Received: 31 March 2022 | Revised: 27 April 2022 | Accepted: 22 May 2022

DOI: 10.1002/jcla.24538

RESEARCH ARTICLE

WILEY

Establishing an 8-gene immune prognostic model based on TP53 status for lung adenocarcinoma

Guodong Wu | **Youyu Wang** | **Yanhui Wan**

Thoracic Surgery, Shenzhen Second People's Hospital, Shenzhen, China

Correspondence
Yanhui Wan, Thoracic Surgery, Shenzhen Second People's Hospital, 3002 Sungang West Road, Futian District, Shenzhen, 440300, China.
Email: wisewyh@163.com

Funding information
This study received no funding or other financial support

Abstract

Background: Lung adenocarcinoma (LUAD) results in a majority of cancer burden worldwide. TP53 is the most commonly mutated in LUAD. This study aimed to reveal the relation between TP53 and tumor microenvironment (TME) for improving LUAD treatment.

Methods: Differentially expressed genes (DEGs) related to immunity were analyzed between TP53-WT and TP53-MUT groups. Least absolute shrinkage and selection operator (LASSO) Cox regression was applied to screen prognostic DEGs. Two independent datasets were included to evaluate the robustness of the prognostic model.

Results: An 8-gene prognostic model containing ANLN, CCNB1, DLGAP5, FAM83A, GJB2, NAPSA, SFTPB, and SLC2A1 was established based on DEGs. LUAD samples were classified into high- and low-risk groups with differential overall survival in the two datasets. M0 macrophages, M1 macrophages, and activated memory CD4 T cells were more enriched in high-risk group. Immune checkpoints of PDCD1, LAG3, and CD274 were also high-expressed in high-risk group.

Conclusion: The study improved the understanding of the role of TP53 in the TME modulation. The 8-gene model had robust performance to predict LUAD prognosis in clinical practice. In addition, the eight prognostic genes may also serve as potential targets for designing therapeutic drugs for LUAD patients.

KEYWORDS
bioinformatics analysis, lung adenocarcinoma, prognostic genes, TP53, tumor microenvironment

1 | INTRODUCTION

Lung cancer is the most commonly diagnosed cancer type, contributing to a majority of cancer burden worldwide.[1] Non-small cell lung cancer (NSCLC) consists of approximately 85% of all lung cancer cases, of which lung adenocarcinoma (LUAD) is the most common histologic subtype. The 5-year survival rate of lung cancer patients is lower than 20%, as most of patients were already at advanced stage when diagnosed.[2] Smoking or smoke exposure is a major risk factor resulting in the development of lung cancer, with male population presenting a higher incidence rate. In addition, females have more favorable outcomes than males in the US.[3] Apart from non-genetic

Guodong Wu and Youyu Wang equal contribution.

risk factors, a number of studies have demonstrated that specific gene mutations affect the development, progression, and treatment efficiency of lung cancer.[4]

Tumor mutation burden (TMB) has been considered as a biomarker in resected NSCLC, and the patients with high nonsynonymous TMB had a better prognosis.[5] Meanwhile, patients with low nonsynonymous TMB were more sensitive to benefit from adjuvant chemotherapy,[5] which provided a direction for personalized therapy. Up to now, large amount of evidences have illustrated that EGFR, ALK, and KRAS were frequently mutated in lung cancer patients.[4,6–8] Based on the close relations between these mutations and prognosis, several inhibitors targeting these genes have been explored and achieved encouraging outcome.

Tumor protein p53 (TP53) is an oncogene and its mutations are the most commonly detected in all cancer types including lung cancer. TP53 gene alteration is associated with poor overall survival of NSCLC patients.[9] The status of TP53 mutations differs according to different stages and pathological types,[9] indicating that TP53 mutations play an important role in tumorigenesis and cancer progression in lung cancer. Tumor microenvironment is verified as a critical factor in determining cancer development or the response to immunotherapy such as immune checkpoint blockade.[10] Lin et al. discovered a different TME between TP53-mutated and TP53-wild type groups in LUAD patients,[11] suggesting that TP53 status could affect the modulation of TME and immune-related pathways.

In this study, we further described molecular features of TP53-mutated and TP53-wild type groups and identify differentially expressed genes (DEGs) related to immunity between the two groups. We constructed an 8-gene model based on DEGs related to immunity and validated the effectiveness and robustness of the model for predicting LUAD prognosis. The model could classify LUAD patients into high- and low-risk groups with differential overall survival, TME, and enriched pathways. Overall, the relation between TP53 and immunity was further demonstrated and cleared. A nomogram based on the model and clinical stages could be applied in clinical practice.

2 | MATERIALS AND METHODS

2.1 | Data source and data preprocessing

The expression profiles and clinical information of 522 samples and mutation of somatic cells of 516 samples in TCGA-LUAD dataset were obtained from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/repository). Ensembl ID was transferred to gene symbol according to gff3 (v22 and v33) file downloaded from GENCODE (https://www.gencodegenes.org/human). Median expression value was taken when multiple gene symbols matched to one gene. Log2 transform was performed for all expression data. GSE30219 dataset containing 293 LUAD samples, and 14 normal samples was downloaded from Gene Expression Omnibus (GEO) dataset (https://www.ncbi.nlm.nih.gov/geo/).

2.2 | Gene set enrichment analysis (GSEA)

GSEA is a methodology for evaluating the enrichment of a pathway or a group of cells based on gene sets.[12] GSEA software (v3.0) was used to calculate the enrichment score of TP53-WT and TP53-MUT groups based on the gene set of "c7.immunesigdb.v7.4.symbols.gmt" downloaded from Molecular Signatures Database (MsigDB, http://www.gsea-msigdb.org/gsea/downloads.jsp). The minimum and maximum number of genes was set as 5 and 5000, respectively.

2.3 | Identification of differentially expressed genes

Limma R package (v3.40.6) was applied to select DEGs between TP53-WT and TP53-MUT groups.[13] Specifically, multiple linear regression was performed by lmFit function. eBays function was conducted to analyze moderated t-statistics, moderated F-statistic, and log odds of differential expression. $|\text{Log2 (fold change [FC])}| > 2$ and $P < 0.05$ were determined to screen DEGs.

2.4 | Establishing an immune prognostic model (IPM)

A total of 486 samples with expression data and TP53 mutation data in TCGA-LUAD dataset were included in the following analysis. Univariate Cox regression analysis was performed to assess the prognostic significance of immune-related DEGs, and $P < 0.05$ was confirmed to screen prognostic DEGs. Least absolute shrinkage and selection operator (LASSO) Cox regression in glmnet R package were employed to reduce the number of prognostic DEGs.[14] Then multivariate Cox regression was used to calculate coefficients of the remained prognostic genes. The IPM was defined as: risk score = coefficient 1*gene 1 + coefficient 2*gene 2 + … + coefficient n*gene n. GSE30219 was used as a validation dataset. Risk score was calculated for each sample. Survminer R package (http://www.sthda.com/english/rpkgs/survminer/) was implemented to determine the optimal cutoff of risk score for classifying samples into high-risk and low-risk groups.

