

Identification of Key Prognosis-Associated Genes in the Lung Adenocarcinoma

HongBo Ma

Fuling central hospital of chongqing city

XiaoLi Wu

fuling central hospital of chongqing city

MiaoMiao Tao

fuling central hospital of chongqing city

Uthaman Jithin

2nd affiliated hospital of chongqing medical university

GenDou Zhou

Fuling central hospital of Chongqing city

YanYan Li

Fuling central hospital of chongqing city

Qi Zhou (✉ qizhou112@163.com)

Fuling central hospital of chongqing city

Research article

Keywords: Lung adenocarcinoma, bioinformatics, differentially expressed genes, TCGA

DOI: <https://doi.org/10.21203/rs.3.rs-21405/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Non-small cell lung cancer (NSCLC) is one of the most common causes of cancer-related death globally, and lung adenocarcinoma (LUAD) accounts for almost 40% of all lung cancer cases. In recent years, despite better understanding of the pathogenesis of the disease and achievements in the multimodal treatment of tumors, there is an urgent need to identify new diagnostic and prognostic biomarkers. In this study, we aim to identify the potential key genes related to the pathogenesis and prognosis of LUAD by using comprehensive bioinformatics analysis. The gene expression profile was downloaded from The Cancer Genome Atlas (TCGA) database, and we calculated the LUAD immune scores and stromal scores by using the ESTIMATE algorithm. Based on these scores, we further quantified the immune and stromal components and obtained the differentially expressed genes (DEGs) in the tumor. Overall survival analysis could better reflect the impact of genes related to immune and stromal cells on the prognosis. Moreover, Gene Ontology (GO) functional enrichment analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment were conducted, while protein-protein interaction (PPI) network was obtained from STRING. Analysis of the correlation revealed that these genes are mainly involved in the immune/inflammatory response. In conclusion, our study showed that 11 prognostic genes (CD33, IRF8, CD80, CD53, IL16, LY86, CD79B, TYROBP, CD1E, CD1C, and CD1B) might show a potentially good performance in predicting overall survival in patients with LUAD. In summary, we identified the key genes related to the microenvironment, which can further serve as the prognostic biomarkers and therapeutic targets for LUAD.

Introduction

Lung cancer is one of the malignant tumors that seriously affect the health of people. In 2018, an estimated 2.09 million new cases and 1.76 million lung cancer-related cancer deaths worldwide^[1]. The 5-year relative survival for NSCLC of all stages is 23%^[2]. LUAD is the main type of NSCLC. The causes of LUAD include gene aberrations. Compared with histological classification, a gene plays an important role in the diagnosis, treatment, and prognosis of LUAD. In recent years, we have learned more about the epidemiology, pathology, and molecular mechanisms of LUAD, and we have more options for implementing emerging treatments (such as targeted and immune-based therapies). However, LUAD is still considered a complex disease. Despite certain biomarkers, some patients do not respond to existing molecularly targeted agents^[3]. To date, due to its biological complexity, the exact molecular mechanism of LUAD is still unclear and based on the poor prognosis and high recurrence rate of LUAD, there is an urgent need to explore novel biomarkers.

Tumor cell-intrinsic genes determine the occurrence, development, initiation, and evolution of LUAD. The tumor microenvironment not only has a high proportion of specific gene expression profiles, but it also plays a very important role in the biological and clinical behavior of cancer cells^[4]. The tumor microenvironment is a dynamically balanced cellular environment composed of tumor cells, peripheral blood vessels, immune cells, stromal cells, mesenchymal cells, endothelial cells, and extracellular matrix

(ECM) molecules, along with different types of normal cells^[5, 6]. Among these cells, immune cells and stromal cells are the two main types of non-malignant cells that contribute to the diagnosis and prognosis assessment of tumors.

ESTIMATE (Estimation of STromal and Immune cells in MAlignant Tumor tissues using Expression data) is a tool for predicting tumor purity and the presence of infiltrating stromal/immune cells in tumor tissues using the gene expression data. The ESTIMATE algorithm is based on a single sample Gene Set Enrichment Analysis, and it generates immune cells and stromal scores^[7]. In this algorithm, the authors analyzed the specific gene expression profile of immune and stromal cells and obtained immune and stromal scores to predict the infiltration of non-malignant cells. There are reports of application of the ESTIMATE algorithm to prostate cancer^[8], breast cancer^[9], colon cancer^[10], glioblastoma^[11], and acute myeloid leukemia^[12]. However, no detailed evaluation of the immune and/or stromal scores of LUAD has been conducted.

In the present study, we downloaded the complete gene expression profiles for all LUAD patients from The Cancer Genome Atlas (TCGA) database, while using the ESTIMATE algorithm to obtain the immune and/or stromal scores.

We further used a series of bioinformatics tools (GO and KEGG) to analyze the main biological functions regulated by DEGs. Finally, to extract the prognostic indicators for LUAD patients, we identified the key genes affecting the incidence and prognosis through protein-protein interaction (PPI) networks and survival analysis.

Materials And Methods

Database

The mRNA gene expression profile and the corresponding clinical information of all 486 adult LUAD patients with the preliminary pathologic diagnosis were downloaded from TCGA (<https://gdc.nci.nih.gov/>) in November 2019. The compiled clinical data mainly included gender, age, clinical stage, overall survival, and prognosis. No exclusion criteria were set for all obtained data.

Data preprocessing and identification of DEGs based on immune scores and stromal scores

The matrix data for each mRNA data set were normalized using the Perl language, and the gene ID was converted. The Limma package was applied(1.

<http://www.bioconductor.org/packages/devel/bioc/html/limma.html>; 2. Limma powers differential expression analyses for RNA-sequencing and microarray studies). Immune scores and stromal scores were calculated by ESTIMATE packages

(<https://bioinformatics.mdanderson.org/estimate/rpackage.html>)in the R software. Eventually, gene expression values and clinical information of samples from different data sets were integrated. All LUAD

patients were classified into high- and low-score groups according to their immune/stromal scores. In this study, genes with a p-value < 0.05 and |fold change| > 1.5 were defined as DEGs. The heatmap and Venn diagrams of the DEGs were constructed using the package edge R. All statistical analyses were performed using R software (version 3.6.1).

