Anoikis-related gene signature is associated with immune infiltration and predicts the prognosis of non-small cell lung cancer

Yixuan Wu[1,2,*], Zhou Zhou[1,2,*], Qianyi Qi[1,2,*], Shirong Xu[3], Lin Chen[4], Feng Wang[1,2]

[1]Research Center of Clinical Medicine, Affiliated Hospital of Nantong University, Medical School of Nantong University, Nantong 226001, China
[2]Department of Laboratory Medicine, Affiliated Hospital of Nantong University, Medical School of Nantong University, Nantong 226001, China
[3]Department of Laboratory Medicine, Taizhou Second People's Hospital, Taizhou 225511, China
[4]Nantong Institute of Liver Diseases, Nantong Third People's Hospital Affiliated Nantong Hospital 3 of Nantong University, Nantong 226006, China
*Equal contribution

Correspondence to: Lin Chen, Feng Wang; email: orange303@ntu.edu.cn, wangfengfy@ntu.edu.cn
Keywords: NSCLC, anoikis, gene signature, HMGA1, immunotherapy
Received: September 28, 2023   Accepted: December 26, 2023   Published: February 7, 2024

ABSTRACT

Non-small cell lung cancer (NSCLC) is the most common histological type of lung cancer. With the in-depth exploration of cell death manners, numerous studies found that anoikis is an important mechanism that associated with treatment. Therefore, we aimed to explore the prognostic value and treatment guidance of anoikis in NSCLC patients. In the current study, we first constructed a prognostic model based on the anoikis-related genes based on bulk RNA-sequencing and single-cell RNA-sequencing (scRNA-seq) dataset. Then, immuno-correlations of anoikis-related risk scores (ARGRS) were analyzed. In addition, HMGA1, a risky gene in ARGRS, was further explored to define its expression and immuno-correlation. Results showed that patients with higher ARGRS had worse clinical outcomes. Moreover, the five genes in the prognostic model were all highly expressed on tumor cells. Moreover, further analysis found that the ARGRS was negatively correlated with ImmuneScore, but positively with tumor purity. Besides, patients in the ARGRS-high group had lower levels of immunological characteristics, such as the immune-related signaling pathways and subpopulations. Additionally, in the immunotherapy cohorts, patients with the ARGRS-high phenotype were more resistant to immunotherapy and tended to not achieve remission after treatment. Last, HMGA1 was chosen as the representative biomarker, and analysis of the in-house cohort showed that HMGA1 was highly expressed in tumor tissues and correlated with decreased T cell infiltration. To sum up, ARGRS was correlated with a desert tumor microenvironment and identified immune-cold tumors, which can be a novel biomarker for the recognition of immunological characteristics and an immunotherapeutic response in NSCLC.

INTRODUCTION

With over 1,600,000 newly diagnosed patients each year, lung cancer is an extraordinarily heterogeneous illness and the acknowledged leading cause of the majority of cancer-related mortalities globally [1]. Non-small cell lung cancer (NSCLC), containing lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) [2], is the most common histological type of lung cancer, accounting for around 85% of all occurrences [3]. Despite tremendous advances in clinical screening and therapeutic therapies for NSCLC, the limited therapeutic benefit of first-line treatment resulted in a low overall cure and survival rate for NSCLC, particularly in metastatic disease. Despite tremendous advances in clinical screening and

therapeutic therapies for NSCLC, the limited therapeutic benefit of first-line treatment resulted in a low overall cure and survival rate for NSCLC, particularly in metastatic disease [4]. As a result, additional research is needed to uncover more relevant biomarkers in order to extend clinical advantages to a larger patient population and enhance NSCLC outcomes.

Immunotherapy, which strengthens the patient's immune system to fight malignancies, has received massive attention in recent years in the context of cancer treatment. Immunosurveillance, or the ability of immune cells in the tumor microenvironment (TME) to recognize and destroy cancer cells under normal conditions, is now widely acknowledged [5, 6]. Nevertheless, further investigation showed that cancer cells can control the host immune system to avoid immune monitoring by enlisting immunosuppressive cell populations and reducing the immunogenicity of tumors [7, 8]. In addition, tumors with different phenotype have distinct therapeutic responses. To be specific, hot tumors, featured by T-cell inflammation, showed a favorite therapeutic response to immunotherapy, while cold tumors are resistant to many treatments [9–12]. Thus, it is crucial to investigate the alteration of the TME to guide the personalized immunotherapy.