2.5 | CIBERSORT analysis

CIBERSORT was applied to estimate the proportion of 22 immune cells for each sample.[15] It supports the calculation of the enrichment score for a group of immune cells based on the corresponding markers or a gene set, and it has been widely used for cancer data analysis. For each sample, a total proportion of all immune cells were set as 100%.

2.6 | Functional analysis

ClusterProfiler (v3.14.3) R package was introduced to annotate the terms of gene ontology (GO) and Kyoto Encyclopedia of Genes and

Genomes (KEGG) pathways in high-risk and low-risk groups.[16] The gene set of "c5.go.mf.v7.4.symbols.gmt" downloaded from MsigDB was used as a basis to annotate GO terms. The latest information of KEGG pathways (https://www.kegg.jp/kegg/rest/keggapi.html) was used. For GO terms and KEGG pathways, the minimum and maximum number of genes were set as 5 and 5000, respectively. $P < 0.05$ was considered as significant.

2.7 | Constructing a nomogram for application in clinic

Firstly, univariate and multivariate Cox regression analyses were performed on risk score and clinical features for identifying independent risk factors (hazard ratio [HR] > 1, $P < 0.05$). Then, risk score and stage served as a basis to construct a nomogram for predicting 1-year, 3-year, and 5-year overall survival using rms R package (http://CRAN.R-project.org/package=rms). Concordance index (C-index) and receiver operation curve analysis were conducted to evaluate the effectiveness of the nomogram through comparing it with other predictors.

2.8 | Statistical analysis

All statistical analysis was performed in R software (v4.1.0). Two-tailed Student *t* test was conducted to compare the gene expression of two groups. Log-rank test was used in Kaplan–Meier survival analysis and Cox regression analysis. $P < 0.05$ was considered as significant. Parameters of software or packages not shown were set as default.

3 | RESULTS

3.1 | Differential molecular features between TP53-MUT and TP53-WT groups

In TCGA-LUAD dataset, the frequencies of different types of single-nucleotide variations (SNVs) were calculated (Figure 1A). In these types, C-to-A and C-to-T contributed over half of all SNVs. Transversions (Tv) were almost twice as much as transitions (Ti). Of the top 20 mutated genes in the samples, TP53 was the most mutated gene (48% mutated in 565 samples), and missense mutations consisted of the majority of mutations (Figure 1B). Compared with other mutated genes, TP53 had a higher proportion of nonsense mutations and frame-shift deletions.

Then TCGA-LUAD samples were grouped into two groups according to whether TP53 mutated. TP53-WT referred to the group without mutation of TP53 (333 samples), while TP53-MUT indicated the group with mutations of TP53 (241 samples). GSEA on the two groups showed that immune-related gene sets were significantly enriched in TP53-WT group. Eight biological processes related to immunity were visualized (Figure 1C), suggesting a relation between immunity and TP53 mutation. Differential expression analysis verified that 119 DEGs related to immunity were identified between the two groups (Figure S1, $P < 0.05$, $|\log2FC| > 2$).

3.2 | Constructing a prognostic model based on DEGs

Univariate Cox regression analysis was performed on 119 DEGs, and 89 DEGs were screened to be significantly associated with LUAD prognosis. Then LASSO Cox regression was used to identify DEGs closely associated with overall survival. At last, eight DEGs including ANLN, CCNB1, DLGAP5, FAM83A, GJB2, NAPSA, SFTPB, and SLC2A1 remained. An 8-gene prognostic model named as IPM was established based on the normalized expression and multivariate Cox coefficients of these genes. Risk score was calculated for each sample based on the prognostic model. In TCGA-LUAD dataset, samples were classified into high-risk and low-risk groups with differential overall survival (Figure 2A, p = 1.8e-10, HR = 2.58 [1.91–3.49]). Dead samples were more accumulated in high-risk group (Figure 2B). Except for NAPSA and SFTPB showing a relatively higher expressing in low-risk group, other six prognostic genes were all relatively higher expressed in high-risk group (Figure 2B). ROC analysis demonstrated that IPM had favorable area under ROC curve (AUC) in predicting 1-year, 3-year, and 5-year overall survival, with an AUC of 0.73, 0.68 and, 0.70, respectively (Figure 2C). In GSE30219 dataset, we observed the similar results that two groups manifested distinct prognosis (p = 6.9e-12, HR = 3.83) although a lower AUC was presented (Figure 2D–F). Compared with an 8-gene signature proposed by Li et al.,[17] our model had better scores of both C-index and AUC in predicting 1- to 5-year overall survival (Figure 2G,H).

3.3 | The performance of IPM in classifying TP53-based groups

On the overall survival in TCGA-LUAD dataset, TP53-MUT and TP53-WT groups showed no significant difference (Figure 3A, p = 0.31). However, IPM could classify LUAD samples into high- and low-risk groups for TP53-MUT and TP53-WT groups, respectively (Figure 3B,C, $p < 0.0001$). Pearson correlation analysis revealed a significantly negative correlation between IPM risk score and survival time in TP53-MUT group (Figure 3D, p = 1.3e-3, R = −0.21), but there was no obvious correlation in TP53-WT group. Univariate and multivariate Cox regression analysis on IPM risk score and TP53 status exhibited that risk score was an independent risk factor in univariate analysis (Figure 3E, HR = 3.77, 95%confidence interval [CI] = 2.55–5.56, $p < 0.001$). In addition, we analyzed whether different mutation types of TP53 had different influence on overall

