Revealing prognostic insights of programmed cell death (PCD)-associated genes in advanced non-small cell lung cancer

Weiwei Dong[1,*], He Zhang[2,*], Li Han[3,*], Huixia Zhao[2], Yue Zhang[3], Siyao Liu[3], Jiali Zhang[3], Beifang Niu[4,5], Wenhua Xiao[1]

[1]Senior Department of Oncology, The Fifth Medical Center of PLA General Hospital, Beijing 100071, P.R. China
[2]Department of Oncology, The Forth Medical Center of PLA General Hospital, Beijing 100048, P.R. China
[3]Beijing ChosenMed Clinical Laboratory Co. Ltd., Beijing 100176, P.R. China
[4]Computer Network Information Center, Chinese Academy of Sciences, Beijing 100083, P.R. China
[5]University of the Chinese Academy of Sciences, Beijing 100049, P.R. China
*Equal contribution and share the first authorship

Correspondence to: Beifang Niu, Wenhua Xiao; email: niubf@cnic.cn; w_hxiao@hotmail.com, https://orcid.org/0000-0002-1193-9342

ABSTRACT

The management of patients with advanced non-small cell lung cancer (NSCLC) presents significant challenges due to cancer cells' intricate and heterogeneous nature. Programmed cell death (PCD) pathways are crucial in diverse biological processes. Nevertheless, the prognostic significance of cell death in NSCLC remains incompletely understood. Our study aims to investigate the prognostic importance of PCD genes and their ability to precisely stratify and evaluate the survival outcomes of patients with advanced NSCLC. We employed Weighted Gene Co-expression Network Analysis (WGCNA), Least Absolute Shrinkage and Selection Operator (LASSO), univariate and multivariate Cox regression analyses for prognostic gene screening. Ultimately, we identified seven PCD-related genes to establish the PCD-related risk score for the advanced NSCLC model (PRAN), effectively stratifying overall survival (OS) in patients with advanced NSCLC. Multivariate Cox regression analysis revealed that the PRAN was the independent prognostic factor than clinical baseline factors. It was positively related to specific metabolic pathways, including hexosamine biosynthesis pathways, which play crucial roles in reprogramming cancer cell metabolism. Furthermore, drug prediction for different PRAN risk groups identified several sensitive drugs explicitly targeting the cell death pathway. Molecular docking analysis suggested the potential therapeutic efficacy of navitoclax in NSCLC, as it demonstrated strong binding with the amino acid residues of C-C motif chemokine ligand 14 (CCL14), carboxypeptidase A3 (CPA3), and C-X3-C motif chemokine receptor 1 (CX3CR1) proteins. The PRAN provides a robust personalized treatment and survival assessment tool in advanced NSCLC patients. Furthermore, identifying sensitive drugs for distinct PRAN risk groups holds promise for advancing targeted therapies in NSCLC.

INTRODUCTION

According to the 8th edition TNM staging proposed by the International Association for the Study of Lung Cancer (IASLC), lung cancer patients with stage IIIA or more advanced stages have been reported to exhibit a 5-year survival rate of less than 36% [1]. Recent global cancer data released in 2020 revealed that lung cancer remains the leading cause of cancer-related mortality [2]. This dismal outcome in lung cancers is due, in part

to the fact that more than half of the patients, about 55%, presented with metastatic lung cancer at the time of diagnosis [3]. Among lung cancer cases, non-small cell lung cancer (NSCLC) accounts for approximately 85% of cases, with lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) being the most prevalent subtypes [1]. Survival outcomes are notably compromised for patients with advanced NSCLC, including those with metastatic or recurrent disease after initial definitive treatment, particularly in stages III and IV [4]. The significantly low survival rate in advanced NSCLC underscores the formidable challenges its intricate nature and heterogeneity pose.

The advent of immune checkpoint inhibitors (ICIs) has significantly transformed the treatment landscape of lung cancer, particularly for advanced metastatic NSCLC. PD-(L)1 expression, tumor mutational burden (TMB) level, and microsatellite instability/defective mismatch repair (MSI/dMMR) status have emerged as important indicators to screen patients for potential long-term benefits from ICI therapy (https://www.nccn.org/). However, the tumor immune microenvironment (TIME) presents complex and heterogeneous characteristics across tumor types, especially for advanced cancers [5]. Due to the effectiveness of ICI therapy in improving the survival benefits of advanced NSCLC patients, there is a continuous need to explore more precise molecular biomarkers that can indicate the efficacy of immunotherapy.

Recently, numerous prognosis signatures have been developed to stratify cancer patients accurately. In breast cancer, Oncotype DX and MammaPrint are two widely used prognostic and predictive biomarkers that assist in optimizing therapy decisions and evaluating prognosis. The MINDACT study demonstrated that patients classified as low risk of recurrence based on MammaPrint but high risk according to clinico-pathological criteria achieved a 94.7% 5-year distant metastasis-free survival [6]. Wu et al. developed an immune-related prognosis signature comprising 21 immune-related genes, enabling more accurate survival risk stratification for early-stage LUAD [7]. Feng et al. developed a CD8+ T cell-related risk model to predict immunotherapy outcomes and evaluate the survival of stage III LUAD patients [8]. In conclusion, mRNA expression is closely associated with cancer patients' prognosis and therapeutic efficacy.

Resisting cell death is a hallmark characteristic of cancer [9]. Programmed cell death (PCD) encompasses various pathways, including apoptosis, ferroptosis, and autophagy, and so on. These PCD pathways play pivotal roles in eliminating malignant and infected cells [10].

Epigenetic modifications can regulate these pathways, thus influencing tumor progression and treatment resistance [11]. Ferroptosis is a distinctive PCD pathway that regulates cell death by accumulating iron and lipid reactive oxygen species (ROS) within cells. It holds significant potential as a target for drug therapy. Ferroptosis inhibitors could regulate and activate P53 and its downstream genes and further induce ROS accumulation and ferroptosis [12, 13]. Ku et al. discovered that JI017 can induce cell autophagy and apoptosis by enhancing ROS levels, reducing tumor size in lung cancer. This remarkable application of cell death regulation in cancer therapy exemplifies the potential of this approach in treating malignant tumors [14]. Furthermore, the induction of autophagy has been demonstrated to effectively overcome the resistance to third-generation EGFR-TKI treatment in patients with NSCLC [15].

However, the implication of PCD pathway genes in the advanced NSCLC is still unknown. We need to elucidate the molecular significance of PCD-related genes in the tumor microenvironment and their correlation with patient survival outcomes and treatment efficacy assessment. This knowledge will enable better guidance for clinicians in precise evaluation. Furthermore, it can provide a theoretical basis for developing corresponding targeted drugs. This study also explored the prognosis value of PCD-related genes for survival and prediction ability for immunotherapy response in NSCLC. The Cancer Genome Atlas (TCGA) [16] dataset and Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) [17] database were utilized to develop and validate the performance of the model. As a result, seven genes associated with prognosis were identified. A risk model was constructed to assess the prognosis and accurately stratify advanced NSCLC patients to enhance survival outcomes.

RESULTS

Multi-omics analysis of PCD-related genes

The workflow for this study is shown in Figure 1. The PCD-related genes contained the essential regulatory genes of twelve PCD pathways collected from the studies of Zou et al. [18]. Finally, 1178 PCD-related genes were brought into the analysis (Supplementary Table 1). By comparing the expression profiles of 1178 genes associated with PCD between advanced NSCLC and normal tissues from the TCGA-Advanced dataset, we have successfully identified 278 differentially expressed genes (DEGs) (P adjust. < 0.05 and | log2FC | > 1). In the TCGA-Advanced NSCLC cohort, 105 genes were upregulated, and 173 genes were downregulated compared to normal tissues (Figure 2A).

The altered expression patterns of PCD-related genes in advanced NSCLC provide valuable insights. To assess the prognostic significance of these DEGs, we analyzed their correlation with the survival status of advanced NSCLC patients. Survival analysis revealed that 12 genes among all DEGs significantly influenced patient prognosis, including adaptor related protein complex 1 subunit sigma 1 (*AP1S1*), bruton tyrosine kinase (*BTK*), cholecystokinin (*CCK*), cyclin dependent kinase inhibitor 2A *(CDKN2A)*, C-X3-C motif chemokine ligand 1 (*CX3CL1*), *CX3CR1*, interleukin 20 receptor subunit alpha (*IL20RA*), NLR family pyrin domain containing 7 (*NLRP7*), platelet factor 4 (*PF4*), rhotekin 2 (*RTKN2*), tribbles pseudokinase 3 (*TRIB3*),
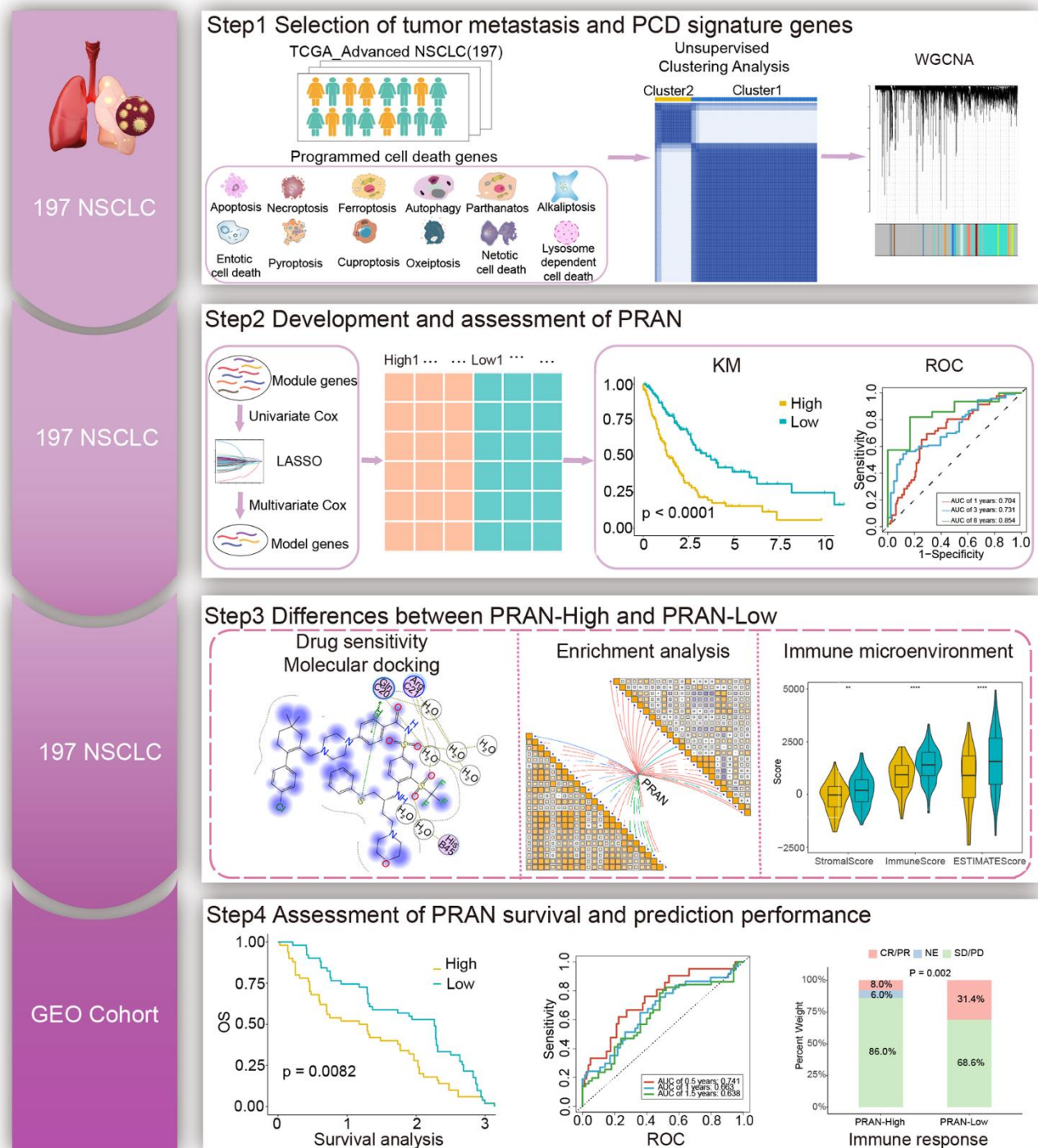

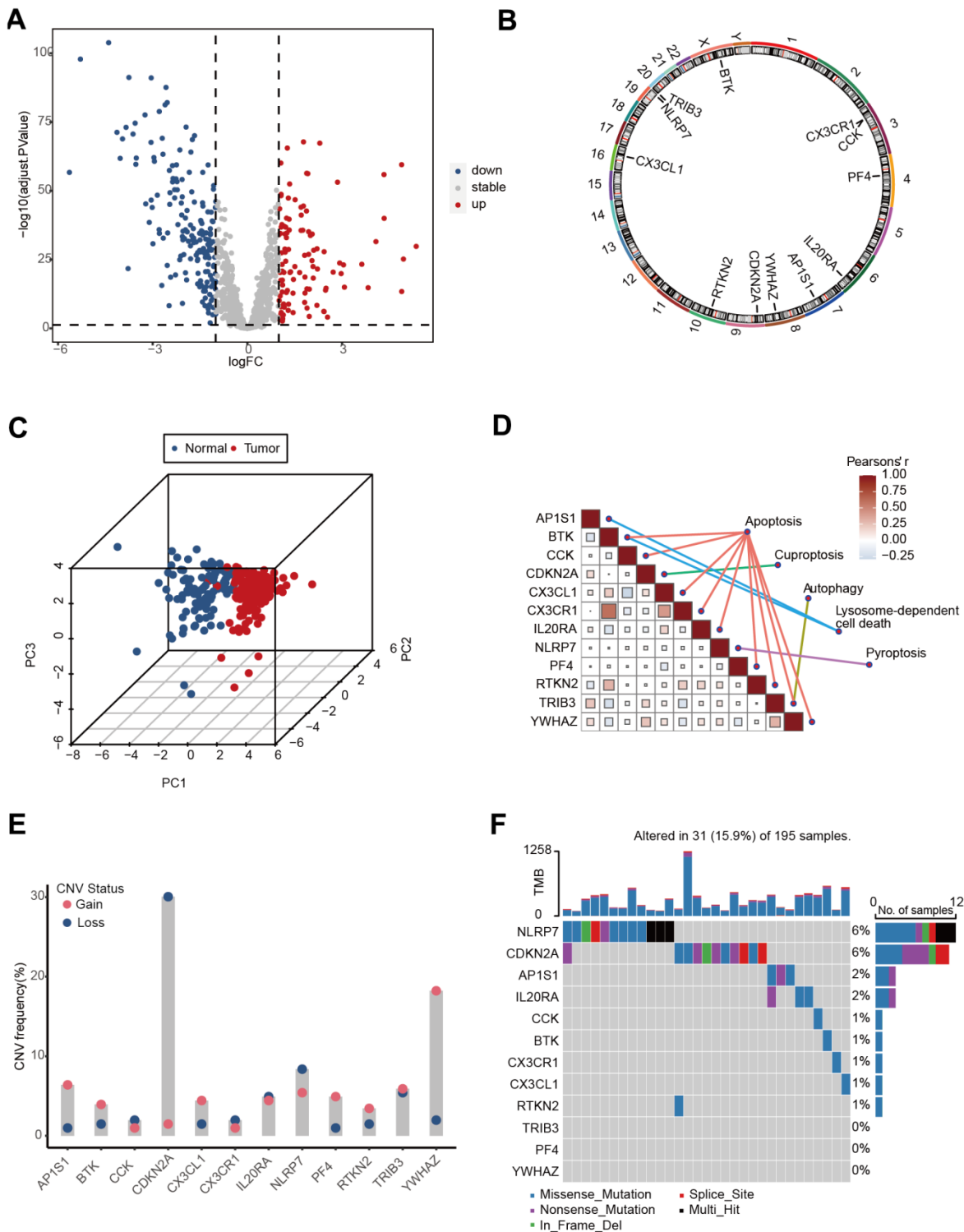
Figure 1. The workflow of the study.

Figure 2. The landscape of genetic and transcriptional alterations of PCD genes in TCGA-Advanced NSCLC. (**A**) Volcano plot depicting the differential expression of PCD genes between tumor and normal samples. (**B**) Circos plot illustrating the chromosomal distribution of 12 prognosis related differential PCD genes. Each outer circle represents a chromosome, and the connecting lines display the genomic location of the PCD genes. (**C**) Principal Component Analysis (PCA) plot of 12 PCD-related genes. (**D**) Correlation analysis heatmap of 12 PCD-related genes in the TCGA-Advanced NSCLC dataset. The color scale represents the correlation coefficients, with red indicating positive correlation and blue indicating negative correlation. (**E**) Copy Number Variation (CNV) frequencies of 12 PCD-related genes. (**F**) Mutation frequencies of 12 PCD-associated differential genes in the TCGA-Advanced NSCLC cohort. The column height represents the frequency of mutations, and different types of mutations are distinguished by color.

and tyrosine 3-monooxygenase/tryptophan 5-mono-oxygenase activation protein zeta (*YWHAZ*) (Supplementary Figure 1). The chromosomal positions of these 12 PCD genes are depicted in Figure 2B. Principal component analysis (PCA) revealed that the expression of the 12 PCD genes could effectively discriminate advanced NSCLC samples from normal samples (Figure 2C). These 12 PCD genes participated in various PCD processes, including apoptosis, cuproptosis, autophagy, pyroptosis, and lysosome-dependent cell death, and exhibited close interactions (Figure 2D).

