

Original Article

Identification and validation of the prognostic value of immune-related genes in non-small cell lung cancer

Rui Li¹, Xiao Liu¹, Xi-Jia Zhou¹, Xiao Chen^{1,2}, Jian-Ping Li¹, Yun-Hong Yin³, Yi-Qing Qu³

¹Department of Pulmonary and Critical Care Medicine, Qilu Hospital, Cheeloo College of Medicine, Shandong University, Jinan 250012, China; ²Department of Respiratory Medicine, Tai'an City Central Hospital, Tai'an 271000, China; ³Department of Pulmonary and Critical Care Medicine, Qilu Hospital of Shandong University, Jinan 250012, China

Received June 15, 2020; Accepted August 7, 2020; Epub September 15, 2020; Published September 30, 2020

Abstract: Immune-related genes play a significant role in predicting the overall survival and monitoring the status of the cancer immune microenvironment. The aim of this research study was to identify differentially expressed immune-related genes (DEIRGs) and establish a Cox prediction model for the evaluation of prognosis in patients with non-small cell lung cancer (NSCLC). Transcription expression data, immune gene data, and tumor transcription factor data from The Cancer Genome Atlas (TCGA), the Immunology Database and Analysis Portal, and the Cistrome Cancer database were analyzed to detect differentially expressed genes (DEGs), DEIRGs, and differentially expressed transcription factors (DETFs). Multivariate Cox regression analysis was used to obtain potential DEIRGs as independent prognostic factors. Oncomine, The Human Protein Atlas (HPA), TIMER databases were performed to validate the mRNA and protein expression level of DEIRGs. TIMER database was performed to explore the immunocytes infiltration of DEIRGs. In total, 7448 DEGs, 536 DEIRGs, 87 DETFs were identified from 1,037 NSCLC tissues and 108 normal tissues in TCGA database. Fifteen-DEIRG signatures (THBS1, S100P, S100A16, DLL4, CD70, DKK1, IL33, NRTN, PDGFB, STC2, VGF, GCGR, HTR3A, LGR4, SHC3) could be perceived as independent prognostic factors for predicting the overall survival of patients with NSCLC ($P = 4.89e-09$). Immune cell correlation analysis showed that neutrophils and b cells were positively and negatively correlated with the riskscore of the prediction model, respectively. Our study identified a Cox prediction model based on DEIRGs to predict the overall survival of patients with NSCLC. The immunocyte infiltration analysis provided a novel horizon for monitoring the status of the NSCLC immune microenvironment.

Keywords: Non-small cell lung cancer, immune-related genes, transcription factors, tumor immune microenvironment, overall survival, prognostic biomarkers

Introduction

Lung cancer remains the leading cause of cancer-related mortality worldwide [1, 2]. Patients with non-small cell lung cancer (NSCLC) account for approximately 85% of lung cancer cases, mainly comprised of lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) [3, 4]. With the increasing incidence and mortality associated with NSCLC, the survival rates of patients with LUAD and LUSC remain suboptimal [5]. For the clinical tumor-node-metastasis (TNM) stage IIIB and IV NSCLCs, the 5-year survival is 7%, and 2%, respectively [6]. High throughput technology, microarray, and RNA sequencing can be applied to identify data profiles related to prognostic biomarkers [7-9].

Recently, some studies have shown that the prediction model based on RNA sequencing data could precisely predict the survival of patients with cancers [9-12]. The immune microenvironment (IME), including immune cells associated with immune-related genes (IRGs), has a significant impact on predicting the prognosis of cancers [13]. A study has shown that an immune-related prognostic model in the context of the TP-53 status could forecast the prognosis of colorectal cancer [14]. Another study identified an immune signature prediction model to forecast the prognosis of LUAD [7]. Recently, a study revealed that the clinical immune signature could be perceived as a prognostic biomarker for forecasting the prognosis of non-squamous NSCLC [15]. However, a Cox predic-

tive model in the context of IRGs and clinical relevance in LUAD and LUSC is currently lacking.

This research mainly aimed to acquire differentially expressed genes (DEGs), differentially expressed immune-related genes (DEIRGs), differentially expressed transcription factors (DETFs) based on The Cancer Genome Atlas (TCGA), Immunology Database and Analysis Portal (ImmPort), and Cistrome Cancer databases. Subsequently, a Cox prediction model based on DEIRGs was constructed for predicting the prognosis of LUAD and LUSC. Furthermore, the DEIRGs-related DETFs regulatory network reveals the potential mechanism of DEIRGs in patients with LUSC and LUAD. Moreover, correlation analysis between the riskScore of the prediction model and immunocyte infiltration was performed to estimate the status of the tumor microenvironment.

Materials and methods

Clinical patients and data collection

We downloaded the fragments per kilobase of exon model per million reads mapped data for LUSC and LUAD from the transcriptome RNA-Sequence data in TCGA database, incorporating 1,037 LUSC and LUAD tissues and 108 normal tissues. We acquired IRGs in the ImmPort (<http://www.immport.org/>) [16]. Furthermore, we used the Cistrome Project (<http://www.cistrome.org/>) to download the cancer transcription factor targets [17].

Differential gene expression analysis in LUAD and LUSC

We used the “limma R” (<http://www.bioconductor.org/packages/release/bioc/html/limma.html>) to perform the differential gene expression analysis in all transcriptome RNA-Seq gene expression data, cancer TF targets and IRGs, on the basis of absolute fold change (\log_2) > 1 and the thresholds of adjusted false discovery rate ($P < 0.05$). We used the Cistrome Cancer database to obtain DETFs from DEGs. Moreover, DEIRGs were extracted from DEGs using the ImmPort database.

Functional enrichment analysis of DEIRGs in LUSC and LUAD

We further investigated the functions of those DEIRGs based on their expression profiles. The

Database for Annotation, Visualization and Integrated Discovery (<http://www.david.niaid.nih.gov>) was utilized to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, together with the Gene Ontology (GO) analysis [18]. The GO analysis showed that a $P < 0.05$ denoted a statistically significant. The GOCircle plot and GO Chord plot were obtained using the GOplot R package (<https://cran.r-project.org/web/packages/GOplot/citation.html>) [19]. Furthermore, the KEGG pathway analysis was conducted on the “cluster profiler R” package (<http://www.bioconductor.org/packages/release/bioc/html/clusterProfiler.html>) [20]. A $P < 0.05$ denoted statistical significance. In addition, we used the Cytoscape (version 3.6.1) software to establish the DEIRGs-related KEGG pathway network for visual analysis [21].

Establishment of the DEIRG-based prediction Cox model in LUAD and LUSC

We used univariate Cox regression analyses to obtain survival-related DEIRGs. The survival-related DEIRGs were selected as prognostic biomarkers to perform multivariate Cox regression analyses ($P < 0.05$). In the context of the median riskScore value, Patients with LUAD and LUSC were classified into low-risk and high-risk groups. More importantly, receiver operating characteristic (ROC) analysis was performed to estimate the signature of DEIRGs (low-risk vs. high-risk) based on the overall survival (OS). The area under the receiver operating characteristic curve values were calculated to evaluate the prediction Cox model for explore the prognostic biomarkers in patients with LUAD and LUSC.

DEIRG-mediated DETF regulatory network in LUAD and LUSC

TFs act as significant molecules that can regulate the expression levels of genes. Thus, investigation of the mechanism of TFs involved in regulating the expression of OS-related DEIRGs is of great importance. The Cistrome Cancer database provides the regulatory relationship between the transcriptome and TFs in TCGA profilers. Therefore, we constructed DEIRG-mediated DETF regulatory network for visual analysis with the thresholds of filters $P < 0.001$ and standard correlation coefficient > 0.4 to investigate the potential mechanism of DEIRGs.

Independent prognosis analysis

We used the univariate and multivariate Cox independent prognosis analyses to further investigate whether the fifteen-DEIRG signatures could be used as independent prognostic factors. We screened 668 LUAD and LUSC patients with full clinical characteristics information and meanwhile integrated the expression and riskScore of fifteen DEIRG signatures to perform the univariate and multivariate Cox independent prognosis analyses. $P < 0.05$ denoted statistical significance.

Clinical relevance analysis between prognosis-related DEIRGs and clinical characteristics in LUAD and LUSC

Correlation analysis was conducted to further identify the interaction between prognosis-related DEIRGs and clinical characteristics in LUSC and LUAD. The clinical characteristics included age (≤ 65 / > 65 years), sex (female/male), pathological T stage (T3&T4/T1&T2), pathological TNM stage (III&IV/I&II), pathological M stage (M1/M0), and pathological N stage (N1-N3/N0). $P < 0.05$ denoted statistical significance.

Oncomine, the Human Protein Atlas (HPA), TIMER validation

Oncomine is an online database, which could automatically detect differentially expressed profiles for cancer types and subtypes to analyze a gene of interest across the whole datasets [22]. We used the Oncomine to analyze the expression level of prognosis-related DEIRGs in 20 cancer types to visualize the differential expression level. The Human Protein Atlas (HPA) database could be used as a tool for identifying potential novel cancer biomarkers using the protein profiles with in silico-based methods [23]. We used the immunohistochemistry data of HPA to validate the expression level of prognosis-related DEIRGs. TIMER 2.0 database is an online database, which could analyze the immunocytes infiltration, gene mutation, differential expression in all genes [24]. We used the TIMER 2.0 to validate the expression level of prognosis-related DEIRGs in all kinds of cancers from the Diff Exp section.

Survival analysis of prognosis-related DEIRGs

The Kaplan-Meier plotter (www.kmplot.com), including clinical data and gene expression

data of lung cancer [25], gastric cancer [26], breast cancer [27], and ovarian cancer [28]. Kaplan-Meier plotter database with log-rank test was performed to conduct the survival analysis of prognosis-related DEIRGs.

Infiltration of immune cells

Tumor Immune Estimation Resource (TIMER) is an online database, which can analyze and visualize the levels of tumor-infiltrating immunocytes [29]. TIMER was used to reanalyze the gene expression data of 10,897 samples from 32 types of cancers in TCGA to reveal correlations between tumor-infiltrating immunocytes, including macrophages, neutrophils, dendritic cells, B cells, CD8 T cells, and CD4 T cells (<https://zenodo.org/record/57669#.Xeezu9V5uMo>) and other characteristics. We used the TIMER online database to download the TIMER immune Estimation file. Subsequently, we interacted the TIMER immune Estimation file with the immune riskScore file to perform the correlation analysis.

Results

Acquisition of DEGs, DEIRGs, and DETFs

In total, 7,448 DEGs, 536 DEIRGs, and 87 DETFs were obtained using the limma R package according to the set standards ($P < 0.05$, $|\log_{2}FC| > 1$) between 1,037 LUAD and LUSC specimens and 108 adjacent normal specimens. Based on the criteria, we identified 5,552 up-regulated DEGs and 1,898 down-regulated DEGs, 340 up-regulated DEIRGs and 196 down-regulated DEIRGs, as well as 53 up-regulated DETFs and 34 down-regulated DETFs. The top 50 most statistically significant DEGs and DEIRGs and all DETFs are shown in [Tables S1, S2, S3](#), respectively. A volcano plot and heat map plot were visualized to further exhibit the distribution of all DEGs, DEIRGs, and DETFs (**Figure 1A-F**). The flow diagram of the whole study is illustrated in **Figure 1G**.

Functional annotation and pathway enrichment analysis of DEIRGs

We performed GO and KEGG pathway analyses to further verify the function of DEIRGs in LUAD and LUSC. In the GO analysis, the top significant difference was the cellular component groups. The identification of the termed named extracellular region was GO:0005576 (adjust-

Prognostic value of immune-related genes in non-small cell lung cancer

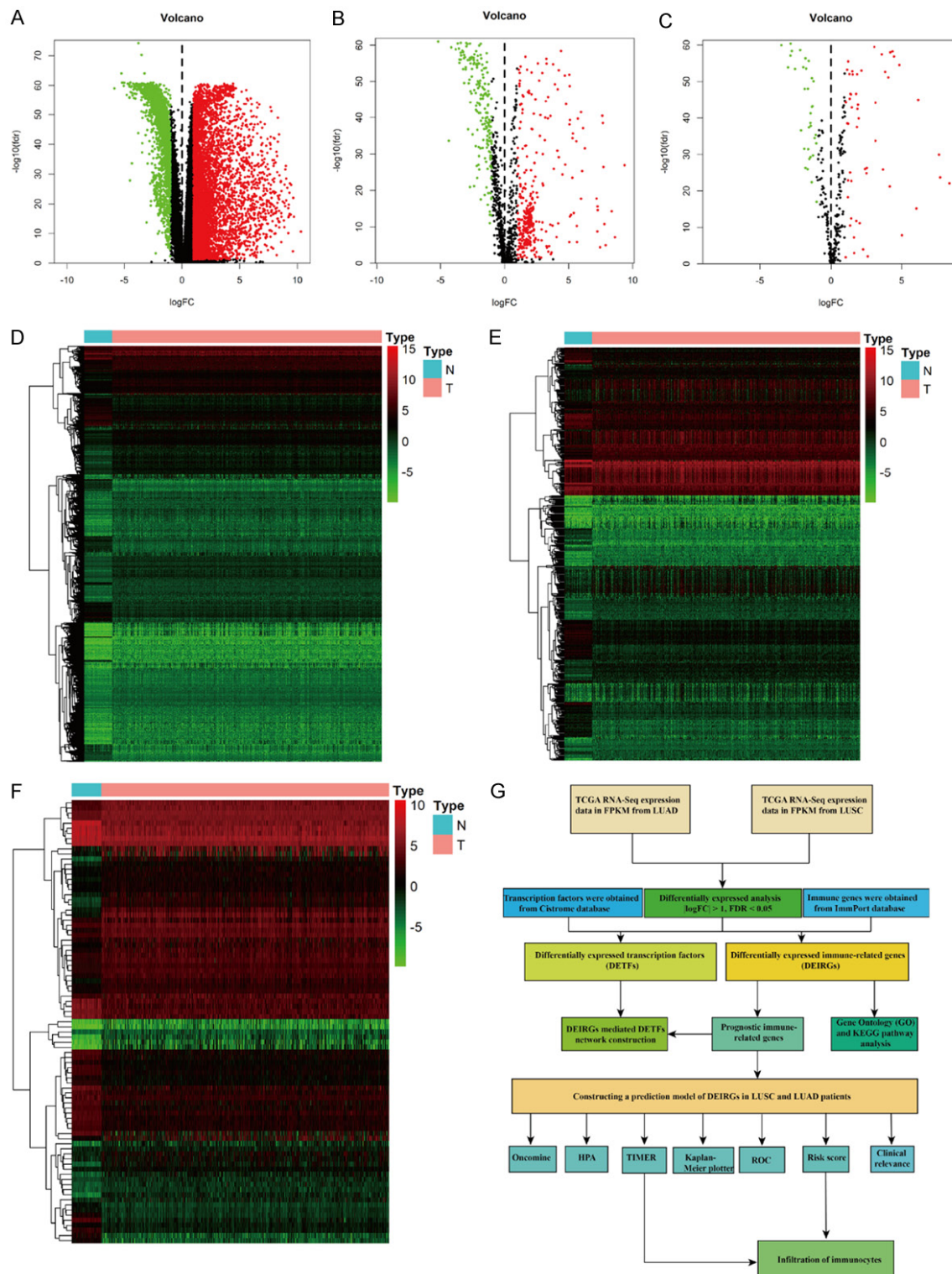


Figure 1. DEGs, DEIRGs, and DETFs between LUAD and LUSC specimens and adjacent normal specimens. A-C. Volcano plot of DEGs, DEIRGs, and DETFs. D-F. Heatmap of DEGs, DEIRGs, and DETFs. G. Flow chart of the research study.