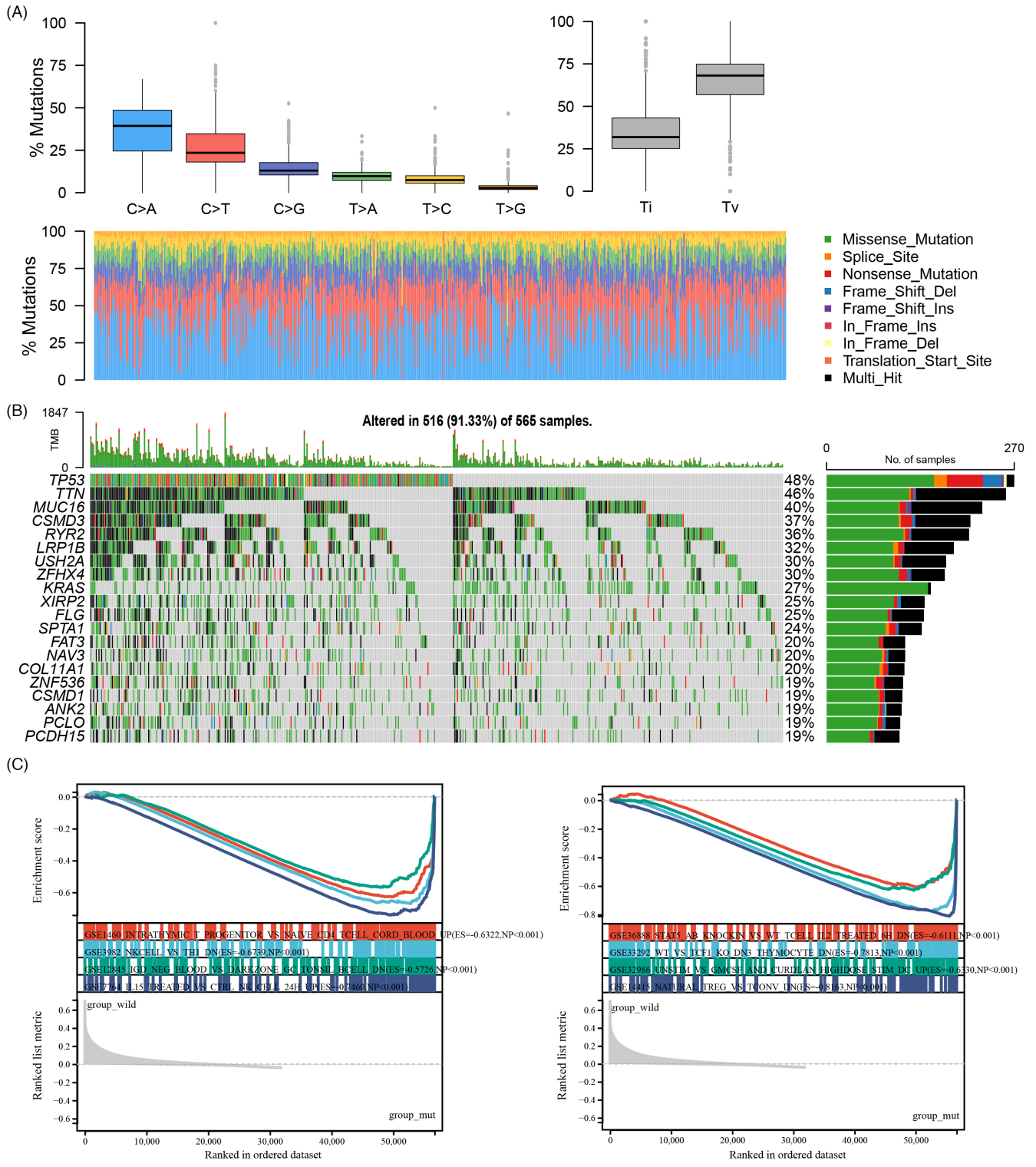

FIGURE 1 Classification of LUAD samples based on TP53 status in TCGA-LUAD dataset. (A) Mutation frequencies of different SNVs in TP53. Ti and Tv indicate transitions and transversions respectively. (B) The top 20 mutated genes. Horizontal axis indicates samples and vertical axis indicates genes. The right line indicates sample numbers of different types of mutations labeled in different colors. (C) GSEA for TP53-MUT and TP53-WT groups, and eight immune-related pathways of TP53-WT group were visualized. Ti, transition Tv, transversion

survival. The result showed no significant difference among them (Figure 3F, $p = 0.63$). Noteworthy, IPM was also effective to divide samples with missense and nonsense mutations into high- and low-risk groups (Figure 3G,H, $p = 4.64e\text{-}4$ and $p = 0.01$, respectively), suggesting that our 8-gene signature was robust to different types of LUAD samples.