Furthermore, copy-number alterations (CNA) analysis revealed frequent copy number variations changes in these 12 PCD genes, with a predominant occurrence of *YWHAZ* amplification and *CDKN2A* deletion (Figure 2E). Further exploration of somatic mutation of 12 PCD genes exhibited a 21% mutation rate in advanced NSCLC. Notably, *NLRP7* and *CDKN2A* showed the highest mutation frequency, accounting for 6% of the observed mutations (Figure 2F).

These findings indicated the potential utility of 12 PCD genes as prognostic markers in advanced NSCLC.

Identification and enrichment analysis of PCD-related subtypes

The TCGA-Advanced NSCLC cohort underwent clustering analysis based on 12 prognostic-related PCD gene expressions using the non-negative matrix factorization (NMF) algorithm. The research identified two distinct clusters (Figure 3A, 3B), with patients in cluster 1 showing significantly worse survival prognosis compared to cluster 2 (P=0.037, Figure 3C). The clinical statistical analysis results between clinical factors and two clusters are shown in Supplementary Table 2. These findings suggest that the expression patterns of these PCD genes can effectively stratify advanced NSCLC patients based on their survival outcomes. A total of 359 DEGs were identified between cluster 1 and cluster 2 (Supplementary Table 3). these DEGs were used for GO enrichment analysis, and numerous metabolic and immune-related biological processes were significantly enriched (Figure 3D). These processes included the icosanoid metabolic process, metabolic hormone process, tertiary alcohol metabolic process, antigen processing and presentation, and leukocyte-mediated cytotoxicity. The KEGG pathway enrichment analysis also demonstrated significant enrichment in 17 pathways (Figure 3E), including cytokine-cytokine receptor interaction, tight junction, arachidonic acid metabolism, and steroid hormone biosynthesis.

Moreover, substantial differences were observed in several hallmark pathways between the two clusters (Figure 3F). These included G2M check-point, MYC targets, oxidative phosphorylation, DNA repair, apoptosis, NOTCH signaling, and KRAS signaling. These findings shed light on the significantly predictive value of PCD genes in advanced NSCLC.

Development of the prognosis risk model for advanced NSCLC based on the PCD-related subtypes

According to WGCNA analysis, the gene co-expression network was constructed to find the critical modules associated with metastatic status and PCD cluster-related subtypes. A total of 39 modules were identified, and the "pink" (PCD cluster: R=0.26, p=2e-04, M stage: R=0.23, p=0.001), "tan" (PCD cluster: R=0.23, p=0.001, M stage: R=0.17, p=0.02), and "brown" (PCD cluster: R=0.25, p=5e-04, M stage: R=0.114, p=0.05) modules were significantly associated with PCD clusters and metastatic status (Figure 4A and Supplementary Figure 2A–2C). Next, 55 genes were initially selected through univariate Cox analysis from 1460 module-related genes (Supplementary Table 4). Subsequently, the LASSO method identified 15 genes (Supplementary Table 4 and Supplementary Figure 3A, 3B). Finally, seven genes (*CCL14*, *CPA3*, *CX3CR1*, IKAROS family zinc finger 3 (*IKZF3*), kinesin family member 21B (*KIF21B*), long intergenic non-protein coding RNA 528 (*LINC00528*), and solute carrier family 16 member 4 (*SLC16A4*) were identified based on the multivariate Cox regression analysis of all different combinations of the above 15 genes (Supplementary Table 4). The multivariate Cox regression analysis demonstrated the significant predictive value of the identified seven genes (Figure 4B). The PCD-related risk score for the Advanced NSCLC (PRAN) model was developed using RNA expression values and multivariate Cox regression coefficients of seven PCD-related genes. The formula was as follows: PRANScore=−0.32×*CCL14*-0.06×*CPA3*-0.22×*CX3CR1*-0.23×*IKZF3*+0.47×*KIF21B*-0.54×*LINC00528*+0.13×*SLC16A4*. Based on the PRAN risk model, advanced NSCLC patients were classified into two groups. The PRAN-Low group exhibited significantly better overall survival (OS) compared to the PRAN-High group (P<0.0001, Figure 4C). The area under the ROC curve (AUC) for predicting 1-, 3-, and 8-year survival using the PRAN risk model were 0.704, 0.731, and 0.854, respectively (Figure 4D). In addition, seven model genes' expression significantly differed between PRAN-High and PRAN-Low groups (Figure 4E). These findings highlight the potential of the PRAN model as a valuable prognostic tool for predicting survival outcomes in advanced NSCLC patients.
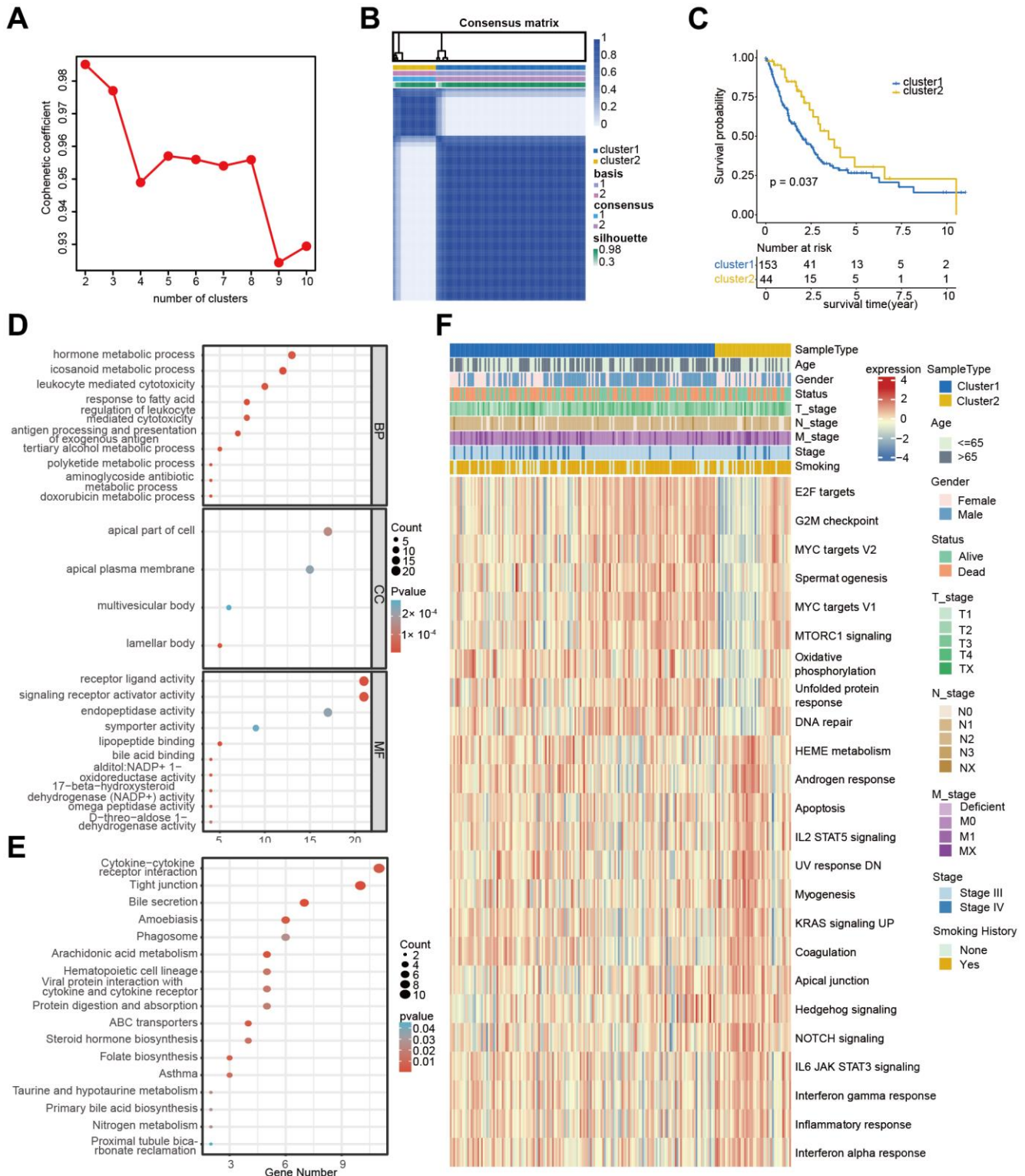
Figure 3. Cluster analysis of 12 PCD-related genes in TCGA-Advanced NSCLC dataset. (**A**) Non-negative Matrix Factorization (NMF) clustering of twelve PCD-related genes. The correlation coefficients at k = 2-10 are presented in the Figure. (**B**) Consistency plot illustrating the stability of NMF clustering results. (**C**) Kaplan-Meier (KM) survival curves of patients in PCD-related clusters. (**D**) Gene Ontology (GO) analysis of differential genes between cluster 1 and cluster 2. (**E**) Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of differential genes between cluster 1 and cluster 2. (**F**) Heatmap displaying the hallmark pathways in different PCD clusters.
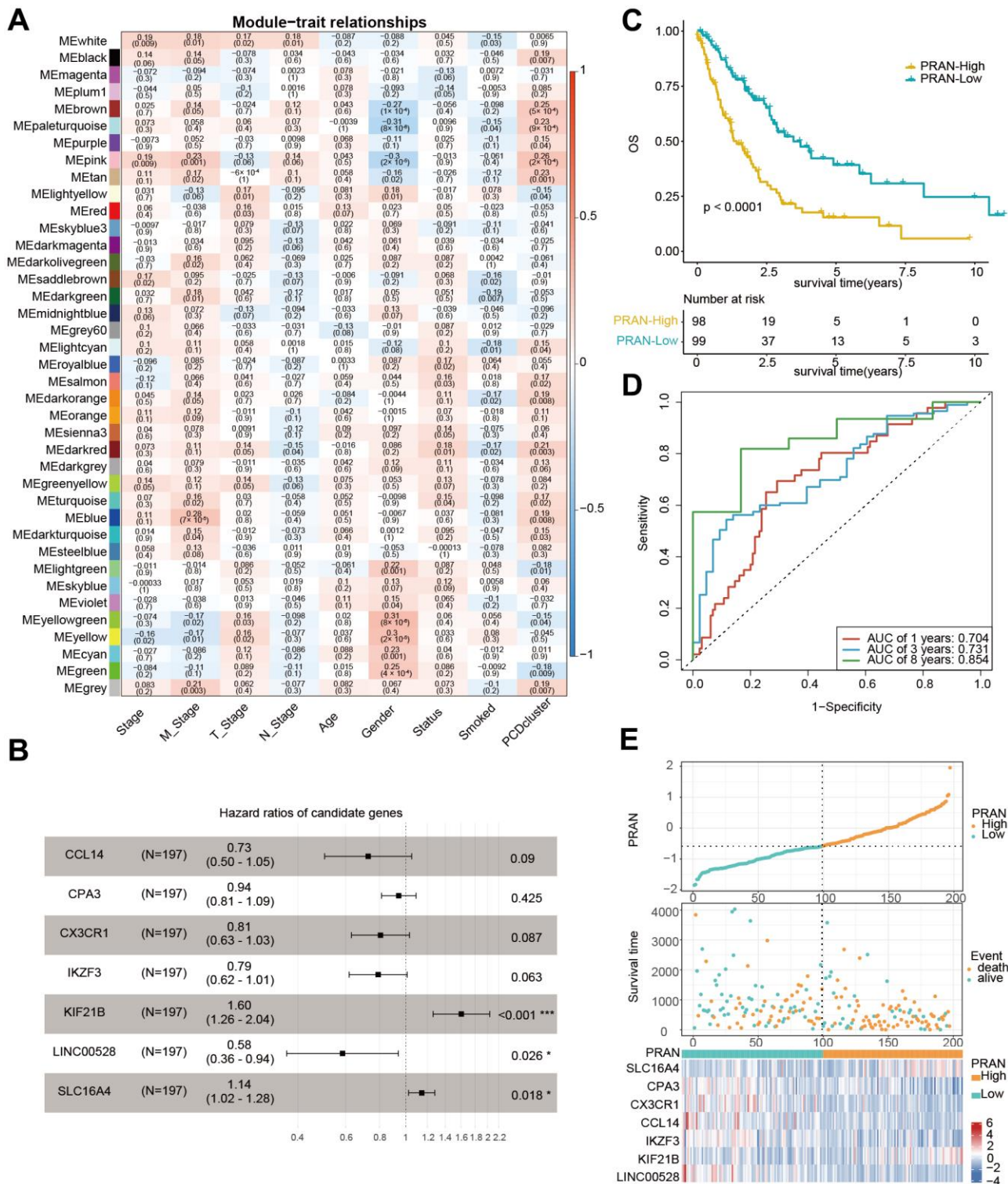
Figure 4. Development of prognostic signature using PCD-related genes in TCGA-Advanced NSCLC dataset. (**A**) Heatmap of correlation between gene modules and clinical traits, each cell containing Pearson's correlation coefficient and p-value. (**B**) Forest plot depicting the associations between the expression levels of seven PCD genes and overall survival (OS) in the training cohort. Hazard Ratio (HR), 95% Confidence Interval (CI), and p-value were determined by multivariate Cox regression analysis. (**C**) Kaplan-Meier (KM) curve analysis of the prognostic model in the training set, showing the survival differences between high-risk and low-risk groups. (**D**) Time-dependent receiver operating characteristic (ROC) curves and area under the curve (AUC) values of the PRAN model for predicting survival status in 1-, 3-, and 8-year. (**E**) Comparison of PRAN scores, survival status, and expression of seven PCD genes between PRAN-High and PRAN-Low groups. P-value: * < 0.05; *** < 0.001.

The correlation analysis between the PCD subtype and the PRAN risk model revealed that the majority of patients in cluster 1 were classified into the PRAN-High group, and the PRAN score exhibited a significantly higher value in cluster 1 compared to cluster 2, indicating an increased risk in advanced NSCLC (P=0.00047, Figure 5A, 5B). Upon further investigating the correlation between the previously identified 12 PCR-related genes and the PRAN risk model, we observed significant expression differences between the PRAN-High and PRAN-Low groups (Figure 5C). Through comprehensive univariate (Figure 5D) and multivariate Cox (Figure 5E) regression analyses, we systematically investigated the prognostic significance of both clinical characteristics (stage, age, and gender) and the PRAN risk model in advanced NSCLC. Univariate Cox regression indicated PRAN was a risk factor for the overall survival in advanced NSCLC (P=2.5e-11, HR=2.7, 95% CI =2-3.6). Meanwhile, Multivariate Cox regression analysis further confirmed the prognosis significance of PRAN in advanced NSCLC (P=2.5e-12, HR=3.1, 95% CI=2.3-4.3). Our findings demonstrated that the model exhibited remarkable stability and superior predictive value compared with clinical factors.

Somatic mutation profile analysis revealed that mutations of filaggrin (*FLG*) and CUB and Sushi multiple domains 1 (*CSMD1*) frequently occurred in the PRAN-High group, while zinc finger homeobox 4 (*ZFHX4*) and xin actin binding repeat containing 2 (*XIRP2*) mutated more in the PRAN-Low group (Supplementary Figure 3C). TMB different analysis demonstrates no difference between PRAN-High and PRAN-Low groups (Supplementary Figure 3D). However, the PRAN-Low group showed higher IFN-γ response (Supplementary Figure 3E) and TCR diversity (Supplementary Figure 3F, 3G), implying robust anti-tumor activity.

These results validate the robustness and clinical utility of the model as a reliable predictive tool for advanced NSCLC.

Enrichment analysis and drug sensitivity prediction based on PRAN risk model for advanced NSCLC

To further explore the enrichment of molecular biological processes and signaling pathways in the PRAN-High and PRAN-Low groups, we conducted a GSEA analysis. The results revealed significant enrichment of pyrimidine metabolism, spliceosome, and aminoacyl tRNA biosynthesis pathway in the PRAN-High group. In contrast, cell adhesion molecules cams, cytokine-cytokine receptor interaction, T/B cell receptor signaling, cell cycle pathway, and so on, in the PRAN-Low group (Figure 6A).

Next, we analyzed the correlation of PRAN score with cancer immune cycle activity and metabolism-related pathways (Figure 6B). Most immune cycle activity was negatively related to the PRAN score, implying decreased anti-tumor immunity in the higher PRAN score population. Interestingly, diverse correlations with the PRAN score were observed concerning metabolic pathways. Notably, the Hexosamine bio-synthesis pathway demonstrated the strongest positive correlation with the PRAN score, highlighting its critical role in the PRAN-High group. Additionally, several signaling pathways, including Galactose metabolism, Glyoxylate and Dicarboxylate metabolism, Purine metabolism, Pyrimidine metabolism, Cysteine and methionine metabolism, Biotin metabolism, and Ubiquinone and other terpenoid-Quinone biosynthesis were significantly positively associated with the PRAN score. Conversely, ADP-Ribosylation, Arachidonic acid metabolism, linoleic acid metabolism, and alpha-linoleic acid metabolism exhibited significant negative correlations with the PRAN score. These findings provide valuable insights into the complex interplay between the PRAN score and metabolic pathways, contributing to a deeper understanding of the metabolic dysregulation underlying advanced NSCLC.

