Individualized Prognostic Score for Lung Adenocarcinoma Based on Immune Status

Peng Chen (✉ 192388805@qq.com)

Tianjin Medical University Cancer Institute and Hospital

Mihray Abdurazik

Tianjin Medical University Cancer Institute and Hospital

Gengwei Huo

Tianjin Medical University Cancer Institute and Hospital

Ye Du

Tianjin Medical University Cancer Institute and Hospital

Ran Zuo

Tianjin Medical University Cancer Institute and Hospital

Fuyi Zhu

Tianjin Medical University Cancer Institute and Hospital

Jinfang Cui

Tianjin Medical University Cancer Institute and Hospital

Dongqi Yuan

Tianjin Medical University Cancer Institute and Hospital

Research Article

Keywords: Lung adenocarcinoma, immune-related genes, immunogenomic landscape, TCGA, individualized medicine

Posted Date: February 1st, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1258129/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Background: Lung adenocarcinoma (LUAD) is the prevalent non-small cell lung cancer (NSCLC) histological subtype, with rising morbidity and mortality. This study was conducted with the motive to develop a prognostic model by examining the relation between immune-related gene (IRGs) profiles.

Methods: We combined immune-related gene (IRG) expression profiles and progression-free intervals (PFIs) in 594 LUAD patients by using TCGA dataset. The COX regression analysis and computational difference technique were used to predict survival-associated and differentially expressed IRGs in LUAD patients. Computational biology was also used to investigate the characteristics and molecular mechanisms of these LUAD-specific IRGs. Using multivariable COX analysis, a novel prognostic index based on immune-related genes was established.

Results: There was a strong correlation between the clinical outcome of LUAD patients and 58 differentially expressed immune-related genes. These genes were found to be actively involved in a KEGG pathway involving cytokines and cytokine receptors, as demonstrated by functional enrichment analysis. A prognostic signature comprised of IRGs (CRABP1, S100A16, FGF2, RBP2, FURIN, SLC11A1, ANGPTL4, VIPR1, BDNF, IL11, INHA, SEMA4B, TNFRSF11A) performed relatively well in prognostic prediction and was associated with tumour stage, tumour burden, age, number of lesions, and metastasis. Surprisingly, the IRG-based prognostic index effectively predicted immune cell infiltration.

Conclusions: Our findings collectively identified multiple clinically relevant IRGs, identified regulators of the immunological repertoire, as well as established the critical role of a tailored, surveillance, IRG-based immune signature in the diagnosis, and prognosis of LUAD.

Introduction

Globally, lung cancer (LC) is the main cause of cancer-related deaths, causing around 2 million diagnoses as well as 1.8 million deaths with a 21% five-year survival rate in recent years (1, 2). Non-small-cell lung cancer (NSCLC) is the major histological subtype of LC, responsible for 84% of all LC diagnoses apart from the other subtype called small-cell lung cancer (SCLC), with a declining contribution of 13% (3). The heterogeneous class of NSCLC includes adenocarcinoma (LUAD), large cell carcinoma (LCC) and squamous cell carcinoma (SCC) (4).

LUAD is the most frequently diagnosed and fatal subtype among the three, with an average 5-year survival rate of less than 20% (5). It accounts for approximately 60% of all NSCLCs (6), which could be due to the increasing proportion of smokers in the majority of countries. The prognosis in patients with LUAD is difficult, which could be due to the high risk of metastasis (7).

Numerous therapies, including chemotherapy, molecular targeted therapy, surgical resection, and radiotherapy, have been developed in the clinical treatment of LUAD over the last decades, but overall survival time for LUAD patients has not improved significantly, owing in part to a dearth of useful

molecular biomarkers (8). As a result, identifying novel prognostic gene markers for LUAD is critical for increasing LUAD patients' overall survival. Given that the immune system has been implicated in the development and progression of cancer, examining the differential expression of immune-related genes (IRGs) in tumour tissue samples is critical for optimising immunotherapy, understanding the tumour microenvironment, predicting patient prognosis and improving clinical diagnosis. Despite widespread recognition of IRGs' importance in cancer growth and treatment, no rigorous genome-wide profiling investigation into their clinical impact and molecular processes on LAUD has been conducted (9).

This systematic, integrated investigation of IRGs in LUAD contributes to the understanding of their clinical relevance and sheds light on possible molecular features. The large number of LUAD samples available for this study aided in the generation of robust results. We identified many IRGs that play a critical role in the beginning and progression of LUAD. These IRGs have the potential to be useful clinical biomarkers. Bioinformatics systems will permit a more in-depth examination of the molecular mechanisms behind these proteins. Most significantly, a customised immune prognostic signature based on differentially expressed IRGs was proposed to predict clinical outcomes.