In the recent years, with the intensive investigation of different manners of cell death, numerous researchers found that anoikis is an important mechanism that can be introduced to treatment. Anoikis is a type of programmed cell death that takes place when cells separate from the proper extracellular matrix. This mechanism is essential for maintaining plastic cell development and attachment [13]. Notably, cancer cells are resistant to anoikis because they do not depend on extracellular matrix adherence for survival and growth [14], indicating that malignancies are a better example of anoikis resistance. Therefore, understanding the NSCLC anoikis regulators helps researchers find new treatments, particularly for cancer metastasis [15, 16]. For instance, through altering the JAK2/STAT3 and SHP2/Grb2 signaling pathways, the TGF-1/SH2B3 axis can control lung cancer cells' anoikis resistance and EMT [17]. In addition, the enhancement of anoikis sensitivity could enhance immune cell-mediated cytotoxicity [18]. However, the association between anoikis and TME features in NSCLC is still unclear.

At present, several studies explored the correlation between the anoikis feature and immunological characteristics. Here, in this study, we first recognized the up-regulated anoikis-related genes (ARG) in NSCLC tumor tissues, and then construct a prognostic model based on these genes. Subsequently, we confirmed that the ARGs in the prognostic model were expressed on tumor cells at the single-cell level. Finally, further analysis was performed to explore the correlation between ARGs and immunological characteristics and the predictive value of ARGs model in immunotherapeutic response. Taken together, our study provided a new perspective to understand the clinical and immunological-related functions of ARGs, which contributes to the advancement of more precise and precise treatment strategies.

MATERIALS AND METHODS

Data collection and processing

The gene expression matrices of NSCLC patients were downloaded from public online databases —the UCSC Xena website and the Gene Expression Omnibus (GEO) portal. The transcriptional omics and clinical annotations of tumors and paracancerous of NSCLC patients in the TCGA-LUAD and TCGA-LUSC cohorts were obtained from the UCSC Xena. In the GEO database, we identified three NSCLC cohorts (GSE30219 [19], GSE37745 [20], and GSE3141 [21]) with prognosis information. Besides, two clinical cohorts (GSE126044 [22] and GSE135222 [23]) of NSCLC patients received immunotherapy, were also obtained from the GEO database. Diagnostic patients with follow-up information, including survival out-comes or therapeutic responses were chosen for further analysis.

Establishment of the anoikis-related gene model

To establish the anoikis-related gene (ARGs) model, we collected the ARGs from the the genecards website (https://www.genecards.org/, accessed on 12 October 2022) [24] and the Harmonizome portals (https://maayanlab.cloud/Harmonizome/, accessed on 12 October 2022) [25] firstly. Then, after performing differential expressions analysis by "limma" package, ARGs up-regulated in tumor samples (fold-change (FC) $\geq$ 1.5 and adjusted P-values < 0.05) were selected. The univariable COX regression analysis was used to identify genes that were significantly (P-value < 0.05) linked to OS. Next, the least absolute shrinkage and selection operator (LASSO) regression algorithm was performed on these OS-related genes to further screen prognostic parameters and construct the model. The risk score of the prognostic model based on NSCLC-related ARGs (ARGRS) of patients was assess according to the linear combination of the expression values of NSCLC-related ARGs multiplied by the relevant LASSO coefficients. To validate the risk score's predictive power, the 50% ARGRS cutoff was used to divide the NSCLC patients into high- and ARGRS-low groups.

Assessment of biological functions

The R package "clusterProfiler" [26] was used to assess the biological functions of gene signatures in terms of Gene Ontology (GO) [27] and Kyoto Encyclopedia of Genes and Genomes (KEGG) [28] pathways. The top ten enriched pathways with the most significantly P-values were displayed.