The results above indicate a significant correlation between the PRAN score, immune microenvironment activity, and metabolic processes in advanced NSCLC. Furthermore, we predicted the potential drug sensitivities for patients in different PRAN risk groups (Figure 6C). The complete sensitive drugs and statistical data are presented in Supplementary Table 5. The smaller IC_{50} value indicates that inhibitors function more effectively with targeted compounds [19]. Remarkably, the PRAN-Low group demonstrated notable sensitivity to sunitinib, AZD1480, regorafenib, and five PCD-targeted drugs, including necrosulfonamide (necrosis apoptosis inhibitor), Serdemetan (HDM2 inhibitor, delayed apoptosis), navitoclax: PLX-4032 (target BCL2 and BRAF, apoptosis inhibitor), MG-132 (inducing apoptosis, activating autophagy), and ML162 (inducing ferroptosis), whereas the PRAN-High group showed significant sensitivity to pevonedistat (causing cell death) and EX-527 (enhancing autophagy). These results underscore the clinical implications of the PRAN model in guiding personalized treatment strategies in advanced NSCLC patients.

Molecular docking was used to screen potential drug candidates and elucidate the molecular mechanisms involved. We employed MOE software to simulate the binding modes of small molecule drugs to seven PCD-related genes**.** As revealed by molecular docking
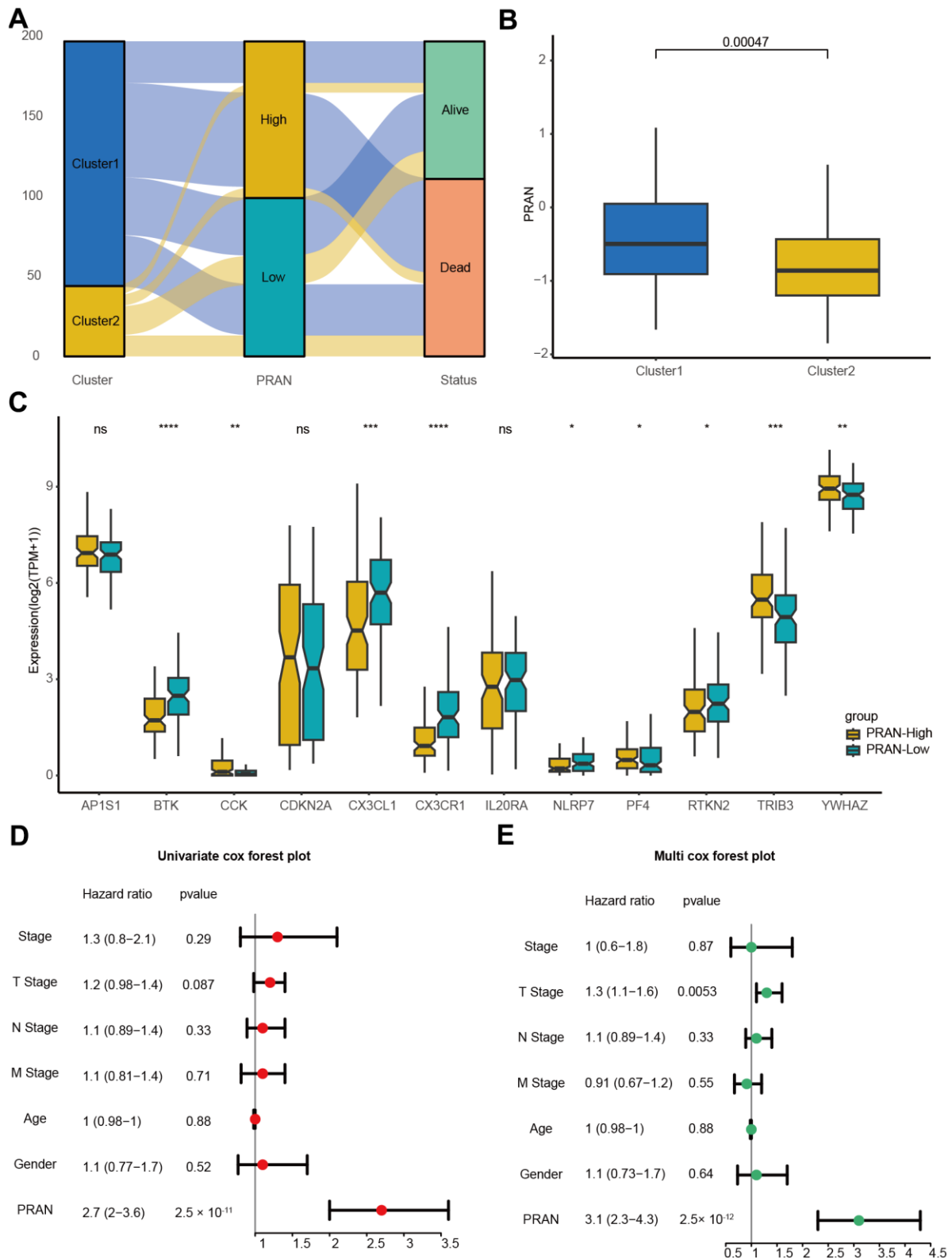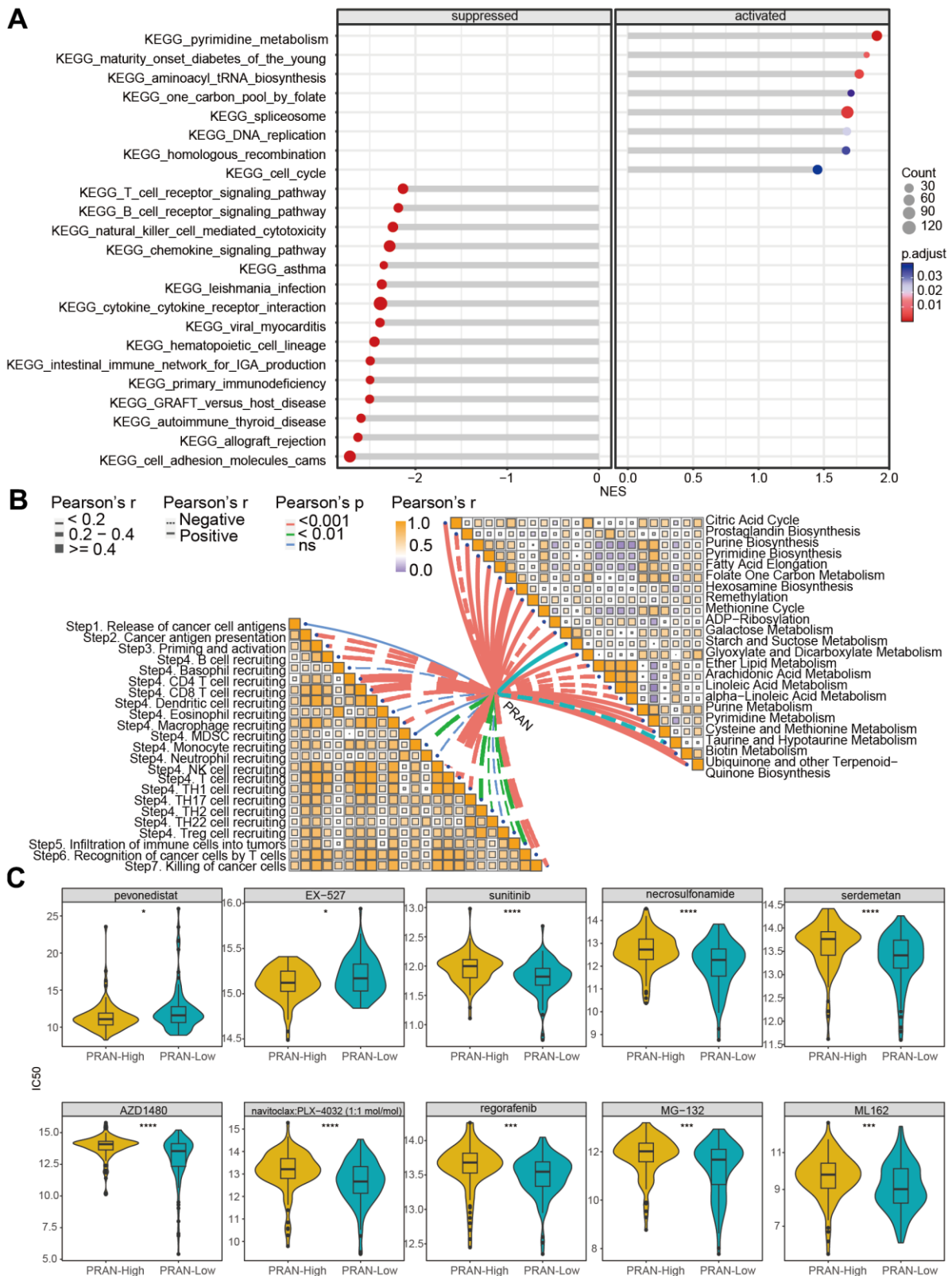
Figure 5. The correlation validation of PRAN risk model in TCGA-Advanced NSCLC dataset. (**A**) The Sankey plot illustrates the distribution of PCD risk groups, PCD clusters, and survival outcomes. (**B**) Box plots depicting the relationship between PCD clusters and PRAN risk groups. (**C**) Box plots showing the expression levels of 12 PCD-related prognosis genes between PRAN-High and PRAN-Low groups. (**D**) Univariate Cox regression analysis of PCD risk scores and clinical variables. (**E**) Multivariate Cox regression analysis of PCD risk scores and clinical variables. P-value: ns >=0.05; * < 0.05; ** < 0.01; *** < 0.001; **** < 0.0001.

Figure 6. Biological function analysis and drug-susceptibility analysis of PRAN risk model in TCGA-Advanced NSCLC dataset. (**A**) The lollipop plot displays the top 10 significantly enriched suppressed pathways and all activated pathways in the PRAN-High and PRAN-Low groups. (**B**) Pearson correlation analysis demonstrates the relationship between PCD scores and cancer immune cycle activity (left) and metabolism-related pathways (right). (**C**) Sensitivity analysis of anti-tumor drugs between PRAN-High and PRAN-Low groups. P-value: ns >=0.05; * < 0.05; ** < 0.01; *** < 0.001; **** < 0.0001.

---

analyses, CCL14, CPA3, CX3CR1, IKZF3, and KIF21B demonstrated the potential to interact with these small molecule drugs (Supplementary Table 6). Furthermore, based on these analyses, we identified the top-ranked drug for each target protein with the highest binding affinity Figure 7A–7E). The navitoclax closely bonds with the CCL14 protein by the amino acid residue of Gln-C20 (Figure 7A). The navitoclax closely bonds with the CPA3 protein by the amino acid residue of Glu-B136 (Figure 7B). The navitoclax closely bonds with CX3CR1 protein by the amino acid residue of Leu-184 (Figure 7C). The AZD1480 closely bonds with the IKZF3 protein by the amino acid residues of Arg-B418, Gly-B416, and Trp-A380 (Figure 7D). The MG-132 closely bonds with the KIF21B protein by the amino acid residue of Glu-A411 and Arg-A105 (Figure 7E).

Tumor immune microenvironment (TIME) characteristics assessment based on the PRAN risk model for advanced NSCLC

We utilized various algorithms to analyze immune cell abundance across different PRAN subgroups in the TCGA-Advanced NSCLC dataset. The MCP counter analysis demonstrated significantly higher enrichment scores of T cells, B lineage cells, monocytic lineage cells, myeloid dendritic cells, and endothelial cells in the PRAN-Low group (Figure 8A). Consistently, ssGSEA analysis revealed enhanced immune cell infiltration in the PRAN-Low group, corroborating the MCP counter findings (Figure 8B). Furthermore, CIBERSORT algorithm results confirmed a higher infiltration content of CD8 T cells and macrophages M1 cells in the PRAN-Low group, further supporting the robust anti-tumor immune activity associated with the PRAN-Low group. Interestingly, the PRAN-High group exhibited significantly higher infiltration content of macrophages M0, T cell follicular helper cells, and activated dendritic cells, suggesting that the survival advantage conferred by these three immune cell populations may be offset by other factors within the TIME (Figure 8C). The PRAN-High group exhibited significantly lower ESTIMATE scores, as well as lower stromal scores and immune scores, indicating higher tumor purity and increased infiltration of tumor cells (Supplementary Figure 4A). The expression of HLA (Supplementary Figure 4B) and immune inhibitors genes (Supplementary Figure 4C) were higher in the PRAN-Low group compared with the PRAN-High group. At the same time, the expression of immunostimulatory genes was also significantly upregulated against the antigen presentation and immune response (Figure 8D). These findings provide comprehensive insights into the intricate interplay between TIME characteristics and the PRAN model, underscoring the complex dynamics of the tumor immune microenvironment in advanced NSCLC.

Validation of the PRAN risk model for advanced NSCLC

The extensive analyses performed in this study provide compelling evidence supporting the PRAN risk model's ability to accurately predict survival prognosis and TIME characteristics of patients with advanced NSCLC. We conducted rigorous validation across multiple independent validation cohorts to further validate the predicted performance. Survival analysis conducted in two independent NSCLC validation cohorts, namely GSE61676 (P=0.034, Figure 9A) and GSE13213 (P=0.022, Figure 9B), revealed a significant difference in overall survival (OS) between the PRAN-High and PRAN-Low groups. The respective optimal AUC values for predicting performance were 0.648 (Figure 9D) and 0.672 (Figure 9E). These results further validate the predictive utility of the PRAN model in accurately stratifying patients with advanced NSCLC.

What's more, consistent with the analysis results of the advanced NSCLC dataset, the patients with early NSCLC from the GSE74777 and GSE50081cohort and those in the PRAN-High showed worse prognoses than those in the PRAN-Low (Figure 9G and Supplementary Figure 5A). In addition, the predicted survival ROC curve confirmed the precise predictive capacity of the risk model, with area under the ROC curve (AUC) values of 0.608, 0.707, 0.729 and 0.585, 0.599, 0.603 for 1-, 2-, and 3-year survival, respectively (Figure 9H and Supplementary Figure 5B).

We sought to assess the PRAN risk model's predictive capacity in determining patients' immunotherapy responses. Survival analyses were performed in two immunotherapy cohorts of lung cancer (GSE135222 and GSE93157) and one melanoma (GSE91061). Due to the limited sample size in the GSE135222 (Supplementary Figure 5C) and GSE93157 (Supplementary Figure 5D) cohorts, we did not observe a significant difference in PFS between the two PRAN subgroups. However, we observed a slightly better prognosis in the PRAN-Low group than in the PRAN-High group. In contrast, in the independent melanoma immunotherapy cohort GSE91061, the PRAN-Low group demonstrated a significantly superior OS benefit compared to the PRAN-High group (P=0.0082, Figure 9C), with an optimal AUC of 0.741 (Figure 9F) for predicting performance. Notably, the PRAN-High group exhibited a significantly lower immunotherapy response rate than the PRAN-Low group in the GSE91061 NSCLC cohort (P=0.002, Figure 9I). These findings further reinforced the value of PRAN model as a predictive tool in guiding treatment decisions for patients with advanced NSCLC.
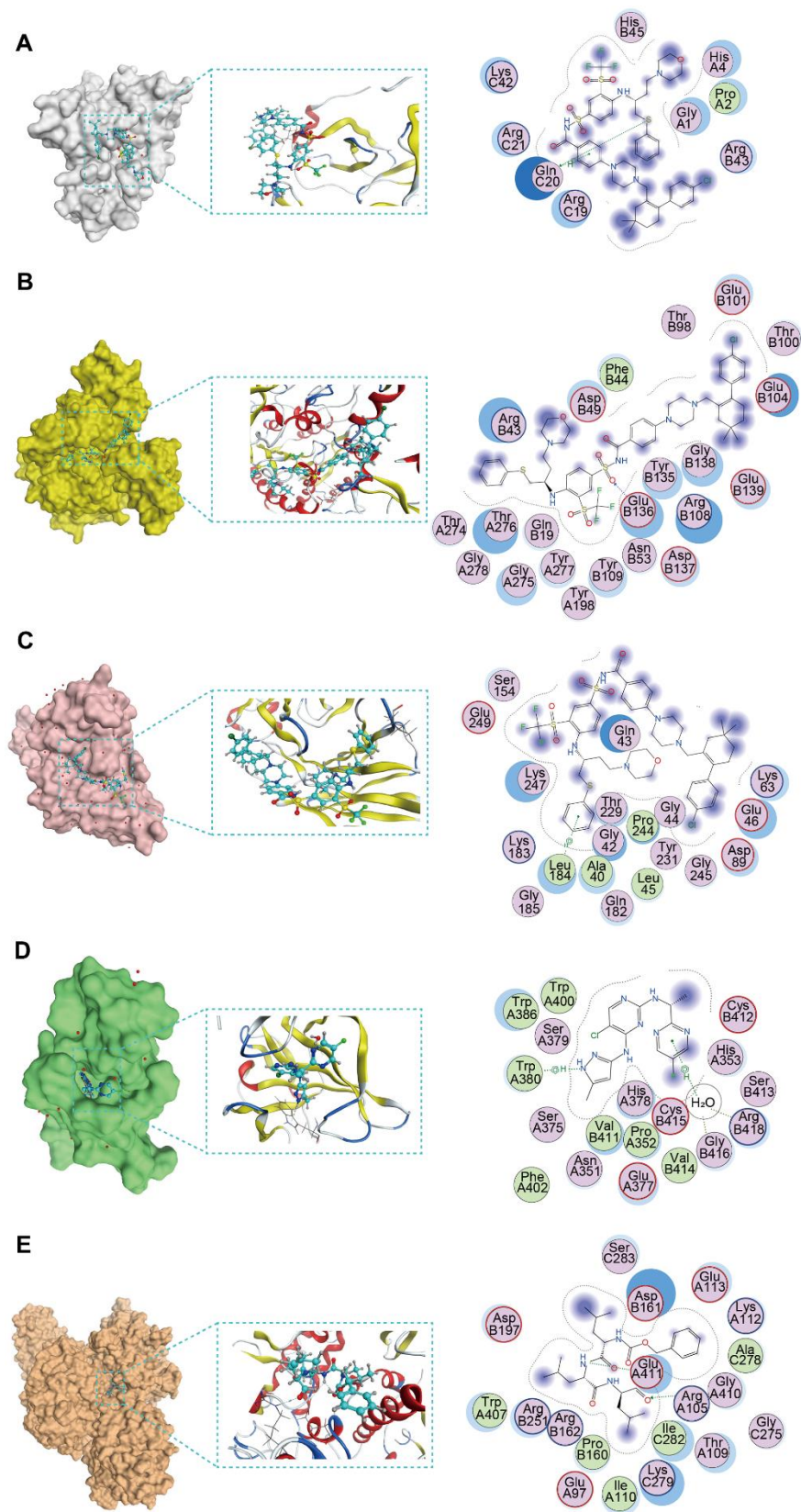
Figure 7. The molecular docking posture predicting for the sensitive anti-tumor drugs and targeted PCD genes. (**A**) Docking position of *CCL14* active pocket with navitoclax. (**B**) Docking position of *CPA3* active pocket with navitoclax. (**C**) Docking position of *CX3CR1* with navitoclax. (**D**) Docking position of *IKZF3* with AZD1480. (**E**) Docking position of *KIF21B* with MG-132.

