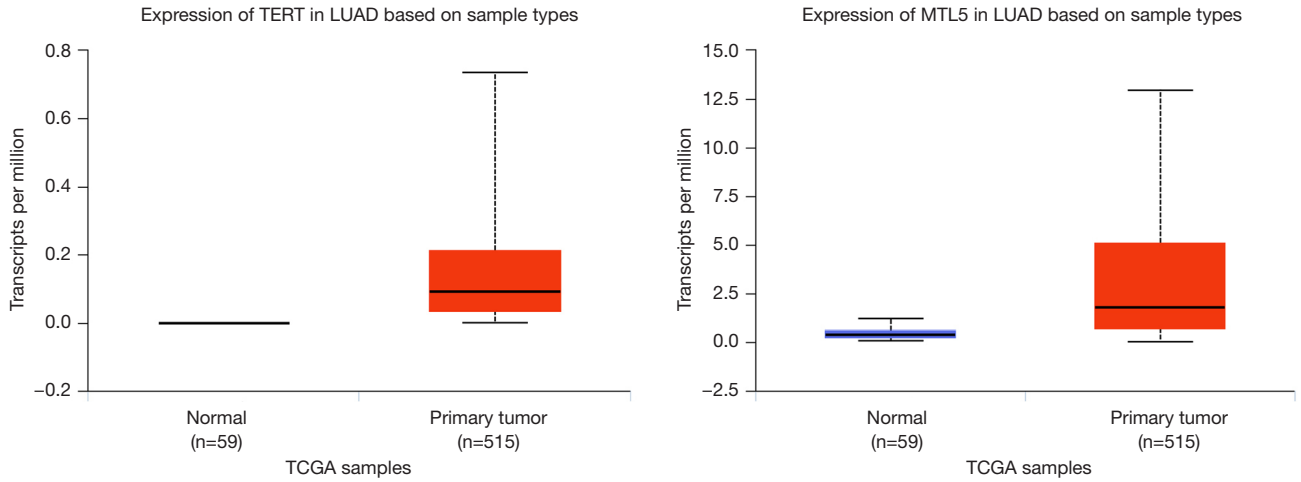


Figure 4 The expression levels of the four key prognosis-related genes in the LUAD tissues. (A) The expression levels of the four key prognosis-related genes in the LUAD and normal tissues in the GEPIA database. (B) The expression levels of the four key prognosis-related genes in the LUAD and normal tissues in the TIMER database. (C) The expression levels of the four key prognosis-related genes in the LUAD and normal tissues in the UALCAN database. *, P<0.05; **, P<0.01; ***, P<0.001. LUAD, lung adenocarcinoma; TPM, transcripts per million; TCGA, The Cancer Genome Atlas; GEPIA, Gene Expression Profiling Interactive Analysis; TIMER, Tumor IMmune Estimation Resource; UALCAN, University of ALabama at Birmingham CANcer.

LUAD patients. To improve the diagnosis and prognosis of LUAD patients, the molecular mechanism of its occurrence and development, needs to be better understood. With the development of high-throughput sequencing technology, we are able to detect novel genetic mutations that contribute to disease progression, and thus identify new genetic targets for the diagnosis, treatment, and prognosis predictions of LUAD.

In the present study, we used TCGA-LUAD data and combined the enrichment, expression, mutation, and immune infiltration analysis methods to examine the data. A univariate Cox analysis with a RF model was used to construct a network based on the correlation between the DEGs and prognostic genes. We identified four key genes related to the diagnosis and prognosis of LUAD. These four genes were highly expressed in LUAD, and these LUAD patients had worse OS. To further verify our results, we performed univariate Cox regression and RF model analyses to examine the relationship between the four genes and OS in TCGA-LUAD data. The results showed that *CENPH* and *TESMIN* exerted a protective effect in patients with LUAD, which suggested that high expression of *CENPH* and *TESMIN* primarily plays a protective role in LUAD.