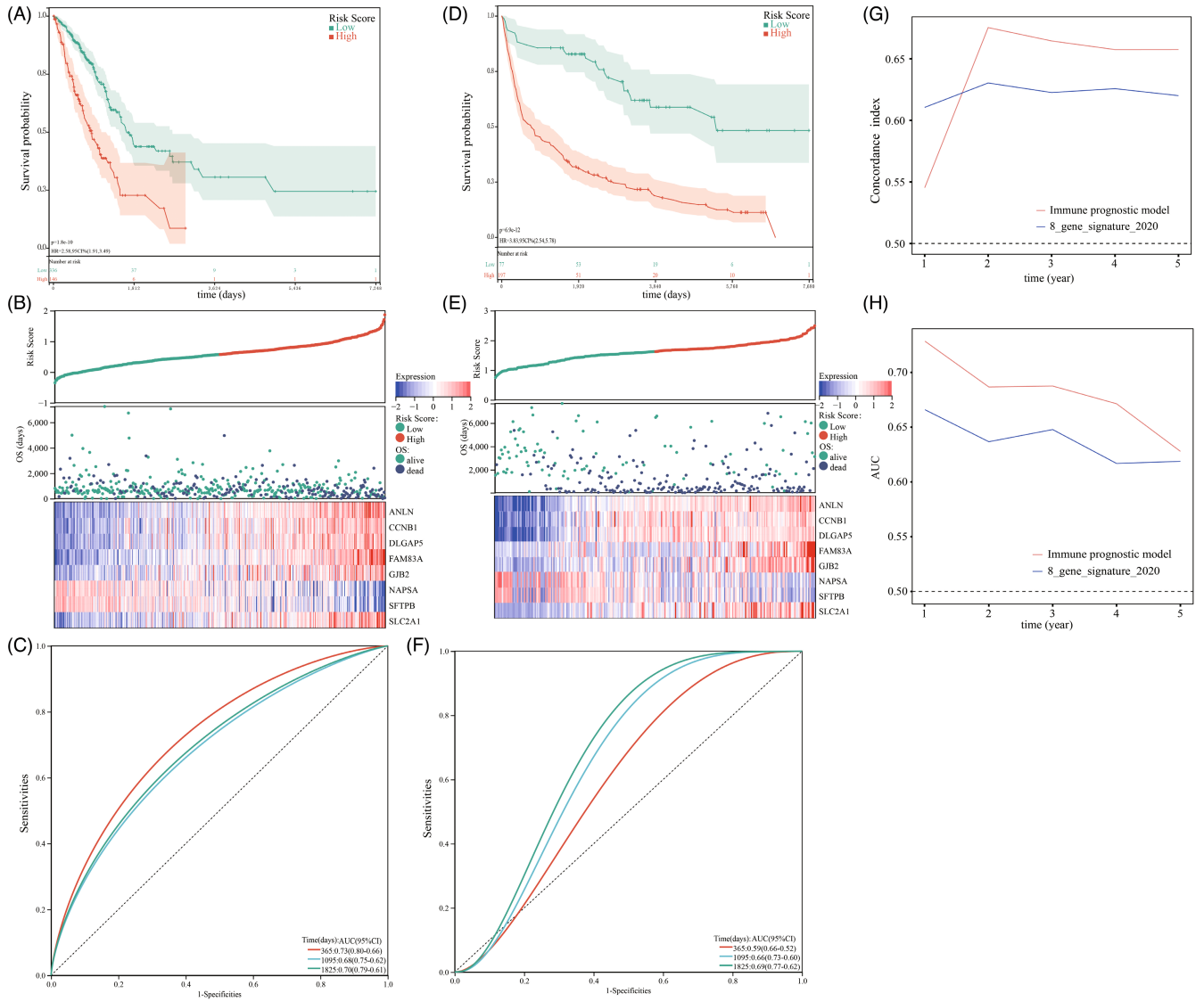

FIGURE 2 Establishment and verification of IPM. (A) Kaplan–Meier survival plot of high- and low-risk groups in TCGA-LUAD dataset. Log-rank test was conducted. (B) Survival status and expression of eight prognostic genes ranking by risk score in TCGA-LUAD dataset. (C) ROC analysis on the effectiveness of IPM for predicting 1-year, 3-year, and 5-year overall survival in TCGA-LUAD dataset. (D-F) Verification of IPM in GSE30219 dataset. (G) C-index of IPM and 8-gene signature from Li et al. (H) AUC of IPM and 8-gene signature from Li et al. ROC, receiver operation curve. AUC, area under ROC curve

3.4 | Comparison of TME between high- and low-risk groups

To understand the difference of TME between high- and low-risk groups, we applied CIBERSORT to estimate the proportion of 22 immune cells in TCGA-LUAD dataset. The distribution of 22 immune cells of 574 LUAD samples was delineated (Figure 4A). Pearson correlation analysis on the proportion of different immune cells revealed obvious correlations among some immune cells, such as M1 macrophages and CD8 T cells, M1 macrophages and activated memory CD4 T cells, resting memory CD4 T cells and CD8 T cells (Figure 4B). We observed that the expression of M0 macrophages, M1 macrophages, activated memory CD4 T cells and resting memory CD4 T cells were differentially distributed between high- and low-risk groups, with the former three cell types showing greater enrichment in high-risk group (Figure 4C, $p < 0.0001$). Furthermore, we examined the expression of six critical immune checkpoints (CTLA4, PDCD1, IDO1, LAG3, CD274, and HAVCR2), and found close correlations between these checkpoints and risk score (Figure 4D). Three immune checkpoints including CD274, LAG3, and PDCD1 exhibited differential expression levels between high-risk and low-risk groups, with high-risk group presenting higher expression of all three checkpoints (Figure 4E, $p < 0.05$). This indicated that two groups may respond differently to immune checkpoint blockade therapy.

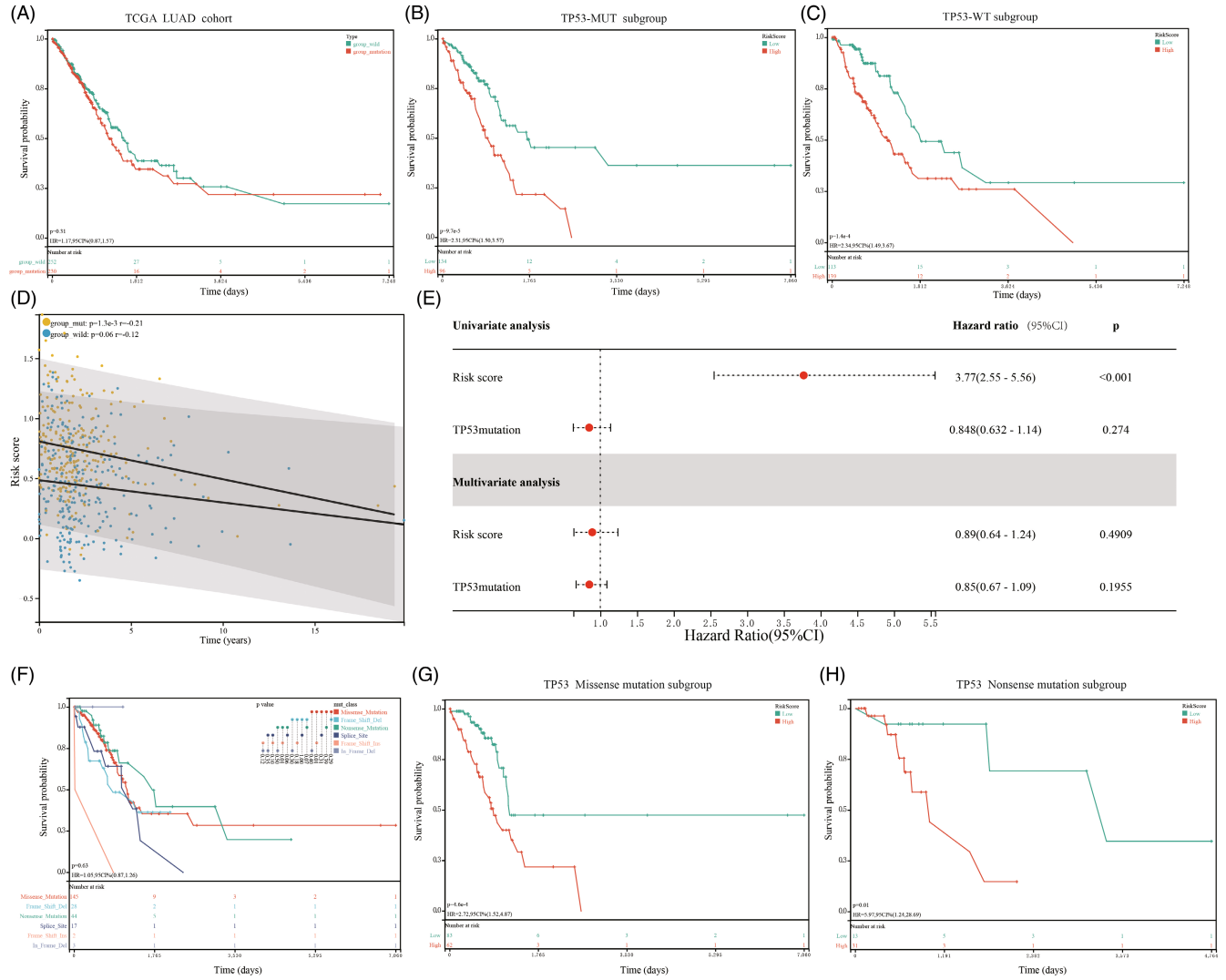

FIGURE 3 IPM performance in groups classified by TP53 status in TCGA-LUAD dataset. (A) Kaplan–Meier survival plot of TP53-WT and TP53-MUT groups. (B-C) IPM performance in classifying high- and low-risk groups for TP53-WT (B) and TP53-MUT (C) groups, respectively. (D) Pearson correlation analysis between risk score and TP53 status (MUT and WT). (E) Univariate and multivariate Cox regression analysis on risk score and TP53 status. (F) Kaplan–Meier survival plot of samples grouped by different types of TP53 mutations. (G, H) Kaplan–Meier survival plot of IPM performance in TP53-missense and TP53-nonsense groups. Log-rank test was conducted. HR, hazard ratio. CI, confidence interval

3.5 | Differentially expressed genes between high- and low-risk groups

By comparing the expression profiles of high- and low-risk groups, we observed that 200 immune-related genes were differentially expressed ($p < 0.05$, Figure S2). Functional analysis revealed that these genes mostly enriched in GO terms of microtubule binding and tubulin binding ($p < 0.05$, Figure 5A). Of KEGG pathways, cancer-related pathways such as cell cycle, cellular senescence, P53 signaling pathway, and central carbon metabolism in cancer were significantly enriched ($p < 0.05$, Figure 5B). The above results suggested that these differentially expressed genes related to immunity may be involved in the cancer development through modulating TME.