Establishment of a nomogram based on the PRAN risk model and clinical factors for advanced NSCLC

Finally, for better application in clinical practice, we established a nomogram based on the multivariate Cox regression analysis of the PRAN score, and age, stage, smoking status, and gender were included (Figure 10A). Survival analysis demonstrated a significant difference in OS between the two subgroups based on the nomogram score, with the high-risk group exhibiting worse OS outcomes (P<0.0001, Figure 10B). The nomogram showed an optimal AUC value of 0.843, indicating its predictive solid performance (Figure 10C). Furthermore, DCA results confirmed the nomogram's robustness and optimal predictive ability (Figure 10D). Calibration curves were generated to assess the accuracy of the nomogram in predicting the 1-, 3-, and 8-year survival rates, further validating its predictive accuracy (Figure 10E). The nomogram could be a practical tool for prognostic assessment and aid in clinical decision-making.

The mRNA and protein expression difference of seven PRAN-genes between tumor and normal cells in NSCLC

The mRNA expression of seven PRAN genes (*CCL14, CPA3, CX3CR1, IKZF3, KIF21B, LINC00528,* and *SLC16A4*) exhibited noticeable difference between normal and tumor in NSCLC (Figure 11A). According to the RNA expression validation in two human NSCLC cell lines, A549 and NCI-H1299, and one human lung
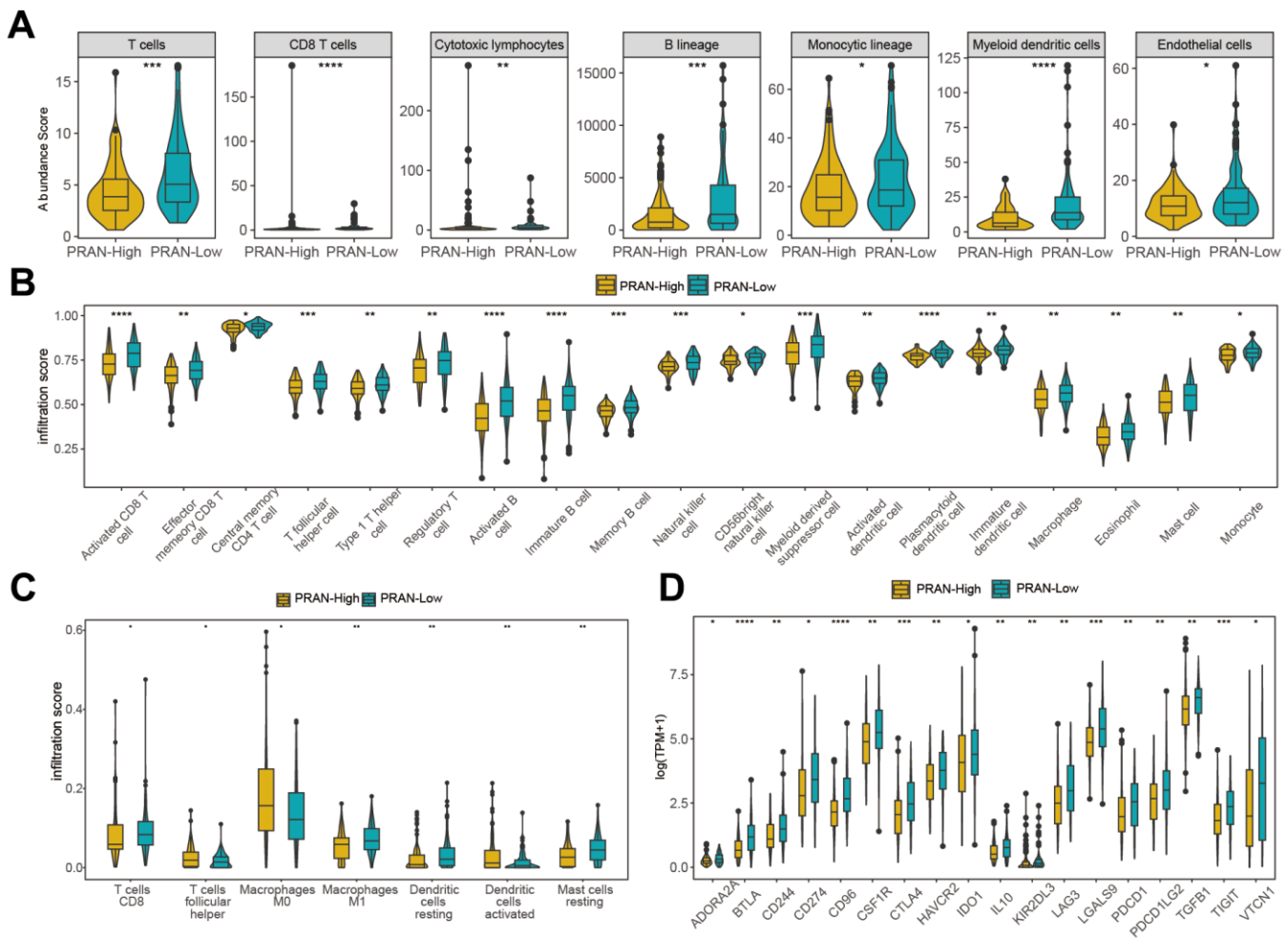

Figure 8. Immune infiltration analysis of PRAN-High and PRAN-Low groups TCGA-Advanced NSCLC dataset. (**A**) Violin plot showing the immune cell enrichment scores of PRAN-High and PRAN-Low groups, assessed using MCPcounter. (**B**) The ssGSEA algorithm shows the immune cell infiltration of immune-related functions and pathways in PRAN-High and PRAN-Low groups. (**C**) Immune cell infiltration content of PRAN-High and PRAN-Low groups was analyzed using the CIBERSORT algorithm. The analysis provides insights into the proportion of different immune cell types in each risk group. (**D**) Expression of immunoinhibitor genes between PRAN-High and PRAN-Low groups. P-value: * < 0.05; ** < 0.01; *** < 0.001; **** < 0.0001.

bronchial epithelial cell line (BEAS-2B), significant expression differences were observed between normal and tumor cells (Figure 11B). H1299 was a p53 deficient cell line, and the expression of *CCL14, CPA3,* *CX3CR1, KIF21B*, and *SLC16A4* was lower than in A549 and Beas-2b. The expression of *IKZF3* was higher in H1299 cells than in A549 and Beas-2b. The expression of *IKZF3, CPA3,* and *LINC00528* were
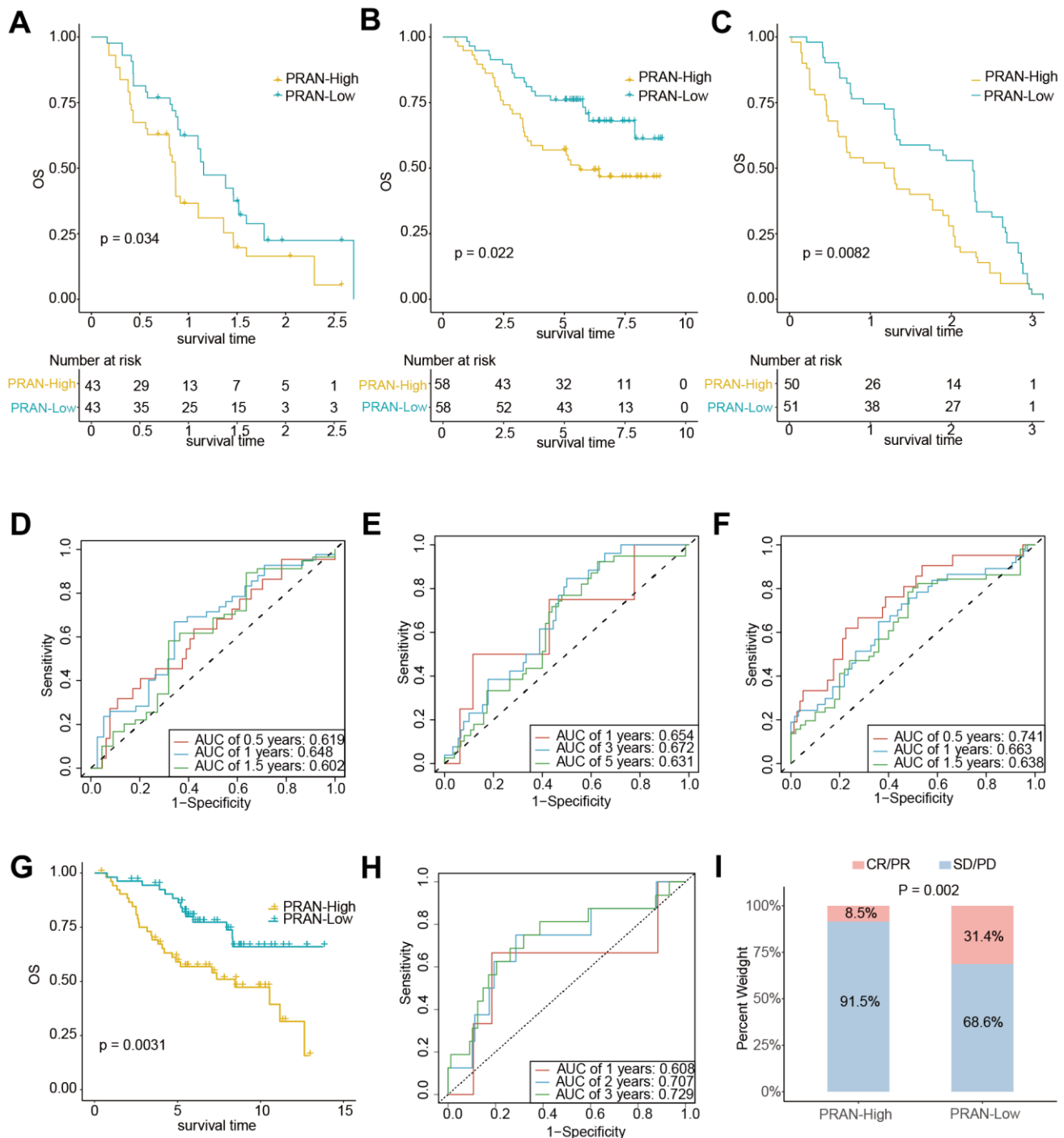

Figure 9. The predicting performance validation of the PRAN risk model in multiple GEO cohorts. (**A**–**C**) Kaplan-Meier survival analysis in NSCLC validation cohort GSE61676, GSE13213, and GSE91061. (**D**–**F**) Time-dependent ROC curves between PRAN-High and PRAN-Low groups in validation cohort GSE61676, GSE13213, and GSE377453. (**G**) Kaplan-Meier survival analysis in validation cohort GSE74777. (**H**) Time-dependent ROC curves between PRAN-High and PRAN-Low groups in validation cohort GSE74777. (**I**) The predicting performance of immunotherapeutic efficiency of PRAN risk model in the GSE91061 cohort.

down-regulated in A549 cells than in normal cells, while the expression of *CCL14, CX3CR1, KIF21B*, and *SLC16A4* were up-regulated in A549 cell than in normal cell. The protein expression results acquired from the Human Protein Atlas (HPA) database were consistent with the mRNA results, showing the significantly different between NSCLC tumor and normal tissues (Figure 11C and Supplementary Table 7).

DISCUSSION

As a common cancer with high morbidity and mortality, lung cancer still needs excellent attention [20]. Advanced NSCLC poses treatment challenges and a bleak prognosis, especially patients with metastatic NSCLC often experience low survival rates [21]. In recent years, ICIs have garnered widespread utilization

in lung cancer treatment, particularly advanced metastatic NSCLC, yielding promising outcomes [22, 23]. Hence, investigating molecular markers associated with NSCLC prognosis and immunotherapy is imperative to inform treatment decisions for advanced NSCLC patients. Multiple studies support the use of risk models to predict tumor immunotherapy responses and guide personalized medicine [24–26].

We rigorously examined the prognostic relevance of tumor programmed cell death processes in advanced NSCLC. We identified twelve DEGs, all showing significant predictive value in advanced NSCLC. According to RNA expression of these twelve genes, molecular subtyping analysis identified two distinct subtypes with divergent survival outcomes. Among them, *CDKN2A* exhibited the highest mutation
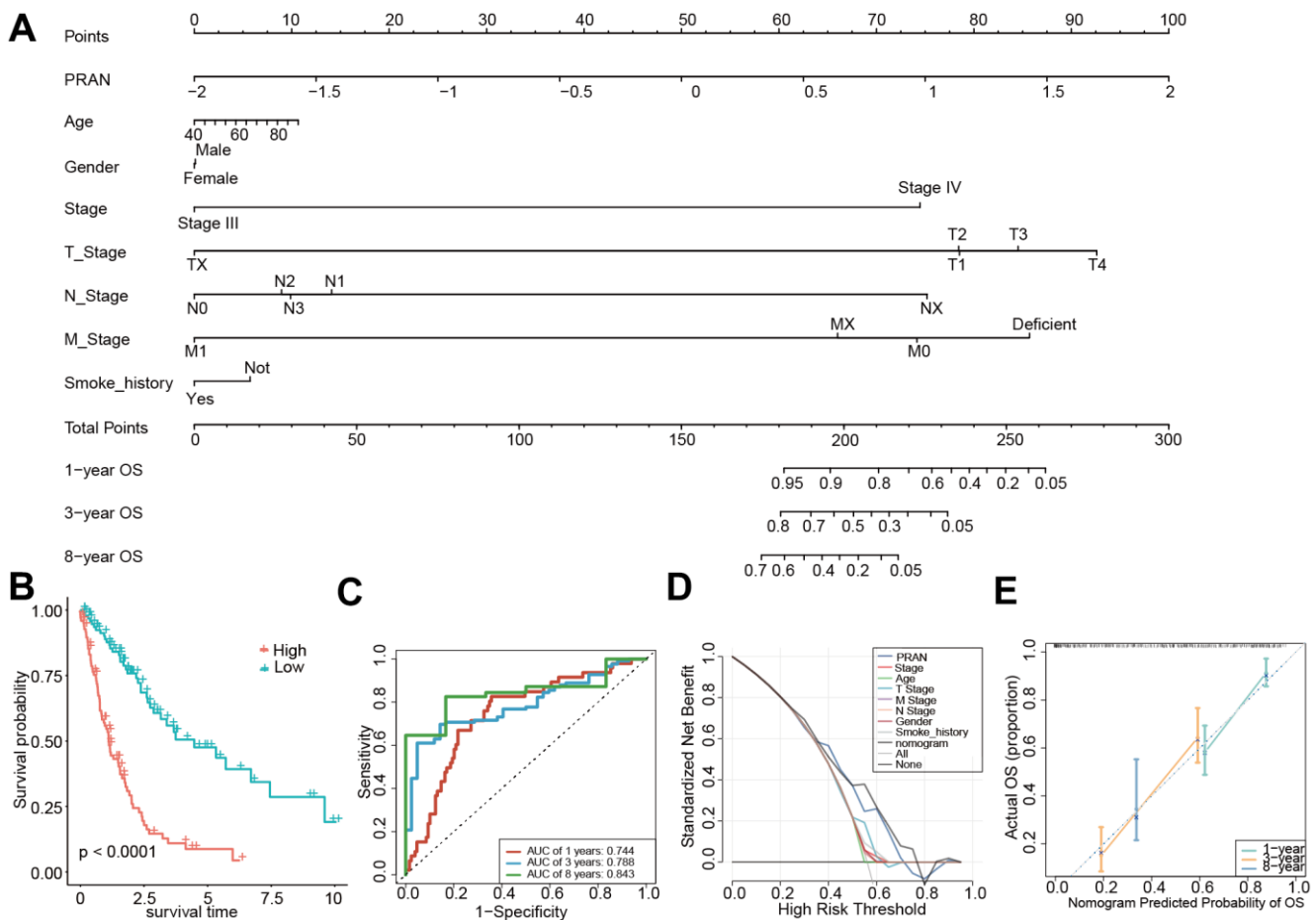

Figure 10. Construction and validation of the PRAN score-based nomogram. (**A**) The nomogram plot was constructed in the training cohort with incorporation of PRAN and clinical characteristics. (**B**) Kaplan-Meier survival curves based on PRAN scores calculated using the nomogram. (**C**) ROC curves for predicting 1-year, 3-year and 8-year OS for the nomogram. (**D**) Decision curve analysis of nomogram, PRAN risk model and clinical characteristics. The black line in this Figure indicates the assumption of no patient death. (**E**) Nomogram calibration plot based on the agreement between predicted and observed values at 1, 3, and 8 years. X-axis is nomogram predicted overall survival, y-axis is actual overall survival, dashed line is ideal performance of nomogram, and 95% confidence interval is represented by closed vertical line.

frequency and CNV deletion. *CDKN2A*, as a cell cycle regulation gene, is associated with cuproptosis processes and is shown to impact cancer patients' prognosis [27].