Functional enrichment analysis of DEGs

To clarify the underlying biological processes, molecular functions, and cellular components associated with overlapping DEGs, we used the color space software package, stringi software package, and ggplot2 in R software to perform functional enrichment analysis of GO and KEGG to clarify the correlation with overlapping DEGs. A p-value < 0.05 and an adjusted p-value < 0.05 were defined as the cut-off criteria.

PPI network for functional enrichment analysis

The STRING online tool (<https://string-db.org/cgi/input.pl>) was applied to identify potential protein-protein interactions among the overlapping DEGs. PPIs with a confidence score ≥ 0.95 were reserved.

Survival analysis

The Perl language was used to integrate DEGs, survival time and state, samples, and other matrix data. The survival software package was used in R software to generate Kaplan-Meier plots to clarify the correlation between the expression of DEGs and the overall survival of patients with LUAD. The statistical significance of the correlation was tested by a log-rank test.

Results

Immune conditions are related to the clinical characteristics of LUAD patients

A total of 486 LUAD patients with complete gene expression profiles and clinical information were obtained from the TCGA database. The clinicopathological characteristics, including age, gender, grade, and clinical stage, are shown in Table 1. Among them, 222 (45.7%) patients were male and 264 (54.3%) patients were female. The age at the initial pathological diagnosis ranged from 33 years to 88 years, and 19 patients were of unknown age. The diagnosis of four pathological stages in these patients included cases of Stage I (262, 53.9%), Stage II (112, 23%), Stage III (79, 16.3%), and Stage IV (25, 5.1%), while 8 (1.6%) patients had an unknown stage (Table 1). Based on the ESTIMATE algorithm, we calculated immune scores (ranging from - 1789 to 2097) and stromal scores (ranging from - 942 to 3442) for all the patients (**supplementary scores.xls**). We plotted the distribution of immune scores and stromal scores based on the clinical information. As shown in Fig. 1C, T4 cases had the lowest average immune scores of all 4 T stages, and cases with distant metastasis (M1) had the lowest stromal scores (Fig. 1H). Besides, the immune scores (Fig. 1I) and stromal scores were significantly associated with the gender subtype (Fig. 1J). No significant differences in the subtype (stage and N staging) were found for both immune scores and stromal scores (Fig. 1A, B, E, F).

Table 1
characteristics of LUAD samples

characteristics	Number of patients	percentage(%)
age		
<60	136	27.90%
≥ 60	331	68.10%
Unknown	19	4%
gender		
male	222	45.70%
female	264	54.30%
Stage		
Stage I	262	53.90%
Stage II	112	23%
Stage III	79	16.2%
Stage IV	25	5.1%
unknown	8	1.6%
T		
T1	163	33.5%
T2	260	53.4%
T3	41	8.4%
T4	19	3.9%
Tx	3	0.6%

To determine the potential relationship between the overall survival and immune scores with/or stromal scores, we classified the 486 LUAD patients into 2 groups (high vs low score groups). From the Kaplan-Meier survival curves, it was concluded that patients with higher median immune scores had a longer overall survival compared to those with lower immune scores (Fig. 1K, $p = 0.025$). Although not statistically significant, from the K-M curve, we could still determine that patients with higher stromal scores had a longer median overall survival than patients with lower stromal scores (Fig. 1L, $p = 0.101$).

Identification of DEGs based on immune scores and stromal scores in LUAD

To further explore the correlation between gene expression profiles and immune scores with/or stromal scores, we compared the RNA-Seq data for all 486 LUAD cases downloaded from the TCGA database. The cut-off criteria were set to $p < 0.05$ and $\log_{2}FC$ filter = 1. DEGs for the high vs low immune score/stromal score groups are illustrated in the heatmap shown in Fig. 2A. Venn diagrams (Fig. 2B) show 62 commonly co-downregulated genes and 317 co-upregulated genes from the immune score/stromal score groups. Therefore, we focused our analysis on these overlapping DEGs.

GO term and KEGG pathway enrichment analysis of DEGs

The GO and KEGG pathway enrichment analyses were performed to elucidate the potential biological functions of 379 overlapping DEGs. As shown in Fig. 3A, the top GO terms identified included regulation of leukocyte activation, T cell activation and leukocyte differentiation (BP), external side of the plasma membrane (CC), carbohydrate-binding, and cytokine activity (MF). In addition, we conducted KEGG pathway enrichment and interrelation analysis. As shown in Fig. 3B, enrichment of DEGs mainly participated in the regulation of pathways related to environmental information processing and tumor progression, such as Cytokine-cytokine receptor interaction, Chemokine signaling pathway, Viral protein interaction with cytokine and cytokine receptor, and Hematopoietic cell lineage.

PPI network construction among genes of prognostic value

Based on the STRING online tool, we performed a PPI network to explore the interactions between the identified overlapping DEGs. When mapping the intersect genes into the STRING database, a network made up of 376 nodes and 165 edges was obtained (Fig. 3C). We analyzed the degree distribution of the nodes in the network, and hub genes with the top 30 node degrees were identified to construct the PPI network of target genes (Fig. 3D).

Survival analysis of DEGs

We generated Kaplan-Meier plots to further clarify the correlation between the expression of DEGs and overall survival of LUAD patients. Among the 379 overlapping DEGs, 129 genes were associated with overall survival according to the log-rank test ($p < 0.05$) (**supplementary 129 genes survival.xls**). Among the top 30 genes with the most edges in PPI, we found 11 representative genes related to prognosis (Fig. 4). We could conclude that high expression of these 11 genes had a better prognosis than their low expression. Therefore, these 11 genes may be involved in the molecular mechanism related to the microenvironment of LUAD and they can be used as predictors and prognostic indicators of LUAD.