3.6 | Application of IPM in LUAD patients

Compared with other clinical factors, IPM risk score was a better independent risk factor, with HR = 3.756 (95% CI = 2.397–5.885) in univariate Cox regression and HR = 3.11 (95% CI = 1.96–4.93) in multivariate Cox regression analysis ($p < 0.001$, Figure 6A). C-index analysis on T stage, N stage, Stage, sex, age, and risk score showed that risk score had the highest C-index in predicting short-term or long-term survival (Figure 6B). Based on risk score and stage, we then constructed a nomogram for more precisely and conveniently predicting LUAD prognosis in clinical practice (Figure 6C). The sum of two points could be obtained according to risk score and stage status, with each total point corresponding to predicted 1-year, 3-year, and 5-year overall survival. The predicted overall survival by

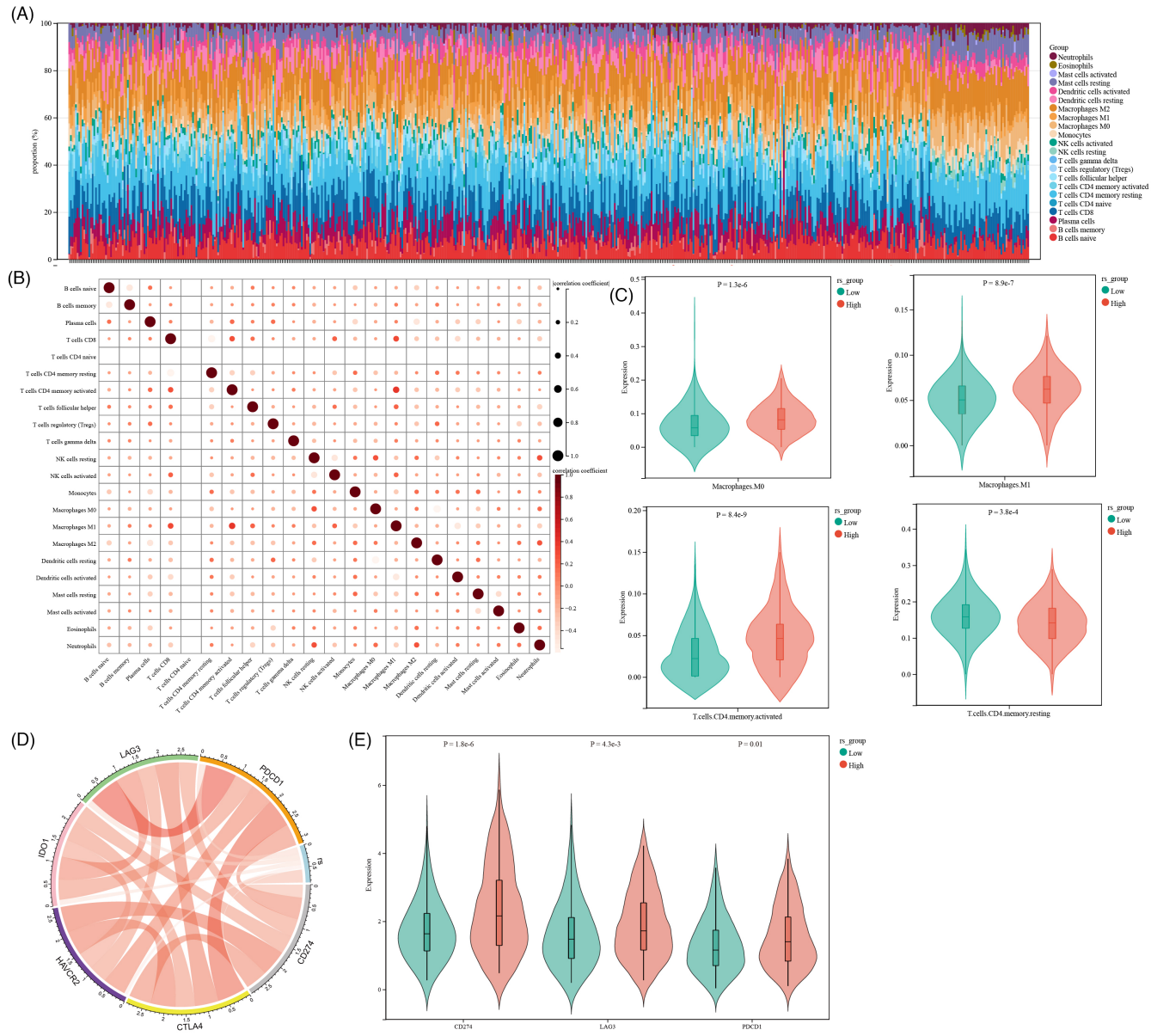

FIGURE 4 TME of high-risk and low-risk groups in TCGA-LUAD dataset. (A) CIBERSORT analysis on 22 immune cells. Horizontal axis indicates samples and vertical axis indicates estimated proportions. (B) Pearson correlation analysis between either of two types of immune cells. Dot size indicates |correlation coefficient| and the shade of color indicates correlation coefficient. (C) Comparing the proportion of M0 macrophages, M2 macrophages, activated memory CD4 T cells, and resting memory CD4 T cells between high- and low-risk groups. (D) The correlation between risk score and six immune checkpoints. (E) The expression of CD274, LAG3, and PDCD1 in high- and low-risk groups. Student *t* test was performed. rs, risk score

the nomogram was corrected by the observed overall survival, with 3-year survival prediction showing a higher accuracy (Figure 6D). ROC analysis also demonstrated that the nomogram was effective to predict 1-year, 3-year, and 5-year prognosis (Figure 6E, AUC = 0.76, 0.74, and 0.72, respectively).