Single-cell RNA sequencing datasets analysis

To further elaborate the tumor-specific of gene signatures in the prognostic model, single-cell RNA sequencing (scRNA-seq) datasets (GSE150660 [29] and GSE148071 [30]) were downloaded from the GEO protocol (All additional analyses were using the Seurat R toolkit [31].

For each sample, we summarized the expressed percentage of mitochondrial genes (percent.mt). Cells with percent.mt < 10% and $200 < nFeature\_RNA < 5,000$ were preserved. Then, for each dataset, the "RunHarmony" function [32] was applied to minimize the batch effects and integrate the transcriptional gene profiles from different patients. Principal component analysis (PCA) was performed based on the top 4,000 genes with highest variability. Then, the high dimensionality of data was reduced based on the top 30 PCs. The cells were unsupervised clustered via shared nearest neighbor (SNN) algorithm with one resolution. To annotate cell types, many well-known signatures were collected from previous studies [33, 34], such as EPCAM for tumor cells, CD3E for T cells, CD14 for macrophages, and CD1C for dendritic cells.

Cell-cell communication analysis

"CellPhoneDB" software [35] was utilized to deconstruct the cell-cell communications among cell types at the single cell level. To define the ARGRS+ and ARGRS- tumor cells, "AddModuleScore" function was applied to estimate the enrichment scores of the ARGRS based on the transcriptional level of the five genes in the model. Then, tumor cells with ARGRS > 0 were defined as ARGRS+ tumor cells, and other were ARGRS- tumor cells. The ligand-receptor pairs with P-value < 0.05 were included in further analysis.

Immune infiltration analysis

The gene signatures of immunomodulators, immune cell subpopulations and immunological signaling pathways were obtained from previous studies [36–38]. R package "estimate" was employed to examine the abundance of infiltrating immune cells (ImmuneScore) and tumor purity (TumorPurity). The R package "GSVA" (version 1.46.0)

[39] was utilized to assess the enrichment scores of characteristics based on the gene signatures.

Clinical samples, immunohistochemistry, and quantitative evaluation

The NSCLC tissue microarray (TMA) section (HLugC120PT01) was obtained from Outdo Biotech (Shanghai, China). Ethical approval (YB-M-05-02) for the study of the TMA was granted by the Clinical Research Ethics Committee, Outdo Biotech (Shanghai, China). The TMA was used for immunohistochemistry (IHC) assay to measure the expression of HMGA1 protein in tumor and non-tumor breast tissues. IHC assay was performed on these sections according to the established steps. These sections were given three 5-minute xylene washes and rehydrated using a series of washes in ethanol grades of 70%, 90%, and 100%. For twenty minutes, endogenous peroxidase activity was inhibited using hydrogen peroxidase. The antigen retrieval solution is EDTA. The primary antibody utilized in the study was anti-HMGA1 (1:200 dilution, Cat. sc-393213, Santa Cruz) and anti-CD8A (ready-to-use, Cat. PA577, Abcarta). Antibody staining was visualized with DAB and hematoxylin counterstain. Using the immunoreactivity score standard, two senior pathologists assessed stained TMA to determine HMGA1 expression [40]. Two senior pathologists used The Cancer Genome Atlas Network's criteria to determine the CD8 score for tumor-infiltrating CD8$^+$ T cell assessment [41]. For every sample, a CD8 score, which is the total of the density and distribution scores (0–6), was determined. Samples classified as immune-cold are those with a CD8 score ≤ 2 (0, 2) and immune-hot samples with a score ≥ 3 (3, 4, 5, 6).

Statistical analysis

R-4.2.2 was used to perform all statistical analyses. The chi-square test was used to compare categorical variables, while the Wilcoxon rank-sum test was for continuous variables between the two groups. The log-rank test was used to assess the prognostic values. A two-paired p-value of less than 0.05 was considered statistically significant for all analyses, and the results were categorized as follows: *p-value < 0.05, **p-value $\leq$ 0.01, ***p-value $\leq$ 0.001, and ****p-value $\leq$ 0.0001.