Through a comprehensive approach, our study aimed to gain a profound understanding of the intricate relationship between the identified genes and the specific risk associated with advanced NSCLC. The mRNA expression of seven PRAN genes (*CCL14, CPA3, CX3CR1, IKZF3, KIF21B, LINC00528*, and *SLC16A4*) showed significant differences between normal and tumor samples in the advanced NSCLC cohort. The immunohistochemistry (IHC) stained images obtained from the HPA dataset also demonstrate differential protein expression. These findings were further validated through RT-qPCR

experiments in A549 and NCI-H1299 (p53 deficient), and BEAS-2B cell lines, consistently showing significant differences in gene expression between normal and tumor cells. However, the expression of these genes in A549 and H1299 cells exhibited some inconsistency, potentially due to the underlying heterogeneity of NSCLC tumors, as A549 and H1299 represent different subtypes with distinct genomic aberrations [28]. We speculate that this difference may arise from specific molecular mechanisms of cell lines, such as cell line origin, differentiation status, gene mutation, epigenetic regulation, and heterogeneity of transcription factors. In addition, A549 and H1299 are cell lines of adenocarcinoma and large cell carcinoma respectively, and they also have significant differences in cell morphology and molecular characteristics. This difference may affect the activation or inhibition of
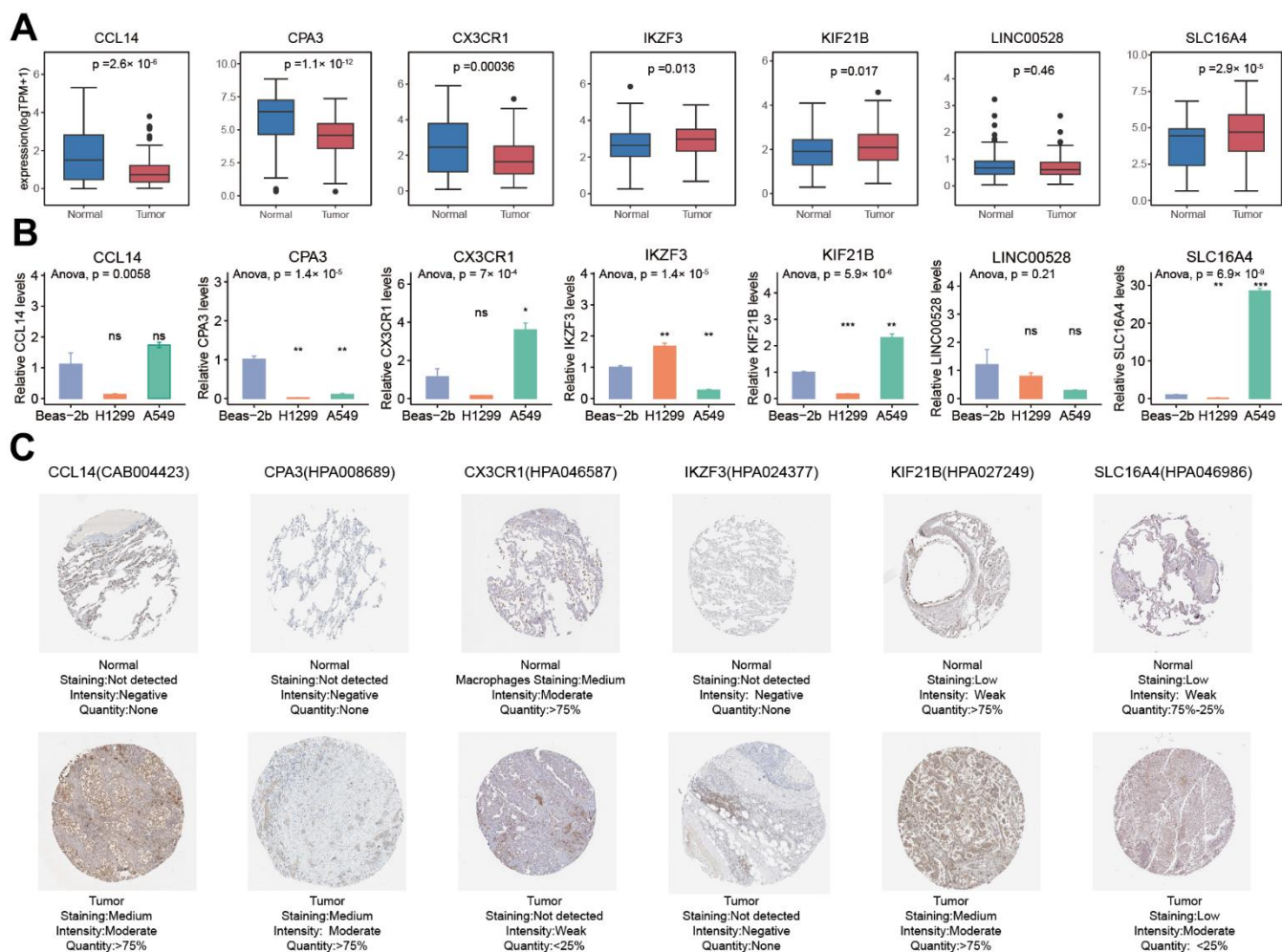

Figure 11. Comparison analysis of the expression of seven PRAN model genes between NSCLC tumor and normal samples at RNA and protein levels. (**A**) RNA expression differences of seven PRAN model genes between tumor and normal samples in TCGA-Advanced NSCLC. (**B**) RNA expression differences of seven PRAN model genes between two tumor cell lines and one normal cell line. (**C**) The immunohistochemistry image of CCL4 (CAB004423), CPA3 (HPA008689), CX3CR1 (HPA046587), IKZF3 (HPA024377), KIF21B (HPA027249), and SLC16A4 (HPA046986) from HPA database. The URLs of the source of each image were shown in Supplementary Table 7.

intracellular signaling pathways, resulting in differences in gene expression levels.

The PRAN-High group is associated with pyrimidine, spliceosome processes, and various metabolic pathways (e.g., hexosamine biosynthesis) based on biological function and signaling pathway enrichment. Pyrimidine synthesis enhances Notch signaling and upregulates c-Myc expression at the transcriptional level, leading to an increase in key glycolytic enzymes. The key enzymes involved in up-regulating pyrimidine synthesis, CAD and DHODH, are implicated in enhancing the chemoresistance of gastric cancer by accelerating glycolysis. Conversely, they sensitize cancer cells to chemotherapy *in vitro* and *in vivo* by inhibiting the pyrimidine biosynthesis pathway [29]. Pyrimidine synthesis is demonstrated up-regulation in Glioblastoma stem cells (GSCs) and is associated with poor survival in glioma patients [30]. The spliceosome components and splicing factor play an essential role in cancers. Notably, splicing alterations can profoundly affect downstream cellular signaling, contributing to cancer development and progression [31]. Metabolic pathways have established crucial roles in cancer development and proliferation over the decades. Moreover, it has become increasingly evident that metabolic pathways closely interact with immune cells during reprogramming cancer cell metabolism [9, 32]. The hexosamine biosynthesis pathway, a vital part of glucose metabolism, is consistently enhanced in cancers during progression [33]. These remarkable insights into the intricate connections between metabolism pathways and cancer biology have revealed the underlying biological mechanism of the PRAN risk model.

Interestingly, the PRAN-High group observed significant enrichment of T-cell follicular helper cells and activated dendritic cells. The T follicular helper cell is related to better survival and therapeutic outcomes in NSCLC, which may be because the T follicular helper cell serves as a critical physiological source of IL-21 for CD8 T cell infiltrating the tumor [34]. Activated dendritic cells serve as an antigen presentation vector, and the related neoantigen vaccine has become an effective therapeutic drug that could improve the survival status of patients with advanced lung cancer [35]. When immune-activating and immune-stimulating cells coexist, their mutual interactions finally have a poor impact on the survival outcomes of patients in the PRAN-High group, and the survival advantage may be offset [36].

To facilitate the practical application of the PRAN risk model in clinical practice, we have combined the model with patients' baseline characteristics to construct a more intuitive nomogram. Its superior performance enhances clinical transformation and aids in decision-making for therapeutic strategies.

Furthermore, we performed sensitive drug prediction on the PRAN risk model. Notably, seven of these drugs specifically target the cell death pathway, leveraging the regulatory role of cell death in tumor biology processes to improve patient survival outcomes. Necrosulfonamide, as an inhibitor of necroptosis, exerts its action by targeting mixed lineage kinase domain-like protein (MLKL), which is pivotal in initiating necroptosis and triggering apoptosis in cancer cells [37]. Ferris et al. uncovered the mechanism of pevonedistat in inducting cell death for colorectal cancer (CRC) therapy, and the responses were effective for both p53 wild-type and mutant mCRC, implying the applicable potent of pevonedistat in other cancers [38]. EX-527 could effectively reverse the AICAR-induced downregulation of c-Myc and metadherin (*MTDH*) expression. Moreover, without *MTDH*, treatment with EX-527 will significantly increase breast cancer cell death [39]. Collectively, these studies offer compelling evidence and elucidate the molecular mechanisms associated with cell death inhibitors, suggesting their promising potential in advanced NSCLC therapy. Moreover, molecule docking analysis further explores the optimal binding position between target PRAN genes and small molecule drugs. As a BCL-2 inhibitor, the navitoclax closely binds with proteins and amino acid residues of hub genes *CCL14, CPA3*, and *CX3CR1*. Having been demonstrated synergy therapeutic efficient when combined with targeted therapies in hematologic malignancies, the navitoclax exhibited compelling clinical efficacy when combined with Osimertinib for *EGFR*-mutant NSCLC patients in a phase IB clinical study (NCT02520778) [40]. In addition, navitoclax also shows increased efficacy in NSCLC patients with a poor response to taxane chemotherapy [41]. These findings shed new light and a solid foundation for the clinical therapy application of navitoclax.

There is still a limitation in our study. We performed the validation experiment using NSCLC cell lines, which cannot reveal the heterogeneous characteristics and prognosis significance underlying advanced NSCLC tumor cells. Further studies across paired advanced NSCLC and normal tissues are warranted to fully elucidate the heterogeneity and characterize subtype-specific expression profiles related to NSCLC pathogenesis.

In conclusion, our findings contribute to advancing precision medicine approaches in NSCLC management. The identified prognostic and immunotherapy biomarkers offer promising avenues for tailoring treatment strategies, optimizing patient

outcomes, and enhancing the overall quality of life for those afflicted with advanced NSCLC. As we continue to uncover the complex molecular landscape of this disease, our predictive model presents a valuable tool to guide clinical decision-making and improve patient care. Further prospective studies are warranted to realize this model's potential in real-world clinical practice fully.

MATERIALS AND METHODS

Data collection and processing

The expression data, somatic mutation information, and copy number variation (CNV) profiles of TCGA-LUAD and TCGA-LUSC from the TCGA database were downloaded from UCSC Xena (https://xenabrowser.net/datapages/), and the corresponding clinical information data were obtained using the R package "TCGAbiolinks". One hundred ninety-seven advanced tumor samples were selected as the training set and named TCGA-Advanced NSCLC (selection criteria for advanced patients: Stage III-IV) (Supplementary Table 2). Three public datasets (GSE13213, GSE61676, GSE91061, GSE74777 and GSE50081) with available overall survival (OS) data and two datasets (GSE135222 and GSE93157) with progression-free survival (PFS) data were downloaded from the GEO database as validation sets. GSE61676 is advanced NSCLC dataset without immunotherapy, GSE13213 is NSCLC dataset without immunotherapy, GSE74777 is early LUSC dataset without immuno-therapy, GSE50081 is early NSCLC dataset without immunotherapy, GSE91061 is melanoma dataset treated with anti-CTLA4 and PD-1, GSE135222 is advanced NSCLC dataset treated with anti-PD-1/PD-L1, and GSE93157 is lung cancer dataset treated with anti-PD-1. All included GEO cohorts' sample information is presented in Supplementary Table 8.

Identification of the expression and variation levels of PCD-related genes

We collected raw transcriptome data from 197 TCGA-Advanced NSCLC patients and 108 normal tissues from the TCGA-LUAD and TCGA-LUSC cohorts. The "limma" package was utilized to identify differentially expressed genes (DEGs) following the criteria of P adjust. < 0.05 and | log2FC| > 1. Kaplan-Meier analysis was performed to assess the impact of each DEG on overall survival (OS) time. The "maftools" package explored somatic mutations among TCGA-Advanced NSCLC patients. We visualized the distinct characteristics of PCD-related genes using a circos plot created with the "RCircos" R package.

Unsupervised clustering and functional analysis

To explore the prognosis significance of PCD pathways, we performed the consensus unsupervised clustering analysis using the NMF package in R based on twelve PCD-related DEGs. The number of clusters k was set from 2 to 10. The stability of clusters obtained through NMF was assessed by the cophenetic correlation, which ranged from 0 and 1. A higher value indicated more excellent cluster stability. DEGs between different PCD clusters were identified using the "limma" R package with cut-off criteria of |logFC| > 1 and P adjust. < 0.05. The "clusterProfiler" R package was used to perform enrichment analysis on the DEGs.

Selection of modules associated with PCD clusters traits by WGCNA

We performed Weighted Gene Co-expression Network Analysis (WGCNA) using the R package "WGCNA" to identify PCD cluster-related genes based on the TPM data. We evaluated different soft thresholds (β values from 1 to 20) for scale independence and average connectivity to construct a scale-free network. The appropriate β value was selected when the scale independence exceeded 0.9 and the average connectivity remained relatively low. Next, we categorized the genes into distinct modules by calculating the topological overlap matrix (TOM). Due to metastatic invariably occurring as a late event in tumor progression and the complex relationship between metastasis and cell death, some tumor cells may increase their ability of survival and migration by inhibiting apoptosis or inducing abnormal cell death in the process of metastasis [42]. And once the tumor cells successfully metastasized to other parts, some cells may suffer cell death due to the influence of heterogeneous environment [43, 44]. Correlations between modules and clinical characteristics were assessed, and modules significantly associated with metastatic status and PCD cluster were further analyzed (p < 0.05).

Development and validation of the risk model

Univariate Cox regression analysis was used to evaluate the prognosis value of the critical module genes and screen predictive genes (p < 0.05). The Least Absolute Shrinkage and Selection Operator (LASSO) regression analysis was performed on candidate genes significantly associated with prognosis. Then, multivariate Cox regression was conducted on all combinations of 15 candidate genes screened by LASSO analysis in the TCGA Advanced NSCLC. The risk scores selected gene combinations were calculated using the multivariate Cox regression

coefficients. Finally, we constructed a predictive model with the highest accuracy and calculated risk scores as follows:

$$\text{Risk score} = \sum_{i=1}^{m} \text{Expression}_i \times \text{Coefficient}_i$$

The m is the number of signature genes for constructing the model; the "Expression$_i$" indicates the expression value of signature gene in the sample; the Coefficient is the multivariate Cox regression coefficient of gene.

The median risk score was used as cutoff value for patients' classification. The receiver operating characteristic (ROC) curve was used to test the accuracy of the risk model. Therefore, the "pROC" package was applied to perform ROC analysis and calculate the area under the ROC curve (AUC) to evaluate the predictive performance of the PCD-related prognostic model. ROC curves for risk scores at 1, 3, and 8 years were plotted using the R package "timeROC".

We constructed a nomogram prediction model using the "RMS" package, incorporating the PCD risk score and clinical factors. Calibration and ROC curves were used to evaluate the prediction accuracy of the nomogram. Additionally, we conducted a decision curve analysis (DCA) to test the nomogram's accuracy.