Previous studies (16-23) have also reported some of the functions of these key genes. *CENPH* plays an important role in the organization and function of the active centromere-kinetochore complex, and is related to the prognosis and progression of gastric carcinoma, colorectal cancer, renal cell carcinoma, endometrial cancer, LUAD, and other cancers. Liao *et al.* reported that the overexpression of *CENPH* was significantly associated with breast cancer progression and patient OS (24). Li *et al.* and Liao *et al.* also found that *CENPH* was more highly expressed in LUAD tissues than in normal tissues and indicated a poor prognosis. Our analysis results are consistent with those of previous studies (20,25).

Telomerase (*TERT*) is a ribonucleoprotein polymerase that plays a role in cellular senescence, and is normally repressed in postnatal somatic cells, resulting in the progressive shortening of telomeres. The expression of *TERT* mRNA is positively correlated with telomerase activity, which suggests that the expression of *TERT* plays a crucial role in the process of carcinogenesis (26,27). In particular, the GABPA/B tetramer selectively binds and activates the mutant TERT promoter, thus inducing overexpression of the *TERT* gene in several tumor types, including glioblastoma, melanoma, hepatocellular carcinoma and bladder cancer (28). In addition, the knockdown of *TERT* expression can prevent the proliferation and migration of cervical cancer cells, colon cancer cells, oral squamous cell cancer cells, and osteosarcoma cells (29-32). In the present study, we found