ed P -value = $1.26E-115$). The GO Circle plotting was visualized, showing the top eight GO

terms, including two cellular component terms, one molecular function term, and five biological

process terms ($P < 0.05$) (**Figure 2A, 2B**). The GO Chord plot showed the top 30 significance DEIRGs related to the top eight GO terms (**Figure 2C**). As shown in **Figure 2C**, the top 30 significance DEIRGs were mostly enriched in GO:0005576 extracellular region. IFNG, IGLV6-57, IGHV3-11, IGKV1-5, PPBP, IGKV1-39, IGLV7-43, IGKV4-1, IGLV3-27, and IGLV3-25 were enriched in GO:0006955 immune response. IGLV6-57, IGHV3-11, IGKV1-5, IGKV1-39, IGLV7-43, IGLV3-27, IGKV4-1, and IGLV3-25 were enriched in GO:0050776 regulation of immune response, GO:0006958 complement activation, classical pathway, and GO:0038096 Fc-gamma receptor signaling pathway involved in phagocytosis. KEGG pathway analysis demonstrated that DEIRGs were most enriched in 65 KEGG pathways (**Table S4**). The bar plot showed the top 12 most statistically significant KEGG pathways, including Cytokine-cytokine receptor interaction, Viral protein interaction with cytokine and cytokine receptor, Rheumatoid arthritis, Chemokine signaling pathway, Neuroactive ligand-receptor interaction, JAK-STAT signaling pathway, Hematopoietic cell lineage, Th17 cell differentiation, Natural killer cell mediated cytotoxicity, IL-17 signaling pathway, Graft-versus-host disease, Intestinal immune network for IgA production. DEIRGs were most enriched in the cytokine-cytokine receptor interaction pathway (**Figure 3A**). The dot plot revealed the top 10 KEGG pathways (**Figure 3B**). Furthermore, we constructed the DEIRGs-pathway network using the Cytoscape 3.6.1 software (**Figure 3C**). As shown in **Figure 3C**, there were up-and down regulated DEIRGs were enriched in 65 KEGG pathways.

Cox prediction model based on DEIRGs

We conducted Kaplan-Meier curve analysis to reveal the prognostics value of DEIRGs. We performed univariate and multivariate Cox regression analyses to establish a Cox prediction model. Screening through univariate Cox regression analysis, yielded 33 OS-related DEIRGs ($P < 0.05$). Subsequently, these 33 OS-related DEIRGs were utilized to conduct the multivariate Cox regression analysis. Finally, 15 DEIRGs were selected to establish a Cox prediction model. The weighted relative coefficients were as follows: riskScore value = $(0.09607 \times \text{THBS1 expression} + 0.03865 \times \text{S100P expression} + 0.11711 \times \text{S100A16}$

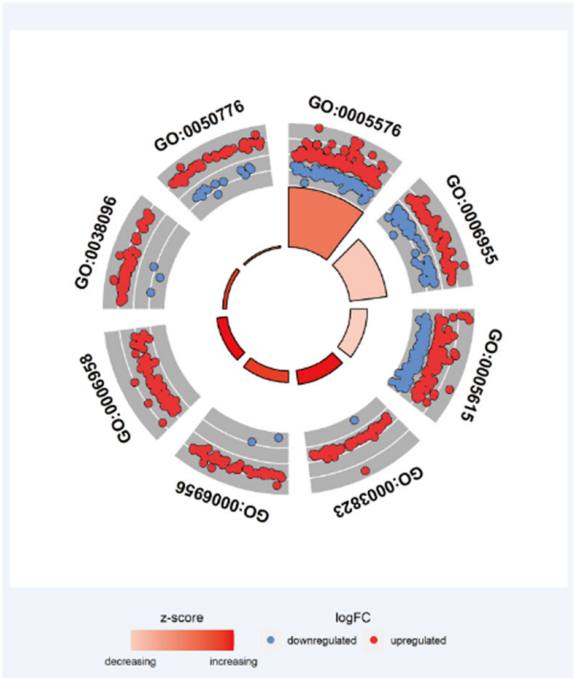
$\text{expression} + 0.07232 \times \text{DKK1 expression} + (-0.14257) \times \text{IL33 expression} + 0.16969 \times \text{CD70 expression} + 0.20890 \times \text{DLL4 expression} + (-0.20366) \times \text{NRTN expression} + 0.12007 \times \text{PDGFB expression} + 0.10827 \times \text{STC2 expression} + 0.13460 \times \text{VGF expression} + (-0.31152) \times \text{GCGR expression} + 0.10294 \times \text{HTR3A expression} + 0.13890 \times \text{LGR4 expression} + (-0.29843) \times \text{SHC3 expression}$). Multivariate Cox regression analyses are shown in **Table S5**. In the context of the median riskScore value, 887 LUAD and LUSC specimens with complete survival status and time data were classified into low- ($n = 444$) and high-risk ($n = 443$) groups. **Figure 4A** illustrates that the high-risk group had a notably poor prognosis compared with the low-risk group ($P = 4.89 \times 10^{-9}$). The area under the receiver operating characteristic curve analysis showed that the Cox prediction model based on DEIRGs achieved better accuracy in the monitoring of survival (**Figure 4B**). The riskScore plot and survival time and status plot are shown in **Figure 4C** and **4D**, respectively. The heatmap plot shows the expression of 15 DEIRGs between the high- and low-risk groups (**Figure 4E**).

Prognostic value of DEIRGs and construction of the DEIRG-related DETF regulatory network

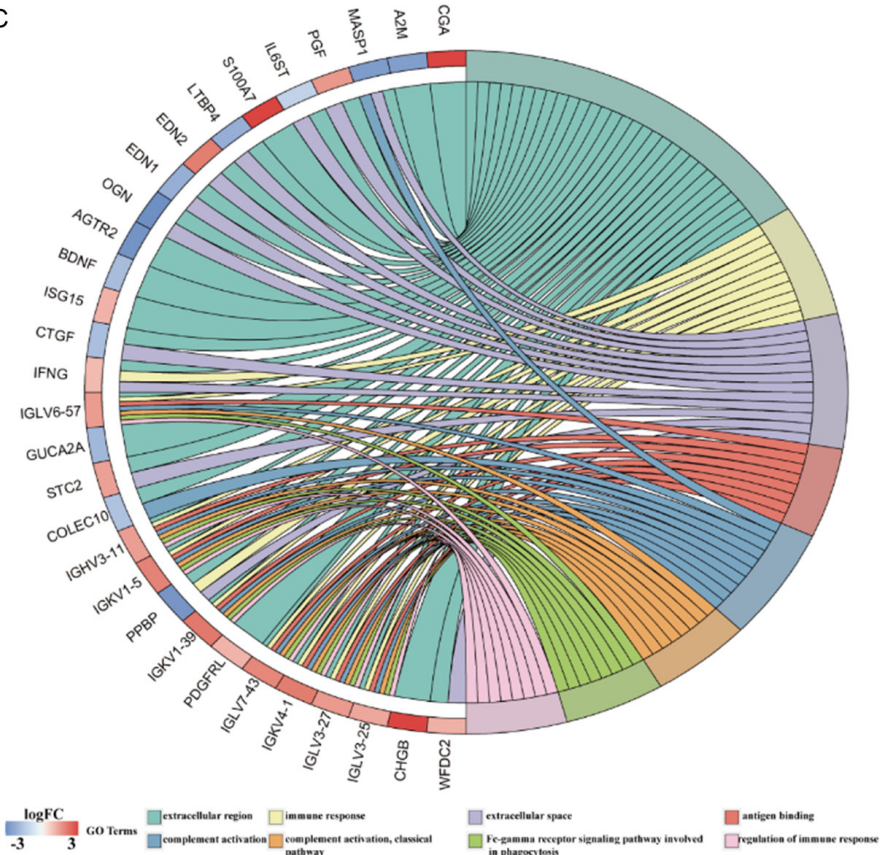
Screening through univariate Cox regression analysis, identified 33 OS-related DEIRGs, including 11 low-risk DEIRGs and 22 high-risk DEIRGs ($P < 0.05$) (**Figure 5A**). To further explore the mechanism of OS-related DEIRGs, we constructed the DEIRG-related DETF network. The network showed that 21 DETFs were associated with 12 DEIRGs. THBS1, HNF4G, DLL4, STC2, IL11, and S100A16 were DEIRGs, which were high-risk DEIRGs in the network. In contrast, ADRB2, IL33, LTB4R2, LTB4R, VIPR1, and SHC3 were low-risk DEIRGs in the network. As shown in **Figure 5B**, the relationship between DEIRG DLL4 and DETF CENPA, DEIRG VIPR1 and DETF NCAPG, DEIRG DLL4 and DETF RARG, DEIRG DLL4 and DETF SANI2, DEIRG DLL4 and DETF TBL1XR1, DEIRG DLL4 and DETF TP63 had a negative regulation in the network. The relationship between DEIRGs and DETFs are shown in **Table S6**. Univariate Cox regression analyses indicated that pathological T stage, pathological TNM stage, risk score, and pathological N stage were related to OS ($P < 0.001$). The pathological M stage was related to OS

Prognostic value of immune-related genes in non-small cell lung cancer

A



C

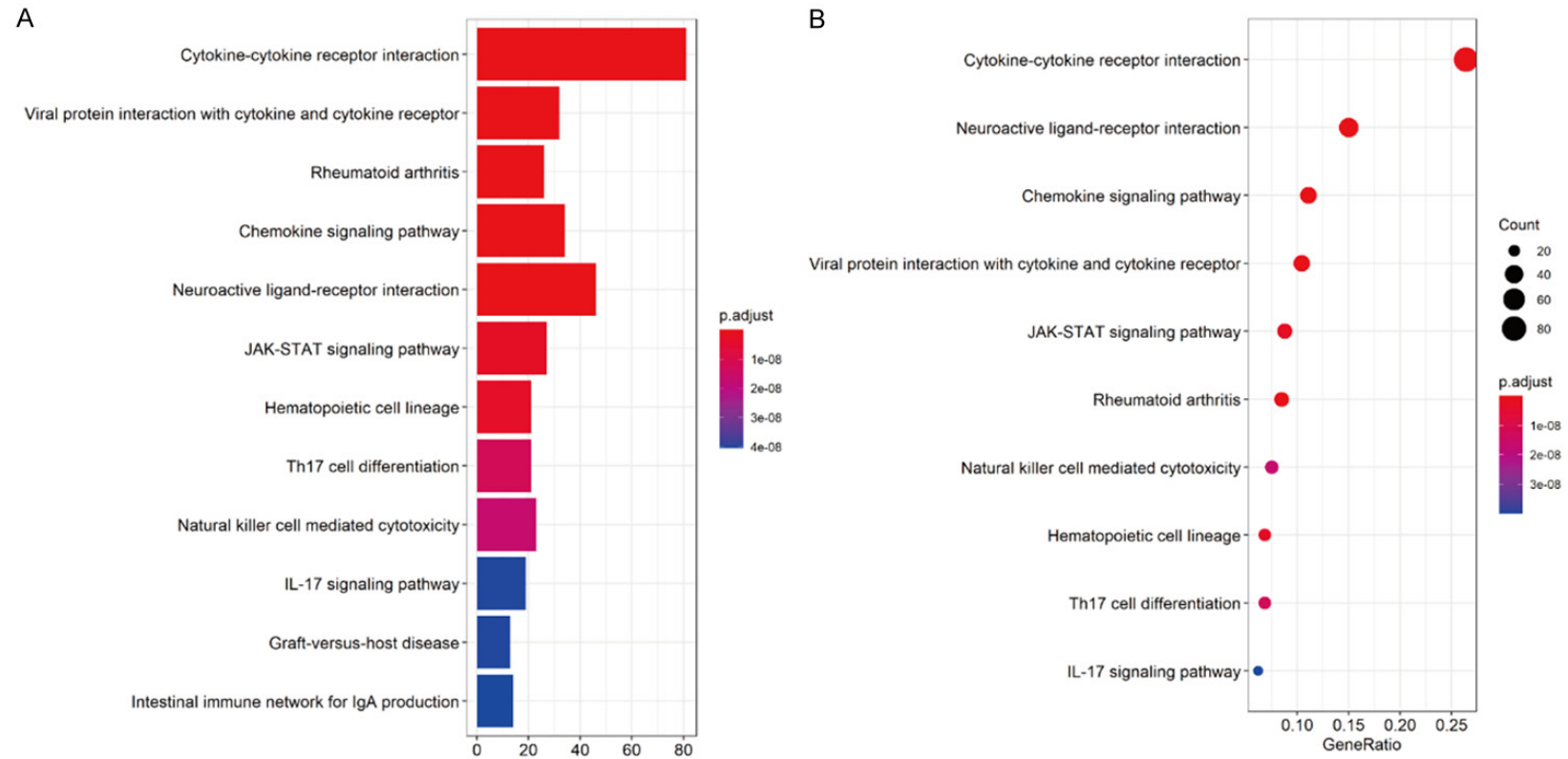


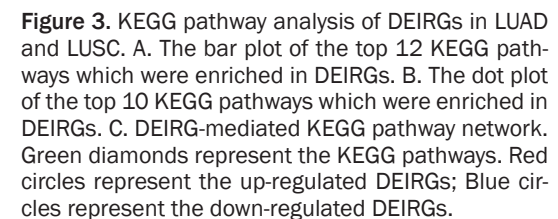
B

Category	ID	Term	adj pval
CC	GO:0005576	extracellular region	1.26E-115
BP	GO:0006955	immune response	3.17E-91
CC	GO:0005615	extracellular space	1.39E-69
MF	GO:0003823	antigen binding	5.51E-67
BP	GO:0006956	complement activation	5.98E-64
BP	GO:0006958	complement activation, classical pathway	8.49E-63
BP	GO:0038096	Fc-gamma receptor signaling pathway involved in phagocytosis	7.74E-56
BP	GO:0050776	regulation of immune response	1.61E-53

Prognostic value of immune-related genes in non-small cell lung cancer

Figure 2. GO analysis of DEIRGs in LUAD and LUSC. A. The inner circle indicates the statistically significant GO terms (log10-adjusted *P* values). The outer circle indicates the logFC of DEIRGs in the GO terms. Blue dots in the GO terms indicate down-regulated DEIRGs. Red dots in the GO terms indicate up-regulated DEIRGs. B. GO Chord indicates the relationship between the top 30 DEIRGs associated with eight enriched GO terms. C. The top eight GO terms.