Correlation analysis between clinical characteristics, mutation, and immune-related characteristics with the PCD risk model

We performed hazard rate analysis on clinicopathological characteristics, including age, gender, stage, TNM stage, and the PCD risk score using univariate and multivariate Cox regression analyses. Somatic mutations were analyzed using the "Maftools" package.

In the TCGA-Advanced NSCLC cohort, the CIBERSORT algorithm [45], single-cell gene set enrichment analysis (ssGSEA) algorithm [46], Micro-environment Cell Populations-counter (MCPcounter) algorithm [47], and Estimation of Stromal and Immune cells in MAlignant Tumors using Expression data (ESTIMATE) algorithm [47] were used for quantify immune cells abundance.

Enrichment analysis, drug susceptibility analysis and molecular docking simulation

Gene set enrichment analysis (GSEA) was performed according to the fold change of all genes in the TCGA-Advanced NSCLC cohort [48]. Differentially enriched KEGG pathways were identified using the Normalized Enrichment Score (NES) and adjusted p-value. The R

package "clusterprofiler" was utilized for this analysis, with the "BH" method employed for p-value correction. Additionally, the cancer immunity cycle and 114 metabolic pathways were quantified using single-sample gene set enrichment analysis (ssGSEA) in the GSVA package.

Cancer Therapeutics Response Portal (CTRP) is a public database containing gene expression of cancer cell line and drug sensitivity information [49, 50]. The CTRP_2 data was downloaded as training sets using the R package "oncoPredict" [51]. The "calcPhenotype" package was employed to construct a ridge regression model that could be applied to the expression profile of the TCGA-Advanced NSCLC cohort and predict the half-maximal inhibitory concentration (IC50) of drugs in lung cancer patients.

Based on functional studies of seven prognostic genes, we screened five protein-coding genes in addition to *LINC00528* and *SLC16A4* for targeted drugs. Drug selection criteria focused on the cell death. Molecular docking simulations are used to predict the formation of stable complexes between large and small molecules. The structure of the proteins was obtained from the Protein Data Bank (https://www.rcsb.org/), and related small molecules were obtained from the zinc database (https://zinc.docking.org/). The Molecular Operating Environment (MOE) software was employed to recognize small molecule drugs and performed the molecular docking simulation to screen optimal docking posture for the anti-tumor drugs and target proteins according to the binding score.

Cell culture and reverse transcription-quantitative polymerase chain reaction (RT-qPCR)

Two human NSCLC cell lines, A549 and NCI-H1299, and one human lung bronchial epithelial cell line (BEAS-2B) were obtained from the Stem Cell Bank, Chinese Academy of Sciences (Beijing, China). The NCI-H1299 cells were cultured in RPMI 1640 medium (Biosharp, Anhui, China); A549 cells were cultured in F12K medium (Biosharp, Anhui, China); and the BEAS-2B cells were cultured in DMEM medium (Biosharp, Anhui, China). All the culture solutions were supplemented with 10% fetal bovine serum (FBS, Gibco). Additionally, trypsin-EDTA (Biosharp, Anhui, China) was also needed to disperse the cells during the cell passage and collection. The cell cultures were maintained at a constant temperature of 37° C in a 5% CO_2 atmosphere, providing optimal cell growth and viability conditions.

RT-qPCR experiments were conducted to validate the differential expression of seven PCD-related genes, including *CCL14, CPA3, CX3CR1, IKZF3, KIF21B,*

Table 1. Primer sequences utilized for RT-qPCR experiment in this study.

Gene	Forward primer	Reverse primer	Size
GAPDH	ATGGGGAAGGTGAAGGTCG	GGGGTCATTGATGGCAACAATA	152bp
SLC16A4	ACCACAAGTCTTACCTCATCCTCTG	TCAACCAGTACAGGCAGTATCAATG	145bp
LINC00528	GCTGAGCTTCTCTCCCTTTCCA	GTTTCCCAGAGCATAGGCAGTG	98bp
KIF21B	AAGAGCCGAGGATCAGAGAAGAC	ACATCATGGAGGACAGCAGGAG	128bp
IKZF3	ACCAAGCCATCAATAACGCCATC	GCTGCTGATAACTGGAACCATCTC	107bp
CX3CR1	CGTGGTCTTTGGGACTGTGTTC	GGCTTCTTGCTGTTGGTGAGG	108bp
CPA3	GAGTCCGAGAAAGAGACGAAAGC	TGGGAGTAGGAATGGAAGGTGATG	92bp
CCL14	TGCTGCTTCACCTACACTACCTAC	CTGGCTGTTGGTCTCATAGTAATCC	72bp

LINC00528, and *SLC16A4* between NSCLC tumors cell lines (A549 and NCI-H1299) and normal lung bronchial epithelial cell line (BEAS-2B), and each. Total RNA was extracted using the TRIzol reagent (Invitrogen, USA) following the manufacturer's protocol. Reverse RNA transcription to cDNA was obtained using PrimeScript® RT reagent Kit (TaKaRa, Shiga, Japan) in a 20 μl reaction mixture. The qPCR was performed with qPCR Kit (SYBR Premix Ex Taq) (TaKaRa, Shiga, Japan) in a 20 μl reaction mixture. Each gene expression reaction was performed in triplicates. The target gene expression was detected using the ABI7300 Fast instrument (Thermo Fisher Scientific, USA). The expression levels of each gene were normalized to the reference gene *GAPDH* and analyzed using the $2^{(-\Delta\Delta CT)}$ method [52], which is a widely accepted approach in scientific research (ΔCT = CT (target gene) − CT (reference gene), $\Delta\Delta CT$ = ΔCT (NSCLC cells) − ΔCT (normal lung cell). The primer sequences are listed in Table 1, and the CT values of RT-qPCR are shown in Supplementary Table 9.

Validation of the protein expression levels of the hub genes via the human protein atlas

To further verify the protein expression levels of *CCL14, CPA3, CX3CR1, IKZF3, KIF21B, LINC00528,* and *SLC16A4* in NSCLC and normal tissues, immunohistochemistry (IHC) data were downloaded from the Human Protein Atlas (HPA, http://www.proteinatlas.org).

Statistical analysis

R software (4.2.1) and corresponding R packages were applied for all statistical analyses. For nonnormally distributed variables, significant quantitative differences between and among groups were determined by the Wilcoxon or Kruskal–Wallis tests, respectively. The Fisher test compared immune responses between other groups, and $p < 0.05$ was used as the significance threshold. Both univariate and multivariate Cox proportional hazards regression analyses were performed using the R package "survival", and regression coefficients, hazards regression, 95% CI, and p-values were calculated. All KM survival curves were plotted using the R package "survminer", and the two-sided log-rank test was used to assess differences between groups.

Abbreviations

AUC: area under the ROC curve; CI: confidence interval; DEGs: differentially expressed genes; HR: hazard ratio; ICIs: immune checkpoint inhibitors; KEGG: Kyoto Encyclopedia of Genes and Genomes; LUAD: lung adenocarcinoma; NSCLC: non-small cell lung cancers; NGS: next-generation sequencing; PD-1: programmed cell death protein-1; PD-L1: programmed cell death protein-1 ligand; TMB: tumor mutational burden; TME: tumor microenvironment; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus GO: Gene Ontology GSEA: gene set enrichment analysis; HPA: Human Protein Atlas; IFN-γ: interferon-gamma; KEGG: Kyoto Encyclopedia of Genes and Genomes; LASSO: least absolute shrinkage and selection operator; PCA: principal component analysis; ROC: receiver operating characteristic; ssGSEA: single sample gene set enrichment analysis; WGCNA: weight gene correlation network analysis.

AUTHOR CONTRIBUTIONS

Conception and design of the study: WD, HZ, LH, and WX. Acquisition of data: HZ, YZ, and SL. Data analysis: HZ and YZ. Visualization: YZ and JZ. Validation: WD. Funding acquisition: WX. Review and editing: BN and WX. Original draft: WD, HZ, LH. Supervision of the study: BN and WX. All authors contributed to the article and approved the submitted version.

CONFLICTS OF INTEREST

FUNDING

REFERENCES

1. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. Nature. 2018; 553:446–54.
https://doi.org/10.1038/nature25183
PMID:29364287

2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin. 2021; 71:209–49.
https://doi.org/10.3322/caac.21660 PMID:33538338

3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin. 2020; 70:7–30.
https://doi.org/10.3322/caac.21590 PMID:31912902

4. Chen R, Manochakian R, James L, Azzouqa AG, Shi H, Zhang Y, Zhao Y, Zhou K, Lou Y. Emerging therapeutic agents for advanced non-small cell lung cancer. J Hematol Oncol. 2020; 13:58.
https://doi.org/10.1186/s13045-020-00881-7
PMID:32448366

5. Pitt JM, Marabelle A, Eggermont A, Soria JC, Kroemer G, Zitvogel L. Targeting the tumor microenvironment: removing obstruction to anticancer immune responses and immunotherapy. Ann Oncol. 2016; 27:1482–92.
https://doi.org/10.1093/annonc/mdw168
PMID:27069014

6. Nicolini A, Ferrari P, Duffy MJ. Prognostic and predictive biomarkers in breast cancer: Past, present and future. Semin Cancer Biol. 2018; 52:56–73.
https://doi.org/10.1016/j.semcancer.2017.08.010
PMID:28882552

7. Wu P, Zheng Y, Wang Y, Wang Y, Liang N. Development and validation of a robust immune-related prognostic signature in early-stage lung adenocarcinoma. J Transl Med. 2020; 18:380.
https://doi.org/10.1186/s12967-020-02545-z
PMID:33028329

8. Feng J, Xu L, Zhang S, Geng L, Zhang T, Yu Y, Yuan R, He Y, Nan Z, Lin M, Guo H. A robust CD8+ T cell-related classifier for predicting the prognosis and efficacy of immunotherapy in stage III lung adenocarcinoma. Front Immunol. 2022; 13:993187.
https://doi.org/10.3389/fimmu.2022.993187
PMID:36119068

9. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022; 12:31–46.
https://doi.org/10.1158/2159-8290.CD-21-1059
PMID:35022204

10. Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, Alnemri ES, Altucci L, Amelio I, Andrews DW, Annicchiarico-Petruzzelli M, Antonov AV, Arama E, et al. Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. Cell Death Differ. 2018; 25:486–541.
https://doi.org/10.1038/s41418-017-0012-4
PMID:29362479

11. Liu L, Li H, Hu D, Wang Y, Shao W, Zhong J, Yang S, Liu J, Zhang J. Insights into N6-methyladenosine and programmed cell death in cancer. Mol Cancer. 2022; 21:32.
https://doi.org/10.1186/s12943-022-01508-w
PMID:35090469

12. Xing N, Du Q, Guo S, Xiang G, Zhang Y, Meng X, Xiang L, Wang S. Ferroptosis in lung cancer: a novel pathway regulating cell death and a promising target for drug therapy. Cell Death Discov. 2023; 9:110.
https://doi.org/10.1038/s41420-023-01407-z
PMID:37005430

13. Li J, Cao F, Yin HL, Huang ZJ, Lin ZT, Mao N, Sun B, Wang G. Ferroptosis: past, present and future. Cell Death Dis. 2020; 11:88.
https://doi.org/10.1038/s41419-020-2298-2
PMID:32015325

14. Ku JM, Kim MJ, Choi YJ, Lee SY, Im JY, Jo YK, Yoon S, Kim JH, Cha JW, Shin YC, Ko SG. JI017 Induces Cell Autophagy and Apoptosis via Elevated Levels of Reactive Oxygen Species in Human Lung Cancer Cells. Int J Mol Sci. 2023; 24:7528.
https://doi.org/10.3390/ijms24087528
PMID:37108692

15. Ning Y, Zheng H, Yang Y, Zang H, Wang W, Zhan Y, Wang H, Luo J, Wen Q, Peng J, Xiang J, Fan S. YAP1 synergize with YY1 transcriptional co-repress DUSP1 to induce osimertinib resistant by activating the EGFR/MAPK pathway and abrogating autophagy in non-small cell lung cancer. Int J Biol Sci. 2023; 19:2458–74.
https://doi.org/10.7150/ijbs.79965
PMID:37215986

16. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn). 2015; 19:A68–77. https://doi.org/10.5114/wo.2014.47136 PMID:25691825

17. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, et al. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res. 2013; 41:D991–5. https://doi.org/10.1093/nar/gks1193 PMID:23193258

18. Zou Y, Xie J, Zheng S, Liu W, Tang Y, Tian W, Deng X, Wu L, Zhang Y, Wong CW, Tan D, Liu Q, Xie X. Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery. Int J Surg. 2022; 107:106936. https://doi.org/10.1016/j.ijsu.2022.106936 PMID:36341760

19. Neubig RR, Spedding M, Kenakin T, Christopoulos A, and International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. International Union of Pharmacology Committee on Receptor Nomenclature and Drug Classification. XXXVIII. Update on terms and symbols in quantitative pharmacology. Pharmacol Rev. 2003; 55:597–606. https://doi.org/10.1124/pr.55.4.4 PMID:14657418

20. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021; 71:7–33. https://doi.org/10.3322/caac.21654 PMID:33433946

21. Arbour KC, Riely GJ. Systemic Therapy for Locally Advanced and Metastatic Non-Small Cell Lung Cancer: A Review. JAMA. 2019; 322:764–74. https://doi.org/10.1001/jama.2019.11058 PMID:31454018

22. Langer CJ, Gadgeel SM, Borghaei H, Papadimitrakopoulou VA, Patnaik A, Powell SF, Gentzler RD, Martins RG, Stevenson JP, Jalal SI, Panwalkar A, Yang JC, Gubens M, et al., and KEYNOTE-021 investigators. Carboplatin and pemetrexed with or without pembrolizumab for advanced, non-squamous non-small-cell lung cancer: a randomised, phase 2 cohort of the open-label KEYNOTE-021 study. Lancet Oncol. 2016; 17:1497–508. https://doi.org/10.1016/S1470-2045(16)30498-3 PMID:27745820

23. Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csőszi T, Fülöp A, Gottfried M, Peled N, Tafreshi A, Cuffe S, O'Brien M, Rao S, Hotta K, et al., and KEYNOTE-024 Investigators. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. N Engl J Med. 2016; 375:1823–33. https://doi.org/10.1056/NEJMoa1606774 PMID:27718847

24. Wu J, Li L, Zhang H, Zhao Y, Zhang H, Wu S, Xu B. A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients with lung adenocarcinoma. Oncogene. 2021; 40:4413–24. https://doi.org/10.1038/s41388-021-01853-y PMID:34108619

25. Prat A, Navarro A, Paré L, Reguart N, Galván P, Pascual T, Martínez A, Nuciforo P, Comerma L, Alos L, Pardo N, Cedrés S, Fan C, et al. Immune-Related Gene Expression Profiling After PD-1 Blockade in Non-Small Cell Lung Carcinoma, Head and Neck Squamous Cell Carcinoma, and Melanoma. Cancer Res. 2017; 77:3540–50. https://doi.org/10.1158/0008-5472.CAN-16-3556 PMID:28487385

26. Li N, Wang J, Zhan X. Identification of Immune-Related Gene Signatures in Lung Adenocarcinoma and Lung Squamous Cell Carcinoma. Front Immunol. 2021; 12:752643. https://doi.org/10.3389/fimmu.2021.752643 PMID:34887858

27. Tsvetkov P, Coy S, Petrova B, Dreishpoon M, Verma A, Abdusamad M, Rossen J, Joesch-Cohen L, Humeidi R, Spangler RD, Eaton JK, Frenkel E, Kocak M, et al. Copper induces cell death by targeting lipoylated TCA cycle proteins. Science. 2022; 375:1254–61. https://doi.org/10.1126/science.abf0529 PMID:35298263

28. Pustovalova M, Alhaddad L, Smetanina N, Chigasova A, Blokhina T, Chuprov-Netochin R, Osipov AN, Leonov S. The p53-53BP1-Related Survival of A549 and H1299 Human Lung Cancer Cells after Multifractionated Radiotherapy Demonstrated Different Response to Additional Acute X-ray Exposure. Int J Mol Sci. 2020; 21:3342. https://doi.org/10.3390/ijms21093342 PMID:32397297

29. He D, Chen M, Chang L, Gu J, Liu F, Gao X, Ruan Y. De novo pyrimidine synthesis fuels glycolysis and confers chemoresistance in gastric cancer. Cancer Lett. 2022; 549:215837. https://doi.org/10.1016/j.canlet.2022.215837 PMID:35921972

30. Wang X, Yang K, Wu Q, Kim LJ, Morton AR, Gimple RC, Prager BC, Shi Y, Zhou W, Bhargava S, Zhu Z, Jiang L, Tao W, et al. Targeting pyrimidine synthesis accentuates molecular therapy response in glioblastoma stem cells. Sci Transl Med. 2019; 11:eaau4972. https://doi.org/10.1126/scitranslmed.aau4972

PMID:31391321

31. Yang H, Beutler B, Zhang D. Emerging roles of spliceosome in cancer and immunity. Protein Cell. 2022; 13:559–79.
https://doi.org/10.1007/s13238-021-00856-5
PMID:34196950

32. Leone RD, Powell JD. Metabolism of immune cells in cancer. Nat Rev Cancer. 2020; 20:516–31.
https://doi.org/10.1038/s41568-020-0273-y
PMID:32632251