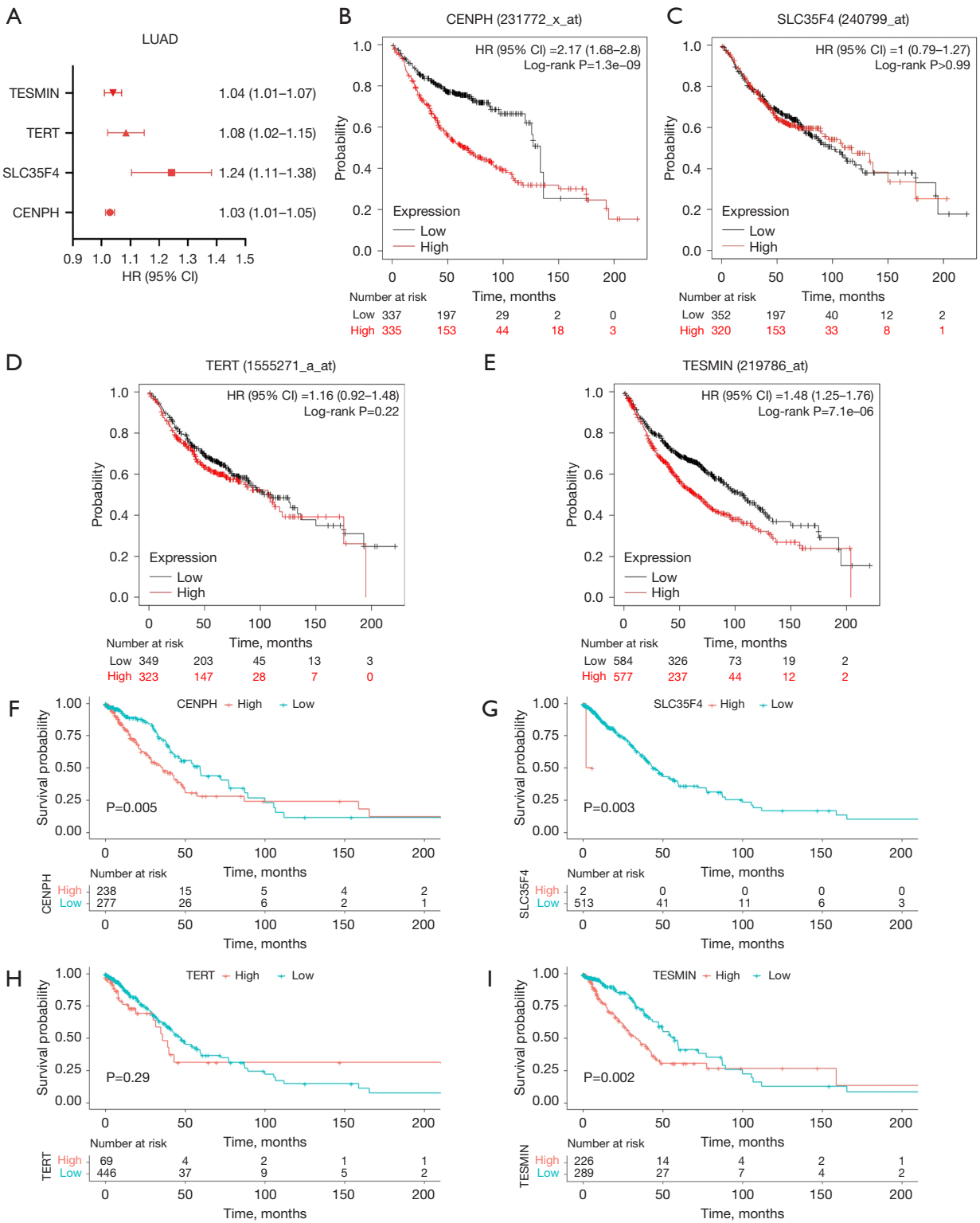


Figure 5 Survival analysis results of the four key prognostic DEGs (A) Prognostic HRs of the four key genes in LUAD. (B-E) Survival analysis of the four key prognostic DEGs by Kaplan-Meier plotter. (F-I) Survival analysis of the four key prognostic DEGs using TCGA-LUAD data. HR, hazard ratio; CI, confidence interval. LUAD, lung adenocarcinoma; DEGs, differentially expressed genes; TCGA, The Cancer Genome Atlas.