4 | DISCUSSION

TP53 status was associated with the prognosis of lung cancer patients. In LUAD patients, TP53-WT and TP53-MUT groups have differential TME, with more activated immune infiltration in TP53-MUT group.[11] Growing evidence has demonstrated that silencing TP53 in fibroblasts leads to the overexpression of Endothelial Nitric Oxide Synthase (eNOS), subsequently inducing an imbalanced state in TME.[18–20] Microenvironment homeostasis is disrupted following with oxidative stress damaging DNA in stromal and tumor cells.[21] In an animal model lacking TP53 expression, decreased amount of CD8 T cells and increased number of myeloid-derived suppressor cells and leukocytes were discovered when compared with wild type,[22] further supporting the evidence of TP53 in modulating TME.

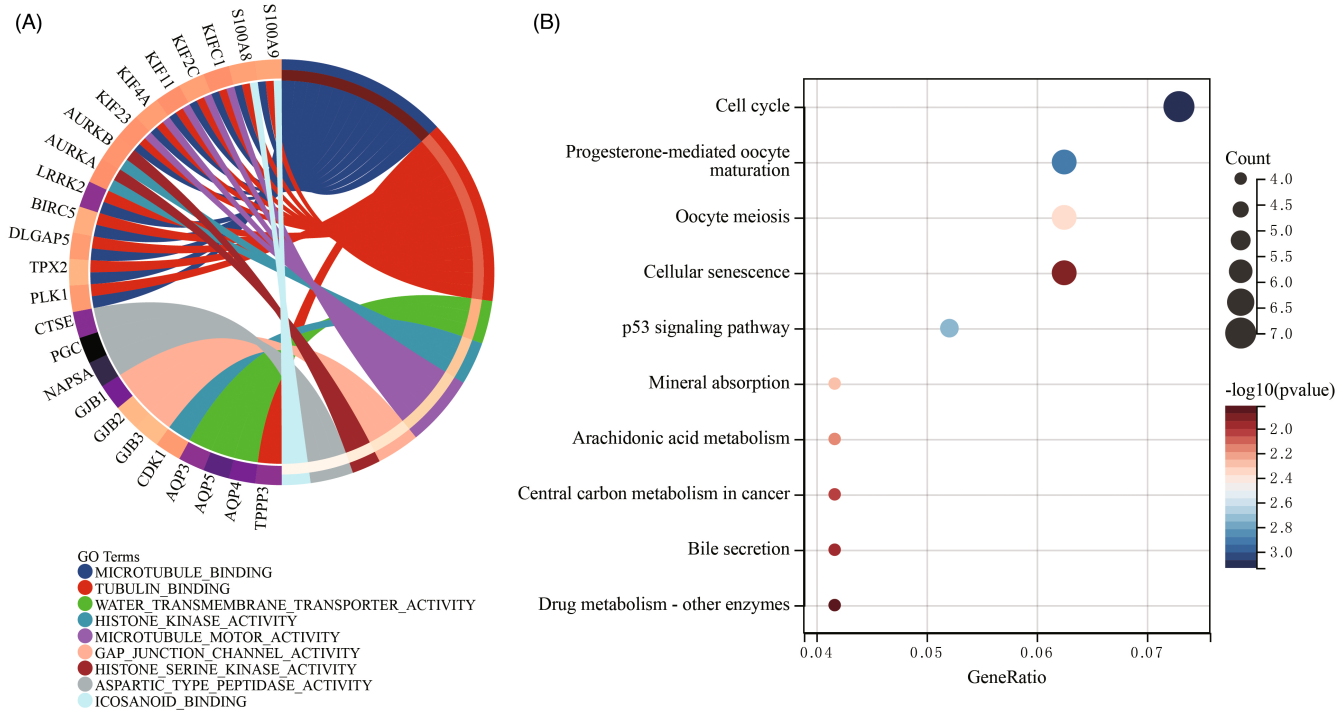

FIGURE 5 Functional analysis of immune-related DEGs in TCGA-LUAD dataset. (A) Circle plot of enriched biological process. Left part indicates genes and right part indicates GO terms. (B) The top 10 enriched KEGG pathways of DEGs. Count represents gene counts in one pathway. DEGs, differentially expressed genes

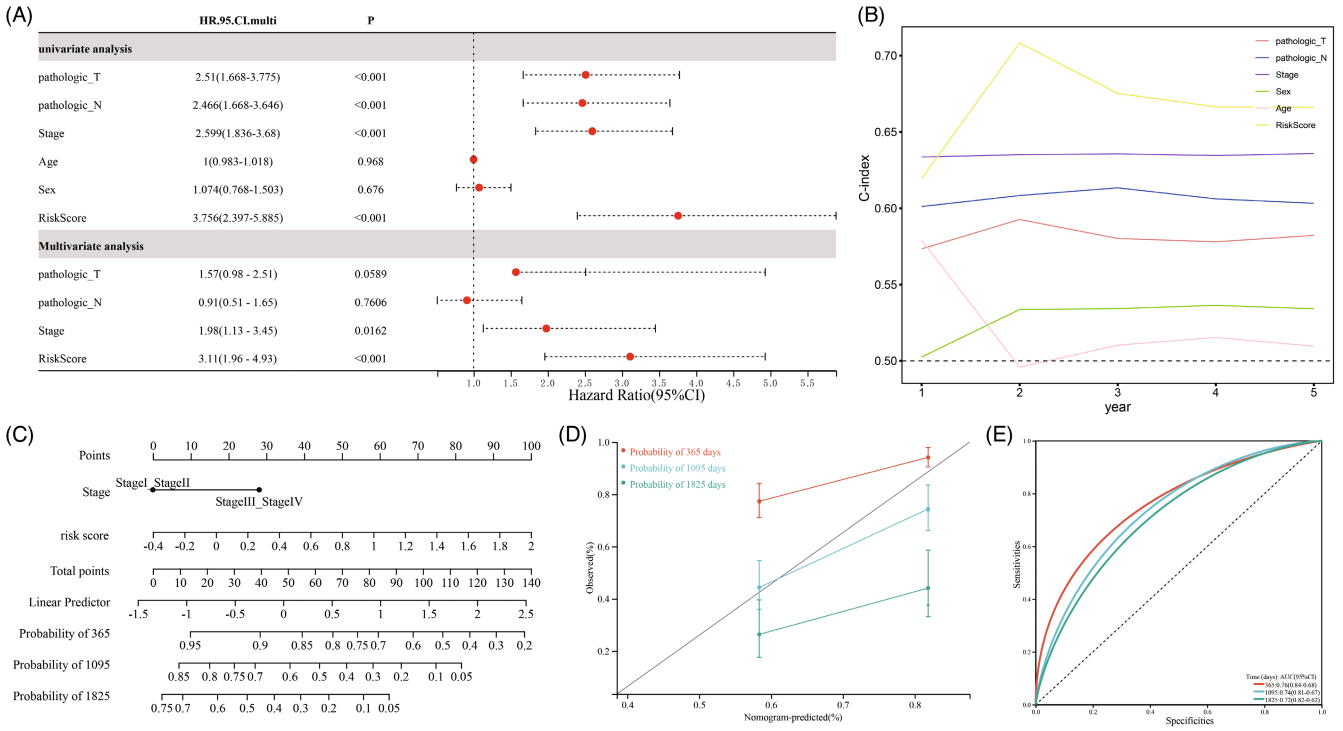

FIGURE 6 Application of IPM in clinic. (A) Univariate and multivariate Cox regression analysis on IPM and clinical features. (B) C-index of IPM and clinical features. (C) A nomogram based on IPM and stage for predicting 1-year, 3-year, and 5-year overall survival. (D) The correction for nomogram-predicted survival based on the observed survival. (E) ROC curve of the nomogram for predicting 1-year, 3-year, and 5-year overall survival. HR, hazard ratio. CI, confidence interval