Discussion

Cancer is a highly heterogeneous and complex disease involving a variety of genetic factors. The hallmarks of cancer include maintaining proliferative signaling, evading growth inhibitors, resisting cell death, immortalizing replication, inducing angiogenesis, and activating cancer cell invasion and metastasis^[13]. LUAD is one of the most invasive subtypes of lung cancer with a high recurrence rate,

short DFS, and easy distant metastasis^[14]. LUAD is composed of small airway epithelial, type II alveolar cells secreting mucus, and other substances^[15]. The current treatment of LUAD ranges from surgery to chemotherapy to radiotherapy to targeted therapy as well as immunotherapy^[16]. The rapid development of gene chip detection technology has great advantages for functionally related differential gene screening. An in-depth discussion of specific mutations related to LUAD can better provide theoretical basis for clinical individualized treatment of tumors. Some current studies have found that tumor heterogeneity and molecular differences in different individuals may lead to poor outcomes or even failure of treatment^[13]. Comprehensive bioinformatics analysis can be widely used to identify potential biomarkers related to the diagnosis, treatment, and prognosis of LUAD. These analysis methods include screening DEGs, finding network-based hub nodes, and performing survival analysis on DEGs. In our study, to understand the pathogenesis of LUAD at the molecular level, we analyzed the expression profiles of cancer-related genes based on the TCGA database, and a variety of bioinformatics methods, such as GO term analysis, KEGG enrichment analysis, and PPI network construction, were utilized to deeply explore the DEGs associated with the tumor microenvironment. Moreover, survival curves were drawn for common DEGs identified in the high vs low immune/stromal score groups.

First, based on the ESTIMATE algorithm, the immune scores and stromal scores of LUAD were different from statistically significant genders. In addition, the immune scores were meaningful for overall survival and T staging. In contrast, stromal scores were significantly related to distant metastasis. Many studies have shown that immune cells and stromal cells are part of the tumor microenvironment and have an impact on tumor progression, metastasis, and therapeutic resistance^[17, 18]. Importantly, stromal cells can be reprogrammed by tumor epithelial cells, and they acquire an “activate” protumorigenic phenotype.

A total of 379 overlapping genes were included in the overall survival analysis, of which 129 genes were correlated with the prognosis of LUAD patients ($p < 0.05$). Furthermore, 30 potential target genes with more hub nodes were selected and analyzed for survival. Eleven target genes with diagnostic and prognostic markers were extracted ultimately (CD33, IRF8, CD80, CD53, IL16, LY86, CD79B, TYROBP, CD1E, CD1C, and CD1B).

The CD33 is a well-known target in Myeloid-derived suppressor cells (MDSCs), which may constitute the innate system and facilitates efficient interaction within the immune cells. MDSCs can influence tumor progression and metastasis, and tumor microenvironment^[19]. Furthermore, a previous study has found that M-MDSCs were significantly increased after surgery. Besides, CD33 is involved in inhibiting T cells proliferation and affecting the development of Treg cells, and it is also involved in identifying tumor cells^[20].

IRF8, Interferon regulatory factor 8, is a member of the IRF family of transcription factors. IRF8 regulates the development and function of a variety of immune cells, linking innate and adaptive immunity^[21]. It plays a regulatory role in the cell's immune response to cancer and can mediate the expression of Fas, Bax, JAK, and STAT, leading to the apoptosis of solid tumor cells and has been shown to have antitumor

activity in a variety of tumor cells^[22]. IRF8 inhibits AKT signaling, affects apoptotic genes, and results in the senescence of lung cancer cells. The expression of IRF8 in tumor cells leads to the regression of lung cancer nodules and has a tumor-suppressive effect. It can be a potent lung tumor suppressor gene^[23].

CD80 is one of the Dendritic cells (DCs) mature markers^[24]. The expression level of CD80 on the surface of tumor-infiltrating dendritic cells (TIDCs) in tumor tissue was significantly lower than that in normal tissues and adjacent tissues, which may affect the antigen presentation of DC cells. As a co-stimulatory molecule, CD80 expression is involved in the activation and proliferation of T lymphocytes and is affected by the tumor microenvironment. It can also cause tumor cells to escape immune surveillance and promote tumor cell development^[25, 26]. In addition, lower CD80 expression independently predicts a poor prognosis in patients with gastric adenocarcinoma^[27].

CD53 is one of the earliest members of the tetraspanin/transmembrane-4 superfamily. CD53 can influence cell migration and adhesion^[28]. Lack of CD53 protein expression can lead to repeated infections in the body, indicating the important role of CD53 in immunity. CD53 is abnormally expressed in radiation-resistant tumor cells. Under some circumstances, CD53 antigen stimulation may protect against programmed cell death^[29]. CD53 may also activate the effector functions of NK cells and stimulate their proliferation^[30].

In recent years, it has been shown that IL-16 activates the stress-activated protein kinase (SAPK) signaling pathway in macrophages^[31]. IL-16 is a pro-inflammatory cytokine that stimulates the release of different types of pro-inflammatory factors, such as IL1 β , IL6, IL15, and Tumor Necrosis Factor- α (TNF- α), through monocytes. These cytokines play an important role in tumor formation^[32]. IL-16 may also regulate the expression of cytochrome P450 enzymes through microRNAs, further causing activation of various carcinogens and damaging the DNA. Besides, IL-16 mutations can cause changes in cell viability and function, and ultimately cause individuals to be susceptible to tumors^[33]. In NSCLC, patients with IL-16 gene mutations may show a higher risk of death and they severely affect patient survival^[34].

LY86, the "lymphocyte antigen 86" also known as protein Myeloid differentiation protein 1 (MD-1), is a well-known secreted glycoprotein that forms the radioprotective protein 105 (RP105) /MD-1 complex in a variety of pathological conditions, further targeting the TLR4 pathway^[35]. The RP105/MD-1 complex can be expressed on a variety of immune cells, including B cells, macrophages, and DCs^[36]. RP105/MD-1 RNA was found to be strongly up-regulated in tumor-associated macrophages (TAMs) as compared to normal macrophages, while tumor cells again showed no expression^[37].

CD79B is a component of the B-cell antigen receptor and is almost universal on the surface of most B-cell derived malignancies, making it an attractive therapeutic target. CD79B is mainly expressed in most patients with non-Hodgkin lymphoma (NHL) and chronic lymphocytic leukemia (CLL)^[38]. This gene has not been previously reported to be associated with the lung cancer microenvironment, but it may serve as a potential biomarker.