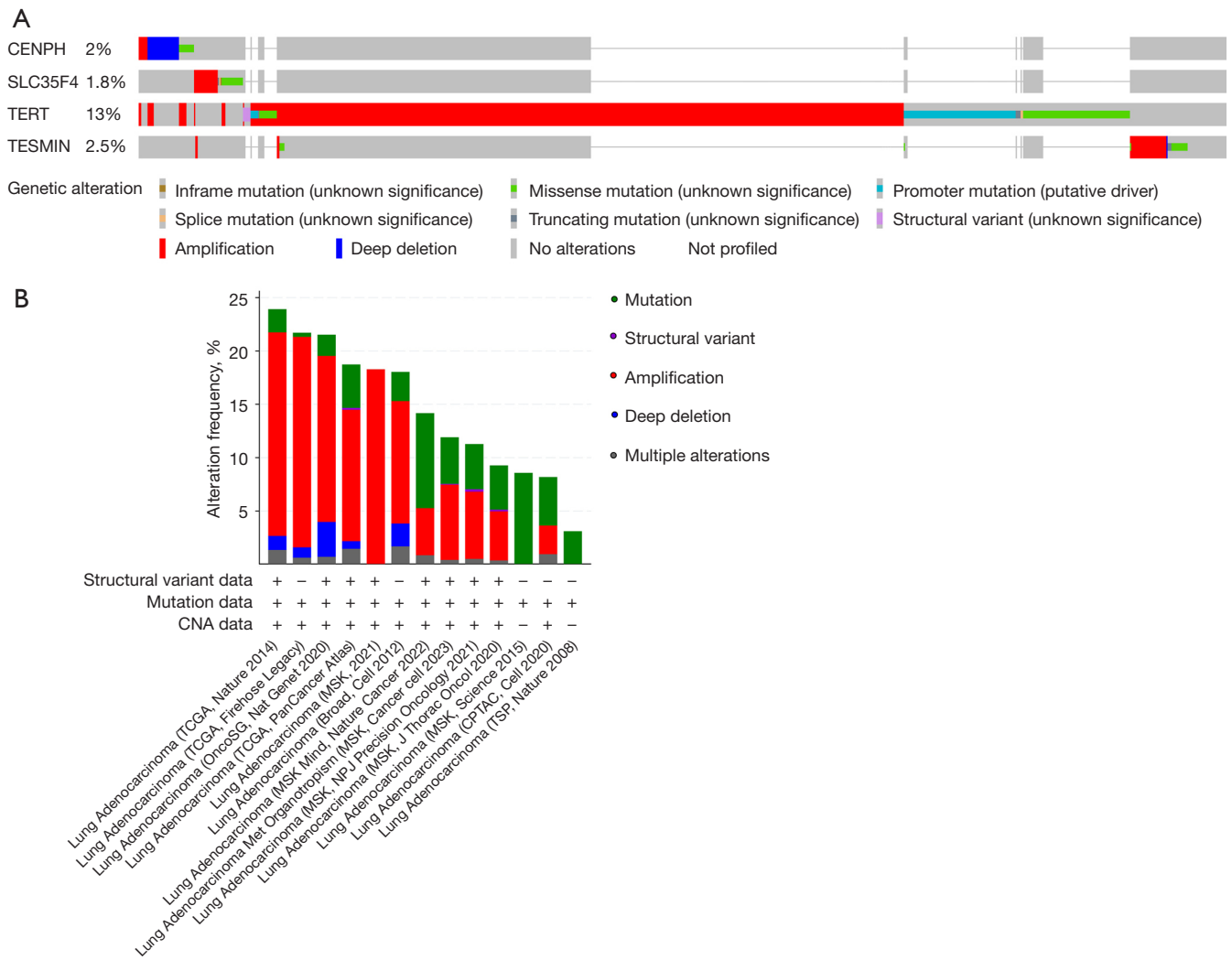


Figure 6 Genomic alterations in the four key prognostic DEGs in LUAD analyzed by the cBioPortal database. (A) OncoPrint of the gene alterations in the four key prognostic DEGs in the LUAD cohort (the different colors indicate different types of genetic alterations, of which amplification accounts for the largest proportion). (B) Details of the gene alteration types in the four key prognosis-related DEGs in the LUAD cohort. CNA, copy number alteration; TCGA, The Cancer Genome Atlas; MSK, Memorial Sloan Kettering; CPTAC, Clinical Proteomic Tumor Analysis Consortium; TSP, Tumour Sequencing Project; DEGs, differentially expressed genes; LUAD, lung adenocarcinoma.

similar results; that is, the high expression of *TERT* in LUAD led to a poor prognosis.

TESMIN (which is also known as *MTL5*) plays a central role in the regulation of cell growth and differentiation and is involved in spermatogenesis (33). TESMIN was first found in mouse testicular tissue, and rat testicular and ovarian tissue during the meiosis of male and female germ cells (34,35). *TESMIN* mRNA expression is not only expressed in the ovary and testis of mice, it is also expressed in kidney, brain, liver and myocardial tissue. Thus far, *TESMIN* expression has been observed in adult prostate, gastric cancers, and NSCLC (36,37). Huang *et al.* reported that *TESMIN* expression was correlated with a poor prognosis and promoted the proliferation of cervical squamous cell carcinoma (38). It has been reported that the expression of *MTL5* is higher in lung cancer cells than in control cells, and its expression may help predict prognosis (39). Our results are consistent with those of