Prognostic value of immune-related genes in non-small cell lung cancer

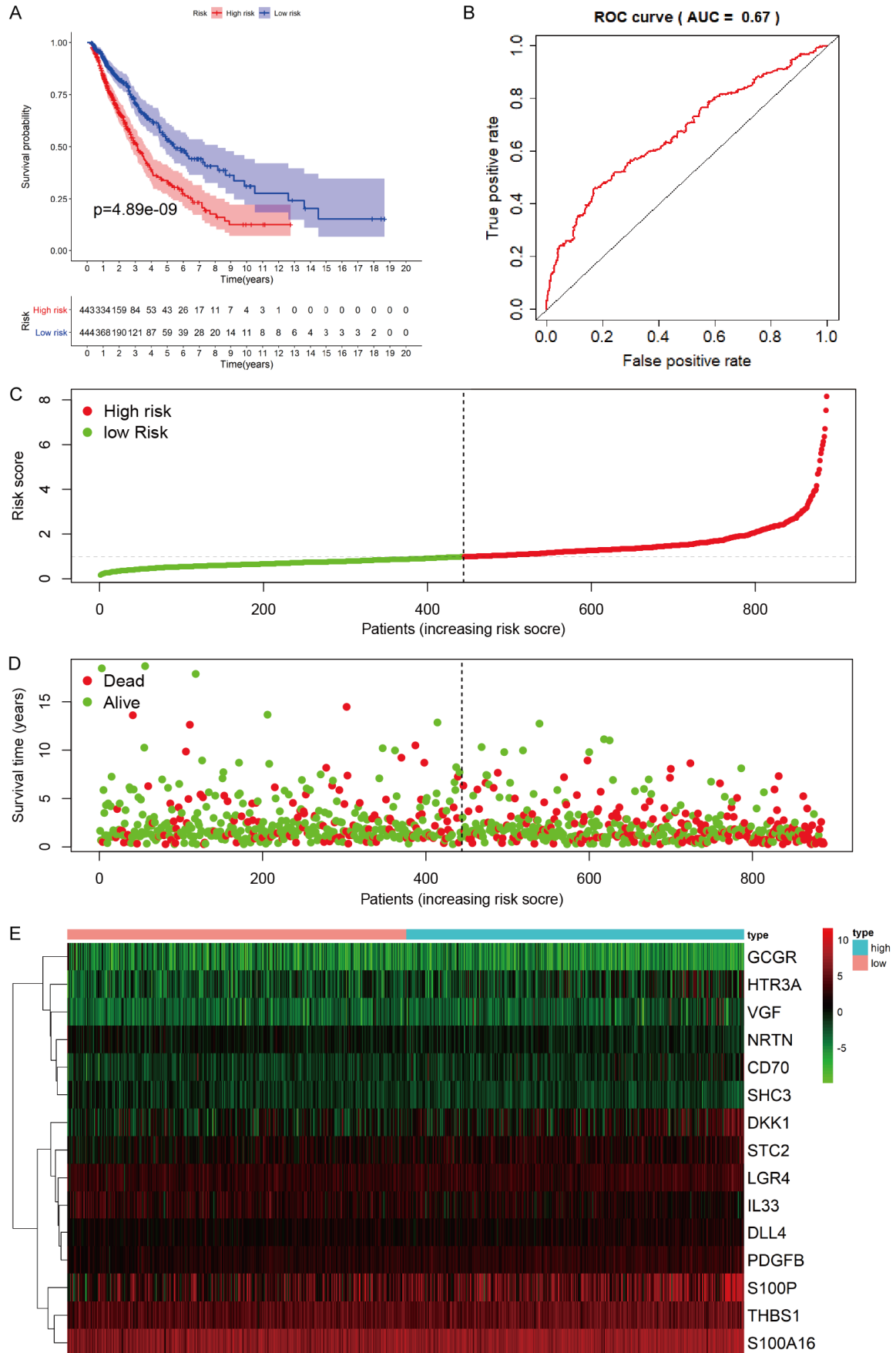


Figure 4. Prognostic model based on DEIRGs in LUAD and LUSC. A. Kaplan-Meier analyses of OS in patients with LUAD and LUSC based on 15 DEIRG signatures. B. ROC analysis of 15 DEIRG signatures. C. The plot of the riskScore based on the prediction model in high- versus low-risk groups. D. The plot of survival status based on the prediction model in low- versus high-risk groups. Red and green dots represent the high- and low-risk groups, respectively. E. The heatmap of DEIRGs in the prediction model between high- and low-risk groups, respectively.

($P < 0.05$) (**Figure 5C**). Multivariate Cox independent prognosis analysis indicated that pathological T stage and riskScore could be perceived as independent prognostic factors for patients with LUAD and LUSC ($P < 0.05$) ([Table S7](#); **Figure 5D**).

Correlation analysis of DEIRGs and clinical pathological characteristics in patients with LUAD and LUSC

We performed a correlation analysis between DEIRGs in the predictive model and the clinical characteristics of LUAD and LUSC specimens' clinical characteristics (pathological TNM stage, age, pathological T stage, sex, pathological N stage, pathological M stage). Nine DEIRGs exhibited significant differences in clinical features ($P < 0.05$). The expression levels of CD70, GCGR, and STC2 were significantly different in patients with lung cancer at pathological M0 stage compared with those at pathological M1 stage ($P < 0.05$) (**Figure 6A-C**). The expression of DLL4, SHC3, STC2, and S100A16 were significantly different in patients with lung cancer (both sexes) ($P < 0.05$) (**Figure 6B, 6E-G**). Differences in the expression levels of IL33, PDGFB and SHC3 were statistically significant between pathological N0 stage and pathological N1-N3 stage ($P < 0.05$) (**Figure 6H-J**). DLL4 expression was significantly different in patients with lung cancer between pathological T3-T4 stage and pathological T1-T2 stage ($P = 0.045$) (**Figure 6K**). The expression of S100A16 were statistically significant in patients with lung cancer aged ≤ 65 and > 65 years ($P = 0.011$) (**Figure 6L**). Differences in the expression levels of S100A16 and S100P were statistically significant between pathological TNM stages III&IV and pathological TNM stages I&II ($P < 0.05$) (**Figure 6M, 6N**). The riskScore was significantly different in pathological N0, N1-N3 stages and pathological TNM stages III&IV and pathological TNM stages I&II ($P < 0.05$) (**Figure 6O, 6P**) ([Table S8](#)).

Validation of prognosis-related DEIRGs

We used the Oncomine and TIMER 2.0 database to validate the mRNA expression level of

prognosis-related DEIRGs. HPA database was performed to verify the protein expression level of prognosis-related DEIRGs. **Figure 7A** showed the heatmap of the mRNA expression level of 15 DEIRGs in 20 cancer types. As we can see from the **Figure 7A**, the mRNA expression levels of S100P, S100A16, DKK1, VGF, HTR3A, and LGR4 were high expression in the heatmap. **Figure 7B** showed the protein expression levels of THBS1, S100P, S100A16, DLL4, PDGFB, LGR4, and SHC between LUAD, LUSC and normal tissues. As shown in **Figure 7B**, the expression levels of S100P, S100A16, and LGR4 were higher in LUAD, LUSC tissues than in normal tissues. Compared with the normal tissues, the expression levels of THBS1, DLL4, PDGFB, and SHC3 were lower in LUSC and LUAD tissues, which is consistent with our results. **Figure 8** showed the expression levels of prognosis-related DEIRGs in 20 cancer types. As shown in **Figure 8**, the mRNA expression levels of S100P, S100A16, NRTN, VGF, GCGR, HTR3A, and LGR4 were significantly highly expressed in LUAD, LUSC tissues than in normal tissues (**Figure 8A-G**). Compared with the normal tissues, the mRNA expression levels of THBS1, DLL4, IL33, PDGFB, SHC3 were significantly lower in LUAD, LUSC tissues (**Figure 8H-L**).

Kaplan-Meier plotter analysis of prognosis-related DEIRGs

We used the Kaplan-Meier plotter database to perform the survival analysis of prognosis-related DEIRGs. **Figure 9** showed that 13 prognosis-related DEIRGs were associated with overall survival ($P < 0.05$), while the high-risk DEIRG LGR4 and the low-risk DEIRG NRTN were not correlated with overall survival in NSCLC ($P > 0.05$).

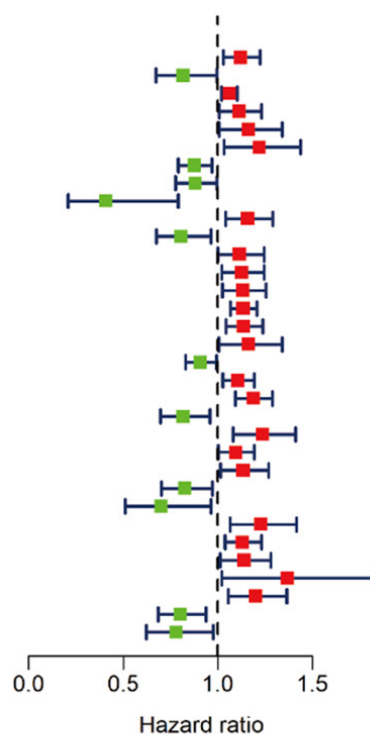
Correlation analysis of DEIRGs in the prediction model and infiltration of immunocytes

To further investigate whether the DEIRGs identified in the prediction model reflected the status of lung cancer IME, we performed a correlation analysis between immunocyte infiltration and the DEIRGs riskScore. As shown in **Figure**

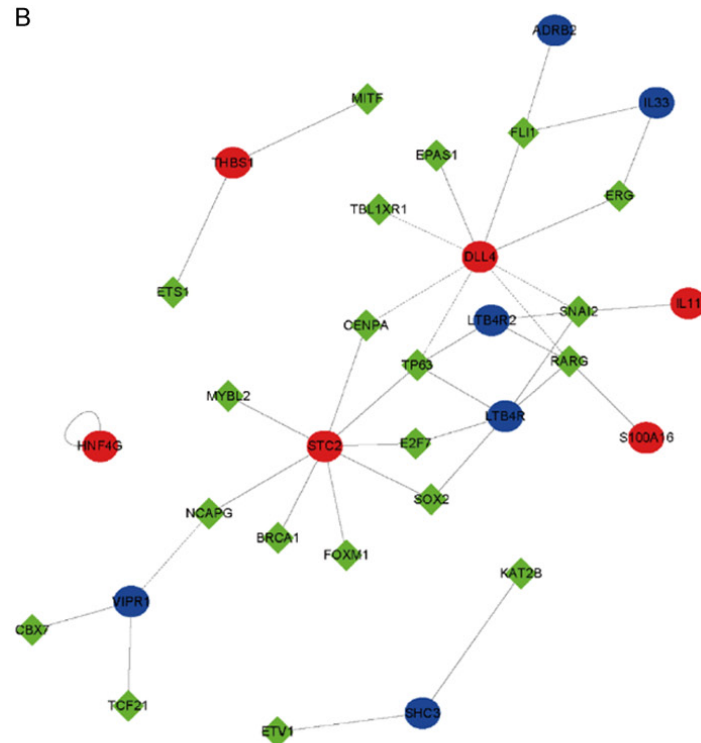
Prognostic value of immune-related genes in non-small cell lung cancer

A

	pvalue	Hazard ratio
THBS1	0.009	1.122(1.029-1.223)
CAMP	0.044	0.819(0.674-0.995)
S100P	0.003	1.062(1.020-1.105)
S100A16	0.032	1.114(1.009-1.230)
DLL4	0.037	1.163(1.009-1.341)
TNFSF11	0.017	1.220(1.035-1.437)
LTB4R	0.011	0.877(0.793-0.970)
SYTL1	0.045	0.882(0.779-0.997)
PIK3R2	0.008	0.407(0.209-0.792)
SEMA3A	0.007	1.159(1.042-1.290)
LTB4R2	0.018	0.808(0.677-0.964)
APLN	0.049	1.116(1.000-1.245)
BMP2	0.017	1.128(1.022-1.245)
CD70	0.016	1.134(1.024-1.255)
DKK1	<0.001	1.135(1.067-1.206)
GREM1	0.003	1.137(1.044-1.239)
IL11	0.042	1.162(1.006-1.342)
IL33	0.034	0.908(0.830-0.993)
INHA	0.008	1.106(1.027-1.192)
INSL4	<0.001	1.188(1.095-1.289)
NRTN	0.014	0.819(0.698-0.961)
PDGFB	0.002	1.237(1.083-1.411)
STC2	0.040	1.095(1.004-1.193)
VEG	0.027	1.135(1.015-1.269)
ADRB2	0.022	0.828(0.704-0.973)
GCGR	0.029	0.702(0.511-0.964)
HNF4G	0.005	1.228(1.066-1.416)
HTR3A	0.005	1.130(1.038-1.231)
LGR4	0.029	1.140(1.013-1.282)
MCHR1	0.035	1.369(1.023-1.833)
TUBB3	0.005	1.201(1.056-1.366)
VIPR1	0.006	0.804(0.688-0.939)
SHC3	0.031	0.781(0.624-0.978)

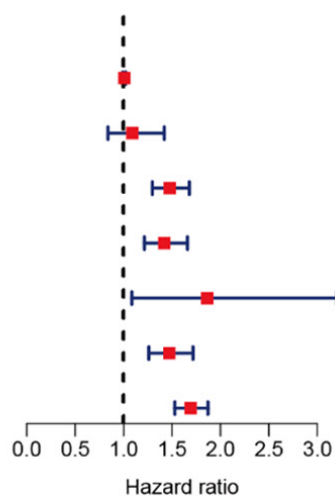


B



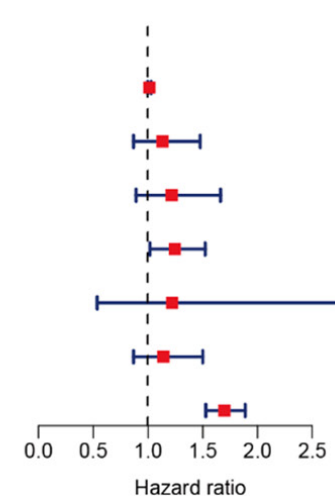
C

	pvalue	Hazard ratio
age	0.328	1.007(0.993-1.021)
gender	0.517	1.091(0.839-1.419)
stage	<0.001	1.476(1.297-1.679)
T	<0.001	1.420(1.215-1.659)
M	0.024	1.864(1.085-3.204)
N	<0.001	1.471(1.260-1.717)
riskScore	<0.001	1.691(1.529-1.871)