Availability of data and materials

The TCGA data are openly available at https://portal.gdc.cancer.gov/, and the GEO data are openly available at https://www.ncbi.nlm.nih.gov/gds.

Consent for publication

All authors are consent for publication.

RESULTS

Identification of differentially expressed overlapping anoikis-related genes

From the genecards websiteand the Harmonizome portals, a total of 244 anoikis-related genes (ARGs) were obtained. By performing PCA on the transcriptional matrix of anoikis-related genes, the normal samples and NSCLC patients clearly differed (Figure 1A). Therefore, a differential expression analysis was carried out to better identify NSCLC patients who had anoikis (Figure 1B). 33 anoikis-related genes were shown to be up-regulated in NSCLC patients, according to the results (Figure 1C–1E). Additionally, given the correlation between genomic variants and tumorigenesis, we summarized the mutation rate and copy number variation (CNV) of these genes. Results showed most of genes carried genomic mutations or amplifications in NSCLC patients (Figure 1F–1G), indicating the potential malignant functions of these in tumorigenesis.

Establishment of a prognostic model of anoikis-related genes in NSCLC

By performing univariable COX regression analysis and LASSO analysis, five independent prognostic genes (HMGA1, PLK1, ETV4, PHLDA2, and ITGB4) were screened for establishing the prognostic model (Figure 2A–2B and Supplementary Figure 1). The risk scores based on the five anoikis-related genes (ARGRS) were calculated according to the combination of the expression levels of these genes multiplied by the corresponding coefficients. As shown in Figure 2C, the mortalities were centralized in the ARGRS-high group, while the living patients were enriched in the ARGRS-low group. Besides, in patients with higher ARGRS, signatures (HMGA1, PLK1, PHLDA2, and ITGB4) associated with poor OS were significantly highly expressed, while ETV4, correlated with favorable clinical outcomes showed the opposite (Figure 2C). Consistently, compared with the ARGRS-low group, patients with high ARGRS showed a worse prognosis (Figure 2D). In addition, further analysis showed that patients with higher pathological stages or grades had higher levels of signatures in the prognostic model and ARGRS (Figure 2E, 2F and Supplementary Figure 2), which further supported the finding that ARGRS was associated with worse clinical outcomes in NSCLC.

In addition, due to the consequence of genomic alterations in diagnosis and therapeutic guidelines, we also compared the mutant frequencies of genes between the high and low ARGRS groups. Compared with the

ARGRS-low group, the ARGRS-high group had more complex genomic mutations (Supplementary Figure 3). Some well-known mutations associated with tumorigenesis and progression, such as TP53 [42, 43] and PTEN [44], were enriched in the ARGRS-high group, further supporting the association between ARGRS and worse clinical outcomes in NSCLC.

Validation of the prognostic model in NSCLC

We further confirmed this finding in additional independent cohorts after discovering the prognostic predictive utility of ARGRS in the TCGA cohort. Patients in the GSE30219 cohort who had higher ARGRS had worse OS than those with lower ARGRS, which is consistent with the findings from the TCGA cohort (Figure 3A, 3B). ARGRS levels were also higher in patients with higher clinical stages and grades (Figure 3C), indicating its role in the malignancy of NSCLC. Patients in the ARGRS-high group similarly exhibited inferior clinical outcomes in the GSE37745 and the GSE3141 cohorts (Figure 3D, 3E). Based on the findings from the TCGA and additional independent cohorts, the ARGRS was linked to a worse OS for NSCLC patients and may have contributed to tumor progression.

The five anoikis-related genes in the model expressed specifically in tumor cells at the single-cell level

To further clarify the tumor-specificity of signatures in the model, scRNA-seq datasets were include. A scRNA-seq dataset contained five NSCLC patients were collected and integrated firstly (Supplementary Figure 4A, 4B). After quality control, integration and unsupervised clustering, the cells were divided into 19 clusters (Supplementary Figure 4B). Then, based on the expression distribution of canonical signatures of different cell types, the cells were identified as tumor cells, T cells, macrophages and cDC (Figure 4A–4C). Then, we explored the distribution and expression levels of five crucial genes. Results showed that compared with non-tumor cells, tumor cells had higher levels of ARGRS (Figure 4D, 4E). Besides, the signatures in the model were specific highly expressed in the tumor cells (Figure 4F), suggesting the tumor-specificity of the five genes, and implying that the risk score can represent the status of tumor cells.