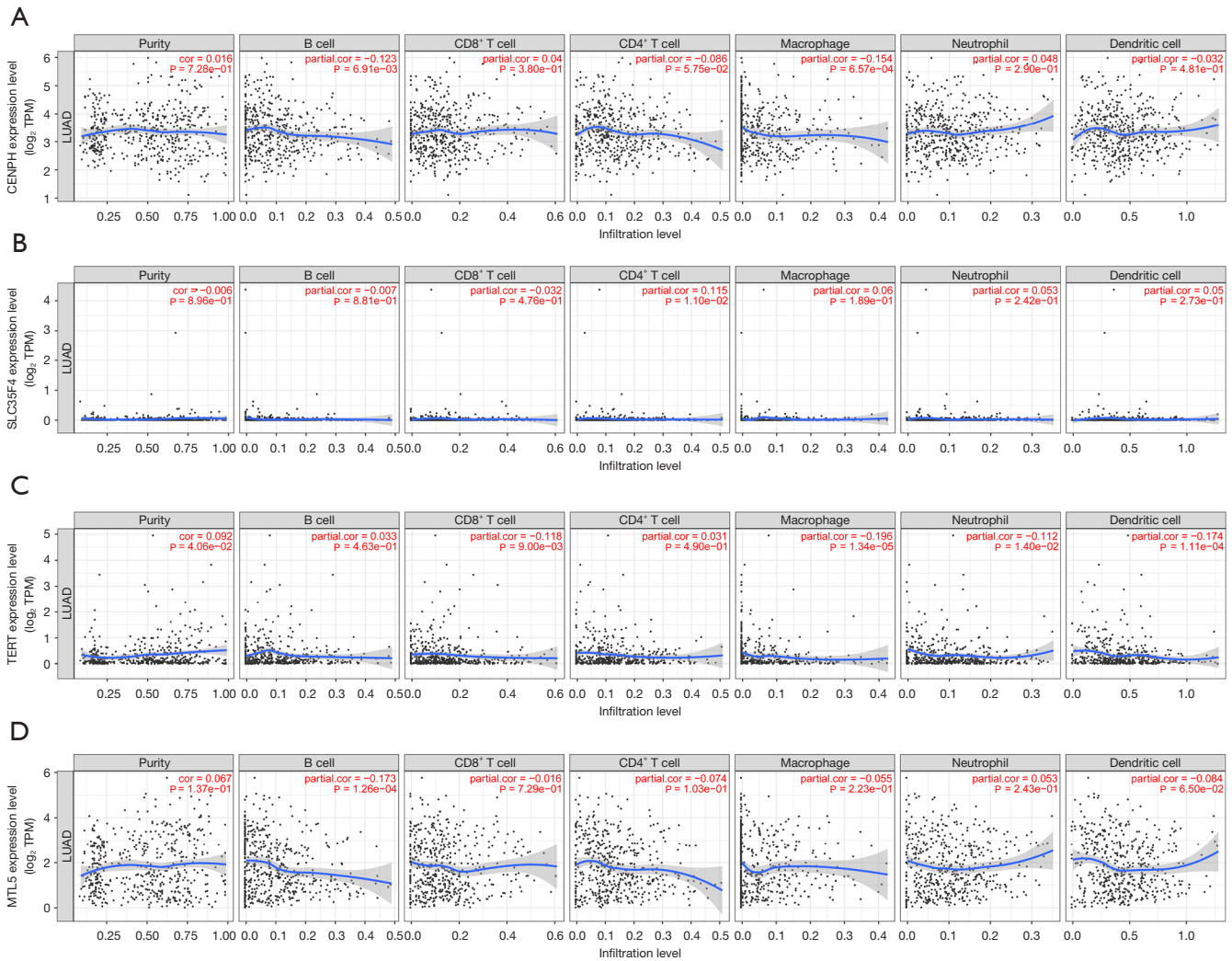


Figure 7 Correlation analysis of the prognosis-related DEGs and immune cell infiltration (TIMER). The correlation between the abundance of immune cell and the expression of (A) CENPH, (B) SLC35F4, (C) TERT, and (D) MTL5 in LUAD. "Purity" represents the purity of the tumor cells in the sample. TPM, transcripts per million; LUAD, lung adenocarcinoma; DEGs, differentially expressed genes; TIMER, Tumor IMmune Estimation Resource; MTL5, metallothionein-like 5 protein, also known as TESMIN.

previous studies.

SLC35F4 is thought to enable transmembrane transporter activity, to be involved in transmembrane transport, and to be an integral component of the membrane. To date, little has been reported about the role of *SLC35F4* in cancer. Indeed, this is the first study to report its role in LUAD cancer.