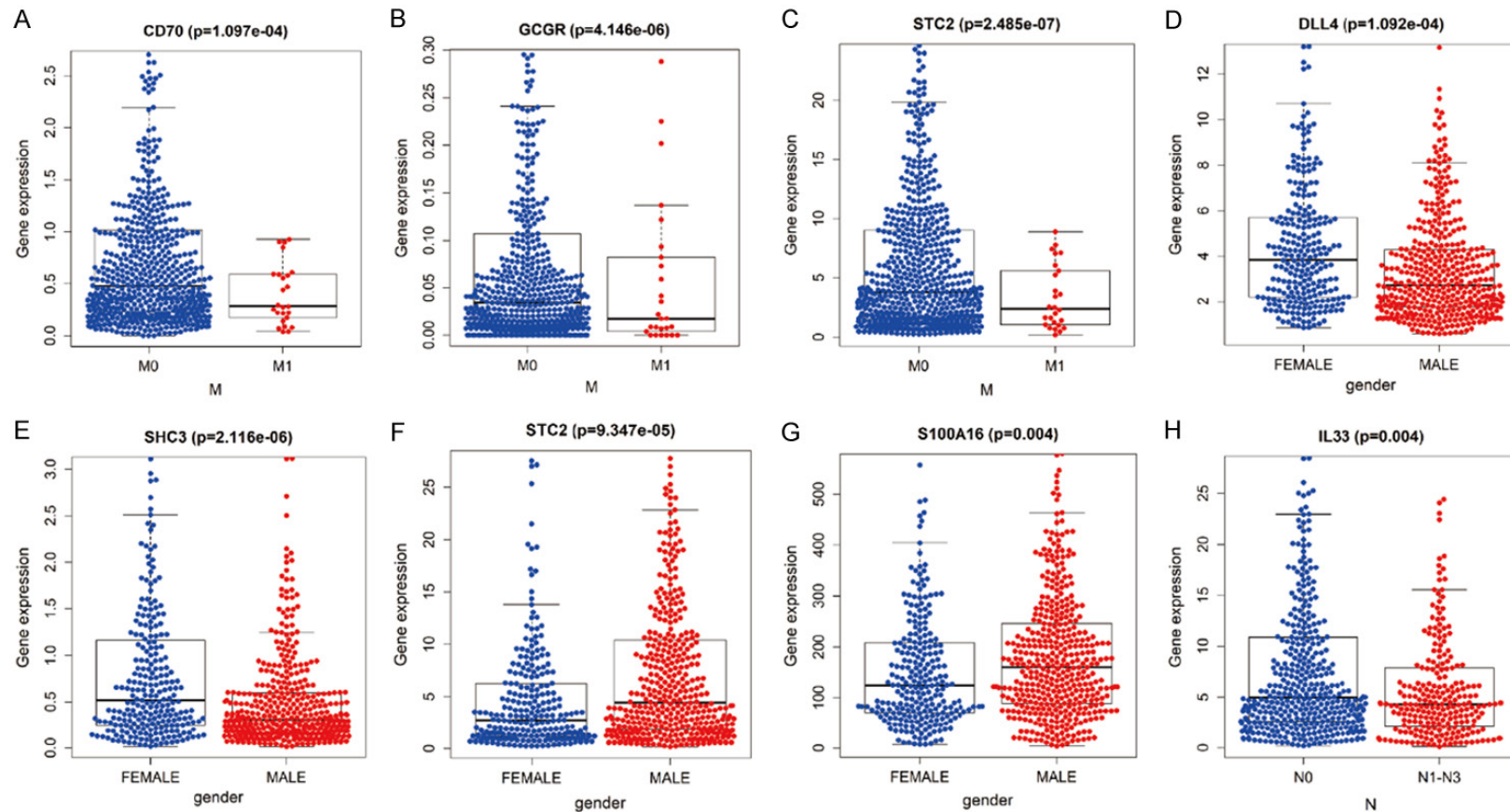
D

	pvalue	Hazard ratio
age	0.117	1.012(0.997-1.027)
gender	0.364	1.131(0.867-1.476)
stage	0.220	1.216(0.890-1.662)
T	0.033	1.245(1.018-1.522)
M	0.636	1.220(0.536-2.774)
N	0.347	1.140(0.867-1.500)
riskScore	<0.001	1.697(1.527-1.887)



Prognostic value of immune-related genes in non-small cell lung cancer

Figure 5. Prognostic value of 15 DEIRG signatures in LUAD and LUSC. A. Visualization of the forest plot of prognosis-related DEIRGs. B. DEIRG-related DETF regulatory network. Green diamonds indicate DETFs. Blue circles represent downregulated DEIRGs. Red circles represent upregulated DEIRGs. Solid and dashed lines represent the positive and negative regulatory relationships, respectively. C. Visualization of the forest map with single factor independent prognosis. D. Visualization of the forest map with multi-factor independent prognosis. The red of the forest map represents that the clinical-pathological characteristic could be used as a high-risk factor. The green of the forest map represents that the clinical-pathological characteristic could be used as a low-risk factor.



Prognostic value of immune-related genes in non-small cell lung cancer

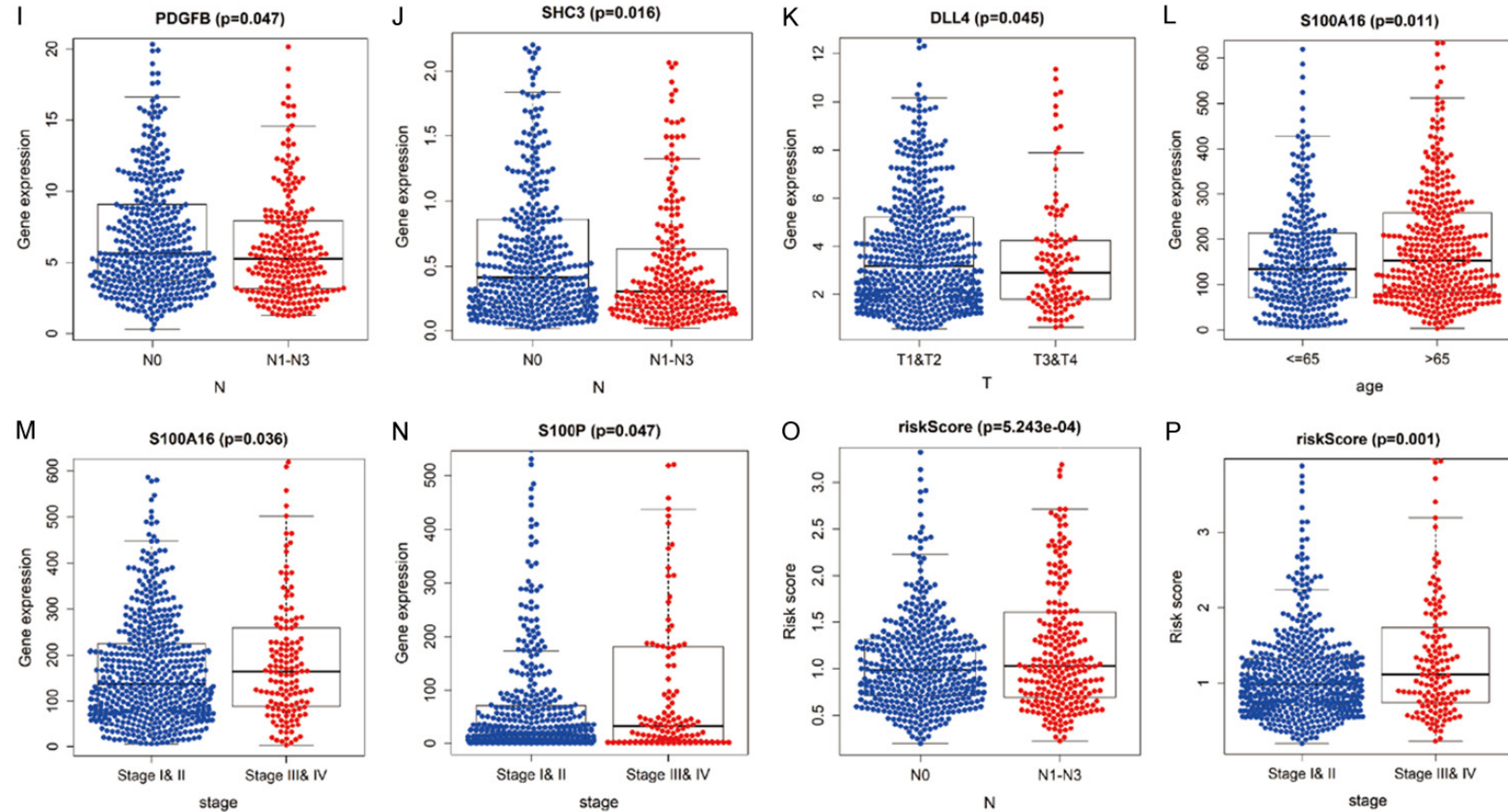


Figure 6. Comparison of DEIRG expression levels in different pathological features in LUAD and LUSC. A-C. Differences in the expression of DEIRGs between the pathological M0/M1 stages. D-G. Differences in the expression of DEIRGs between the sexes (male/female). H-J. Differences in the expression of DEIRGs between the pathological N0/N1-N3 stages. K. Differences in the expression of DEIRGs between the pathological T1-T2/T3-T4 stages. L. Differences in the expression of DEIRGs in terms of age (≤ 65 / > 65 years). M, N. Differences in the expression of DEIRGs between the pathological TNM stages (I&II/III&IV). O, P. Differences in the risk score of DEIRGs in pathological N0/N1-N3 stages and pathological TNM stages (I&II/III&IV).

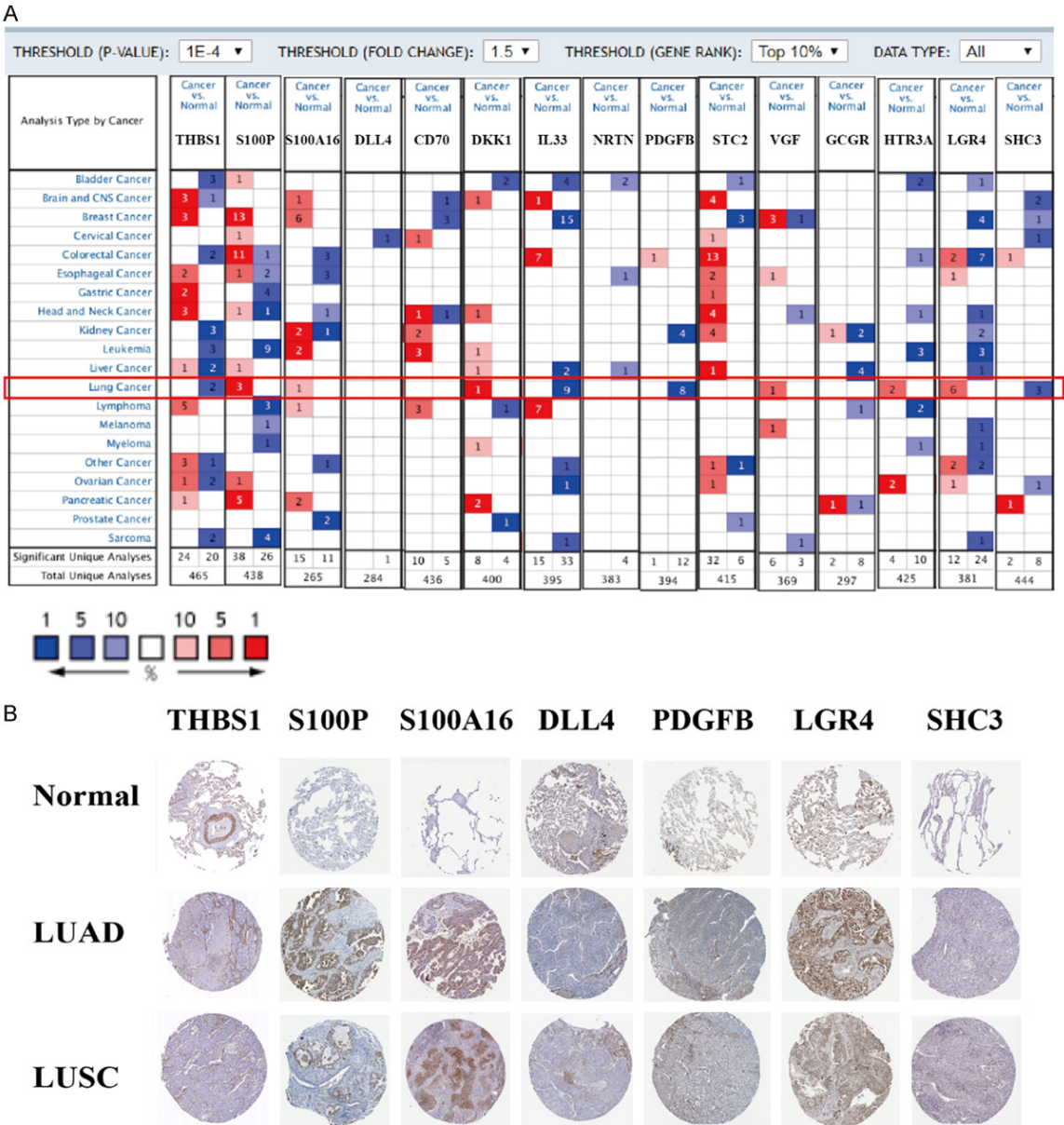


Figure 7. Validation of prognosis-related DEIRGs of NSCLC at mRNA and protein level. A. The heatmap of gene summary of mRNA expression levels in 20 cancer types from Oncomine database. Red indicate up-regulated, blue indicate down-regulated. B. The protein expression levels of THBS1, S100P, S100A16, DLL4, PDGFB, LGR4, and SHC3 for NSCLC tissues and normal tissues in immunohistology assay from HPA database.

10, B cells had a negative relationship with the DEIRGs riskScore ($P < 0.05$). In contrast, neutrophils had a positive relationship with the DEIRGs riskScore ($P < 0.05$), which may provide a novel horizon for investigating the lung cancer IME.