Materials And Methods

Data Collection

The transcriptome RNA sequencing datasets as well as clinical follow-up data concerning patients with lung adenocarcinoma (LUAD) were obtained first from TCGA Data Portal (<https://cancergenome.nih.gov/>). HTSeq-FPKM was used to collect the expression data, which included 535 primary LUAD cells and 59 non-tumor samples, as well as clinical information for these patients. A detailed list of IRGs was obtained from the Immunology Database and Analysis Portal (ImmPort) database (<https://www.immport.org/shared/home>), which reliably and timely updates immunology data, facilitating the exchange of immunology data among cancer researchers and giving a list of IRGs for cancer researchers to use. The GSE31210 dataset was a second set of transcriptome profiling data from LUAD patients that was acquired from the GEO database and used to validate the predictive IRG model. It was derived and retrieved from the GEO database (<https://www.ncbi.nlm.nih.gov/geo>). TCGA and GEO belong to public databases. The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open source data, so there are no ethical issues and other conflicts of interest. The open source data from TCGA and GEO were performed in accordance with the Declaration of Helsinki.

Differential Gene Analysis and Filtering

The training set consisted of instances from the TCGA database. The Wilcox Test was used to determine whether IRGs were differentially expressed between LUAD and nearby non-tumor lung tissues. All transcriptional data were subjected to differential gene analysis with a false discovery rate (FDR) of 0.05 and a $\log_2$ |fold change| greater than one as cutoff values. Next, differential IRGs were filtered out of all

differentially expressed genes to remove those that were not. Univariate Cox regression analysis was performed to find IRGs that were associated with survival time using the R software survival package.

Functional Enrichment Analysis

The functional enrichment of IRGs was studied using Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses and Gene Ontology (GO). The threshold for false discovery rate (FDR) was set to 0.05.

Investigation of the Molecular Mechanisms Governing the Function of Immune Genes

It is very well established that transcription factors (TFs) play a critical role in determining the extent of gene expression. Correlation analysis was used to determine the relationship between prognosis-associated IRGs and differential TFs, and 318 TFs were retrieved from the Cistrome Cancer database (<http://www.cistrome.org/>). The corFilter was set to 0.4 and the P-value was 0.001 for the analysis. Additionally, Cytoscape programme version 3.7.2 was used to obtain the regulatory network for these IRGs and TFs.

IRGs-Based Prognostic Risk Model Construction and Validation

Multivariate Cox regression analysis was used to analyse the IRGs, as well as the weighted score was used to calculate the risk score on every respective patient. The predictive model's score was calculated as follows: $(\text{exprIRG1coef1}) + (\text{exprIRG2coef2}) + \dots + (\text{exprIRGncoefn})$. LUAD patients were classified into a low-risk and even a high-risk group using the median risk score as the cut-off number. Following a cut-off point of the median risk score, individuals with LUAD were divided into two groups: those at low risk and those at high risk. The R packages "survival" was used to perform Kaplan-Meier analysis and "survminer" was used to visualise the survival curves. Additionally, a receiver operating characteristic curve called ROC curve was produced and the area under the curve (AUC) was determined to determine the predictive potential of this risk model. The "survivalROC" R package was used to perform these calculations. Following the trend of risk scores, the survival status of patients was determined, as well as the expression of the IRGs involved in this model. Additionally, multivariate and univariate studies were conducted to see whether the risk score could be used independently as a prognostic predictor. Furthermore, the link between IRG expression and other clinicopathological variables was explored.

Immune Cells Infiltration and Correlation Analysis

LUAD patients were evaluated for the prevalence of tumor-infiltrating immune cells using the CIBERSORT method. Each sample with abundance ratio matrix of 22 immune cell types was calculated using the gene expression data, which include different types of T cell (T follicular helper cells, naïve CD4 T cells, activated memory CD4 T cells, resting memory CD4 T cells, CD8 T cells, $\gamma\delta$ T cells, and Tregs), neutrophils, macrophages (M0, M1, M2), memory B cells, resting dendritic cells, activated DCs, naïve B cells, resting natural killer cells, activated NK cells, plasma cells, resting mast cells, eosinophils, activated

mast cells, resting dendritic cells, and monocytes (10). The CIBERSORT results for samples with a P 0.05 suggested that the inferred fractions of immune cell populations were accurate and suitable for further study.

Statistical Analysis

All of the analyses were carried out with the help of the R programming language (version 3.6.2). The Wilcoxon test was used to determine whether or not there was any difference in gene expression between the TCGA cohorts. Cox regression was used to conduct univariate and multivariate analyses. The log-rank test was used to conduct the survival analysis. Pearson correlation analysis was used to determine the correlation. Student's t-tests were used to determine differences in clinicopathologic characteristics. In all statistical tests, P values less than 0.05 were considered significant.