The possible genetic alterations of the four hub genes in LUAD were examined using cBioPortal. Our research showed the most common mutations in the four key genes. Mutations in transcription factors have long been thought to contribute to tumor development, and previous research suggests that overexpressed oncogenic transcription factors may alter the core self-regulatory circuits of cells (40-42) and may be associated with multiple types of cancer. Additionally, genetic alterations were also observed in *CENPH*, *SLC35F4*, *TERT*, and *TESMIN*, including an in-frame mutation, missense mutation, promoter mutation, splice mutation, truncating mutation, structural variant, amplification, and deep deletion. However, the clinical potential of these gene mutations needs to be confirmed with larger sample sizes, and the exact mechanism of these

gene mutations needs to be validated *in vitro* and *in vivo*.

Finally, we analyzed the relationship between the four key genes and immune infiltration. Gemelli *et al*. have shown that immune infiltration affects the occurrence and development of lung cancer (43). Recent evidence suggests that B cells play a key role in lung cancer, and a study has shown that proliferating B cells are observed in 35% of lung cancers, and are present at all stages of lung cancer development, with differences only observed between the clinical stages and histological subtypes (44). Tumor-associated macrophages are important immune cells in the tumor microenvironment, infiltrated in tumor tissue, and can promote tumor cell proliferation, induce tumor immune tolerance, stimulate tumor angiogenesis, and increase tumor cell invasion and metastasis (45). Tumor-infiltrating lymphocytes (TILs) have been shown to be independent predictors of prognosis and immunotherapy outcomes in cancer patients (46). $CD4^+$ T cell ($CD4^+$ T) cell and ($CD8^+$ T cell) $CD8^+$ T cell are common TILs. In our research, we found that the four key genes were related to B cells, macrophages, and TILs, which shows the potential immune function of the four genes in LUAD. Genetic mutations play a crucial role in the development and progression of LUAD, impacting the immune microenvironment and influencing the choice of treatment regimen, as well as the efficacy and prognosis of patients (47).

However, even though we used multiple databases and R (version 4.3.0) for our comprehensive analysis, there are some limitations in this study. First, the data differences between the databases made the data less specific. Second, we did not conduct any *in vitro* or *in vivo* tests. Third, while we found that the four genes were associated with immunity, our findings have not been verified with further studies. Therefore, we lack specific experimental evidence that the four genes can be used as targets for lung cancer diagnosis and prognosis. In our upcoming research, we intend to examine whether the four genes can serve as targets of LUAD.

Conclusions

We conducted a comprehensive analysis and identified four key genes related to the prognosis of LUAD. Of these four genes, we found that *CENPH* and *TESMIN* could potentially serve as diagnostic and prognostic biomarkers for LUAD with immunomodulatory effects.

Acknowledgments

Funding: None.

Footnote

Reporting Checklist: The authors have completed the STREGA reporting checklist. Available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-1530/rc

Peer Review File: Available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-1530/prf

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://jtd.amegroups.com/article/view/10.21037/jtd-24-1530/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Siegel RL, Miller KD, Wagle NS, et al. Cancer statistics, 2023. CA Cancer J Clin 2023;73:17-48.
2. Lahiri A, Maji A, Potdar PD, et al. Lung cancer immunotherapy: progress, pitfalls, and promises. Mol Cancer 2023;22:40.
3. Thai AA, Solomon BJ, Sequist LV, et al. Lung cancer. Lancet 2021;398:535-54.
4. Miller M, Hanna N. Advances in systemic therapy for non-small cell lung cancer. BMJ 2021;375:n2363.
5. Grant MJ, Herbst RS, Goldberg SB. Selecting the optimal