TYROBP, also known as DAP12; KARAP; PLOSL; PLOSL1, is expressed in natural killer cells, monocyte-macrophages, dendritic cells, and neutrophils, and it plays a pivotal role in the immune receptor tyrosine activation motif (ITAM) pathway^[39]. The DAP12/ITAM pathway might participate in the polarization of macrophages, which secrete TGF- β to act on tumor tissues, This pathway can also promote tumor progression and metastasis through TGF- β -mediated changes in the microenvironment of the tumor, such as EMT, changes in anti-apoptotic genes, increased escape from immune surveillance, and immune suppression^[40]. DAP12 is expressed in breast cancer (BRC) and hepatocellular carcinoma cells and is closely related to the malignant progression of tumors, distant metastases (such as bone metastases, liver metastases), and prognosis^[41]. A previous study suggested that TYROBP might be a target gene for diagnosing kidney cancer^[42].

CD1 molecules are usually expressed on malignant cells and can contribute to anti-tumor immune responses by presenting tumor-derived lipid and glycolipid antigens to T cells and NKT cells^[43]. CD1d has been demonstrated in not only hematological malignancies (such as AML, B-ALL, CLL, lymphomas, and multiple myeloma), but also in some solid tumors (glioma, medulloblastoma, renal cell carcinoma, and breast and prostate cancers)^[44]. CD1d has been shown to reduce lung tumor burden and enhance the anti-tumor immunity of DC cells^[45].

Conclusion

LUAD tissue not only changes the intracellular pathway but also the tumor microenvironment, which is different from normal lung tissue. Because the microenvironment of lung cancer is quite complicated, more experiments and comprehensive analysis are needed in a larger research population in the future. Fortunately, the development of bioinformatics and high-throughput tumor databases (such as TCGA) has made it possible to analyze the "big data" of large-scale LUAD populations. This study conducted a bioinformatics analysis on the TCGA's LUAD data set, with a focus on the immune microenvironment. We obtained 11 potential biomarkers for the treatment of LUAD, which revealed the pathogenesis of LUAD to a certain extent, and provided a theoretical basis for the study of LUAD treatment targets. Taken together, we identified the TME-related DEGs and predicted the prognosis-related hub genes in LUAD. However, it is necessary to further validate our findings using basic cell and animal experiments. Thus, further clinical researches are needed to uncover the roles in prognosis and therapy.

Abbreviations

NSCLC Non-small cell lung cancer

LUAD lung adenocarcinoma

TCGA The Cancer Genome Atlas

ESTIMATE Estimation of STromal and Immune cells in Malignant Tumours using Expression data

DEGs Differentially expressed genes

GO Gene Ontology

KEGG Kyoto Encyclopedia of Genes and Genomes

PPI Protein-protein interaction

ECM Extracellular matrix

BP Biological processes

CC Cellular component

MF Molecular function

MDSCs Myeloid-derived suppressor cells

IRF8 Interferon regulatory factor 8

DCs Dendritic cells

TIDCs Tumor-Infiltrating Dendritic cells

NK cells Natural killer cells

SAPK Stress-activated protein kinase

TNF- α Tumor Necrosis Factor- α

LY86 Lymphocyte antigen 86

MD-1 Myeloid differentiation protein 1

RP105 Radioprotective protein 105

TAMs tumor-associated macrophages

NHL non-Hodgkin lymphoma

CLL Chronic lymphocytic leukemia

Declarations

Competing interests:

The authors declare that they have no competing interests.

Funding:

Not Applicable.

Authors' contributions

Qi Zhou and YanYan Li conceived and designed the study. Xiaoli Wu obtained the datasets. GenDou Zhou and MiaoMiao Tao conducted data analysis. Uthaman Jithin made the figures, HongBo Ma provided valuable recommendations, contributing to the study design and wrote the manuscript, Qi Zhou revised the manuscript. All authors reviewed and approved the final manuscript.

Acknowledgments

We thank the TCGA databases for the availability of the data. We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. *Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries*. *CA Cancer J Clin*. 2018. 68(6): 394–424.
2. Miller KD, Nogueira L, Mariotto AB, et al *Cancer treatment and survivorship statistics, 2019*. *CA Cancer J Clin*. 2019. 69(5): 363–385.
3. Mitsudomi T, Yatabe Y. *Mutations of the epidermal growth factor receptor gene and related genes as determinants of epidermal growth factor receptor tyrosine kinase inhibitors sensitivity in lung cancer*. *Cancer Sci*. 2007. 98(12): 1817-24.
4. Mittal V, El Rayes T, Narula N, McGraw TE, Altorki NK, Barcellos-Hoff MH. *The Microenvironment of Lung Cancer and Therapeutic Implications*. *Adv Exp Med Biol*. 2016. 890: 75–110.
5. Hui L, Chen Y. *Tumor microenvironment: Sanctuary of the devil*. *Cancer Lett*. 2015. 368(1): 7–13.
6. Hanahan D, Coussens LM. *Accessories to the crime: functions of cells recruited to the tumor microenvironment*. *Cancer Cell*. 2012. 21(3): 309 – 22.
7. Yoshihara K, Shahmoradgoli M, Martínez E, et al *Inferring tumour purity and stromal and immune cell admixture from expression data*. *Nat Commun*. 2013. 4: 2612.
8. Shah N, Wang P, Wongvipat J, et al *Regulation of the glucocorticoid receptor via a BET-dependent enhancer drives antiandrogen resistance in prostate cancer*. *Elife*. 2017. 6.
9. Priedigkeit N, Watters RJ, Lucas PC, et al. *Exome-capture RNA sequencing of decade-old breast cancers and matched decalcified bone metastases*. *JCI Insight*. 2017. 2(17)..
10. Alonso MH, Aussó S, Lopez-Doriga A, et al *Comprehensive analysis of copy number aberrations in microsatellite stable colon cancer in view of stromal component*. *Br J Cancer*. 2017. 117(3): 421–431.
11. Jia D, Li S, Li D, Xue H, Yang D, Liu Y. *Mining TCGA database for genes of prognostic value in glioblastoma microenvironment*. *Aging (Albany NY)*. 2018. 10(4): 592–605.
12. Yan H, Qu J, Cao W, et al *Identification of prognostic genes in the acute myeloid leukemia immune microenvironment based on TCGA data analysis*. *Cancer Immunol Immunother*. 2019. 68(12): 1971–

1978.