In our study, we also observed a number of immune-related genes differentially expressed between TP53-WT and TP53-MUT groups, indicating that TP53 mutations could influence immune regulation in LUAD. Based on immune-related DEGs, we established an 8-gene prognostic model consisted of ANLN, CCNB1, DLGAP5, FAM83A, GJB2, NAPSA, SFTPB, and SLC2A1. The prognostic model was stable to divide LUAD samples into high- and low-risk groups with distinct overall survival in two independent datasets. In addition, the model manifested a favorable performance in classifying samples with different TP53 status.

By comparing high-risk to low-risk groups based on TME features, we discovered that M0 macrophages, M1 macrophages, and activated memory CD4 T cells were more accumulated in high-risk group. Moreover, some critical immune checkpoints including CD274, PDCD1, and LAG3 all showed higher expression in high-risk group. These observations supported that TME could be regulated by TP53 status and therefore may have a difference in the efficiency of immune checkpoint blockade therapy. Furthermore, compared to the 8-gene model (KIF2C, MKI67, CCNA2, CCNB2, TTK, HMMR, ASPM, and CDCA8) developed by Li et al., our model exhibited a stronger performance with higher C-index and AUC.

ANLN has been identified as a potential prognostic biomarker in lung cancer in previous research.[23,24] Overexpression of ANLN is associated with metastasis and poor prognosis,[25] which is consistent with our result that high-risk group had higher expression of ANLN. Suzuki et al. have found that ANLN may induce carcinogenesis through the involvement in PI3K-Akt signaling pathway in NSCLC cells.[26] CCNB1 expression has been reported to be involved in the cancer development in many cancer types including in breast cancer,[27] colorectal cancer,[28] gastric cancer,[29] and lung cancer.[30] DLGAP5 is reported as a mitosis-related gene associated with NSCLC prognosis, and high expression of DLGAP5 was correlated with unfavorable prognosis.[31,32]

FAM83A is considered as a potential target for designing therapeutic drugs for NSCLC.[33] Chen et al. have revealed that FAM83A-AS1 facilitates cancer cell migration and progression through HIF-1α/glycolysis axis in LUAD cells.[34] Hippo signaling and Wnt signaling could be regulated by FAM83A in lung cancer cells.[35] GJB2 is one of connexin genes, whose expression is significantly elevated in NSCLC patients.[36,37] LUAD patients with high GJB2 expression exhibit poorer prognosis compared with LUAD patients with relatively low GJB2 expression.[38] NAPSA is less reported compared with the above genes, but it is identified as one of surfactant metabolism-related genes associated with the prognosis of primary LUAD and LUAD with brain metastasis.[39] Pro-SFTPB is considered as a risk biomarker for predicting the risk of developing lung cancer.[40,41] SLC2A1 encodes a glucose transporter (GLUT1) participating in the metabolism of glucose, which is an energy source for tumor cell growth. High expression of SLC2A1 is associated with diminished prognosis in gastric cancer.[42]

The total of eight prognostic genes related to immunity have been reported to be associated with cancer development and progression. The mechanism of their regulation in TME is complicated possibly through oncogenic pathways such as Hippo, Wnt, PI3K-Akt signaling as previous research reported. The direct interaction between these genes and immunity is less described in the previous studies. Some of prognostic genes are not only related to immunity, but associated with other roles such as metabolism and cell cycle. This study paved a way for uncovering the regulation network among TP53, TME and oncogenic pathways.

Overall, this study focused on TP53 status and identified a number of DEGs related to immunity. Based on the comparison between high-risk and low-risk groups, we further clarified the interaction between TP53 and TME, and the role of eight prognostic genes in cancer development in LUAD patients. The prognostic model was robust to be applied in clinical practice for predicting LUAD prognosis, but also could serve as potential targets for exploring therapeutic drugs for LUAD treatment.

AUTHOR CONTRIBUTIONS

Guodong Wu contributed to data curation and writing—original draft preparation. Youyu Wang contributed to software and validation. Yanhui Wan contributed to conceptualization, supervision, and writing—reviewing and editing.

CONFLICT OF INTEREST

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

DATA AVAILABILITY STATEMENT

The dataset analyzed in this study could be found in [GSE30219] at [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE30219]

ORCID

Yanhui Wan https://orcid.org/0000-0002-9514-2532

REFERENCES

1. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71:209-249.
2. Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. *Transl Lung Cancer Res*. 2016;5:288-300.
3. Wainer Z, Wright GM, Gough K, et al. Sex-dependent staging in non-small-cell lung cancer; analysis of the effect of sex differences in the eighth edition of the tumor, node, metastases staging system. *Clin Lung Cancer*. 2018;19:e933-e944.
4. Chapman AM, Sun KY, Ruestow P, Cowan DM, Madl AK. Lung cancer mutation profile of EGFR, ALK, and KRAS: meta-analysis and comparison of never and ever smokers. *Lung Cancer*. 2016;102:122-134.
5. Devarakonda S, Rotolo F, Tsao MS, et al. Tumor mutation burden as a biomarker in resected non-small-cell lung cancer. *J Clin Oncol*. 2018;36:2995-3006.
6. Tu HY, Ke EE, Yang JJ, et al. A comprehensive review of uncommon EGFR mutations in patients with non-small cell lung cancer. *Lung Cancer*. 2017;114:96-102.