immunotherapy regimen in driver-negative metastatic NSCLC. Nat Rev Clin Oncol 2021;18:625-44.
6. Zhang S, You X, Zheng Y, et al. The UBE2C/CDH1/DEPTOR axis is an oncogene and tumor suppressor cascade in lung cancer cells. J Clin Invest 2023;133:e162434.
7. de Jong D, Das JP, Ma H, et al. Novel Targets, Novel Treatments: The Changing Landscape of Non-Small Cell Lung Cancer. Cancers (Basel) 2023;15:2855.
8. Goodall GJ, Wickramasinghe VO. RNA in cancer. Nat Rev Cancer 2021;21:22-36.
9. Obeng EA, Stewart C, Abdel-Wahab O. Altered RNA Processing in Cancer Pathogenesis and Therapy. Cancer Discov 2019;9:1493-510.
10. Verbeke R, Lentacker I, Breckpot K, et al. Broadening the Message: A Nanovaccine Co-loaded with Messenger RNA and α-GalCer Induces Antitumor Immunity through Conventional and Natural Killer T Cells. ACS Nano 2019;13:1655-69.
11. Lorentzen CL, Haanen JB, Met Ö, et al. Clinical advances and ongoing trials on mRNA vaccines for cancer treatment. Lancet Oncol 2022;23:e450-8.
12. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2:401-4.
13. Li T, Fu J, Zeng Z, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. Nucleic Acids Res 2020;48:W509-14.
14. Passaro A, Jänne PA, Peters S. Antibody-Drug Conjugates in Lung Cancer: Recent Advances and Implementing Strategies. J Clin Oncol 2023;41:3747-61.
15. MacManus M, Hegi-Johnson F. Overcoming immunotherapy resistance in NSCLC. Lancet Oncol 2022;23:191-3.
16. He WL, Li YH, Yang DJ, et al. Combined evaluation of centromere protein H and Ki-67 as prognostic biomarker for patients with gastric carcinoma. Eur J Surg Oncol 2013;39:141-9.
17. Wu W, Wu F, Wang Z, et al. CENPH Inhibits Rapamycin Sensitivity by Regulating GOLPH3-dependent mTOR Signaling Pathway in Colorectal Cancer. J Cancer 2017;8:2163-72.
18. Wu X, Lin Y, Shi L, et al. Upregulation of centromere protein H is associated with progression of renal cell carcinoma. J Mol Histol 2015;46:377-85.
19. Cong R, Kong F, Ma J, et al. The PVT1/miR-612/CENP-H/CDK1 axis promotes malignant progression of advanced endometrial cancer. Am J Cancer Res 2021;11:1480-502.
20. Li Q, Huang X, Li QY, et al. CENPH overexpression promotes the progression, cisplatin resistance, and poor prognosis of lung adenocarcinoma via the AKT and ERK/P38 pathways. Am J Cancer Res 2023;13:1682-97.
21. Zhou H, Bian T, Qian L, et al. Prognostic model of lung adenocarcinoma constructed by the CENPA complex genes is closely related to immune infiltration. Pathol Res Pract 2021;228:153680.
22. Liao WT, Yu CP, Wu DH, et al. Upregulation of CENP-H in tongue cancer correlates with poor prognosis and progression. J Exp Clin Cancer Res 2009;28:74.
23. Lu G, Hou H, Lu X, et al. CENP-H regulates the cell growth of human hepatocellular carcinoma cells through the mitochondrial apoptotic pathway. Oncol Rep 2017;37:3484-92.
24. Liao WT, Feng Y, Li ML, et al. Overexpression of centromere protein H is significantly associated with breast cancer progression and overall patient survival. Chin J Cancer 2011;30:627-37.
25. Liao WT, Wang X, Xu LH, et al. Centromere protein H is a novel prognostic marker for human nonsmall cell lung cancer progression and overall patient survival. Cancer 2009;115:1507-17.
26. Sayed ME, Yuan L, Robin JD, et al. NOVA1 directs PTBP1 to hTERT pre-mRNA and promotes telomerase activity in cancer cells. Oncogene 2019;38:2937-52.
27. Hasanau TN, Pisarev EP, Kisil OV, et al. The TERT Promoter: A Key Player in the Fight for Cancer Cell Immortality. Biochemistry (Mosc) 2023;88:S21-38.
28. Tornesello ML, Cerasuolo A, Starita N, et al. Reactivation of telomerase reverse transcriptase expression in cancer: the role of TERT promoter mutations. Front Cell Dev Biol 2023;11:1286683.
29. Shi YA, Zhao Q, Zhang LH, et al. Knockdown of hTERT by siRNA inhibits cervical cancer cell growth in vitro and in vivo. Int J Oncol 2014;45:1216-24.
30. Li Y, Li M, Yao G, et al. Telomerase inhibition strategies by siRNAs against either hTR or hTERT in oral squamous cell carcinoma. Cancer Gene Ther 2011;18:318-25.
31. Liu AQ, Ge LY, Lu XL, et al. Silencing of the hTERT gene by shRNA inhibits colon cancer SW480 cell growth in vitro and in vivo. PLoS One 2014;9:e107019.
32. Xie L, Yin W, Tang F, et al. Pan-Cancer analysis of TERT and Validation in Osteosarcoma Cell Lines. Biochem Biophys Res Commun 2023;639:106-16.