Results

Screening and Identification of IRGs

The TCGA data portal was used to obtain transcriptome RNA sequencing data for LUAD samples (<https://cancergenome.nih.gov/>). It was obtained using the HTSeq-FPKM method the expression data for 535 primary LUAD and 59 non-tumor tissues. Comparing tumour and non-tumorous tissues' gene expression using LUAD transcriptome RNA data revealed 6778 differentially expressed genes, of which 5176 were up-regulated in tumours and 1600 were down-regulated in tumors (Figures 1A and 1C). From this list of genes, we were able to identify 505 differentially expressed IRGs, with 338 of them being up-regulated and 167 being down-regulated (Figures 1B and 1D). Due to the fact that survival duration and survival status are the primary clinical outcomes for patients, we concentrated our efforts on identifying molecular biomarkers that could serve as effective prognostic indications. Using univariate Cox regression analysis, 54 prognosis-associated IRGs (p-IRGs) were identified by merging the differential immune gene matrix and the clinical outcome matrix (Figure 1E).

Function Enrichment Analysis

The Gene Ontology (GO) is now divided into three categories, which are as follows: Molecular Function (MF), Biological Process (BP), and Cellular Component (CC) (CC). On 505 differentially expressed IRGs, we applied functional enrichment analysis and discovered 1,263 GO keywords and 45 significant KEGG pathways. The analysis of gene functional enrichment revealed that the inflammatory pathways were the ones that were most frequently implicated. It was cellular components, biological processes, and molecular activities that were the most frequently mentioned in biological terms, with "humoral immune response", "immunoglobulin complex," and "signalling receptor activity" being the most frequently mentioned (Figure 2A). When it comes to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways, differentially expressed IRGs were shown to be the most commonly enriched in cytokine-cytokine receptor interactions (Figure 2B). Following that, we ran a similar enrichment study on p-IRGs and discovered that the immune microenvironment was a significant factor (Figure 2C). Additionally, the

most often discovered KEGG pathways were the Cytokine-cytokine receptor interaction, the Chemokine signalling route, and the Ras signalling pathway (Figure 2D).

Association Between Differential IRGs and Transcription Factors

Transcription factors (TFs) are responsible for the regulation of gene expression. So we investigated putative molecular mechanisms that could explain the clinical importance of IRGs and TFs retrieved from databases. After following the analysis of 318 transcription factors (TFs), we discovered that 70% were differently expressed between the LUAD and the non-tumor lung samples (Figures 3A, B). According to the results of the correlation study between TFs and p-IRGs, it can be determined that 29 TFs are involved, which includes TTF2, TCF21, SOX9, RXRG, PRKDC, PDX1, NCAP, MYH11, MYBL2, LMO, LMNB1, KLF, IRF4, H2AFX, FOXP3, FOXM1, FOS, FLI1, EZH2, ETV1, ETS1, ERG, EPAS1, E2F7, E2F1, CENPA, CBX7, CBX3, and BRCA1 were found to be mutually mediated with 26 immunological genes. Following that, we constructed a regulatory network based on the foregoing findings, which we then illustrated using the Cytoscape software programme (Figure 3C).

IRGs Risk Model Construction and Validation for LUAD Patients

The 54 p-IRGs identified by merging the differential immune gene matrix and the clinical outcome matrix are not the typical driver genes for LUAD, so we hypothesised that aberrant expression of these IRGs would be a predictive factor.

We then entered 54 p-IRGs into a multivariate Cox regression analysis which acquired 13 risk-related immunological genes (r-IRGs) and their coefficients that were used in the creation of the risk model, which was then used in the formulation of the risk model (Table 1).

TABLE 1

The names and coefficients of risk immune-related genes (r-IRGS) in the risk model.

TABLE 1

id	coef	p-value
S100A16	0.001773111	0.008469147
CRABP1	0.003985264	0.002726465
RBP2	0.066658916	6.94E-05
FURIN	0.000782017	0.025374477
FGF2	0.266568609	0.000919108
SLC11A1	-0.051733135	0.063630963
SEMA4B	0.00527517	0.046001982
BDNF	0.202206651	0.039089523
IL11	0.121242355	0.035304811
INHA	0.007145753	0.010420721
ANGPTL4	0.005490665	0.022431459
TNFRSF11A	0.23831301	1.13E-06
VIPR1	-0.100145639	0.049630993