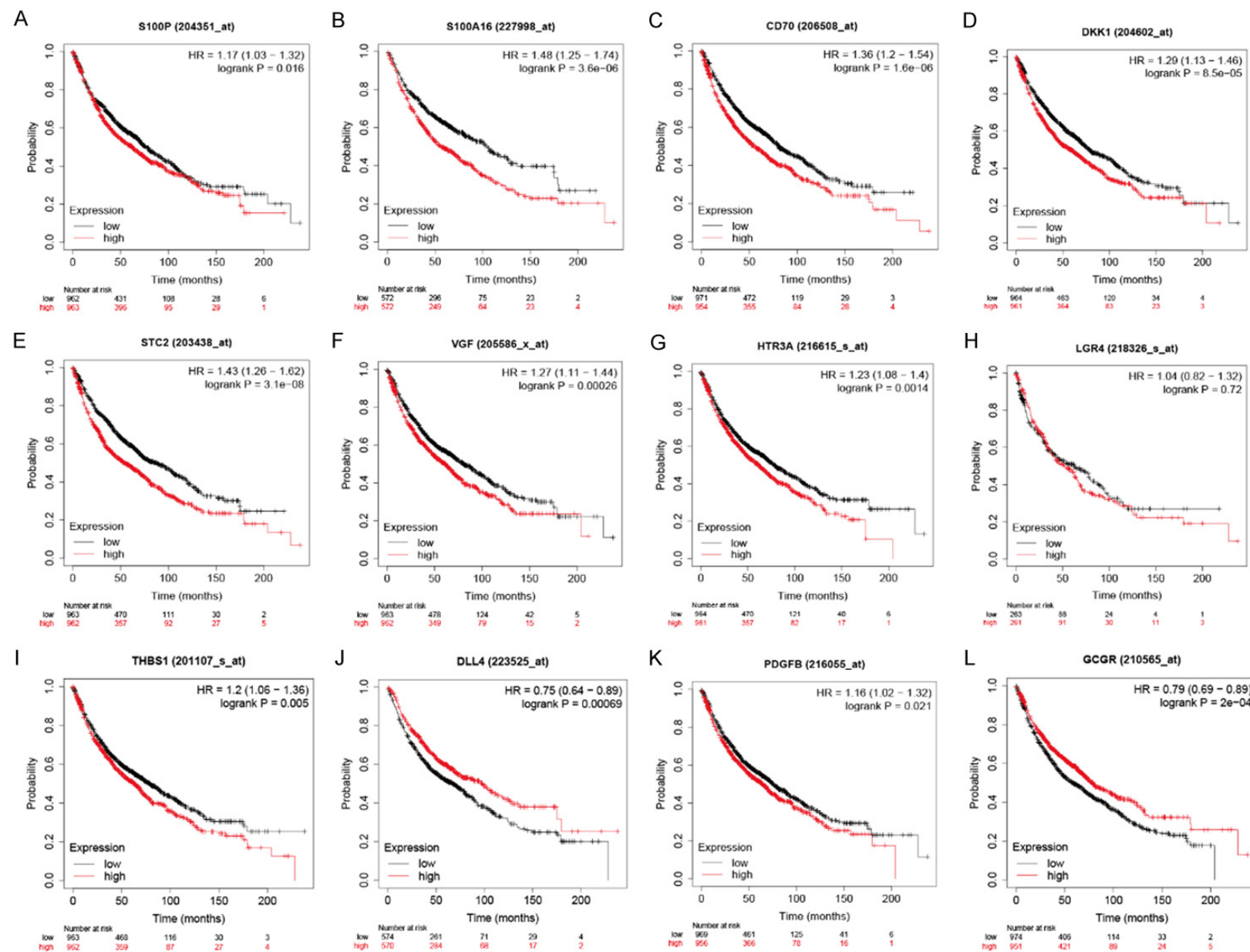
Discussion

Recently, the great significance of IRGs in the development of tumors and immunology has

been recognized [30-33]. However, RNA sequencing associated with DEIRGs and the construction of the prediction model has not been explored thus far. Therefore, it is of crucial importance to verify underlying biomarkers for predicting the OS of patients with LUAD and LUSC. In the present study, we first used the limma R package to identify DEGs, followed by analyses of the ImmPort and Cistrome Cancer databases to obtain DEIRGs and DETFs, respectively. Subsequently, the DEIRG-medi-

Figure 8. Validation of prognosis-related DEIRGs in TIMER 2.0 database. A-F. The mRNA expression levels of S100P, S100A16, NRTN, VGF, GCGR and HTR3A in all kinds of cancer types. G-L. The mRNA expression levels of LGR4, THBS1, DLL4, IL33, PDGFB, SHC3 in all kinds of cancer types. Red in the box plot indicates tumor tissues. Blue in the box plot indicates normal tissues. * represent $P < 0.05$, ** represent $P < 0.01$, *** represent $P < 0.001$.

Prognostic value of immune-related genes in non-small cell lung cancer



Prognostic value of immune-related genes in non-small cell lung cancer

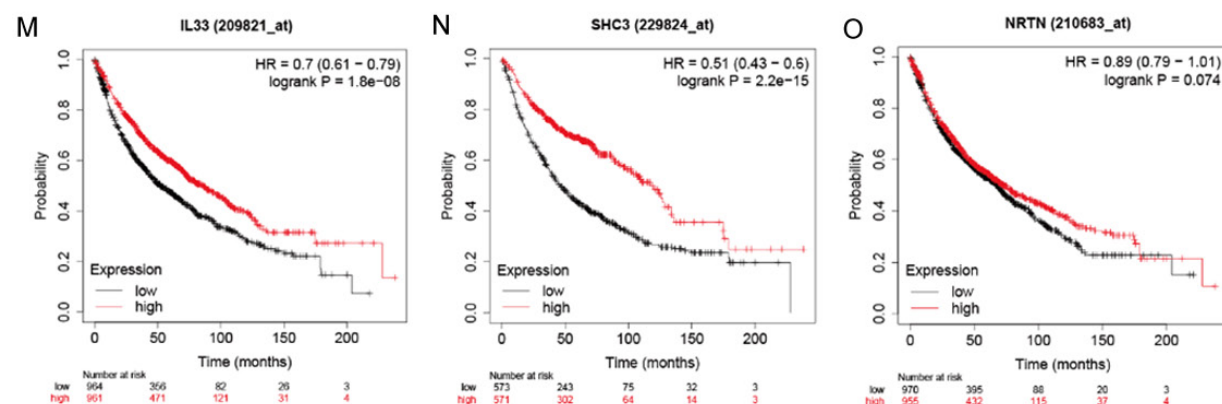


Figure 9. Kaplan-Meier plotter survival analysis of 15 OS-related DEIRGs. A-K. The survival analysis of 11 high-risk DEIRGs. L-O. The survival analysis of 4 low-risk DEIRGs.

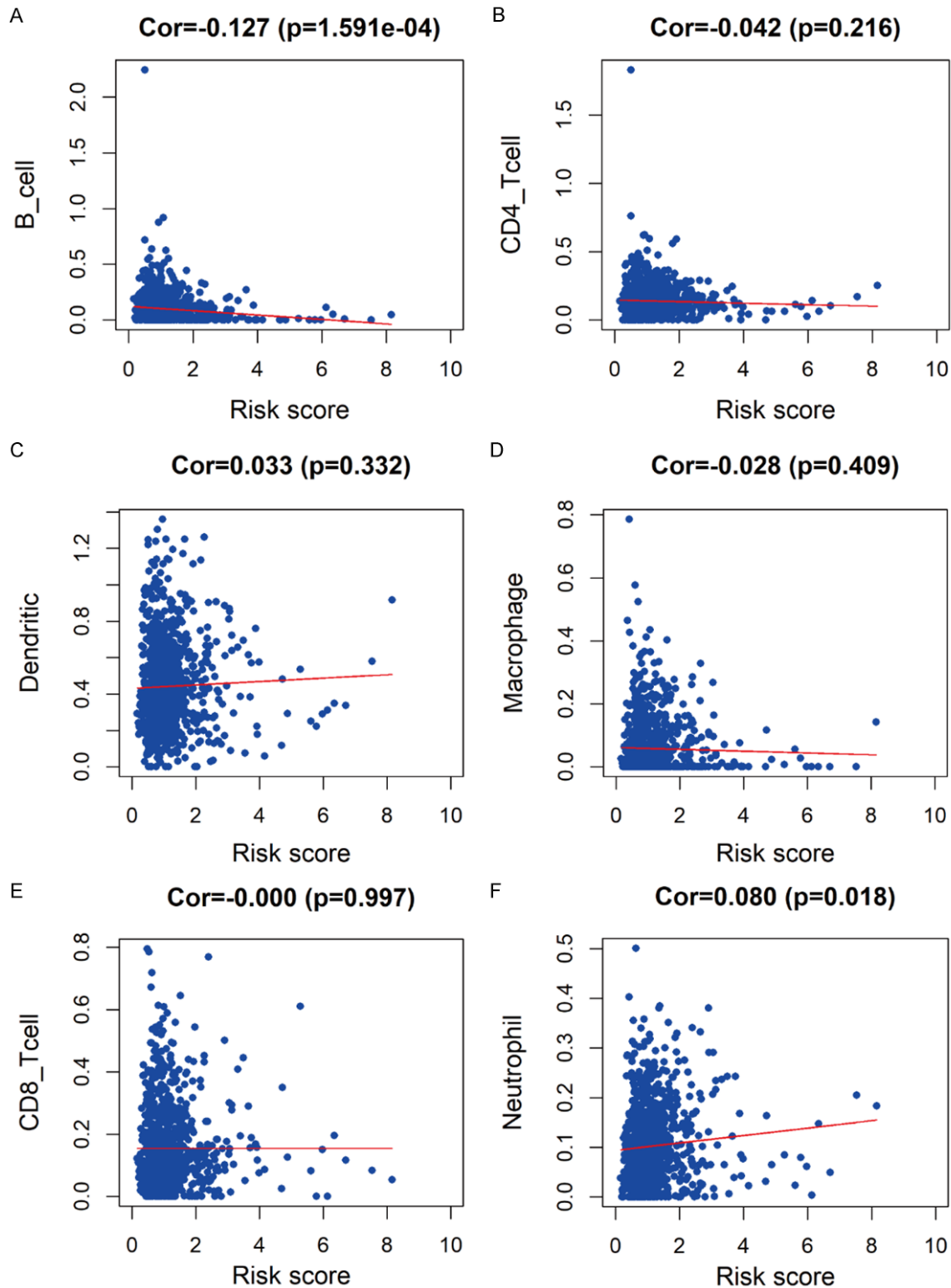


Figure 10. Correlation analysis of the prognostic value and six types of immunocyte infiltration. A. B-cells. B. CD4-T cells. C. Dendritic cells. D. Macrophages. E. CD8-T cells. F. Neutrophils.

ed DETF regulatory network was established. Using univariate and multivariate Cox regres-

sion analyses, a Cox prediction model based on DEIRGs was established to determine whether

these DEIRGs could be used as independent prognostic factors in patients with LUAD and LUSC. Furthermore, immunocyte estimation in the TIMER database was used to perform the correlation analysis between immunocytes and the riskScore of the prediction model. In this study, a prediction model associated with DEIRGs was adopted for monitoring immunocyte infiltration and estimating the prognosis of LUAD and LUSC.

In recent years, the identification of potential molecular targets associated with the tumor IME for the investigation of promising prognostic biomarkers related to OS has attracted considerable attention [26, 34-37]. A study indicated that a prediction model based on four IRGs could estimate the prognosis of LUAD [30]. Recently, it was demonstrated that a prognostic genetic signature associated with the tumor IME could predict the survival of patients with NSCLC [38]. A research study revealed that an immune response related to mutational signatures could predict the prognosis of NSCLC [39]. The expression of PD-L1 expression could be used as a potential prognostic biomarker for predicting responses to immunotherapy and prognosis of NSCLC [40]. The molecular expression subtypes of LUSC and LUAD indicate that determining the tumor expression subtype could be used as a potential prognostic biomarker for immunotherapy [41]. However, the tumor IME related to IRGs and potential prognostic biomarkers based on the prediction model in LUAD and LUSC remain unknown.

In this study, we initially obtained DEGs, DEIRGs, and DETFs. Subsequently, a DEIRG-related DETF regulatory network was constructed. Furthermore, a prediction model based on DEIRGs was established to detect the DEIRGs that could be perceived as independent prognostic factor. Compared with previous studies, in the present study, a Cox prediction model based on DEIRGs was initially constructed, followed by DEIRG-mediated DETF regulatory network to investigate the mechanism of DEIRGs. Meanwhile, the TIMER database provided an estimation of the infiltration of immunocytes.

To further investigate the underlying mechanism of DEIRGs at the molecular level, we constructed a DEIRG-mediated DETF regulatory network, which revealed significant DETFs reg-

ulating DEIRGs in the network. According to the network, we hypothesized that ETS1 and MITF could regulate THBS1, ETV1 and KAT2B could regulate SHC3. In the network, VIPR1 was down-regulated in both LUAD and LUSC. Consistent with our study, a recent study showed that VIPR1 plays an important role in patients with LUAD as a tumor suppressor [42]. In the network, we hypothesized that TF CBX7, TF TCF21 and TF NCAPG could regulate VIPR1, which provided a novel insight for exploring the DEIRG VIPR1 at the molecular level.

In the present study, we conducted univariate and multivariate Cox regression analyses and established a Cox prediction model based on DEIRGs was established. The 15 DEIRG signatures provided a novel horizon to forecast the OS of patients with LUAD and LUSC. Furthermore, an independent prognosis analysis was performed, revealing the 15 DEIRGs (GCCR, HTR3A, VGF, NRTN, CD70, SHC3, DKK1, STC2, LGR4, IL33, DLL4, PDGFB, S100P, THBS1, and S100A16), which could be perceived as independent prognostic factors to forecast the survival of patients with LUSC and LUAD. CD70, which acts as a tumor necrosis factor, may play a pivotal role in forecasting the prognosis of patients with malignant pleural mesothelioma [43]. CD70 is a potential molecular target, which plays a significant role in cancer immunotherapy [44]. STC2 may act as a novel target for potential interventions against glioblastoma [45]. LGR4/GPR48 could be used as a potential unique target in regulating TLR2/4 mediated autoimmune diseases [46]. Compared with previous investigations, our study was the first identify 15 DEIRG signatures that could be used as independent prognostic factors in patients with LUAD and LUSC. This evidence, may provide a new horizon for forecasting the survival of patients with LUAD and LUSC.

According to the correlation analysis of DEIRGs based on the prediction model and clinical features of LUAD and LUSC, a total of nine DEIRGs were associated with the clinical characteristics. More significantly, our study was the first to access the correlation of the riskScore of the prediction model and six types of immunocyte infiltrations in TIMER, which may provide important insight into monitoring the status of the tumor IME.