13. Hanahan D, Weinberg RA. *Hallmarks of cancer: the next generation*. *Cell*. 2011. 144(5): 646 – 74.
14. Kerr KM. *Pulmonary adenocarcinomas: classification and reporting*. *Histopathology*. 2009. 54(1): 12–27.
15. Noguchi M, Morikawa A, Kawasaki M, et al *Small adenocarcinoma of the lung. Histologic characteristics and prognosis*. *Cancer*. 1995. 75(12): 2844-52.
16. Zappa C, Mousa SA. *Non-small cell lung cancer: current treatment and future advances*. *Transl Lung Cancer Res*. 2016. 5(3): 288–300.
17. Gao D, Mittal V. *The role of bone-marrow-derived cells in tumor growth, metastasis initiation and progression*. *Trends Mol Med*. 2009. 15(8): 333 – 43.
18. Junttila MR, de Sauvage FJ. *Influence of tumour micro-environment heterogeneity on therapeutic response*. *Nature*. 2013. 501(7467): 346 – 54.
19. Talmadge JE, Gabrilovich DI. *History of myeloid-derived suppressor cells*. *Nat Rev Cancer*. 2013. 13(10): 739 – 52.
20. Wang J, Yang L, Yu L, et al *Surgery-induced monocytic myeloid-derived suppressor cells expand regulatory T cells in lung cancer*. *Oncotarget*. 2017. 8(10): 17050–17058.
21. Twum DY, Colligan SH, Hoffend NC, et al *IFN regulatory factor-8 expression in macrophages governs an antimetastatic program*. *JCI Insight*. 2019. 4(3).
22. Mattei F, Schiavoni G, Sestili P, et al *IRF-8 controls melanoma progression by regulating the cross talk between cancer and immune cells within the tumor microenvironment*. *Neoplasia*. 2012. 14(12): 1223-35.
23. Liang J, Lu F, Li B, et al *IRF8 induces senescence of lung cancer cells to exert its tumor suppressive function*. *Cell Cycle*. 2019. 18(23): 3300–3312.
24. Dyck L, Mills K. *Immune checkpoints and their inhibition in cancer and infectious diseases*. *Eur J Immunol*. 2017. 47(5): 765–779.
25. Antonia SJ, Muñoz-Antonia T, Soldevila G, Miller J, Flavell RA. *B7-1 expression by a non-antigen presenting cell-derived tumor*. *Cancer Res*. 1995. 55(11): 2253-6.
26. Jia J, Wang Z, Li X, Wang Z, Wang X. *Morphological characteristics and co-stimulatory molecule (CD80, CD86, CD40) expression in tumor infiltrating dendritic cells in human endometrioid adenocarcinoma*. *Eur J Obstet Gynecol Reprod Biol*. 2012. 160(2): 223-7.
27. Feng XY, Lu L, Wang KF, et al *Low expression of CD80 predicts for poor prognosis in patients with gastric adenocarcinoma*. *Future Oncol*. 2019. 15(5): 473–483.
28. Puls KL, Hogquist KA, Reilly N, Wright MD. *CD53, a thymocyte selection marker whose induction requires a lower affinity TCR-MHC interaction than CD69, but is up-regulated with slower kinetics*. *Int Immunol*. 2002. 14(3): 249 – 58.
29. Yunta M, Lazo PA. *Apoptosis protection and survival signal by the CD53 tetraspanin antigen*. *Oncogene*. 2003. 22(8): 1219-24.

30. Todros-Dawda I, Kveberg L, Vaage JT, Inngjerdingen M. The tetraspanin CD53 modulates responses from activating NK cell receptors, promoting LFA-1 activation and dampening NK cell effector functions. *PLoS One*. 2014. 9(5): e97844.
31. Krautwald S. IL-16 activates the SAPK signaling pathway in CD4 + macrophages. *J Immunol*. 1998. 160(12): 5874-9.
32. Mathy NL, Scheuer W, Lanzendörfer M, et al Interleukin-16 stimulates the expression and production of pro-inflammatory cytokines by human monocytes. *Immunology*. 2000. 100(1): 63 – 9.
33. Jones VS, Huang RY, Chen LP, Chen ZS, Fu L, Huang RP. Cytokines in cancer drug resistance: Cues to new therapeutic strategies. *Biochim Biophys Acta*. 2016. 1865(2): 255 – 65.
34. Pérez-Ramírez C, Cañadas-Garre M, Alnatsha A, et al Interleukins as new prognostic genetic biomarkers in non-small cell lung cancer. *Surg Oncol*. 2017. 26(3): 278–285.
35. Nagai Y, Shimazu R, Ogata H, et al Requirement for MD-1 in cell surface expression of RP105/CD180 and B-cell responsiveness to lipopolysaccharide. *Blood*. 2002. 99(5): 1699 – 705.
36. Kikuchi J, Kuroda Y, Koyama D, et al Myeloma Cells Are Activated in Bone Marrow Microenvironment by the CD180/MD-1 Complex, Which Senses Lipopolysaccharide. *Cancer Res*. 2018. 78(7): 1766–1778.
37. Gottfried E, Faust S, Fritsche J, et al Identification of genes expressed in tumor-associated macrophages. *Immunobiology*. 2003. 207(5): 351-9.
38. Palanca-Wessels MC, Czuczman M, Salles G, et al Safety and activity of the anti-CD79B antibody-drug conjugate polatuzumab vedotin in relapsed or refractory B-cell non-Hodgkin lymphoma and chronic lymphocytic leukaemia: a phase 1 study. *Lancet Oncol*. 2015. 16(6): 704 – 15.
39. Töpfer K, Cartellieri M, Michen S, et al DAP12-based activating chimeric antigen receptor for NK cell tumor immunotherapy. *J Immunol*. 2015. 194(7): 3201-12.
40. Takamiya R, Ohtsubo K, Takamatsu S, Taniguchi N, Angata T. The interaction between Siglec-15 and tumor-associated sialyl-Tn antigen enhances TGF- β secretion from monocytes/macrophages through the DAP12-Syk pathway. *Glycobiology*. 2013. 23(2): 178 – 87.
41. Shabo I, Olsson H, Stål O, Svanvik J. Breast cancer expression of DAP12 is associated with skeletal and liver metastases and poor survival. *Clin Breast Cancer*. 2013. 13(5): 371-7.
42. Wang Y, Zheng B, Xu M, et al Prediction and analysis of Hub Genes in Renal Cell Carcinoma based on CFS Gene selection method combined with Adaboost algorithm. *Med Chem*. 2019.
43. Godfrey DI, Uldrich AP, McCluskey J, Rossjohn J, Moody DB. The burgeoning family of unconventional T cells. *Nat Immunol*. 2015. 16(11): 1114-23.
44. Consonni M, Dellabona P, Casorati G. Potential advantages of CD1-restricted T cell immunotherapy in cancer. *Mol Immunol*. 2018. 103: 200–208.
45. Li Y, Zhao C, Liu J, et al CD1d highly expressed on DCs reduces lung tumor burden by enhancing antitumor immunity. *Oncol Rep*. 2019. 41(5): 2679–2688.