7. Du X, Shao Y, Qin HF, Tai YH, Gao HJ. ALK-rearrangement in non-small-cell lung cancer (NSCLC). *Thorac Cancer*. 2018;9:423-430.
8. Ferrer I, Zugazagoitia J, Herbertz S, John W, Paz-Ares L, Schmid-Bindert G. KRAS-mutant non-small cell lung cancer: from biology to therapy. *Lung Cancer*. 2018;124:53-64.
9. Gu J, Zhou Y, Huang L, et al. TP53 mutation is associated with a poor clinical outcome for non-small cell lung cancer: evidence from a meta-analysis. *Mol Clin Oncol*. 2016;5:705-713.
10. Li HY, McSharry M, Bullock B, et al. The tumor microenvironment regulates sensitivity of murine lung tumors to PD-1/PD-L1 antibody blockade. *Cancer Immunol Res*. 2017;5:767-777.
11. Lin X, Wang L, Xie X, et al. Prognostic biomarker TP53 mutations for immune checkpoint blockade therapy and its association with tumor microenvironment of lung adenocarcinoma. *Front Mol Biosci*. 2020;7:602328.
12. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102:15545-15550.
13. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47.
14. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw*. 2010;33:1-22.
15. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol*. 2018;1711:243-259.
16. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics*. 2012;16:284-287.
17. Li Z, Qi F, Li F. Establishment of a gene signature to predict prognosis for patients with lung adenocarcinoma. *Int J Mol Sci*. 2020;21:8479.
18. Palumbo A Jr, Da Costa NO, Bonamino MH, Pinto LF, Nasciutti LE. Genetic instability in the tumor microenvironment: a new look at an old neighbor. *Mol Cancer*. 2015;14:145.
19. Martinez-Outschoorn UE, Balliet RM, Rivadeneira DB, et al. Oxidative stress in cancer associated fibroblasts drives tumor-stroma co-evolution: a new paradigm for understanding tumor metabolism, the field effect and genomic instability in cancer cells. *Cell Cycle*. 2010;9:3256-3276.
20. Moskovits N, Kalinkovich A, Bar J, Lapidot T, Oren M. p53 attenuates cancer cell migration and invasion through repression of SDF-1/CXCL12 expression in stromal fibroblasts. *Cancer Res*. 2006;66:10671-10676.
21. Toullec A, Gerald D, Despouy G, et al. Oxidative stress promotes myofibroblast differentiation and tumour spreading. *EMBO mol Med*. 2010;2:211-230.
22. Guo G, Marrero L, Rodriguez P, Del Valle L, Ochoa A, Cui Y. Trp53 inactivation in the tumor microenvironment promotes tumor progression by expanding the immunosuppressive lymphoid-like stromal network. *Cancer Res*. 2013;73:1668-1675.
23. Tu H, Wu M, Huang W, Wang L. Screening of potential biomarkers and their predictive value in early stage non-small cell lung cancer: a bioinformatics analysis. *Transl Lung Cancer Res*. 2019;8:797-807.
24. Zeng S, Yu X, Ma C, et al. Transcriptome sequencing identifies ANLN as a promising prognostic biomarker in bladder urothelial carcinoma. *Sci Rep*. 2017;7:3151.
25. Xu J, Zheng H, Yuan S, et al. Overexpression of ANLN in lung adenocarcinoma is associated with metastasis. *Thorac Cancer*. 2019;10:1702-1709.
26. Suzuki C, Daigo Y, Ishikawa N, et al. ANLN plays a critical role in human lung carcinogenesis through the activation of RHOA and by involvement in the phosphoinositide 3-kinase/AKT pathway. *Cancer Res*. 2005;65:11314-11325.
27. Ding K, Li W, Zou Z, Zou X, Wang C. CCNB1 is a prognostic biomarker for ER+ breast cancer. *Med Hypotheses*. 2014;83:359-364.
28. Fang Y, Yu H, Liang X, Xu J, Cai X. Chk1-induced CCNB1 overexpression promotes cell proliferation and tumor growth in human colorectal cancer. *Cancer Biol Ther*. 2014;15:1268-1279.
29. Chen EB, Qin X, Peng K, et al. HnRNPR-CCNB1/CENPF axis contributes to gastric cancer proliferation and metastasis. *Aging (Albany NY)*. 2019;11:7473-7491.
30. Kettunen E, Anttila S, Seppänen JK, et al. Differentially expressed genes in nonsmall cell lung cancer: expression profiling of cancer-related genes in squamous cell lung cancer. *Cancer Genet Cytogenet*. 2004;149:98-106.
31. Schneider MA, Christopoulos P, Muley T, et al. AURKA, DLGAP5, TPX2, KIF11 and CKAP5: five specific mitosis-associated genes correlate with poor prognosis for non-small cell lung cancer patients. *Int J Oncol*. 2017;50:365-372.
32. Shi YX, Yin JY, Shen Y, Zhang W, Zhou HH, Liu ZQ. Genome-scale analysis identifies NEK2, DLGAP5 and ECT2 as promising diagnostic and prognostic biomarkers in human lung cancer. *Sci Rep*. 2017;7:8072.
33. Richtmann S, Wilkens D, Warth A, et al. FAM83A and FAM83B as prognostic biomarkers and potential new therapeutic targets in NSCLC. *Cancers (Basel)*. 2019;11:652.
34. Chen Z, Hu Z, Sui Q, et al. LncRNA FAM83A-AS1 facilitates tumor proliferation and the migration via the HIF-1α/ glycolysis axis in lung adenocarcinoma. *Int J Biol Sci*. 2022;18:522-535.
35. Zheng YW, Li ZH, Lei L, et al. FAM83A promotes lung cancer progression by regulating the Wnt and hippo signaling pathways and indicates poor prognosis. *Front Oncol*. 2020;10:180.
36. Aasen T, Sansano I, Montero M, et al. Insight into the role and regulation of gap junction genes in lung cancer and identification of nuclear Cx43 as a putative biomarker of poor prognosis. *Cancers (Basel)*. 2019;11:320.
37. Han SS, Kim WJ, Hong Y, et al. RNA sequencing identifies novel markers of non-small cell lung cancer. *Lung Cancer*. 2014;84:229-235.
38. Tang Y, Zhang YJ, Wu ZH. High GJB2 mRNA expression and its prognostic significance in lung adenocarcinoma: a study based on the TCGA database. *Medicine (Baltimore)*. 2020;99:e19054.
39. Pocha K, Mock A, Rapp C, et al. Surfactant expression defines an inflamed subtype of lung adenocarcinoma brain metastases that correlates with prolonged survival. *Clin Cancer Res*. 2020;26:2231-2243.
40. Taguchi A, Hanash S, Rundle A, et al. Circulating pro-surfactant protein B as a risk biomarker for lung cancer. *Cancer Epidemiol Biomarkers Prev*. 2013;22:1756-1761.
41. Sin DD, Tammemagi CM, Lam S, et al. Pro-surfactant protein B as a biomarker for lung cancer prediction. *J Clin Oncol*. 2013;31:4536-4543.
42. Min KW, Kim DH, Son BK, et al. High SLC2A1 expression associated with suppressing CD8 T cells and B cells promoted cancer survival in gastric cancer. *PLoS One*. 2021;16:e0245075.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Wu G, Wang Y, Wan Y. Establishing an 8-gene immune prognostic model based on TP53 status for lung adenocarcinoma. *J Clin Lab Anal*. 2022;36:e24538. doi: 10.1002/jcla.24538