Conclusions

Our study identified a Cox prediction model based on DEIRGs from TCGA, ImmPort, and Cistrome Cancer databases for predicting the OS of patients with NSCLC. Furthermore, multivariate Cox regression analysis was performed to verify 15 prognosis-related DEIRGs could be used as independent prognostic factors. Moreover, Oncomine, HPA, TIMER, and Kaplan-Meier plotter databases validated the prognosis-related DEIRGs at mRNA and protein level. Finally, the immunocyte infiltration analysis provided a new insight into monitoring the status of the NSCLC IME.

Acknowledgements

We thank the Charlesworth Group's author services for professional language editing. This work was supported by grants from the Major Scientific and Technological Innovation Project of Shandong Province (grant no. 2018CXGC-1212), the Science and Technology Foundation of Shandong Province (grant no. 2014GSF-118084), the CSCO-Qilu Cancer Research Fund (grant no. Y-Q201802-014), the Medical and Health Technology Innovation Plan of Jinan City (grant no. 201805002) and the National Natural Science Foundation of China (grant no. 81372333).

Disclosure of conflict of interest

None.

Abbreviations

DEGs, differentially expressed genes; DEIRGs, differentially expressed immune-related genes; DETF, differentially expressed transcription factor; GO, Gene Ontology; IME, immune microenvironment; ImmPort, Immunology Database and Analysis Portal; IRG, immune-related gene; KEGG, Kyoto Encyclopedia of Genes and Genomes; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; NSCLC, non-small cell lung cancer; OS, overall survival; TCGA, The Cancer Genome Atlas; TF, transcription factor; TIMER, Tumor Immune Estimation Resource.

Address correspondence to: Dr. Yi-Qing Qu, Department of Pulmonary and Critical Care Medicine,

Qilu Hospital of Shandong University, Wenhua Road 107#, Jinan 250012, China. Tel: +86-531-8216-9335; Fax: +86-531-8296-7544; E-mail: quyiqing@sdu.edu.cn

References

- [1] Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J and Jemal A. Global cancer statistics, 2012. *CA Cancer J Clin* 2015; 65: 87-108.
- [2] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016; 66: 115-132.
- [3] Travis WD, Brambilla E, Noguchi M, Nicholson AG, Geisinger K, Yatabe Y, Ishikawa Y, Wistuba I, Flieder DB, Franklin W, Gazdar A, Hasleton PS, Henderson DW, Kerr KM, Petersen I, Roggli V, Thunnissen E and Tsao M. Diagnosis of lung cancer in small biopsies and cytology: implications of the 2011 international association for the study of lung cancer/American thoracic society/European respiratory society classification. *Arch Pathol Lab Med* 2013; 137: 668-684.
- [4] Travis W, Brambilla E, Noguchi M, Nicholson A, Geisinger K, Yatabe Y, Ishikawa Y, Wistuba I, Flieder D, Franklin W, Gazdar A, Hasleton P, Henderson D, Kerr K, Nakatani Y, Petersen I, Roggli V, Thunnissen E and Tsao M. Diagnosis of lung adenocarcinoma in resected specimens implications of the 2011 international association for the study of lung cancer/American thoracic society/European respiratory society classification. *Arch Pathol Lab Med* 2013; 137: 685-705.
- [5] Cagle PT, Allen TC and Olsen RJ. Lung cancer biomarkers: present status and future developments. *Arch Pathol Lab Med* 2013; 137: 1191-1198.
- [6] Goldstraw P, Crowley J, Chansky K, Giroux D, Groome P, Rami-Porta R, Postmus P, Rusch V and Sobin L; International Association for the Study of Lung Cancer International Staging Committee. The IASLC lung cancer staging project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM classification of malignant tumours. *J Thorac Oncol* 2007; 2: 706-714.
- [7] Shi X, Li R, Dong X, Chen AM, Liu X, Lu D, Feng S, Wang H and Cai K. IRGS: an immune-related gene classifier for lung adenocarcinoma prognosis. *J Transl Med* 2020; 18: 55.
- [8] Duan S, Wang P, Liu F, Huang H, An W, Pan S and Wang X. Novel immune-risk score of gastric cancer: a molecular prediction model combining the value of immune-risk status and chemosensitivity. *Cancer Med* 2019; 8: 2675-2685.

- [9] Wang S, Yang L, Ci B, Maclean M, Gerber DE, Xiao G and Xie Y. Development and validation of a nomogram prognostic model for SCLC patients. *J Thorac Oncol* 2018; 13: 1338-1348.
- [10] Chang CF, Huang PW, Chen JS, Chen YY, Lu CH, Chang PH, Hung YS and Chou WC. Prognostic factors for advanced pancreatic cancer treated with gemcitabine Plus S-1: retrospective analysis and development of a prognostic model. *Cancers (Basel)* 2019; 11: 57.
- [11] Liu A, Hou F, Qin Y, Song G, Xie B, Xu J and Jiao W. Predictive value of a prognostic model based on pathologic features in lung invasive adenocarcinoma. *Lung Cancer* 2019; 131: 14-22.
- [12] Ahlberg J, Giragossian C, Li H, Myzithras M, Raymond E, Caviness G, Grimaldi C, Brown SE, Perez R, Yang D, Kroe-Barrett R, Joseph D, Pamulapati C, Coble K, Ruus P, Woska JR, Ganesan R, Hansel S and Mbow ML. Retrospective analysis of model-based predictivity of human pharmacokinetics for anti-IL-36R monoclonal antibody MAB92 using a rat anti-mouse IL-36R monoclonal antibody and RNA expression data (FANTOM5). *MAbs* 2019; 11: 956-964.
- [13] Lu Y, Zhou X, Liu Z, Wang B, Wang W and Fu W. Assessment for risk status of colorectal cancer patients: a novel prediction model based on immune-related genes. *DNA Cell Biol* 2020; 39: 958-964.
- [14] Zhao X, Liu J, Liu S, Yang F and Chen E. Construction and validation of an immune-related prognostic model based on TP53 status in colorectal cancer. *Cancers (Basel)* 2019; 11.
- [15] Song Q, Shang J, Yang Z, Zhang L, Zhang C, Chen J and Wu X. Identification of an immune signature predicting prognosis risk of patients in lung adenocarcinoma. *J Transl Med* 2019; 17: 70.
- [16] Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, Berger P, Desborough V, Smith T, Campbell J, Thomson E, Monteiro R, Guimaraes P, Walters B, Wiser J and Butte AJ. ImmPort: disseminating data to the public for the future of immunology. *Immunol Res* 2014; 58: 234-239.
- [17] Mei S, Meyer CA, Zheng R, Qin Q, Wu Q, Jiang P, Li B, Shi X, Wang B, Fan J, Shih C, Brown M, Zang C and Liu XS. Cistrome cancer: a web resource for integrative gene regulation modeling in cancer. *Cancer Res* 2017; 77: e19-e22.
- [18] Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC and Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol* 2003; 4: P3.
- [19] Walter W, Sánchez-Cabo F and Ricote M. GOplot: an R package for visually combining expression data with functional analysis. *Bioinformatics* 2015; 31: 2912-4.
- [20] Yu G, Wang LG, Han Y and He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012; 16: 284-287.
- [21] Reimand J, Isserlin R, Voisin V, Kucera M, Tanus-Lopes C, Rostamianfar A, Wadi L, Meyer M, Wong J, Xu C, Merico D and Bader GD. Pathway enrichment analysis and visualization of omics data using g: Profiler, GSEA, Cytoscape and EnrichmentMap. *Nat Protoc* 2019; 14: 482-517.
- [22] Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, Barrette TR, Anstet MJ, Kincead-Beal C, Kulkarni P, Varambally S, Ghosh D and Chinnaiyan AM. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia* 2007; 9: 166-180.
- [23] Ponten F, Schwenk JM, Asplund A and Edqvist PH. The human protein atlas as a proteomic resource for biomarker discovery. *J Intern Med* 2011; 270: 428-446.
- [24] Yang S, Liu T, Nan H, Wang Y, Chen H, Zhang X, Zhang Y, Shen B, Qian P, Xu S, Sui J and Liang G. Comprehensive analysis of prognostic immune-related genes in the tumor microenvironment of cutaneous melanoma. *J Cell Physiol* 2020; 235: 1025-1035.
- [25] Sun CC, Zhou Q, Hu W, Li SJ, Zhang F, Chen ZL, Li G, Bi ZY, Bi YY, Gong FY, Bo T, Yuan ZP, Hu WD, Zhan BT, Zhang Q, Tang QZ and Li DJ. Transcriptional E2F1/2/5/8 as potential targets and transcriptional E2F3/6/7 as new biomarkers for carcinoma. *Aging (Albany NY)* 2018; 10: 973-987.
- [26] Pan JH, Zhou H, Cooper L, Huang JL, Zhu SB, Zhao XX, Ding H, Pan YL and Rong L. LAYN is a prognostic biomarker and correlated with immune infiltrates in gastric and colon cancers. *Front Immunol* 2019; 10: 6.
- [27] Sun CC, Li SJ, Hu W, Zhang J, Zhou Q, Liu C, Li LL, Songyang YY, Zhang F, Chen ZL, Li G, Bi ZY, Bi YY, Gong FY, Bo T, Yuan ZP, Hu WD, Zhan BT, Zhang Q, He QQ and Li DJ. Comprehensive analysis of the expression and prognosis for E2Fs in human breast cancer. *Mol Ther* 2019; 27: 1153-1165.
- [28] Gou R, Zhu L, Zheng M, Guo Q, Hu Y, Li X, Liu J and Lin B. Annexin A8 can serve as potential prognostic biomarker and therapeutic target for ovarian cancer: based on the comprehensive analysis of Annexins. *J Transl Med* 2019; 17: 275.
- [29] Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B and Liu XS. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res* 2017; 77: e108-e110.
- [30] Zhang M, Zhu K, Pu H, Wang Z, Zhao H, Zhang J and Wang Y. An immune-related signature

- predicts survival in patients with lung adenocarcinoma. *Front Oncol* 2019; 9: 1314.
- [31] Sundar R, Qamra A, Tan ALK, Zhang S, Ng CCY, Teh BT, Lee J, Kim KM and Tan P. Transcriptional analysis of immune genes in Epstein-Barr virus-associated gastric cancer and association with clinical outcomes. *Gastric Cancer* 2018; 21: 1064-1070.
- [32] Lin P, Guo YN, Shi L, Li XJ, Yang H, He Y, Li Q, Dang YW, Wei KL and Chen G. Development of a prognostic index based on an immunogenomic landscape analysis of papillary thyroid cancer. *Aging (Albany NY)* 2019; 11: 480-500.
- [33] Yang S, Wu Y, Deng Y, Zhou L, Yang P, Zheng Y, Zhang D, Zhai Z, Li N, Hao Q, Song D, Kang H and Dai Z. Identification of a prognostic immune signature for cervical cancer to predict survival and response to immune checkpoint inhibitors. *Oncoimmunology* 2019; 8: e1659094.
- [34] Taube JM, Galon J, Sholl LM, Rodig SJ, Cottrell TR, Giraldo NA, Baras AS, Patel SS, Anders RA, Rimm DL and Cimino-Mathews A. Implications of the tumor immune microenvironment for staging and therapeutics. *Mod Pathol* 2018; 31: 214-234.
- [35] Wagner NB, Weide B, Gries M, Reith M, Tarnanidis K, Schuermans V, Kemper C, Kehrel C, Funder A, Lichtenberger R, Sucker A, Herpel E, Holland-Letz T, Schadendorf D, Garbe C, Umansky V, Utikal J and Gebhardt C. Tumor microenvironment-derived S100A8/A9 is a novel prognostic biomarker for advanced melanoma patients and during immunotherapy with anti-PD-1 antibodies. *J Immunother Cancer* 2019; 7: 343.
- [36] Senbabaoglu Y, Gejman RS, Winer AG, Liu M, Van Allen EM, de Velasco G, Miao D, Ostrovna I, Drill E, Luna A, Weinhold N, Lee W, Manley BJ, Khalil DN, Kaffenberger SD, Chen Y, Danilova L, Voss MH, Coleman JA, Russo P, Reuter VE, Chan TA, Cheng EH, Scheinberg DA, Li MO, Choueiri TK, Hsieh JJ, Sander C and Hakimi AA. Tumor immune microenvironment characterization in clear cell renal cell carcinoma identifies prognostic and immunotherapeutically relevant messenger RNA signatures. *Genome Biol* 2016; 17: 231.
- [37] Lazăr DC, Avram MF, Romosan I, Cornianu M, Tăban S and Goldis A. Prognostic significance of tumor immune microenvironment and immunotherapy: novel insights and future perspectives in gastric cancer. *World J Gastroenterol* 2018; 24: 3583-3616.
- [38] Li J, Li X, Zhang C, Zhang C and Wang H. A signature of tumor immune microenvironment genes associated with the prognosis of non-small cell lung cancer. *Oncol Rep* 2020; 43: 795-806.
- [39] Chen H, Chong W, Teng C, Yao Y, Wang X and Li X. The immune response-related mutational signatures and driver genes in non-small-cell lung cancer. *Cancer Sci* 2019; 110: 2348-2356.
- [40] Takada K, Toyokawa G, Shoji F, Okamoto T and Maehara Y. The significance of the PD-L1 expression in non-small-cell lung cancer: trenchant double swords as predictive and prognostic markers. *Clin Lung Cancer* 2018; 19: 120-129.
- [41] Faruki H, Mayhew GM, Serody JS, Hayes DN, Perou CM and Lai-Goldman M. Lung adenocarcinoma and squamous cell carcinoma gene expression subtypes demonstrate significant differences in tumor immune landscape. *J Thorac Oncol* 2017; 12: 943-953.
- [42] Zhao L, Yu Z and Zhao B. Mechanism of VIPR1 gene regulating human lung adenocarcinoma H1299 cells. *Med Oncol* 2019; 36: 91.
- [43] Inaguma S, Lasota J, Czapiewski P, Langfort R, Rys J, Szpor J, Waloszczyk P, Okon K, Biernat W, Schrupp DS, Hassan R, Kasai K, Miettinen M and Ikeda H. CD70 expression correlates with a worse prognosis in malignant pleural mesothelioma patients via immune evasion and enhanced invasiveness. *J Pathol* 2020; 250: 205-216.
- [44] Jacobs J, Deschoolmeester V, Zwaenepoel K, Rolfo C, Silence K, Rottey S, Lardon F, Smits E and Pauwels P. CD70: an emerging target in cancer immunotherapy. *Pharmacol Ther* 2015; 155: 1-10.
- [45] Tarassishin L, Lim J, Weatherly DB, Angeletti RH and Lee SC. Interleukin-1-induced changes in the glioblastoma secretome suggest its role in tumor progression. *J Proteomics* 2014; 99: 152-168.
- [46] Du B, Luo W, Li R, Tan B, Han H, Lu X, Li D, Qian M, Zhang D, Zhao Y and Liu M. Lgr4/Gpr48 negatively regulates TLR2/4-associated pattern recognition and innate immunity by targeting CD14 expression. *J Biol Chem* 2013; 288: 15131-15141.