Figures

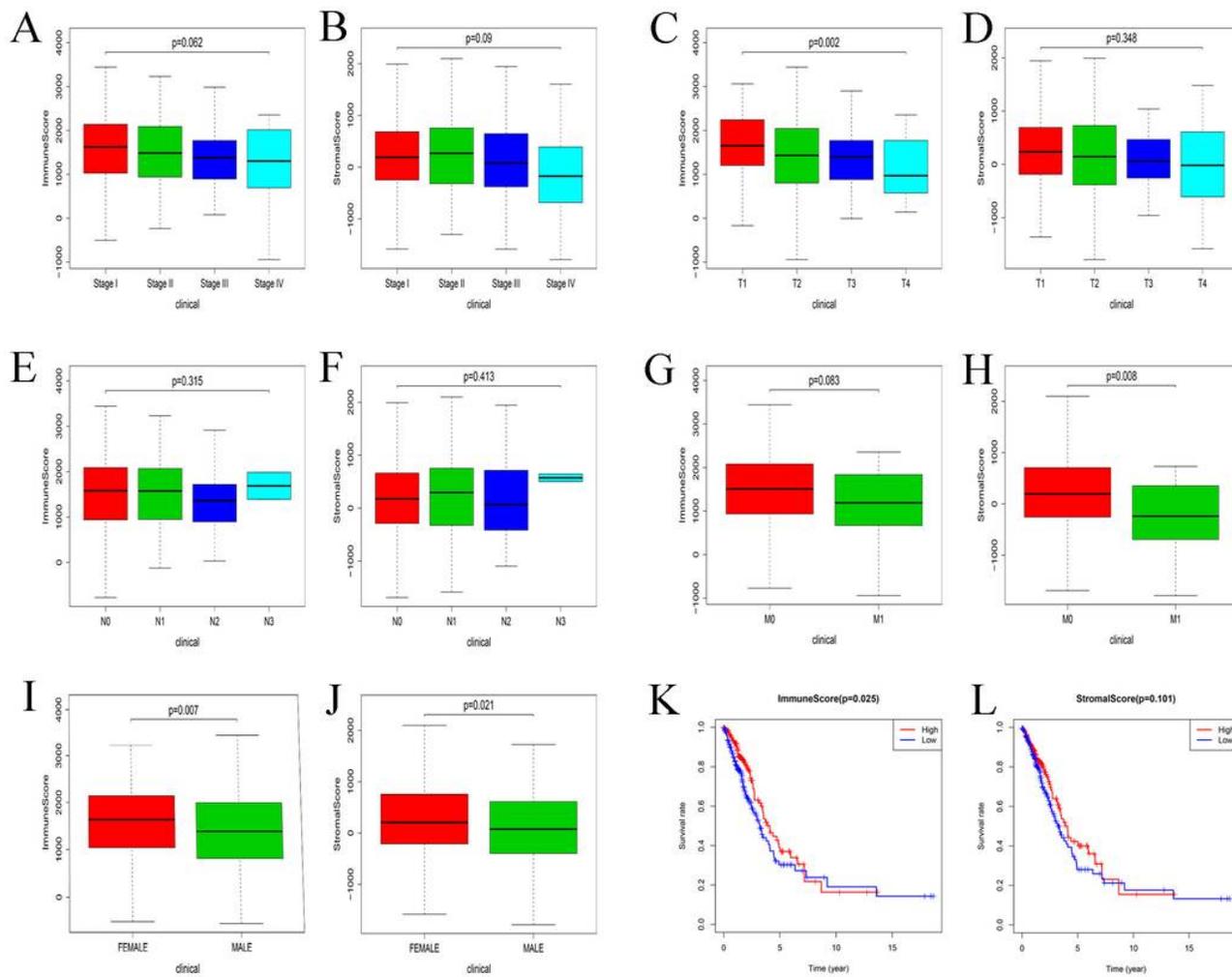


Figure 1

Immune scores and stromal scores are associated with LUAD clinical features and overall survival. (A, B) Distribution of immune scores/stromal scores for LUAD stage. (C, D) Distribution of immune scores/stromal scores for LUAD T stage. (E, F) Distribution of immune scores/stromal scores for LUAD N stage. (G, H) Distribution of immune scores/stromal scores for LUAD M stage. (I, J) Distribution of immune scores/stromal scores for LUAD gender. (K, L) Kaplan-Meier survival curve revealed that higher immune scores are associated with significantly longer overall survival (log-rank test, $p = 0.025$). The high stromal score group showed a longer median overall survival than the low stromal score group, with no significant difference (log-rank test, $p = 0.101$)