The 13 r-IRGs involved VIPR1, FURIN, RBP2, FGF2, SLC11A1, SEMA4B, BDNF, IL11, TNFRSF11A, ANGPTL4, INHA, S100A16, CRABP1. In the model, the risk score was generated using the coefficients expression values. We refer to this risk score as the "immune-based prognostic index" in this section (IPI). On the basis of the median value of IPI, patients were divided into two groups: those at high risk of developing cancer and those at low risk. Figures 4A–C shows the survival status of the discovery cohort (TCGA) as well as the expression of r-IRGs. Based on the median risk score, patients with LUAD in different cohorts were categorised into high- and low-risk groups, respectively. Because of the statistically substantial difference between the survival curves of the high risk and low risk groups reported in both discovery (TCGA) and validation (GSE31210) cohorts, IPI has been shown to be a good predictor of mortality (Figure 5A, Figure 5B). The dependability of this risk model was validated using ROC analysis, which we performed. The area under the receiver operating characteristic (ROC) curve for the discovery cohort (TCGA) and validation cohort (GSE31210) was 0.747 and 0.688, respectively, for the two cohorts (Figure 5C, Figure 5D).

As a result of these findings, IPI has a moderate potential for developing a prognostic signature based on IRGs in the context of survival monitoring. An further study was carried out by the prognostic model,

which was a univariate COX regression analysis. It was demonstrated that the clinical stage, the T stage, the N stage, and the risk score were all strongly connected with the prognosis of LUAD patients ($P < 0.001$) (Figure 6A). Once this was completed, a multivariate COX regression analysis was conducted on both clinical characteristics and the risk score. Only the clinical stage and the risk score were shown to be significantly associated with the prognosis of LUAD patients ($P < 0.05$). (Figure 6B). Finally, IPI has the potential to become a standalone predictor.

Clinical utility of prognostic signature

In order to gain a better understanding of the clinical significance of r-IRGs and IPI, we investigated the link between them and clinical pathology, including age, gender, clinical stage, regional lymph node metastasis (N), tumour burden (T), and distant metastases (M). ANGPTL4 was positively correlated with advanced T-stage cases (Figure 7A, $P = 0.014$). FURIN was higher expressed in male patients (Figure 7B) and positively correlated with lymph node metastases (Figure 7C, $P = 0.034$). INHA was positively correlated with the clinical stage (Figure 7D, $P = 0.046$). RBP2 were correlated with no distant metastasis (Figure 7E, $P = 0.004$). S100A16 positively correlated with lymph node metastases (Figure 7F, $P = 0.003$) and advanced case (Figure 7G, $P = 0.047$). SEMA4B was negatively correlated with distant metastatic cases (Figure 7H, $P = 0.047$) but with lymph node metastases cases positively (Figure 7I, $P = 0.006$). SLC11A1 was negatively correlated with distant metastatic cases (Figure 7J, $P = 0.010$), no lymph node metastases cases (Figure 7K, $P = 0.011$) and advanced cases (Figure 7L, $P = 9.913 \times 10^{-4}$). VIPR1 was positively correlated with the cases without lymph node metastases (Figure 7M, $P = 0.004$).

IRI and Immune Cell Infiltration Correlation

For the purpose of determining the potential of IRI to forecast the condition of the immune microenvironment, the CIBERSORT algorithm was employed to evaluate the quantity of tumor-infiltrating immune cells in the LUAD patients' immune microenvironment. In our study, we discovered that risk score was significantly positively correlated to NK cell resting levels (Figures 8).

Discussion

Infiltrates of lymphocytes are widespread in lung adenocarcinoma (LUAD), the most common kind of non-small-cell lung cancer. This shows that the immune system is actively involved in the formation and progression of the cancer (5). The identification of many important gene makers associated with survival prognosis, as proven by existing research, is critical, particularly for cancers with substantial tumour heterogeneity and poor prognosis. In addition to the present histopathological staging system, "Immunoscore" is a quantitative forecasting tool based on tumour immune contexture that is currently being tested in clinical trials in a variety of cancer types as a supplement (11, 12).

As a result, in this retrospective investigation, we developed a 13-gene survival risk predictive model connected to the overall survival of LUAD patients. By using a multivariate logistic regression analysis, researchers were able to determine the relationships between these prognosis-related genes and clinical survival. In addition, the capacity of the risk model to predict the condition of the immunological

microenvironment was explored. We first assessed LUAD immune infiltration using the TCGA LUAD cohort and then identified the IRGs that were differentially expressed between the LUAD and nearby non-tumor LUAD samples in the LUAD. Using massive RNA sequencing and weighted immune-related gene expression, it is possible to determine the health of tumour cells. According to the KEGG analysis, the most significant pathway was found to be enriched in the process of Cytokine-cytokine receptor interaction. GO terms that were significantly enriched in terms of biological process and molecular function included terms that were highly connected to receptor-ligand action. Exactly as predicted, the functional enrichment analysis revealed that these genes were largely involved in signalling pathways that were relevant to the immune system. Furthermore, these enhanced signalling pathways are typically implicated in tumorigenesis-related processes such as tumour growth, invasion, and metastasis, implying that immune-related genes may be useful in predicting clinical outcomes. The TCGA LUAD cohort was utilised to develop a prognostic gene signature, which was then used to predict the survival outcomes of LUAD patients based on their gene expression profiles. A greater risk score was found to be associated with a poorer overall survival rate.