Additionally, given of the heterogeneity among NSCLC patients, another scRNA-seq dataset was used to further confirm the results found in the GSE150660 dataset. After preprocessing, a total of 9,563 individual cells of the GSE148071 cohort were passed quality control. Subsequently, unsupervised clustering and cell
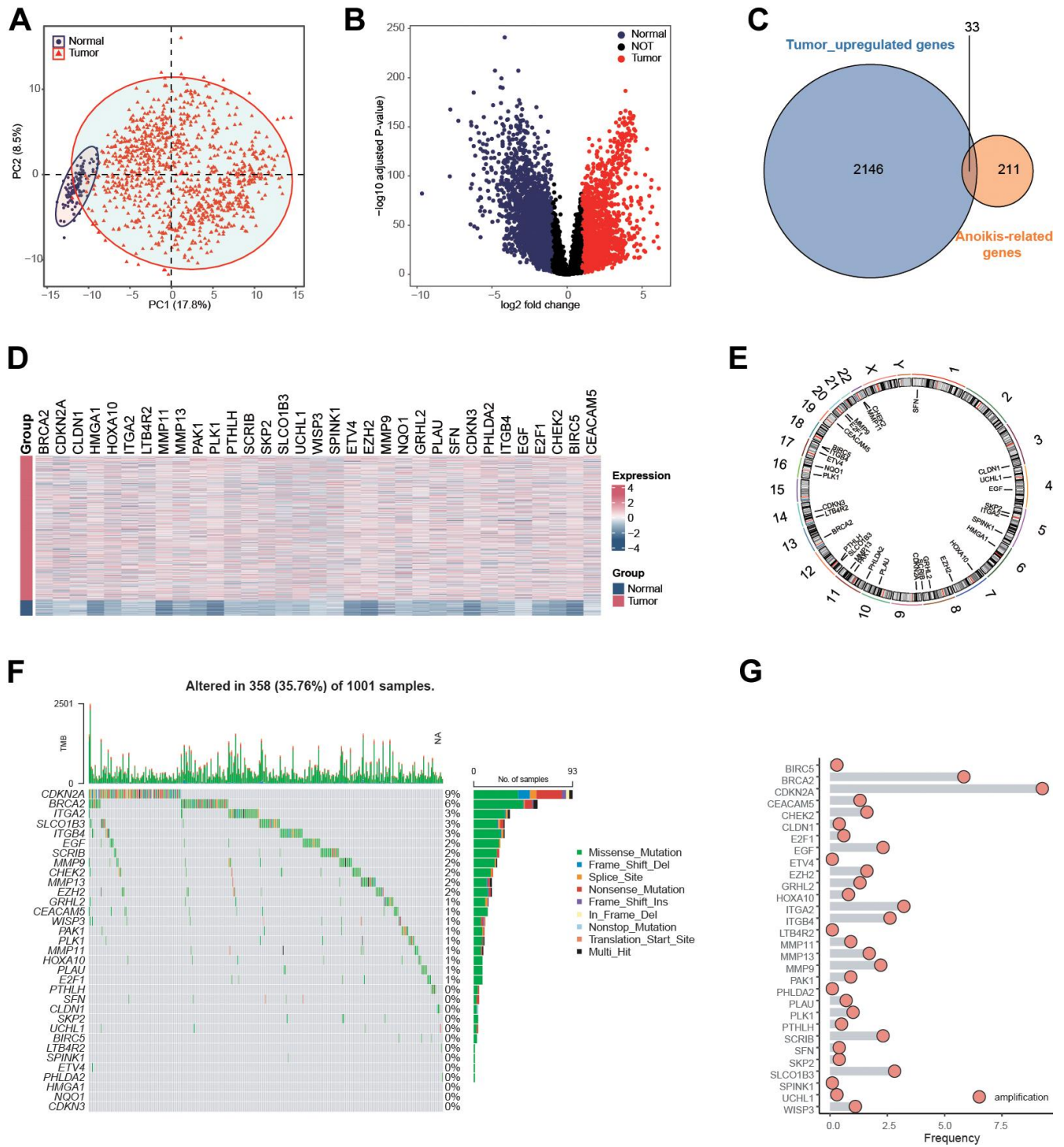
Figure 1. Identification of anoikis-related genes in NSCLC patients in the TCGA NSCLC cohort. (**A**) Principal component analysis of TCGA samples based on the expression matrix of anoikis-related genes obtained from the genecards website and the Harmonizome portals. (**B**) Volcano plot showing the differentially expressed genes (DEGs) for the NSCLC tissues and paracancerous in the TCGA-NSCLC cohort. (**C**) 33 anoikis-related genes were up-regulated in NSCLC patients in the TCGA cohort. (**D**) Heatmap showing the expression values of anoikis-related between the NSCLC tissues and paracancerous in the TCGA-NSCLC cohort. (**E**) The specific location of anoikis-related genes on the human chromosomes. (**F**) Mutation frequency of 33 overlapping signature genes in the TCGA NSCLC cohort. Each column represented a single patient. The upper bar plot showed TMB. The number on the right indicates the