33. Wei S, Zhao Q, Zheng K, Liu P, Sha N, Li Y, Ma C, Li J, Zhuo L, Liu G, Liang W, Jiang Y, Chen T, Zhong N. GFAT1-linked TAB1 glutamylation sustains p38 MAPK activation and promotes lung cancer cell survival under glucose starvation. Cell Discov. 2022; 8:77.
https://doi.org/10.1038/s41421-022-00423-0
PMID:35945223

34. Cui C, Wang J, Fagerberg E, Chen PM, Connolly KA, Damo M, Cheung JF, Mao T, Askari AS, Chen S, Fitzgerald B, Foster GG, Eisenbarth SC, et al. Neoantigen-driven B cell and CD4 T follicular helper cell collaboration promotes anti-tumor CD8 T cell responses. Cell. 2021; 184:6101–18.e13.
https://doi.org/10.1016/j.cell.2021.11.007
PMID:34852236

35. Ding Z, Li Q, Zhang R, Xie L, Shu Y, Gao S, Wang P, Su X, Qin Y, Wang Y, Fang J, Zhu Z, Xia X, et al. Personalized neoantigen pulsed dendritic cell vaccine for advanced lung cancer. Signal Transduct Target Ther. 2021; 6:26.
https://doi.org/10.1038/s41392-020-00448-5
PMID:33473101

36. Sorin M, Rezanejad M, Karimi E, Fiset B, Desharnais L, Perus LJ, Milette S, Yu MW, Maritan SM, Doré S, Pichette É, Enlow W, Gagné A, et al. Single-cell spatial landscapes of the lung tumour immune microenvironment. Nature. 2023; 614:548–54.
https://doi.org/10.1038/s41586-022-05672-3
PMID:36725934

37. Bai Y, Lam HC, Lei X. Dissecting Programmed Cell Death with Small Molecules. Acc Chem Res. 2020; 53:1034–45.
https://doi.org/10.1021/acs.accounts.9b00600
PMID:32297735

38. Ferris J, Espona-Fiedler M, Hamilton C, Holohan C, Crawford N, McIntyre AJ, Roberts JZ, Wappett M, McDade SS, Longley DB, Coyle V. Pevonedistat (MLN4924): mechanism of cell death induction and therapeutic potential in colorectal cancer. Cell Death Discov. 2020; 6:61.
https://doi.org/10.1038/s41420-020-00296-w
PMID:32714568

39. Gollavilli PN, Kanugula AK, Koyyada R, Karnewar S, Neeli PK, Kotamraju S. AMPK inhibits MTDH expression via GSK3β and SIRT1 activation: potential role in triple negative breast cancer cell proliferation. FEBS J. 2015; 282:3971–85.
https://doi.org/10.1111/febs.13391 PMID:26236947

40. Bertino EM, Gentzler RD, Clifford S, Kolesar J, Muzikansky A, Haura EB, Piotrowska Z, Camidge DR, Stinchcombe TE, Hann C, Malhotra J, Villaruz LC, Paweletz CP, et al. Phase IB Study of Osimertinib in Combination with Navitoclax in EGFR-mutant NSCLC Following Resistance to Initial EGFR Therapy (ETCTN 9903). Clin Cancer Res. 2021; 27:1604–11.
https://doi.org/10.1158/1078-0432.CCR-20-4084
PMID:33376097

41. Tan N, Malek M, Zha J, Yue P, Kassees R, Berry L, Fairbrother WJ, Sampath D, Belmont LD. Navitoclax enhances the efficacy of taxanes in non-small cell lung cancer models. Clin Cancer Res. 2011; 17:1394–404.
https://doi.org/10.1158/1078-0432.CCR-10-2353
PMID:21220478

42. Wang K, Chen Q, Shao Y, Yin S, Liu C, Liu Y, Wang R, Wang T, Qiu Y, Yu H. Anticancer activities of TCM and their active components against tumor metastasis. Biomed Pharmacother. 2021; 133:111044.
https://doi.org/10.1016/j.biopha.2020.111044
PMID:33378952

43. Yang A, Herter-Sprie G, Zhang H, Lin EY, Biancur D, Wang X, Deng J, Hai J, Yang S, Wong KK, Kimmelman AC. Autophagy Sustains Pancreatic Cancer Growth through Both Cell-Autonomous and Nonautonomous Mechanisms. Cancer Discov. 2018; 8:276–87.
https://doi.org/10.1158/2159-8290.CD-17-0952
PMID:29317452

44. Zhang Y, Commisso C. Macropinocytosis in Cancer: A Complex Signaling Network. Trends Cancer. 2019; 5:332–4.
https://doi.org/10.1016/j.trecan.2019.04.002
PMID:31208695

45. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015; 12:453–7.
https://doi.org/10.1038/nmeth.3337
PMID:25822800

46. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep. 2017; 18:248–62.
https://doi.org/10.1016/j.celrep.2016.12.019
PMID:28052254

47. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N,

Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, de Reyniès A. Erratum to: Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016; 17:249.
https://doi.org/10.1186/s13059-016-1113-y
PMID:27908289

48. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005; 102:15545–50.
https://doi.org/10.1073/pnas.0506580102
PMID:16199517

49. Seashore-Ludlow B, Rees MG, Cheah JH, Cokol M, Price EV, Coletti ME, Jones V, Bodycombe NE, Soule CK, Gould J, Alexander B, Li A, Montgomery P, et al. Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. Cancer Discov. 2015; 5:1210–23.

https://doi.org/10.1158/2159-8290.CD-15-0235
PMID:26482930

50. Iorio F, Knijnenburg TA, Vis DJ, Bignell GR, Menden MP, Schubert M, Aben N, Gonçalves E, Barthorpe S, Lightfoot H, Cokelaer T, Greninger P, van Dyk E, et al. A Landscape of Pharmacogenomic Interactions in Cancer. Cell. 2016; 166:740–54.
https://doi.org/10.1016/j.cell.2016.06.017
PMID:27397505

51. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting *in vivo* or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform. 2021; 22:bbab260.
https://doi.org/10.1093/bib/bbab260
PMID:34260682

52. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001; 25:402–8.
https://doi.org/10.1006/meth.2001.1262
PMID:11846609
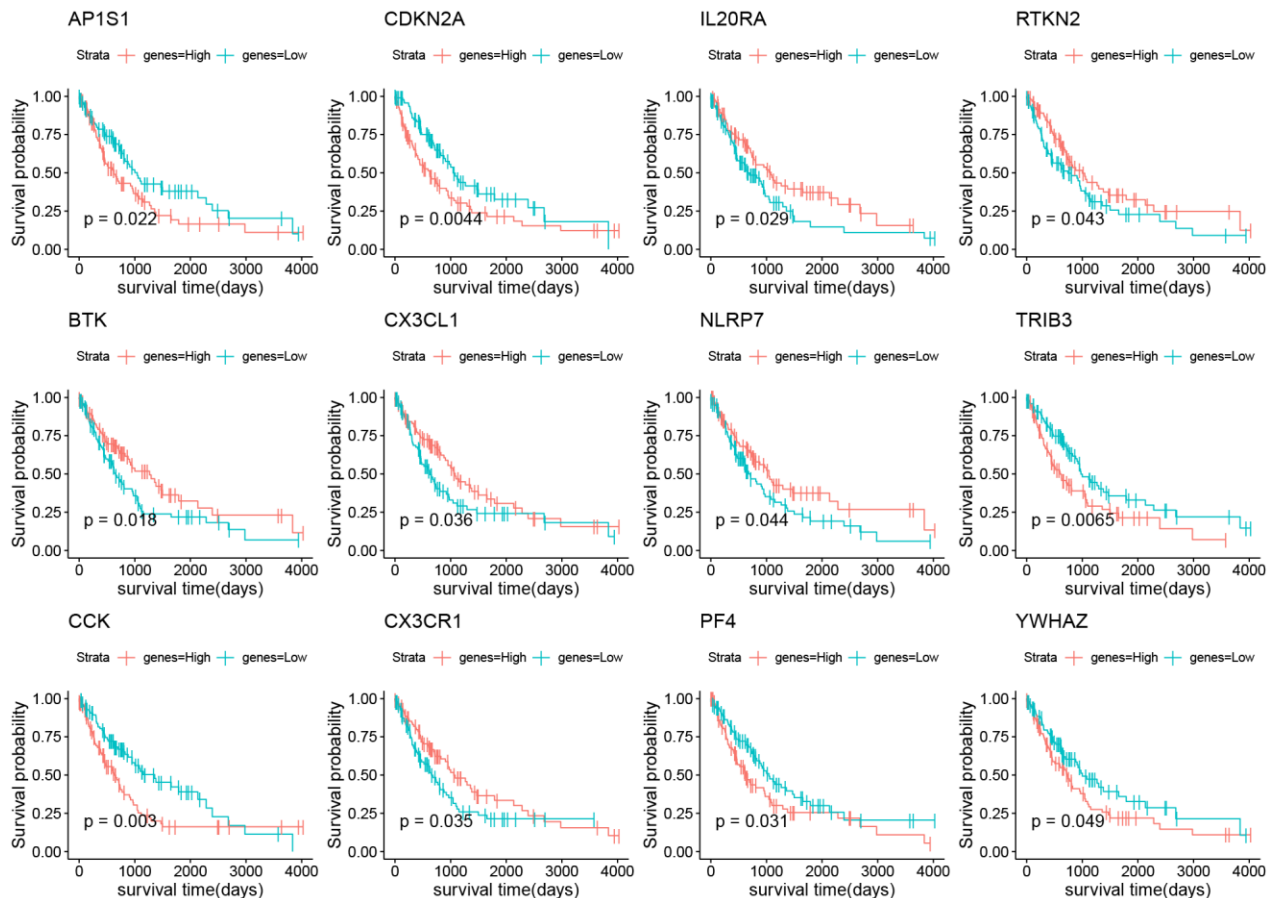
Supplementary Figure 1. The prognostic value of the 12 prognosis-related PCD genes.
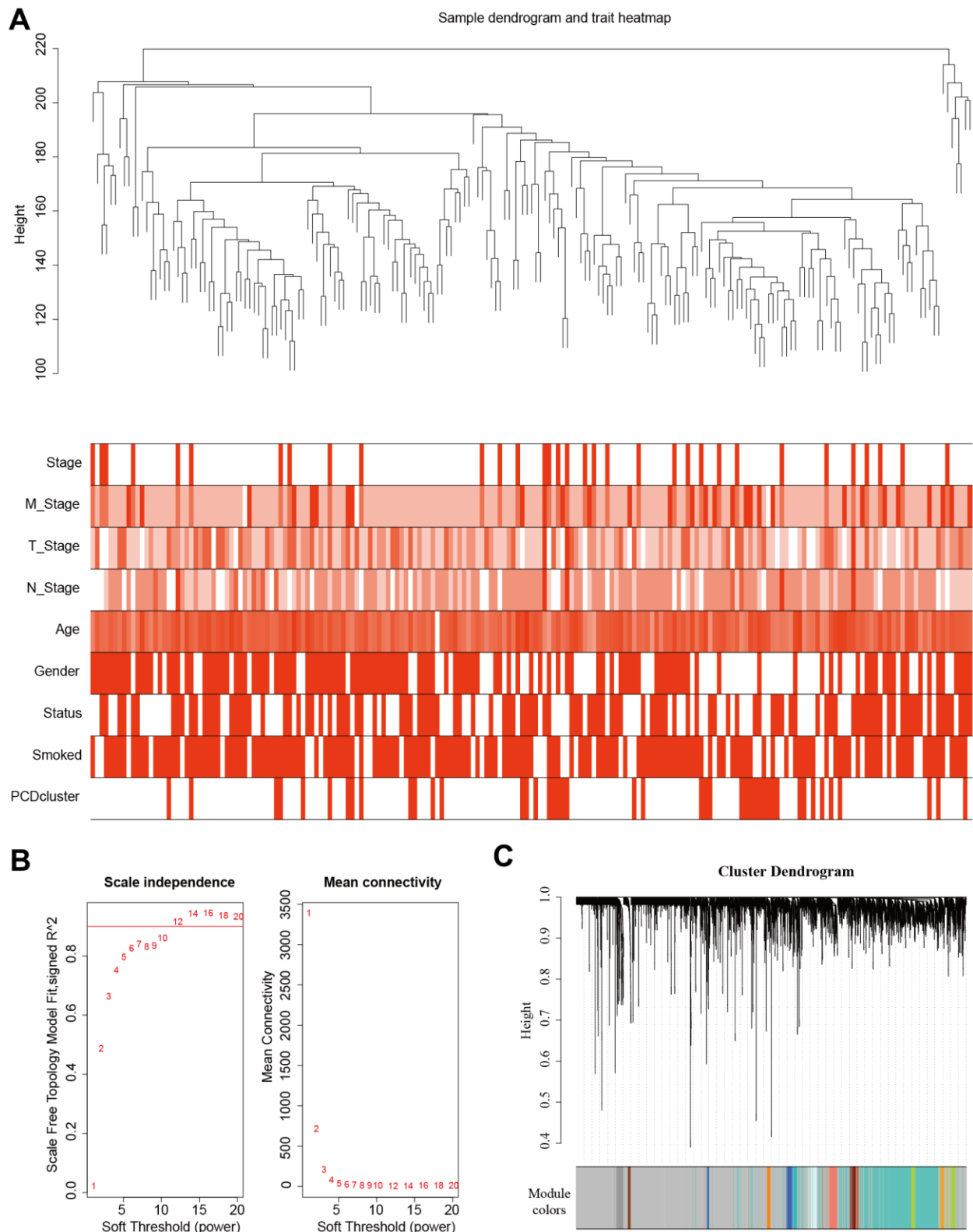
Supplementary Figure 2. Identification of modules by WGCNA. (**A**) Clustering dendrogram of TCGA-Advanced NSCLC samples and corresponding heatmap displaying clinical traits. The color band beneath the dendrogram represents the tissue trait status. (**B**) Evaluation of the scale-free fit index (left) and mean connectivity (right) across various soft threshold power values. The appropriate power value was determined for subsequent module identification. (**C**) Cluster dendrogram representing 39 distinct colored modules, based on a dissimilarity measure (1-TOM).
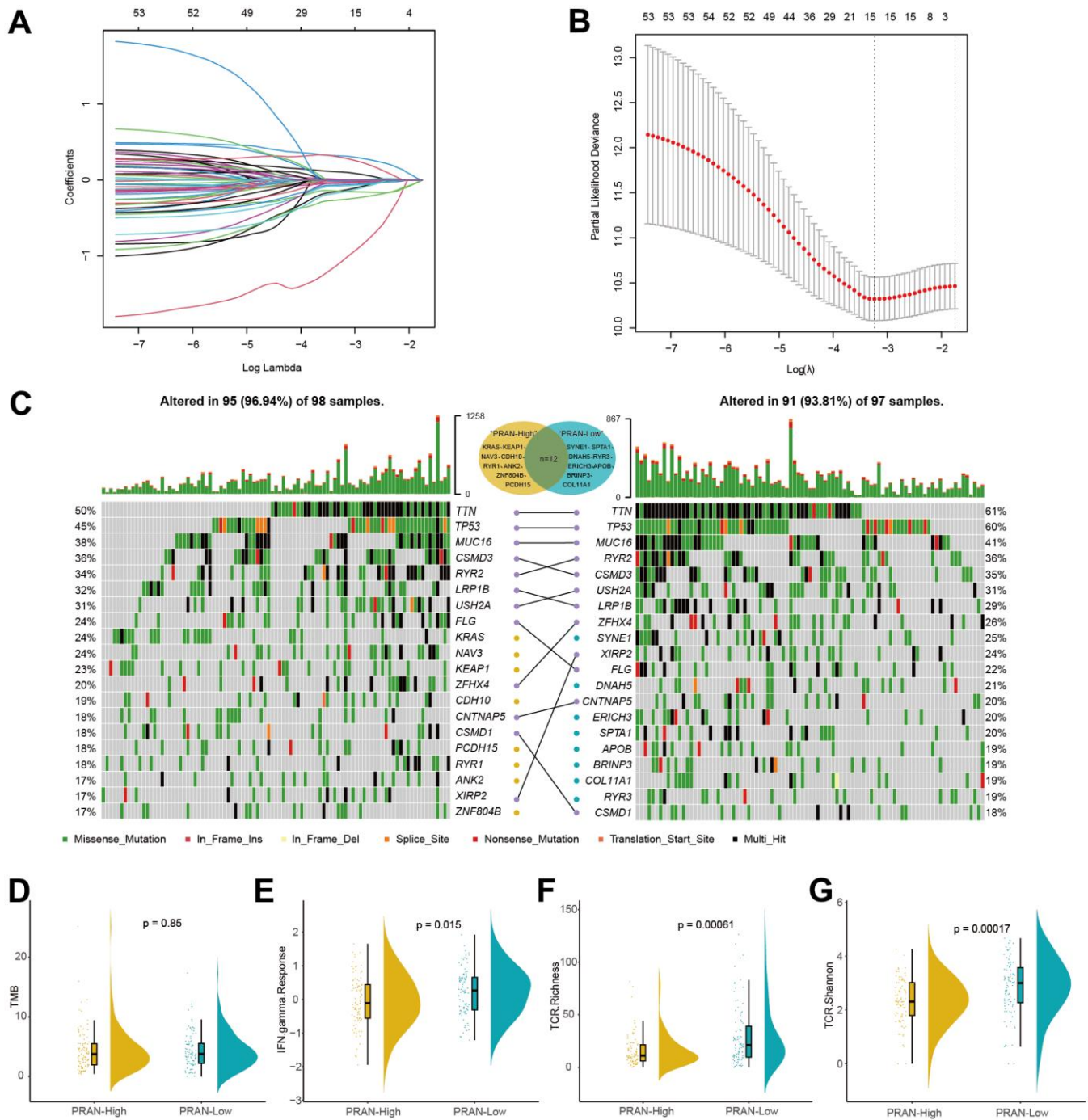
Supplementary Figure 3. The candidate genes screen by LASSO regression and association analysis of mutation profiles and immunological factors with PRAN risk score in the TCGA-Advanced NSCLC cohort. (**A**) The LASSO coefficient profiles of PCD-related genes. (**B**) Two dotted vertical lines are drawn at the optimal values according to the minimum criterion (right) and the Mean-Squared Error criterion (left). The horizontal axis represents the log value of lambda, and the vertical axis represents the Mean-Squared Error criterion of the independent variable. (**C**) The mutation profiles revealing the different mutation frequency between the two PRAN risk groups. (**D**) Box plot representing the comparison of Tumor Mutation Burden (TMB) levels between the two PRAN risk groups. (**E**) Box plot displaying the comparison of Interferon Gamma (IFN-γ) levels between the two PRAN risk groups. (**F**) Box plot showing the comparison of T-cell Receptor (TCR) richness levels between the two PRAN risk groups. (**G**) Box plot illustrating the comparison of T-cell Receptor (TCR) Shannon diversity index between the two PRAN risk groups.
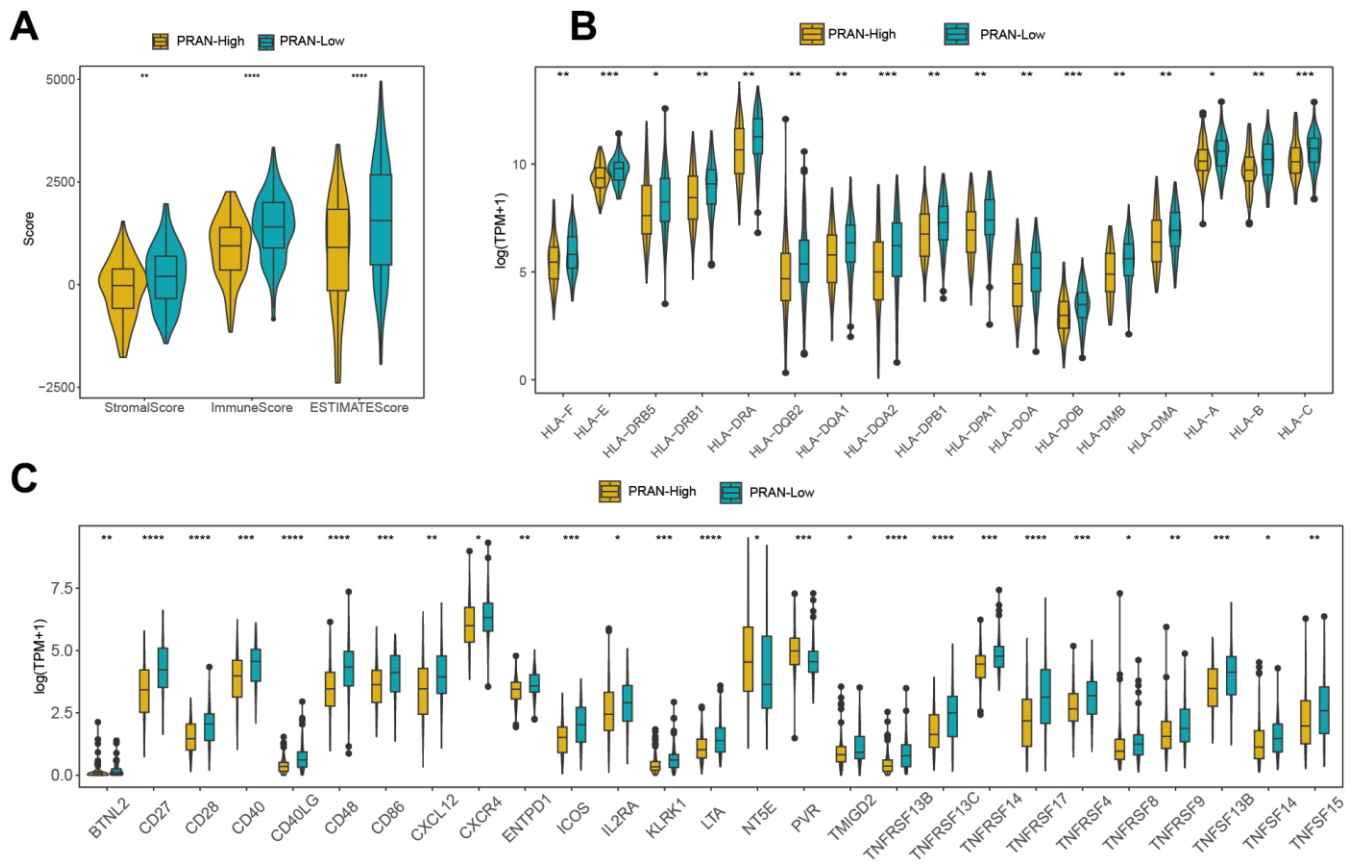
Supplementary Figure 4. Immune cell infiltration analysis between the two PRAN risk groups in the TCGA-Advanced NSCLC cohort. (**A**) Stromal score, immune score, and ESTIMATE score between PRAN-High and PRAN-Low groups. The expression of (**B**) HLA genes, and (**C**) immunostimulatory genes between PRAN-High and PRAN-Low groups. P-value: * < 0.05; ** < 0.01; *** < 0.001; **** < 0.0001.
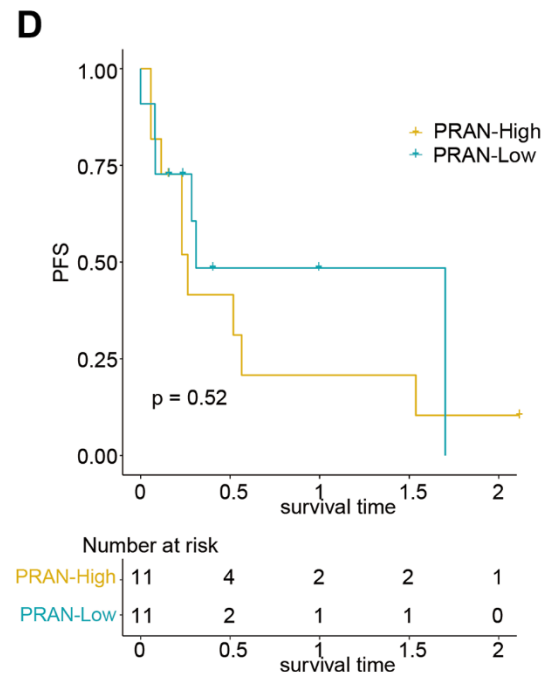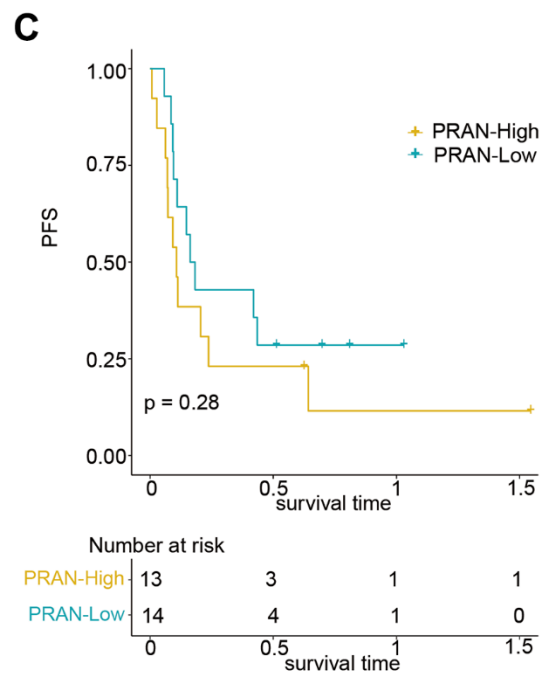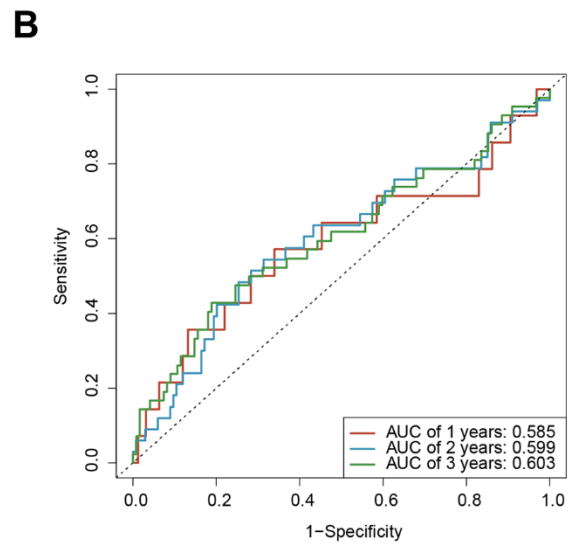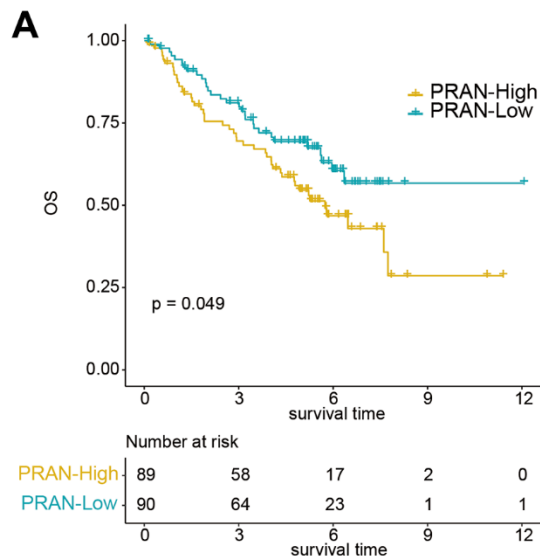
Supplementary Figure 5. The predicting performance validation of the PRAN risk model in multiple GEO cohorts. (**A**) Kaplan-Meier survival analysis in validation cohort GSE50081. (**B**) Time-dependent ROC curves between PRAN-High and PRAN-Low groups in validation cohort GSE50081. (**C**, **D**) Kaplan-Meier survival analysis in validation cohort GSE135222, and GSE93157.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1, 3–5, 7–9.