All of the genes associated with risk model construction are found in S100A16, CRABP1-RBP2, FURIN, FGF2, SLC11A1, SEMA4B, BDNF-IL11, INHA, ANGPTL4, TNFRSF11A, and VIPR1 (see list below). The S100 protein family of genes is one of those that plays a critical role in the development and progression of tumours at various stages. A large number of human pathophysiological processes, including adipogenesis, inflammation, tumour progression, and osteoporosis, have been linked to the S100 gene family. S100A16, the S100 family's latest member has been linked to several human pathophysiological conditions, including tumour progression, inflammation, adipogenic differentiation, and osteoporosis (13–15). When it comes to lung cancer, high S100A16 expression has been demonstrated to be significantly associated with a bad prognosis (16–18). SEMA4B is a member of the semaphorin protein family, which is responsible for largely regulating cell motility and adhesion. In vitro studies have demonstrated that SEMA4B inhibits the invasion of non-small cell lung cancer via inhibiting the PI3K/Ak signalling pathway (19). IL-11 is a member of glycoprotein (GP) 130 cytokine family that also includes IL-6 and IL-27 (20). These cytokines share a receptor component GP130 and can activate the Janus kinase (JAK)/STAT3 pathway (21–24). In the past, the binding of interleukin-11 (IL-11) to its transmembrane co-receptor, IL-11R, and the subsequent signalling cascade have been related with platelet maturation, neurogenesis, osteoclastogenesis, and adipogenesis, among other things (25, 26). Both in mouse models and in humans, interleukin-11 (IL-11) has been linked to a variety of inflammation-associated malignancies, mostly because of its potential to activate the JAK-STAT3 signalling pathway. (27). This protein, also known as angiopoietin-like 4 (ANGPTL4), is a member of the angiopoietin superfamily of secreted proteins that are structurally linked to factors that modulate angiogenesis. ANGPTL4 has been found in a number of solid tumours, indicating that it may play an essential role in cancer growth and progression, altered redox regulation, anoikis resistance, metastasis and angiogenesis (28–30). Hypoxia conditions, which are a significant characteristic of the tumour microenvironment, may serve as a link between ANGPTL4 and carcinogenesis in one instance. Indeed, hypoxia-inducible factor-1 (HIF-1) promotes the overexpression of cyclooxygenase-2 (COX-2) in response to hypoxia (HIF-1) (31–33),

Prostanoids, particularly prostaglandins PGE₂, are produced as a result of this process. Increased levels of PGE₂ activate an intracellular signalling cascade, which results in the stimulation of the ANGPTL4 gene expression in the cells (34). A type I homotrimeric transmembrane protein, tumour necrosis factor receptor superfamily member 11a (TNFRSF11a), shares the highest amount of homology with the CD40 antigen (35). In addition to the lung, skeletal muscle, brain, heart, kidney, skin and liver, TNFRSF11a is found in high concentrations in certain malignancies. Following an earlier work in mice (35–37), it was discovered that TNFRSF11a-mediated intracellular signalling is required for mammary gland development through controlling the expansion of the stem and progenitor cell compartments, as well as some cancers. In a recent study in mice, it was discovered that TNFRSF11a-mediated intracellular signalling is required for mammary gland development via regulating the expansion of the stem and progenitor cell compartments (38). TNFRSF11a overexpression in mice, on the other hand, resulted in aberrant proliferation and poor differentiation, resulting in an increase in the incidence of cancer (39). The vasoactive intestinal peptide (VIP), also known as PACAP, is a critical neuropeptide that regulates lung physiology. It primarily acts via the VAPC1 and VAPC2 receptor subtypes. VIP is an autocrine lung cancer growth factor whose receptor is expressed in a large number of lung cancer cell lines. Several studies have shown that VIPR1 inhibits the growth of several cancers, including prostate cancer, lymphoblastoma, and medulloblastoma, implying that VIPR1 may significantly inhibit the growth and development of cancer cells. These findings shed light on the role of VIPR1 expression in regulating lung adenocarcinoma's biological behaviour (40). The remainder of the LUAD genes' function and regulation mechanism are largely unclear.

In summary, we developed an immunological signature based on IRGs for the purpose of predicting prognosis in patients with LUAD. This model has the potential to be a valuable tool for differentiating LUAD patients based on probable distinct clinical outcomes.