mutation frequency of each regulatory gene. The right bar plot showed the proportion of each variant type. The stacked bar plot below showed a fraction of conversions in each sample. (**G**) CNV frequency of 33 overlapping genes in the TCGA NSCLC cohort. The height of the column represented the alteration frequency. Blue dot: the deletion frequency; Red dot: the amplification frequency.
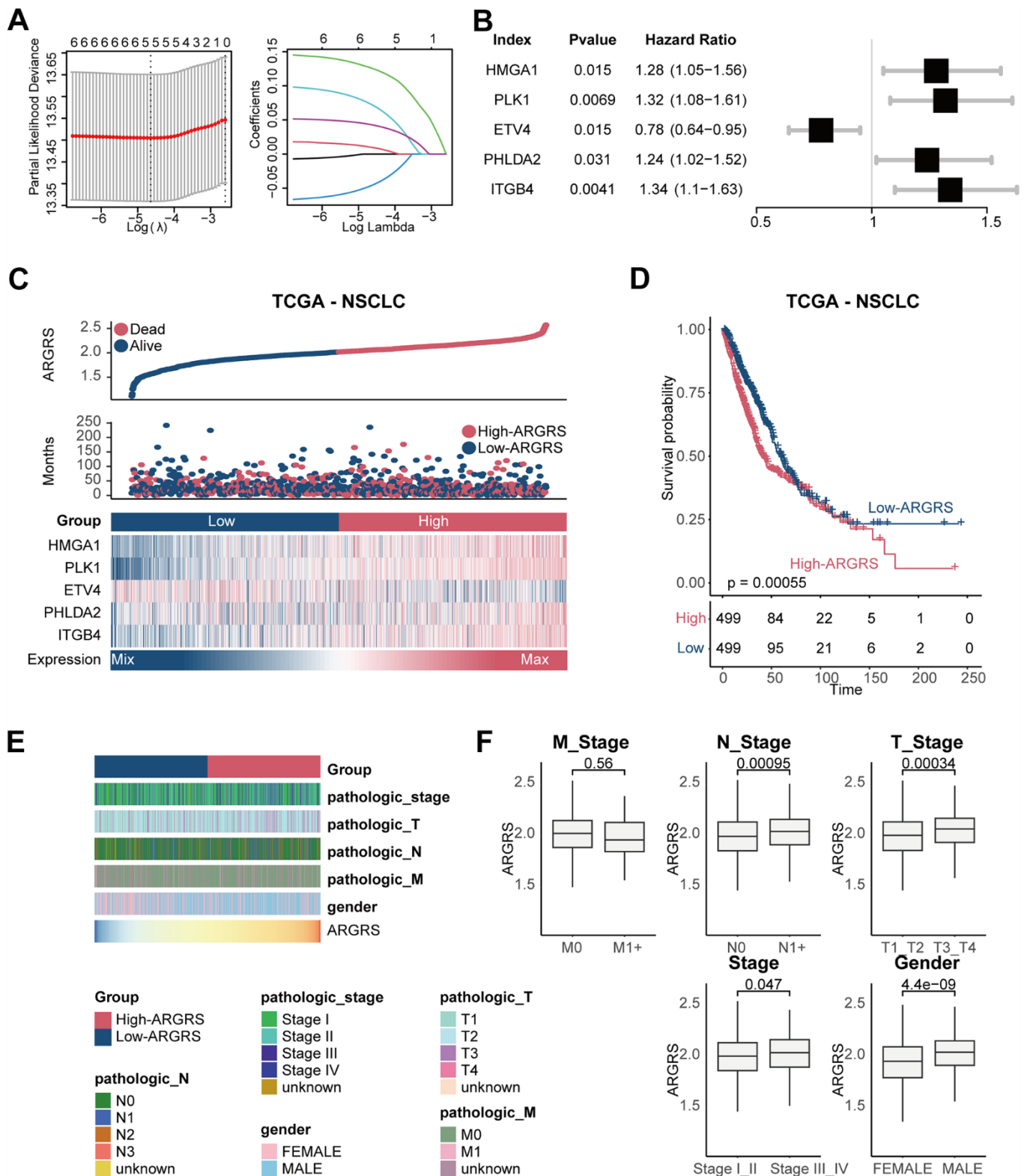
Figure 2. Construction of prognostic model in the TCGA NSCLC cohort. (**A**) The LASSO coefficient profiles were constructed using the 33 anoikis-related genes, and the tuning parameter (λ) was calculated based on the minimum criteria for OS with ten-fold cross validation. Five genes were selected according to the best fit profile. (**B**) Univariable analyses of the expression values of the five genes with overall survival in the TCGA NSCLC cohort. (**C**) Distributions of ARGRS, survival status of NSCLC patients, and expression profiles of the gene signatures. (**D**) Survival analysis showing the prognostic value of ARGRS in the TCGA NSCLC cohort. (**E**) Heatmap showing the distribution of clinicopathologic features between ARGRS-high and low groups. (**F**) Comparison of ARGRS among different clinicopathologic features.