33. Sutou S, Miwa K, Matsuura T, et al. Native tesmin is a 60-kilodalton protein that undergoes dynamic changes in its localization during spermatogenesis in mice. Biol Reprod 2003;68:1861-9.
34. Olesen C, Møller M, Byskov AG. Tesmin transcription is regulated differently during male and female meiosis. Mol Reprod Dev 2004;67:116-26.
35. Matsuura T, Kawasaki Y, Miwa K, et al. Germ cell-specific nucleocytoplasmic shuttling protein, tesmin, responsive to heavy metal stress in mouse testes. J Inorg Biochem 2002;88:183-91.
36. Grzegrzolka J, Olbromski M, Gomulkiewicz A, et al. Role of tesmin expression in non-small cell lung cancer. Oncol Lett 2021;21:48.
37. Oh JH, Yang JO, Hahn Y, et al. Transcriptome analysis of human gastric cancer. Mamm Genome 2005;16:942-54.
38. Huang Y, Wu Q, Tan X. Metallothionein-like 5 expression is correlated with poor prognosis and promotes proliferation of cervical squamous cell carcinoma. Bioengineered 2022;13:6955-65.
39. Grzegrzolka J, Gomulkiewicz A, Olbromski M, et al. Expression of tesmin (MTL5) in non small cell lung cancer: A preliminary study. Oncol Rep 2019;42:253-62.
40. Iranzo J, Martincorena I, Koonin EV. Cancer-mutation network and the number and specificity of driver mutations. Proc Natl Acad Sci U S A 2018;115:E6010-9.
41. Martincorena I, Campbell PJ. Somatic mutation in cancer and normal cells. Science 2015;349:1483-9.
42. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415-21.
43. Gemelli M, Noonan DM, Carlini V, et al. Overcoming Resistance to Checkpoint Inhibitors: Natural Killer Cells in Non-Small Cell Lung Cancer. Front Oncol 2022;12:886440.
44. Leong TL, Bryant VL. B cells in lung cancer-not just a bystander cell: a literature review. Transl Lung Cancer Res 2021;10:2830-41.
45. Qiao Y, Fu E. Advances in the Study of Tumor-associated Macrophages in Lung Cancer. Zhongguo Fei Ai Za Zhi 2022;25:34-9.
46. Wu Y, Yuan M, Wang C, et al. T lymphocyte cell: A pivotal player in lung cancer. Front Immunol 2023;14:1102778.
47. Wang Y, Jiang H, Fu L, et al. Prognostic value and immunological role of PD-L1 gene in pan-cancer. BMC Cancer 2024;24:20.

(English Language Editor: L. Huleatt)

Cite this article as: Peng L, Xia L, Yang M, Wen Y, Zeng Q. Comprehensive bioinformatics analysis of prognosis and immunotherapy in lung adenocarcinoma. J Thorac Dis 2024;16(12):8633-8647. doi: 10.21037/jtd-24-1530