Declarations

Author Contributions Statement

PC contributed to the concept. M.A., G.H., and Y.D. wrote the manuscript. F.Z. and J.C. participated in the study design and helped draft the manuscript. M.A. and R.Z. performed the literature search and collected the data. M.A., G.H. and D.Y. analyzed the data. All authors reviewed the manuscript.

Declaration

The patients involved in the database have obtained ethical approval. Users can download relevant data for free for research and publish relevant articles. Our study is based on open source data, so there are no ethical issues and other conflicts of interest. The open source data from TCGA and GEO were performed in accordance with the Declaration of Helsinki. All methods were carried out in accordance with guidelines and regulations associated with each study.

Acknowledgments

We acknowledge TCGA and GEO database for providing their platforms and contributors for uploading their meaningful datasets.

Conflict of Interests

Authors state no conflict of interest.

References

1. Thandra KC, Barsouk A, Saginala K, Aluru JS, Barsouk A. Epidemiology of lung cancer. *Contemp Oncol (Pozn)*. (2021) 25:45–52. doi: 10.5114/wo.2021.103829
2. Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. *CA Cancer J Clin*. (2017) 67:7-30. doi: 10.3322/caac.21387
3. Giffin MJ, Cooke K, Lobenhofer EK, Estrada J, Zhan J, Deegen P, et al. AMG 757, a Half-Life Extended, DLL3-Targeted Bispecific T-Cell Engager, Shows High Potency and Sensitivity in Preclinical Models of Small-Cell Lung Cancer. *Clin Cancer Res*. (2021) 27:1526–37. doi: 10.1158/1078-0432.CCR-20-2845
4. Gridelli C, Rossi A, Carbone DP, Guarize J, Karachaliou N, Mok T, et al. Non-small-cell lung cancer. *Nat Rev Dis Primers*. (2015) 1:15009. doi: 10.1038/nrdp.2015.9
5. Luo C, Lei M, Zhang Y, Zhang Q, Li L, Lian J, et al. Systematic construction and validation of an immune prognostic model for lung adenocarcinoma. *J Cell Mol Med*. (2020) 24:1233–44. doi: 10.1111/jcmm.14719
6. Dong HX, Wang R, Jin XY, Zeng J, Pan J. LncRNA DGCR5 promotes lung adenocarcinoma (LUAD) progression via inhibiting hsa-mir-22-3p. *J Cell Physiol*. (2018) 233:4126–36. doi: 10.1002/jcp.26215
7. Fu L, Wang H, Wei D, Wang B, Zhang C, Zhu T, et al. The value of CEP55 gene as a diagnostic biomarker and independent prognostic factor in LUAD and LUSC. *PLoS One*. (2020) 15:e0233283. doi: 10.1371/journal.pone.0233283
8. Li R, Liu J, Fang Z, Liang Z, Chen X. Identification of Mutations Related to Cisplatin-Resistance and Prognosis of Patients With Lung Adenocarcinoma. *Front Pharmacol*. (2020) 11:572627. doi: 10.3389/fphar.2020.572627
9. Zhang L, Chen J, Cheng T, Yang H, Li H, Pan C. Identification of the key genes and characterizations of Tumor Immune Microenvironment in Lung Adenocarcinoma (LUAD) and Lung Squamous Cell Carcinoma (LUSC). *J Cancer*. (2020) 11:4965–79. doi: 10.7150/jca.42531
10. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. (2015) 12:453–7. doi: 10.1038/nmeth.3337
11. Angell H, Galon J. From the immune contexture to the Immunoscore: the role of prognostic and predictive immune markers in cancer. *Curr Opin Immunol*. (2013) 25:261–7. doi:

12. Pagès F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet*. (2018) 391:2128–39. doi: 10.1016/S0140-6736(18)30789-X
13. Domínguez B, Pardo BG, Noia M, Millán A, Gómez-Tato A, Martínez P, et al. Microarray analysis of the inflammatory and immune responses in head kidney turbot leucocytes treated with resveratrol. *Int Immunopharmacol*. (2013) 15:588–96. doi: 10.1016/j.intimp.2013.01.024
14. Li D, Zhang R, Zhu W, Xue Y, Zhang Y, Huang Q, et al. S100A16 inhibits osteogenesis but stimulates adipogenesis. *Mol Biol Rep*. (2013) 40:3465–73. doi: 10.1007/s11033-012-2413-2
15. Babini E, Bertini I, Borsi V, Calderone V, Hu X, Luchinat C, et al. Structural characterization of human S100A16, a low-affinity calcium binder. *J Biol Inorg Chem*. (2011) 16:243–56. doi: 10.1007/s00775-010-0721-3
16. Saito K, Kobayashi M, Nagashio R, Ryuge S, Katono K, Nakashima H, et al. S100A16 is a Prognostic Marker for Lung Adenocarcinomas. *Asian Pac J Cancer Prev*. (2015) 16:7039–44. doi: 10.7314/apjcp.2015.16.16.7039
17. Katono K, Sato Y, Kobayashi M, Nagashio R, Ryuge S, Igawa S, et al. S100A16, a promising candidate as a prognostic marker for platinum-based adjuvant chemotherapy in resected lung adenocarcinoma. *Onco Targets Ther*. (2017) 10:5273–9. doi: 10.2147/OTT.S145072
18. Kobayashi M, Nagashio R, Saito K, Aguilar-Bonavides C, Ryuge S, Katono K, et al. Prognostic significance of S100A16 subcellular localization in lung adenocarcinoma. *Hum Pathol*. (2018) 74:148–55. doi: 10.1016/j.humpath.2018.01.001
19. Jian H, Zhao Y, Liu B, Lu S. SEMA4B inhibits growth of non-small cell lung cancer in vitro and in vivo. *Cell Signal*. (2015) 27:1208–13. doi: 10.1016/j.cellsig.2015.02.027
20. Kishimoto T, Akira S, Narazaki M, Taga T. Interleukin-6 family of cytokines and gp130. *Blood*. (1995) 86:1243-54.
21. Bromberg J. Stat proteins and oncogenesis. *J Clin Invest*. (2002) 109:1139–42. doi: 10.1172/JCI15617
22. Bromberg JF, Wrzeszczynska MH, Devgan G, Zhao Y, Pestell RG, Albanese C, et al. Stat3 as an oncogene. *Cell*. (1999) 98:295–303. doi: 10.1016/s0092-8674(00)81959-5
23. Garbers C, Hermanns HM, Schaper F, Müller-Newen G, Grötzinger J, Rose-John S, et al. Plasticity and cross-talk of interleukin 6-type cytokines. *Cytokine Growth Factor Rev*. (2012) 23:85–97. doi: 10.1016/j.cytogfr.2012.04.001
24. Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat Rev Cancer*. (2009) 9:798–809. doi: 10.1038/nrc2734
25. Schwertschlag US, Trepicchio WL, Dykstra KH, Keith JC, Turner KJ, Dorner AJ. Hematopoietic, immunomodulatory and epithelial effects of interleukin-11. *Leukemia*. (1999) 13:1307-15. doi: 10.1038/sj.leu.2401514

26. Teramura M, Kobayashi S, Yoshinaga K, Iwabe K, Mizoguchi H. Effect of interleukin 11 on normal and pathological thrombopoiesis. *Cancer Chemother Pharmacol.* (1996) 38 Suppl:S99-102. doi: 10.1007/s002800051048
27. Xu DH, Zhu Z, Wakefield MR, Xiao H, Bai Q, Fang Y. The role of IL-11 in immunity and cancer. *Cancer Lett.* (2016) 373:156–63. doi: 10.1016/j.canlet.2016.01.004
28. Zhu P, Goh YY, Chin HF, Kersten S, Tan NS. Angiopoietin-like 4: a decade of research. *Biosci Rep.* (2012) 32:211–9. doi: 10.1042/BSR20110102
29. Nakayama T, Hirakawa H, Shibata K, Nazneen A, Abe K, Nagayasu T, et al. Expression of angiopoietin-like 4 (ANGPTL4) in human colorectal cancer: ANGPTL4 promotes venous invasion and distant metastasis. *Oncol Rep.* (2011) 25:929–35. doi: 10.3892/or.2011.1176
30. Huang XF, Han J, Hu XT, He C. Mechanisms involved in biological behavior changes associated with Angptl4 expression in colon cancer cell lines. *Oncol Rep.* (2012) 27:1541–7. doi: 10.3892/or.2012.1672
31. Han S, Xu W, Wang Z, Qi X, Wang Y, Ni Y, et al. Crosstalk between the HIF-1 and Toll-like receptor/nuclear factor- κ B pathways in the oral squamous cell carcinoma microenvironment. *Oncotarget.* (2016) 7:37773–89. doi: 10.18632/oncotarget.9329
32. Fanale D, Bazan V, Caruso S, Castiglia M, Bronte G, Rolfo C, et al. Hypoxia and human genome stability: downregulation of BRCA2 expression in breast cancer cell lines. *Biomed Res Int.* (2013) 2013:746858. doi: 10.1155/2013/746858
33. Fanale D, Bazan V, Corsini LR, Caruso S, Insalaco L, Castiglia M, et al. HIF-1 is involved in the negative regulation of AURKA expression in breast cancer cell lines under hypoxic conditions. *Breast Cancer Res Treat.* (2013) 140:505–17. doi: 10.1007/s10549-013-2649-0
34. Kim SH, Park YY, Kim SW, Lee JS, Wang D, DuBois RN. ANGPTL4 induction by prostaglandin E2 under hypoxic conditions promotes colorectal cancer progression. *Cancer Res.* (2011) 71:7010–20. doi: 10.1158/0008-5472.CAN-11-1262
35. Anderson DM, Maraskovsky E, Billingsley WL, Dougall WC, Tometsko ME, Roux ER, et al. A homologue of the TNF receptor and its ligand enhance T-cell growth and dendritic-cell function. *Nature.* (1997) 390:175–9. doi: 10.1038/36593
36. Wada T, Nakashima T, Hiroshi N, Penninger JM. RANKL-RANK signaling in osteoclastogenesis and bone disease. *Trends Mol Med.* (2006) 12:17–25. doi: 10.1016/j.molmed.2005.11.007
37. Hsu H, Lacey DL, Dunstan CR, Solovyev I, Colombero A, Timms E, et al. Tumor necrosis factor receptor family member RANK mediates osteoclast differentiation and activation induced by osteoprotegerin ligand. *Proc Natl Acad Sci U S A.* (1999) 96:3540–5. doi: 10.1073/pnas.96.7.3540
38. Fata JE, Kong YY, Li J, Sasaki T, Irie-Sasaki J, Moorehead RA, et al. The osteoclast differentiation factor osteoprotegerin-ligand is essential for mammary gland development. *Cell.* (2000) 103:41–50. doi: 10.1016/s0092-8674(00)00103-3
39. Bonifaci N, Palafox M, Pellegrini P, Osorio A, Benítez J, Peterlongo P, et al. Evidence for a link between TNFRSF11A and risk of breast cancer. *Breast Cancer Res Treat.* (2011) 129:947–54. doi:

40. Zhao L, Yu Z, Zhao B. Mechanism of VIPR1 gene regulating human lung adenocarcinoma H1299 cells. Med Oncol. (2019) 36:91. doi: 10.1007/s12032-019-1312-y

Figures

Figure 1

Figure 1

Immune-related genes with differential expression. A heatmap (A) and volcano plot (B) of genes that are differently expressed between lung adenocarcinomas (LUAD) and surrounding non-tumor tissues. The heatmap (C) and volcano plot (D) depict differential expression of immune-related genes (IRGs). The black dots indicate undifferentiated genes, whereas the red and green spots indicate differentiated genes. (E) Forest plot of hazard ratios illustrating immune-related gene prognostic values.

Figure 2

Figure 2

Functional enrichment of IRGs' genes. The analysis of differentially expressed immune-related genes using the Gene Ontology (GO) and the important Kyoto Encyclopedia of Genes and Genomes (KEGG) (B). Analysis of prognostic-associated immune-related genes using the Gene Ontology (GO) and the most significant Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (D).

Figure 3

Figure 3

A regulatory network mediated by transcription factors. The heatmap (A) and volcano plot (B) depict transcription factors that are differentially expressed (TFs). (C) Regulatory network based on clinically significant transcription factors and IRGs.

Figure 4

Figure 4

The development of a prognostic risk model using immune-related genes. (A) Prognostic risk score and group distribution. (B) Survival rates of patients classified into several groups. (C) Heatmap of the included genes' expression profiles.

Figure 5

Figure 5

The risk score's prognostic value. The overall survival (OS) time for patients in the high-risk and low-risk groups in the training data set (A) and the testing data set (B). The receiver operating characteristic curves (ROC) for the 13-gene immune-related risk score. (C) Dataset for training (D) Dataset for testing.

A

B

Figure 6

Figure 6

Using univariate analysis (A) and multivariate analysis (B), a forest plot of hazard ratios depicting the prognostic values of immune-related genes involved in the risk model is shown (B).

Figure 7

Figure 7

The association of a risk immunological gene with clinicopathologic characteristics. The correlation between ANGPTL4 expression and tumour stage (A); The associations between FURIN and the gender of patients (B) and lymph node metastases (C); INHA-clinical stage(D) correlations; RBP2 and distant metastasis(E); S100A16 and lymph node metastases (F) and clinical stage (G); SEMA4B and distant

metastasis(H) and lymph node metastasis(I); SLC11A1 and distant metastases (J), lymph node metastasis (K), and clinical stage (L); VIPR1 and lymph node metastasis: a connection (M)

Figure 8

Figure 8

Correlations between the Immune Gene Risk Index (IRI) and immune cell infiltration abundances. (A) NK cells at rest; (B) CD8 T cells; (C) neutrophils; (D) memory activated CD4 T cells; (E) active NK cells; (F) monocytes.