Supplementary Table 1. The programmed cell death (PCD) genes acquired from Zou et al. 2022.

Supplementary Table 2. The sample information in TCGA-Advanced NSCLC cohort.

	TCGA_Advanced NSCLC					
	Cluster1 (N=153)	Cluster2 (N=44)	P-value	High_risk (N=98)	Low_risk (N=99)	P-value
Age:			0.213			1.000
<=65	65 (42.5%)	24 (54.5%)		44 (44.9%)	45 (45.5%)	
>65	88 (57.5%)	20 (45.5%)		54 (55.1%)	54 (54.5%)	
Gender:			0.010			0.267
Female	52 (34.0%)	25 (56.8%)		34 (34.7%)	43 (43.4%)	
Male	101 (66.0%)	19 (43.2%)		64 (65.3%)	56 (56.6%)	
Stage:			0.953			0.239
III	128 (83.7%)	36 (81.8%)		78 (79.6%)	86 (86.9%)	
IV	25 (16.3%)	8 (18.2%)		20 (20.4%)	13 (13.1%)	
T_stage:			0.665			0.769
T1	14 (9.15%)	6 (13.6%)		9 (9.18%)	11 (11.1%)	
T2	68 (44.4%)	19 (43.2%)		42 (42.9%)	45 (45.5%)	
T3	39 (25.5%)	8 (18.2%)		27 (27.6%)	20 (20.2%)	
T4	30 (19.6%)	10 (22.7%)		19 (19.4%)	21 (21.2%)	
TX	2 (1.31%)	1 (2.27%)		1 (1.02%)	2 (2.02%)	
M_stage:			0.254			0.052
Deficient	2 (1.31%)	1 (2.27%)		3 (3.06%)	0 (0.00%)	
M0	107 (69.9%)	25 (56.8%)		58 (59.2%)	74 (74.7%)	
M1	24 (15.7%)	8 (18.2%)		20 (20.4%)	12 (12.1%)	
MX	20 (13.1%)	10 (22.7%)		17 (17.3%)	13 (13.1%)	
N_stage:			0.896			0.418
N0	22 (14.4%)	8 (18.2%)		16 (16.3%)	14 (14.1%)	
N1	35 (22.9%)	11 (25.0%)		23 (23.5%)	23 (23.2%)	
N2	87 (56.9%)	22 (50.0%)		56 (57.1%)	53 (53.5%)	
N3	5 (3.27%)	2 (4.55%)		1 (1.02%)	6 (6.06%)	
NX	4 (2.61%)	1 (2.27%)		2 (2.04%)	3 (3.03%)	
Tissue_origine:			0.871			0.139
Lower lobe	51 (33.3%)	15 (34.1%)		26 (26.5%)	40 (40.4%)	
Middle lobe	9 (5.88%)	1 (2.27%)		4 (4.08%)	6 (6.06%)	
Other	17 (11.1%)	4 (9.09%)		11 (11.2%)	10 (10.1%)	
Upper lobe	76 (49.7%)	24 (54.5%)		57 (58.2%)	43 (43.4%)	
Race:			0.127			0.897
Asian	2 (1.31%)	3 (6.82%)		2 (2.04%)	3 (3.03%)	
Black or African American	14 (9.15%)	3 (6.82%)		9 (9.18%)	8 (8.08%)	
Unknown	39 (25.5%)	7 (15.9%)		21 (21.4%)	25 (25.3%)	
white	98 (64.1%)	31 (70.5%)		66 (67.3%)	63 (63.6%)	
Smoking_Status:			0.756			0.125
non-smoke	36 (23.5%)	12 (27.3%)		29 (29.6%)	19 (19.2%)	
Smoke	117 (76.5%)	32 (72.7%)		69 (70.4%)	80 (80.8%)	

OS.time	468 [215;951]692 [407;1177]	0.030	389 [144;728]716 [420;1177]	<0.001
OS_status:		0.139		0.003
Alive	62 (40.5%) 24 (54.5%)		32 (32.7%) 54 (54.5%)	
Dead	91 (59.5%) 20 (45.5%)		66 (67.3%) 45 (45.5%)	
TMB:		0.541		0.422
NA	1 (0.65%) 1 (2.27%)		0 (0.00%) 2 (2.02%)	
TMB<10	144 (94.1%) 41 (93.2%)		92 (93.9%) 93 (93.9%)	
TMB>=10	8 (5.23%) 2 (4.55%)		6 (6.12%) 4 (4.04%)	

Supplementary Table 3. The 359 DEGs were identified between cluster 1 and cluster 2.

Supplementary Table 4. The stepwise screening process of PCD hub genes for PRAN model construction, including 55 univariate Cox genes, 15 LASSO genes, and 7 multivariate Cox genes.

Supplementary Table 5. Drug sensitivity prediction based on TCGA-Advanced NSCLC patients.

Supplementary Table 6. Interaction between target protein and candidate small molecules according to the molecular docking analysis.

2Q8R:ZINC150338726(navitoclax)								
Ligand		Receptor				Interaction	Distance	E(kcal/mol)
S	86	OE1	GLN	20	(H)	H-donor	3.87	-0.5
O	100	O	HOH	112	(G)	H-acceptor	2.93	-0.6
O	101	O	HOH	117	(H)	H-acceptor	2.91	-1.3
6-ring		CB	GLN	20	(H)	pi-H	4.26	-0.5
4BD9:ZINC150338726(navitoclax)								
Ligand		Receptor				Interaction	Distance	E(kcal/mol)
O	52	CA	GLU	136	(B)	H-acceptor	3.46	-0.6
5WN9:ZINC150338726(navitoclax)								
Ligand		Receptor				Interaction	Distance	E(kcal/mol)
6-ring		CA	LEU	184	(H)	pi-H	4.01	-0.5
7BQV:ZINC000066099927(AZD1480)								
Ligand		Receptor				Interaction	Distance	E(kcal/mol)
N	37	5-ring	TRP	380	(A)	H-pi	3.71	-0.7
6-ring		O	HOH	602	(B)	pi-H	3.3	-0.9
7X4N:ZINC13476439(MG-132)								
Ligand		Receptor				Interaction	Distance	E(kcal/mol)
N	38	OE2	GLU	411	(A)	H-donor	3.28	-1
O	18	NE	ARG	105	(A)	H-acceptor	3.16	-2

Supplementary Table 7. Representative IHC staining of model genes in normal (top) and tumor (bottom) tissues of the NSCLC.

Supplementary Table 8. The sample information in GEO cohort.

Supplementary Table 9. The RNA expression data of seven PRAN model genes using RT-qPCR experiment.