Prognostic value of immune-related genes in non-small cell lung cancer

Table S1. The top 50 differentially expressed genes

gene	conMean	treatMean	logFC	p Value	fd
AC093787.1	2.001689	0.1459692	-3.7775	2.85E-79	5.78E-75
AL136369.1	2.110536	0.1861224	-3.5033	5.00E-75	5.08E-71
LINC02016	3.234495	0.0848785	-5.252	1.35E-68	9.15E-65
AC105914.2	1.388684	0.1458004	-3.2517	1.94E-68	9.83E-65
LINC01996	2.753144	0.1238178	-4.4748	2.53E-65	1.03E-61
STX11	26.62021	4.0596279	-2.7131	3.18E-65	1.08E-61
PECAM1	116.6206	20.874817	-2.482	4.24E-65	1.08E-61
AL606469.1	4.051778	0.1167581	-5.117	4.26E-65	1.08E-61
AGER	1013.069	27.411292	-5.2078	5.18E-65	1.17E-61
SFTPC	7334.833	225.54502	-5.0233	6.79E-65	1.38E-61
FHL1	58.46862	5.2913938	-3.4659	1.20E-64	2.23E-61
SAPCD2	0.308531	6.6488528	4.4296	2.10E-64	2.70E-61
RGCC	292.693	37.010186	-2.9834	2.26E-64	2.70E-61
TEK	20.6907	1.9495199	-3.4078	2.00E-64	2.70E-61
LINC01082	1.736378	0.092122	-4.2364	1.69E-64	2.70E-61
CLEC3B	97.1204	6.2520069	-3.9574	2.23E-64	2.70E-61
TNNC1	60.34884	4.0070186	-3.9127	2.03E-64	2.70E-61
FAM107A	49.16122	1.8968682	-4.6958	2.47E-64	2.79E-61
CLIC5	43.01635	2.6853318	-4.0017	2.98E-64	2.88E-61
S1PR1	52.39975	6.6234046	-2.9839	2.93E-64	2.88E-61
CAVIN2	113.1793	11.404188	-3.311	2.98E-64	2.88E-61
EFNA3	0.364794	4.6434596	3.67	3.24E-64	2.99E-61
RAMP2	77.11837	10.4974	-2.877	3.43E-64	3.03E-61
EDNRB	33.62722	2.6773744	-3.6507	4.17E-64	3.53E-61
LIMS2	12.49984	1.8552838	-2.7522	4.53E-64	3.68E-61
INMT	79.29264	5.6812733	-3.8029	6.03E-64	4.13E-61
JAM2	10.4103	1.668869	-2.6411	5.91E-64	4.13E-61
ACVRL1	36.89685	5.732853	-2.6862	6.09E-64	4.13E-61
EMP2	167.3807	26.612848	-2.6529	5.47E-64	4.13E-61
CAV1	326.3963	35.274121	-3.2099	5.55E-64	4.13E-61
GPM6A	7.872331	0.3359547	-4.5505	6.54E-64	4.28E-61
RS1	1.385829	0.0804006	-4.1074	7.34E-64	4.34E-61
ADAMTS8	12.32018	0.8337192	-3.8853	7.69E-64	4.34E-61
ECSCR	13.4532	2.3286936	-2.5304	7.12E-64	4.34E-61
PAICS	4.83456	22.322463	2.207	8.18E-64	4.49E-61
EPAS1	314.8134	44.048282	-2.8373	8.71E-64	4.66E-61
AOC3	70.53672	9.3733112	-2.9117	1.01E-63	4.99E-61
LDB2	16.7636	2.5012351	-2.7446	1.03E-63	5.00E-61
SLC39A8	94.80858	13.493621	-2.8127	1.07E-63	5.03E-61
CDH5	45.67183	6.4603903	-2.8216	1.10E-63	5.06E-61
UBE2T	1.828724	23.626699	3.6915	1.15E-63	5.06E-61
ADRB2	12.23144	1.653604	-2.8869	1.12E-63	5.06E-61
ROBO4	21.01313	2.5384781	-3.0493	1.21E-63	5.11E-61
ARHGAP31	17.17005	3.6641871	-2.2283	1.19E-63	5.11E-61
CA4	18.1179	0.7414741	-4.6109	1.35E-63	5.49E-61
TEDC2	0.195647	2.9715621	3.9249	1.57E-63	6.25E-61
PLAC9	19.20497	2.6641495	-2.8497	1.72E-63	6.60E-61
FXYD1	2.544697	0.3049901	-3.0607	1.77E-63	6.65E-61
LINC00968	2.797461	0.2146358	-3.7042	1.85E-63	6.84E-61
GPIHBP1	24.08572	1.3597194	-4.1468	1.92E-63	6.96E-61

Prognostic value of immune-related genes in non-small cell lung cancer

Table S2. The top 50 differentially expressed immune-related genes

ID	logFC	p Value	fdr
AGER	-5.2078	5.18E-65	1.17E-61
TEK	-3.4078	2.00E-64	2.70E-61
S1PR1	-2.9839	2.93E-64	2.88E-61
EDNRB	-3.6507	4.17E-64	3.53E-61
ACVRL1	-2.6862	6.09E-64	4.13E-61
CAT	-1.8583	9.03E-64	4.70E-61
ADRB2	-2.8869	1.12E-63	5.06E-61
TGFBR2	-1.9648	4.75E-63	1.30E-60
NPR1	-2.9642	1.33E-62	2.48E-60
ANGPT4	-4.1518	2.74E-62	4.38E-60
GDF10	-3.5532	8.05E-62	1.03E-59
FABP4	-3.8094	1.42E-61	1.62E-59
SFTPA1	-3.2773	4.64E-61	4.15E-59
BIRC5	4.41184	5.34E-61	4.60E-59
TIE1	-2.2065	5.94E-61	5.03E-59
MARCO	-3.2275	1.07E-60	8.37E-59
DES	-3.419	1.65E-60	1.23E-58
TNFSF12	-1.3927	2.63E-60	1.85E-58
A2M	-2.3498	5.53E-60	3.58E-58
IL3RA	-1.867	8.17E-60	4.98E-58
SEMA3G	-3.1489	8.38E-60	5.08E-58
AGRP	-3.3776	1.30E-59	7.48E-58
FGR	-2.0795	1.74E-59	9.73E-58
PAK1	1.63449	3.05E-59	1.63E-57
ANGPTL1	-2.5513	4.08E-59	2.10E-57
SFTPA2	-3.0862	4.20E-59	2.15E-57
CALCRL	-2.6009	4.59E-59	2.32E-57
SLIT2	-2.4125	6.50E-59	3.18E-57
FGFR4	-2.5926	6.54E-59	3.18E-57
OLR1	-2.6999	7.89E-59	3.76E-57
ADRB1	-3.1491	9.58E-59	4.48E-57
ENG	-1.8112	9.87E-59	4.59E-57
KL	-2.5843	1.26E-58	5.69E-57
TFR2	3.26594	1.28E-58	5.76E-57
RETN	-3.4086	2.11E-58	9.19E-57
ANGPTL7	-3.4945	2.20E-58	9.49E-57
RXFP1	-2.6991	2.75E-58	1.16E-56
MSR1	-2.4042	3.04E-58	1.27E-56
CCL23	-2.8672	3.19E-58	1.32E-56
PTH1R	-2.1071	4.37E-58	1.79E-56
HLA-E	-1.3738	1.26E-57	4.74E-56
ANGPT1	-2.3572	1.94E-57	7.06E-56
HTR3C	-3.5136	2.33E-57	8.35E-56
C5AR1	-2.0062	3.02E-57	1.06E-55
LGR4	2.59821	3.08E-57	1.08E-55
SFTPD	-2.6533	3.39E-57	1.18E-55
DUOX1	-2.1363	3.80E-57	1.31E-55
ADM2	3.27668	3.80E-57	1.31E-55
COLEC12	-1.9653	3.81E-57	1.32E-55
MASP1	-2.4056	4.59E-57	1.56E-55

Prognostic value of immune-related genes in non-small cell lung cancer

Table S3. Differentially expressed transcription factors

ID	conMean	treatMean	logFC	p Value	fdr
ATF3	48.2328	15.46977	-1.64056	3.14E-27	1.12E-26
BCL11A	0.576196	3.288123	2.512633	1.44E-27	5.23E-27
BRCA1	0.701695	2.609126	1.894651	2.17E-45	2.22E-44
CBX2	0.615327	5.46938	3.151951	5.68E-46	6.07E-45
CBX3	20.59131	47.75739	1.213688	8.42E-58	3.25E-56
CBX7	10.32704	3.507957	-1.55772	5.85E-58	2.33E-56
CBX8	1.118802	3.14963	1.493228	7.14E-55	1.74E-53
CDX2	0.00431	0.282431	6.034053	3.08E-16	6.75E-16
CENPA	0.265535	5.469856	4.364532	7.02E-61	5.79E-59
DNMT3A	2.273654	5.041851	1.148941	4.48E-40	3.21E-39
E2F1	2.985909	10.79214	1.853738	3.53E-44	3.29E-43
E2F3	2.356947	6.775923	1.523498	5.85E-54	1.29E-52
E2F7	0.114801	1.913881	4.059292	3.77E-53	7.67E-52
EBF1	1.679026	0.81885	-1.03595	1.81E-34	9.32E-34
EGR1	266.6757	76.74543	-1.79693	1.50E-27	5.44E-27
EGR2	13.98798	5.116766	-1.45088	1.36E-30	5.78E-30
EPAS1	314.8134	44.04828	-2.83734	8.71E-64	4.66E-61
ERG	9.169996	2.038247	-2.16959	2.58E-61	2.54E-59
ETS1	35.80682	14.51961	-1.30223	6.56E-45	6.44E-44
ETV1	9.326728	3.671053	-1.34518	8.80E-45	8.56E-44
EZH2	0.837579	6.948515	3.052408	2.15E-62	3.55E-60
FLI1	8.399317	2.432291	-1.78796	1.00E-58	4.65E-57
FOS	320.7163	87.35255	-1.87638	6.21E-33	2.96E-32
FOXA1	6.972364	17.6119	1.336831	1.90E-11	3.35E-11
FOXA2	16.38792	6.851629	-1.25811	4.34E-31	1.88E-30
FOXM1	0.762025	14.06518	4.206147	1.40E-60	1.05E-58
FOXP3	0.680199	2.409447	1.824673	4.71E-39	3.16E-38
GATA2	9.695165	2.708203	-1.83993	8.51E-50	1.23E-48
GATA6	11.6636	3.025457	-1.94679	4.24E-56	1.24E-54
GREB1	0.137838	0.724053	2.393126	2.81E-27	1.01E-26
GRHL2	4.772569	12.67443	1.409083	3.08E-48	3.91E-47
H2AFX	10.34606	36.89555	1.834366	4.98E-54	1.11E-52
HDAC1	19.53822	42.79464	1.131131	2.42E-55	6.30E-54
HDAC2	4.000933	8.246329	1.043416	4.56E-46	4.92E-45
HNF4G	0.157468	1.400664	3.152985	3.31E-35	1.78E-34
HOXA9	0.094721	0.454657	2.263019	8.86E-05	0.000114
HOXB13	0.007283	1.534419	7.718929	6.63E-25	2.14E-24
HOXB7	1.66708	9.329764	2.484518	4.42E-26	1.51E-25
HOXC11	0.005041	0.999767	7.631637	2.68E-31	1.18E-30
HOXC9	0.10867	1.398361	3.685711	4.26E-27	1.52E-26
JUNB	225.9428	112.8933	-1.001	4.15E-18	9.88E-18
JUND	175.066	65.05175	-1.42824	6.02E-43	5.19E-42
KAT2B	8.481527	3.62125	-1.22784	5.35E-52	9.52E-51
KDM5B	4.596284	11.21934	1.287447	4.28E-54	9.59E-53
KLF4	44.77088	15.35732	-1.54363	5.55E-39	3.70E-38
LEF1	1.243607	3.352382	1.430656	6.19E-24	1.93E-23
LHX2	0.011173	0.80046	6.162801	1.14E-46	1.29E-45

Prognostic value of immune-related genes in non-small cell lung cancer

LIN9	1.219119	2.831627	1.215792	5.30E-40	3.77E-39
LMNB1	3.744825	16.5902	2.147361	4.32E-55	1.07E-53
LMO2	13.05369	3.904208	-1.74136	1.18E-55	3.22E-54
LYL1	3.446202	1.646161	-1.0659	7.89E-37	4.63E-36
MAFF	15.84162	6.176294	-1.35891	1.47E-23	4.48E-23
MEF2B	0.168476	0.428147	1.345562	2.78E-26	9.52E-26
MITF	4.386587	1.726814	-1.34499	6.39E-53	1.26E-51
MYBL2	1.339345	29.60834	4.466404	2.60E-59	1.41E-57
MYH11	24.43855	8.996456	-1.44173	3.26E-40	2.36E-39
NCAPG	0.354363	4.279897	3.594276	5.61E-60	3.62E-58
NR2F1	12.90429	5.819093	-1.14898	6.10E-36	3.41E-35
NR4A1	58.81776	13.66279	-2.106	4.16E-38	2.62E-37
NR5A2	1.261942	0.439854	-1.52055	4.05E-30	1.68E-29
PDX1	0.001203	0.394448	8.357328	2.87E-23	8.65E-23
POLR3G	0.582625	1.338556	1.200038	4.30E-24	1.34E-23
POU5F1	0.645851	1.323458	1.035039	0.015286	0.01728
PPARG	12.40388	4.02222	-1.62473	7.03E-50	1.03E-48
PRKDC	7.130427	19.63949	1.461697	1.28E-44	1.23E-43
RARG	5.386871	12.25974	1.186408	2.26E-19	5.71E-19
RBP2	3.708237	0.455754	-3.02441	4.75E-56	1.38E-54
RFX2	4.930107	1.624115	-1.60197	7.83E-33	3.71E-32
RUNX1T1	0.978201	0.364206	-1.42538	2.24E-39	1.53E-38
RXRG	2.438163	0.423877	-2.52408	1.44E-55	3.89E-54
SALL4	0.061045	0.880307	3.850071	1.06E-54	2.54E-53
SCML2	0.211175	0.848525	2.006518	4.03E-12	7.35E-12
SNAI2	4.699027	14.31572	1.607167	3.80E-08	5.76E-08
SOX17	6.095981	0.843703	-2.85305	1.47E-59	8.32E-58
SOX2	2.589633	46.54031	4.167662	2.13E-22	6.16E-22
SOX4	11.93901	40.79933	1.772862	2.03E-48	2.61E-47
SOX9	3.456263	11.89982	1.783655	9.89E-13	1.85E-12
SPDEF	2.236554	13.38089	2.580825	0.007749	0.008908
SPIB	0.80504	2.484895	1.626053	0.000971	0.001176
TAL1	3.033958	0.452484	-2.74526	5.13E-61	4.46E-59
TBL1XR1	9.435569	23.46339	1.31423	5.40E-47	6.22E-46
TCF21	11.84824	1.062207	-3.47954	3.81E-63	1.12E-60
TFAP2A	0.155655	4.473159	4.844874	9.41E-57	3.05E-55
TFAP2C	2.77184	10.67053	1.944716	1.64E-41	1.29E-40
TP63	1.126625	36.59902	5.021726	9.46E-09	1.47E-08
TP73	0.891178	2.389986	1.423217	1.88E-13	3.63E-13
TTF2	1.143069	2.949746	1.367678	2.10E-48	2.70E-47