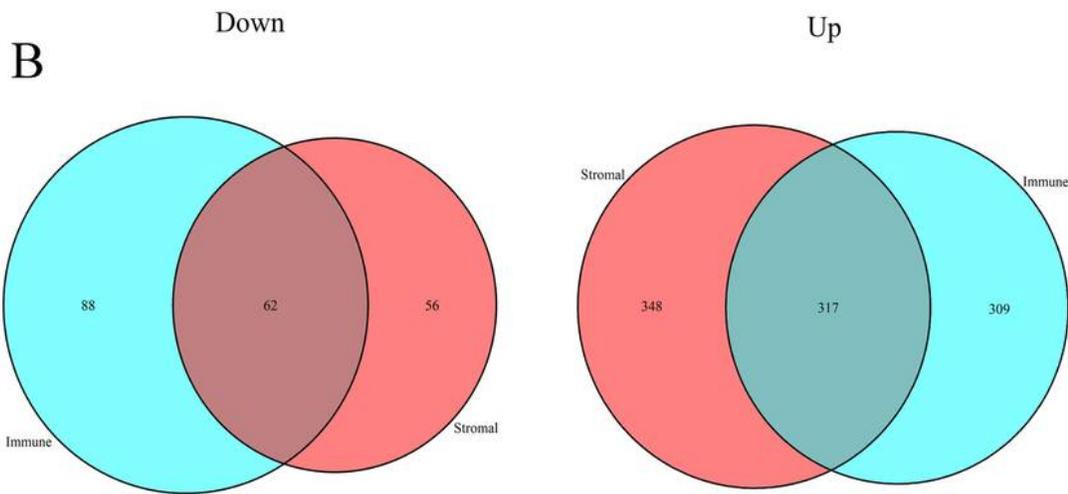
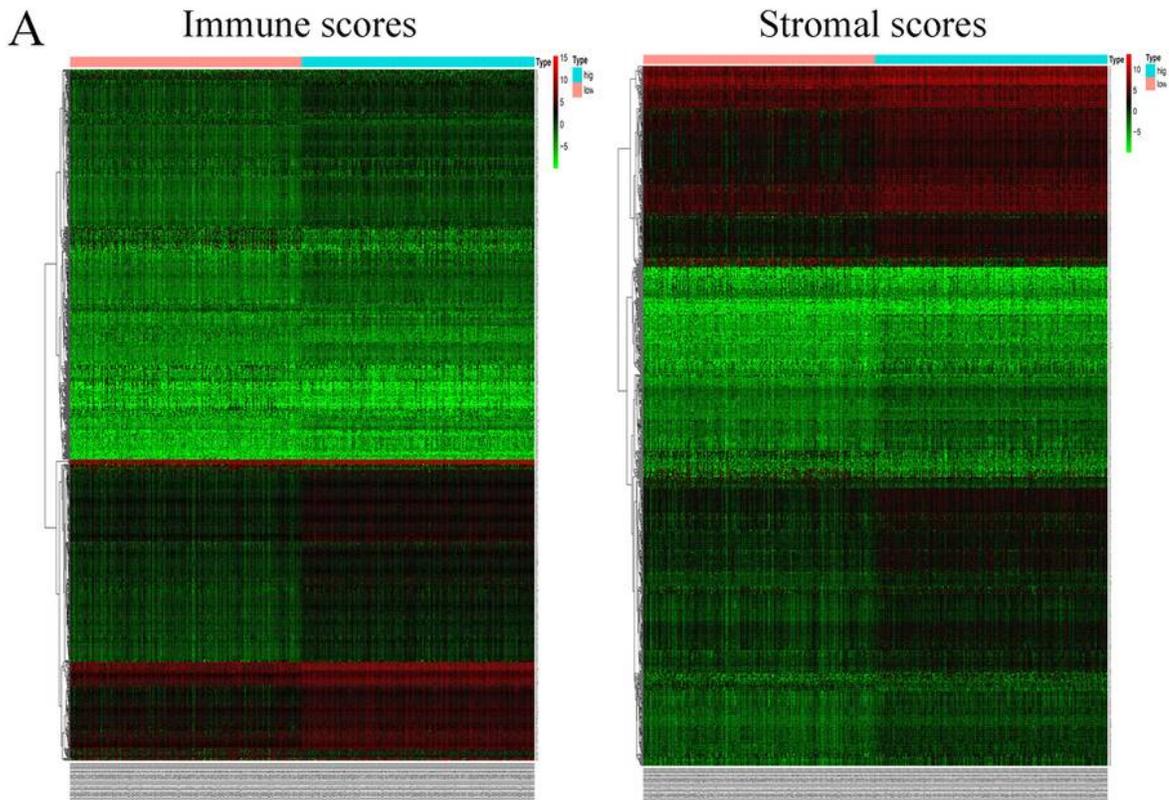


Figure 2

Identification of DEGs based on immune scores and stromal scores. A. Heatmap of DEGs for the immune and stromal score groups. B. Venn diagrams showing the number of commonly changed DEGs in the immune scores and stromal scores (62 downregulated and 317 upregulated genes)