Prognostic value of immune-related genes in non-small cell lung cancer

Table S4. DEIRG-related KEGG pathway

ID	Description	p value	p.adjust	q value	Count
hsa04060	Cytokine-cytokine receptor interaction	3.65E-49	8.11E-47	5.31E-47	81
hsa04061	Viral protein interaction with cytokine and cytokine receptor	1.28E-21	1.42E-19	9.31E-20	32
hsa05323	Rheumatoid arthritis	3.94E-16	2.92E-14	1.91E-14	26
hsa04062	Chemokine signaling pathway	1.98E-14	1.10E-12	7.20E-13	34
hsa04080	Neuroactive ligand-receptor interaction	2.71E-14	1.20E-12	7.87E-13	46
hsa04630	JAK-STAT signaling pathway	7.06E-11	2.61E-09	1.71E-09	27
hsa04640	Hematopoietic cell lineage	9.58E-11	3.04E-09	1.99E-09	21
hsa04659	Th17 cell differentiation	4.45E-10	1.24E-08	8.08E-09	21
hsa04650	Natural killer cell mediated cytotoxicity	6.63E-10	1.64E-08	1.07E-08	23
hsa04657	IL-17 signaling pathway	1.83E-09	4.00E-08	2.62E-08	19
hsa05332	Graft-versus-host disease	1.98E-09	4.00E-08	2.62E-08	13
hsa04672	Intestinal immune network for IgA production	2.18E-09	4.04E-08	2.64E-08	14
hsa04145	Phagosome	2.62E-09	4.47E-08	2.92E-08	24
hsa04658	Th1 and Th2 cell differentiation	8.55E-09	1.31E-07	8.60E-08	18
hsa05330	Allograft rejection	8.88E-09	1.31E-07	8.60E-08	12
hsa05321	Inflammatory bowel disease (IBD)	1.45E-08	2.02E-07	1.32E-07	15
hsa04151	PI3K-Akt signaling pathway	1.88E-08	2.46E-07	1.61E-07	37
hsa04612	Antigen processing and presentation	2.89E-08	3.57E-07	2.33E-07	16
hsa05416	Viral myocarditis	3.83E-08	4.47E-07	2.93E-07	14
hsa05320	Autoimmune thyroid disease	6.24E-08	6.93E-07	4.53E-07	13
hsa05150	Staphylococcus aureus infection	1.05E-07	1.11E-06	7.28E-07	17
hsa05310	Asthma	1.27E-07	1.28E-06	8.38E-07	10
hsa05140	Leishmaniasis	1.63E-07	1.53E-06	1.00E-06	15
hsa05169	Epstein-Barr virus infection	1.65E-07	1.53E-06	1.00E-06	25
hsa04010	MAPK signaling pathway	2.59E-07	2.25E-06	1.47E-06	31
hsa05144	Malaria	2.63E-07	2.25E-06	1.47E-06	12
hsa04940	Type I diabetes mellitus	4.17E-07	3.43E-06	2.24E-06	11
hsa05218	Melanoma	4.35E-07	3.45E-06	2.26E-06	14
hsa04015	Rap1 signaling pathway	1.41E-06	1.08E-05	7.05E-06	24
hsa05164	Influenza A	1.86E-06	1.38E-05	9.03E-06	21
hsa04350	TGF-beta signaling pathway	2.38E-06	1.71E-05	1.12E-05	15
hsa04014	Ras signaling pathway	2.51E-06	1.74E-05	1.14E-05	25
hsa04514	Cell adhesion molecules (CAMs)	2.95E-06	1.99E-05	1.30E-05	19
hsa04668	TNF signaling pathway	4.91E-06	3.21E-05	2.10E-05	16
hsa04064	NF-kappa B signaling pathway	8.64E-06	5.48E-05	3.59E-05	15
hsa05224	Breast cancer	1.16E-05	7.18E-05	4.70E-05	18
hsa04360	Axon guidance	1.80E-05	0.0001081	7.07E-05	20
hsa05418	Fluid shear stress and atherosclerosis	2.07E-05	0.0001212	7.93E-05	17
hsa05145	Toxoplasmosis	2.16E-05	0.0001227	8.03E-05	15
hsa01521	EGFR tyrosine kinase inhibitor resistance	4.10E-05	0.0002273	0.000149	12
hsa05152	Tuberculosis	5.53E-05	0.0002995	0.000196	19
hsa04066	HIF-1 signaling pathway	6.49E-05	0.0003433	0.000225	14
hsa04380	Osteoclast differentiation	0.0001043	0.0005387	0.000352	15
hsa04620	Toll-like receptor signaling pathway	0.0001564	0.0007893	0.000516	13
hsa05226	Gastric cancer	0.0001751	0.0008639	0.000565	16
hsa05167	Kaposi sarcoma-associated herpesvirus infection	0.0002624	0.0012457	0.000815	18
hsa05166	Human T-cell leukemia virus 1 infection	0.0002637	0.0012457	0.000815	20
hsa04933	AGE-RAGE signaling pathway in diabetic complications	0.0004094	0.0018935	0.001239	12
hsa05146	Amoebiasis	0.0004918	0.0022282	0.001458	12
hsa05322	Systemic lupus erythematosus	0.0005424	0.0024082	0.001576	14

Prognostic value of immune-related genes in non-small cell lung cancer

hsa04660	T cell receptor signaling pathway	0.0005877	0.0025584	0.001674	12
hsa05219	Bladder cancer	0.000825	0.0035221	0.002305	7
hsa04928	Parathyroid hormone synthesis, secretion and action	0.0023532	0.0098569	0.00645	11
hsa04024	cAMP signaling pathway	0.0037638	0.0153357	0.010035	17
hsa04926	Relaxin signaling pathway	0.0037994	0.0153357	0.010035	12
hsa04510	Focal adhesion	0.0039655	0.0157202	0.010286	16
hsa01522	Endocrine resistance	0.0041455	0.0161455	0.010565	10
hsa05134	Legionellosis	0.005758	0.0220391	0.014421	7
hsa05235	PD-L1 expression and PD-1 checkpoint pathway in cancer	0.0067928	0.0255595	0.016725	9
hsa04810	Regulation of actin cytoskeleton	0.007586	0.0280682	0.018366	16
hsa04614	Renin-angiotensin system	0.0104891	0.0381734	0.024978	4
hsa05205	Proteoglycans in cancer	0.0115284	0.0412791	0.02701	15
hsa05215	Prostate cancer	0.0117206	0.0413012	0.027025	9
hsa04662	B cell receptor signaling pathway	0.0128595	0.0446063	0.029188	8
hsa05340	Primary immunodeficiency	0.014178	0.0484232	0.031685	5

Table S5. Multivariate Cox regression analyses of 15 DEIRG signatures

IRG	coef	exp (coef)	se (coef)	z	p
THBS1	0.09607	1.10084	0.05391	1.782	0.07471
S100P	0.03865	1.03941	0.02484	1.556	0.11968
S100A16	0.11711	1.12424	0.05834	2.007	0.04472
DLL4	0.2089	1.23232	0.10121	2.064	0.03901
CD70	0.16969	1.18493	0.05873	2.889	0.00386
DKK1	0.07232	1.075	0.03656	1.978	0.04788
IL33	-0.14257	0.86712	0.05208	-2.738	0.00619
NRTN	-0.20366	0.81574	0.08726	-2.334	0.0196
PDGFB	0.12007	1.12757	0.07962	1.508	0.13154
STC2	0.10827	1.11435	0.05072	2.135	0.0328
VGF	0.1346	1.14408	0.06403	2.102	0.03553
GCGR	-0.31152	0.73233	0.1655	-1.882	0.05979
HTR3A	0.10294	1.10842	0.0476	2.163	0.03057
LGR4	0.1389	1.14901	0.06339	2.191	0.02843
SHC3	-0.29843	0.74198	0.13697	-2.179	0.02935

Table S6. The relationship between DEIRGs and DETFs in the network

TF	immuneGene	cor	p value	Regulation
BRCA1	STC2	0.401038	1.34E-35	positive
CBX7	VIPR1	0.473077	1.16E-50	positive
CENPA	DLL4	-0.42843	6.64E-41	negative
CENPA	STC2	0.452618	5.20E-46	positive
E2F7	LTB4R	0.423057	7.98E-40	positive
E2F7	STC2	0.456965	5.68E-47	positive
EPAS1	DLL4	0.439677	3.16E-43	positive
ERG	DLL4	0.412597	8.87E-38	positive
ERG	IL33	0.425236	2.93E-40	positive
ETS1	THBS1	0.458692	2.33E-47	positive
ETV1	SHC3	0.451719	8.20E-46	positive
FLI1	DLL4	0.425396	2.72E-40	positive
FLI1	IL33	0.442618	7.54E-44	positive
FLI1	ADRB2	0.417937	8.17E-39	positive
FOXM1	STC2	0.454838	1.69E-46	positive
HNF4G	HNF4G	0.884451	3.35E-295	positive
KAT2B	SHC3	0.408584	5.17E-37	positive
MITF	THBS1	0.423307	7.11E-40	positive
MYBL2	STC2	0.414803	3.33E-38	positive
NCAPG	STC2	0.406295	1.40E-36	positive
NCAPG	VIPR1	-0.44588	1.52E-44	negative
RARG	S100A16	0.414974	3.09E-38	positive
RARG	DLL4	-0.40664	1.21E-36	negative
RARG	LTB4R	0.476227	2.09E-51	positive
RARG	LTB4R2	0.434939	3.08E-42	positive
SNAI2	DLL4	-0.42041	2.67E-39	negative
SNAI2	LTB4R	0.485926	9.50E-54	positive
SNAI2	LTB4R2	0.427973	8.23E-41	positive
SNAI2	IL11	0.429688	3.69E-41	positive
SOX2	LTB4R	0.445523	1.81E-44	positive
SOX2	STC2	0.456363	7.73E-47	positive
TBL1XR1	DLL4	-0.40307	5.60E-36	negative
TCF21	VIPR1	0.429712	3.65E-41	positive
TP63	DLL4	-0.46008	1.14E-47	negative
TP63	LTB4R	0.584838	1.64E-82	positive
TP63	LTB4R2	0.513958	5.88E-61	positive
TP63	STC2	0.444411	3.13E-44	positive

Table S7. Independent prognosis analysis of 15 DEIRG signatures

Variables	Univariate analysis		Multivariate analysis	
	Hazard ratio (95% CI)	P-value	Hazard ratio (95% CI)	P-value
age	1.007 (0.993-1.021)	0.328	1.012 (0.997-1.027)	0.117
sex	1.091 (0.839-1.419)	0.517	1.131 (0.867-1.476)	0.364
stage	1.476 (1.297-1.679)	3.45E-09	1.216 (0.890-1.662)	0.220
T	1.420 (1.215-1.659)	1.02E-05	1.245 (1.018-1.522)	0.033
M	1.864 (1.085-3.204)	0.024	1.220 (0.536-2.774)	0.636
N	1.471 (1.260-1.717)	1.01E-06	1.140 (0.867-1.500)	0.347
riskScore	1.691 (1.529-1.871)	1.86E-24	1.697 (1.527-1.887)	1.36E-22

Prognostic value of immune-related genes in non-small cell lung cancer

Table S8. Correlation analysis of 15 DEIRGs and LUAD and LUSC clinical characteristics

IRG	Age ($\leq 65 / > 65$)		Gender (male/female)		TNM stage (I-II/III-IV)		T (T1-T2/T3-T4)		M (M0/M1)		N (N0/N1-N3)	
	t	p-value	t	p-value	t	p-value	t	p-value	t	p-value	t	p-value
THBS1	0.425	0.671	1.889	0.059	-0.977	0.33	-1.196	0.234	1.928	0.063	-0.752	0.453
S100P	0.573	0.567	-1.112	0.266	-2.001	0.047	0.518	0.605	-1.49	0.148	-1.43	0.153
S100A16	-2559	0.011	-2.897	0.004	-2.109	0.036	-1.287	0.2	0.02	0.984	-1.561	0.119
DLL4	0.944	0.346	3.899	1.09E-04	-1.896	0.06	2.014	0.045	-0.944	0.354	-1.393	0.165
CD70	-1.139	0.255	1.257	0.209	1.426	0.154	-0.868	0.387	3.917	1.10E-04	0.374	0.708
DKK1	1.634	0.103	0.232	0.816	-1.781	0.077	-0.555	0.58	-0.792	0.436	-0.347	0.729
IL33	-0.814	0.416	0.73	0.466	1.401	0.162	1.068	0.287	0.724	0.474	2.872	0.004
NRTN	0.519	0.604	-1.412	0.158	0.208	0.835	-0.297	0.767	1.646	0.11	-0.945	0.345
PDGFB	-1.424	0.155	-1.546	0.123	1.517	0.13	-1.329	0.186	0.851	0.401	1.988	0.047
STC2	-0.586	0.558	-3.935	9.35E-05	1.477	0.14	-1.225	0.223	5.829	2.49E-07	-1.121	0.263
VGF	1.638	0.103	-0.566	0.572	-1.447	0.15	1.365	0.173	-1.44	0.162	0.809	0.419
GCGR	0.352	0.725	-0.687	0.492	1.248	0.212	0.448	0.655	4.649	4.15E-06	0.457	0.648
HTR3A	1.294	0.196	1.757	0.08	-0.786	0.432	0.936	0.35	-0.389	0.7	-1.089	0.277
LGR4	0.406	0.685	1.53	0.127	-1.305	0.193	0.196	0.845	-0.291	0.773	-0.424	0.671
SHC3	0.865	0.387	4.811	2.12E-06	-0.163	0.871	-0.486	0.627	-0.594	0.557	2.405	0.016
riskScore	-0.81	0.418	0.361	0.718	-3.232	0.001	-0.598	0.551	-0.977	0.337	-3.501	5.24E-04