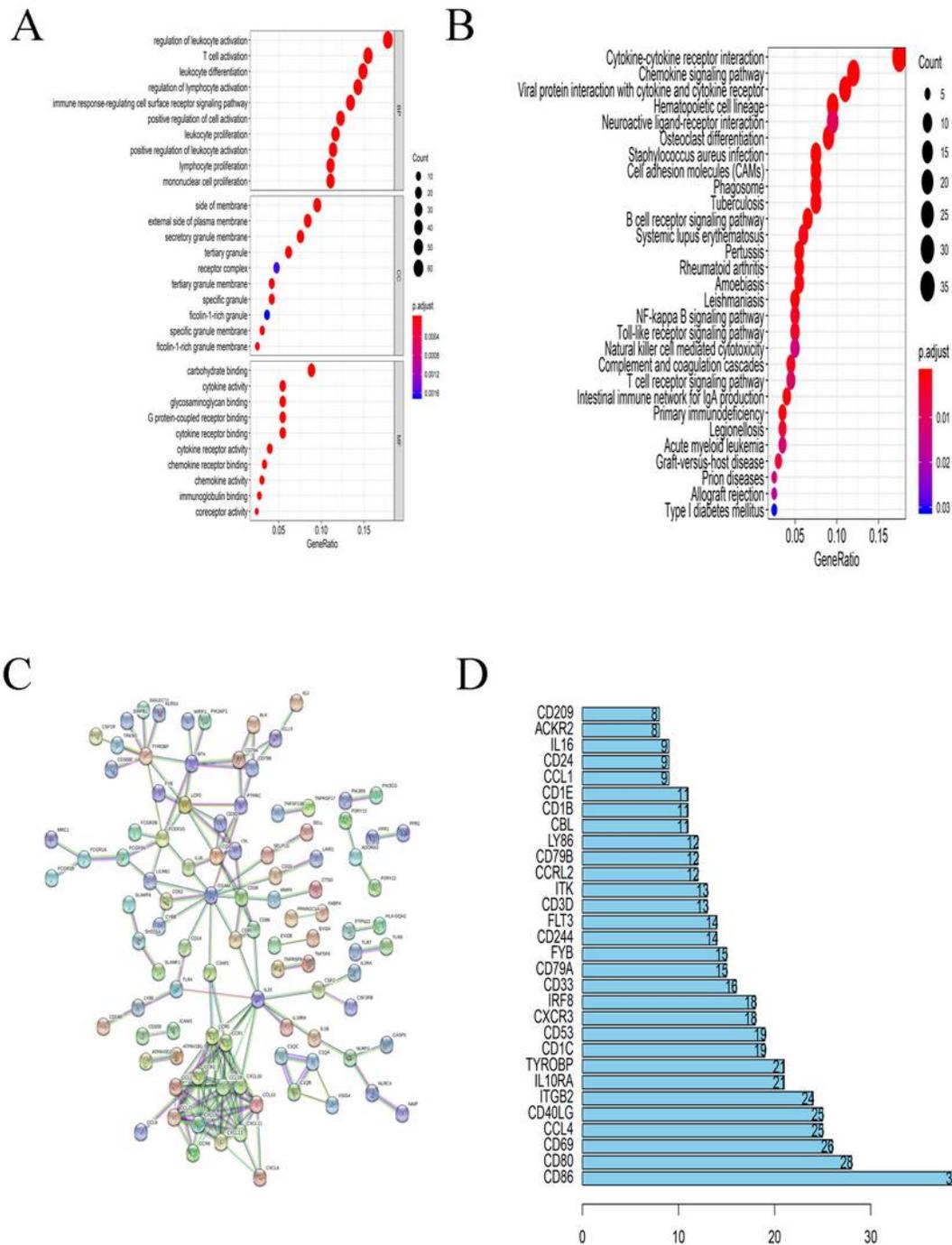


Figure 3

Prognostic markers related to the immune microenvironment of LUAD. (A). Gene ontology (GO) term enrichment analysis of common DEGs, including three subontologies, biological process, cellular component, and molecular function. (B). KEGG enrichment analysis of the 379 genes. (C). Protein interaction networks of the 379 genes. (D) The degree distribution of nodes in the network.

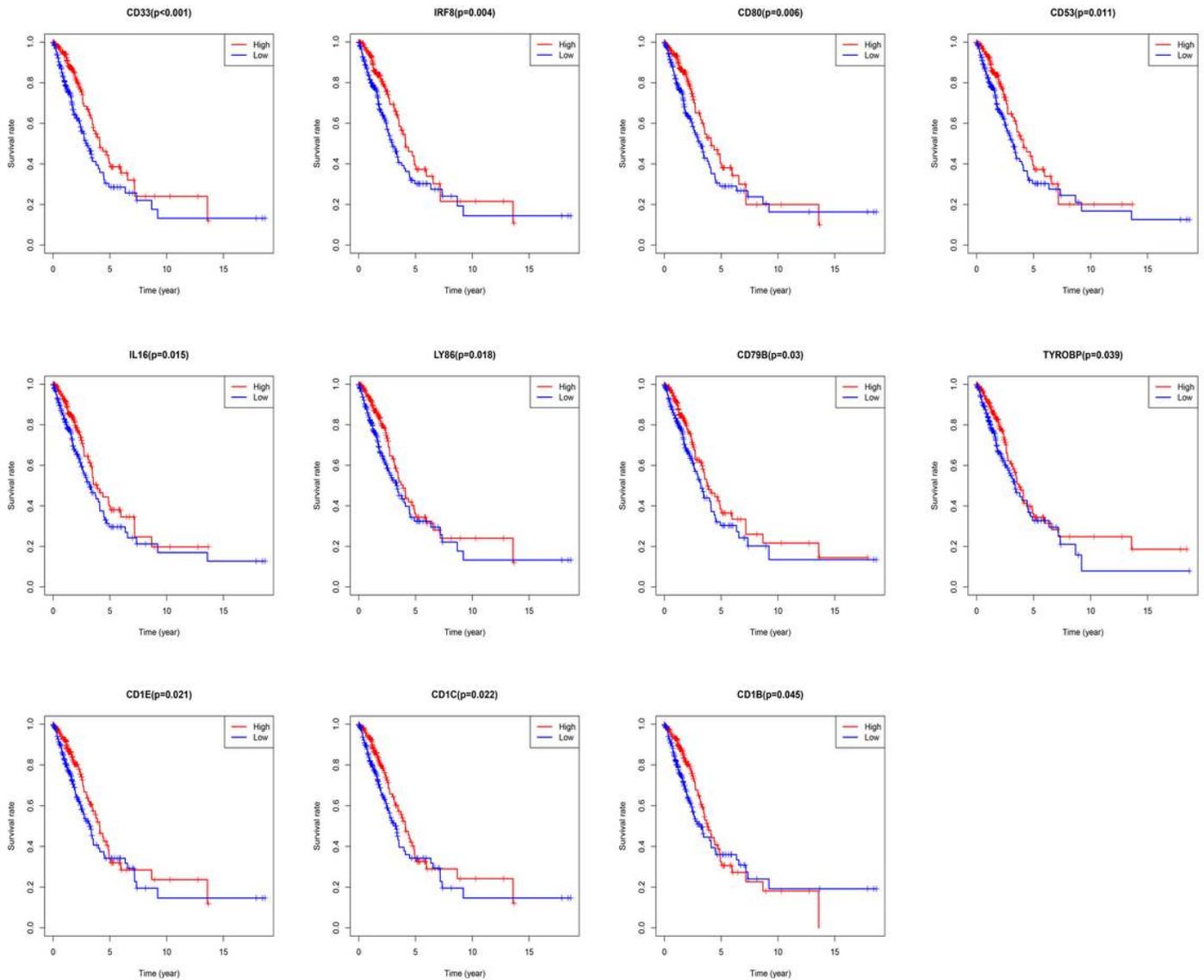


Figure 4

Relationship between microenvironment-related genes and LUAD prognosis in TCGA. Kaplan-Meier survival curves with the log-rank test were performed for individual DEGs.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [renamed77dae.pdf](#)
- [scores.xls](#)