annotation were performed, and these cells were classified into tumor cells, T/NK, macrophages, and fibroblasts (Supplementary Figure 5A–5D). Consistently, tumor cells had higher levels of ARGRS (Supplementary Figure 5E, 5F) and the genes in the prognostic model (Supplementary Figure 5G), further confirming the tumor cell-specificity of the five genes.

Patients with higher ARGRS exhibit low immune infiltration

Having observed the existence of many immune cells in the NCLSC patients at the high dimensional datasets, we next explored the correlation between ARGRS and the fraction of each cell type. Notably, the ARGRS was negatively correlated with the fraction of T cells at the single-cell level ($R^2 = -0.93$, $p = 0.02$, Figure 5A). Also, in the TCGA cohort, patients with higher ARGRS had lower levels of ImmuneScore but higher levels of tumor purity

(Figure 5B), consistent with the results found in the single-cell datasets. In addition, a functional analysis of up-regulated genes in ARGRS-low group was highly related to immunological processes, such as cytokine activity, chemokine receptor binding and activity, and cytokine-cytokine receptor interaction (Figure 5C). Meanwhile, compared with the ARGRS-high group, the ARGRS-low group expressed higher transcriptional levels of immunological signatures, such as the biomarkers of CD8+T cells and NK (Figure 5D and Supplementary Figure 6). Also, the ARGRS-low group had higher fraction of CD8+T cells, and enriched some immune activation characteristics, such as the molecules of HLA, MHC, and immunostimulatory (Figure 5E, 5F). To be sum up, combined with the results found at the single-cell and omics transcriptional datasets, we found that ARGRS was negatively correlation with the fraction of T cells, and patients with the ARGRS-high phenotype exhibited low immune infiltration.
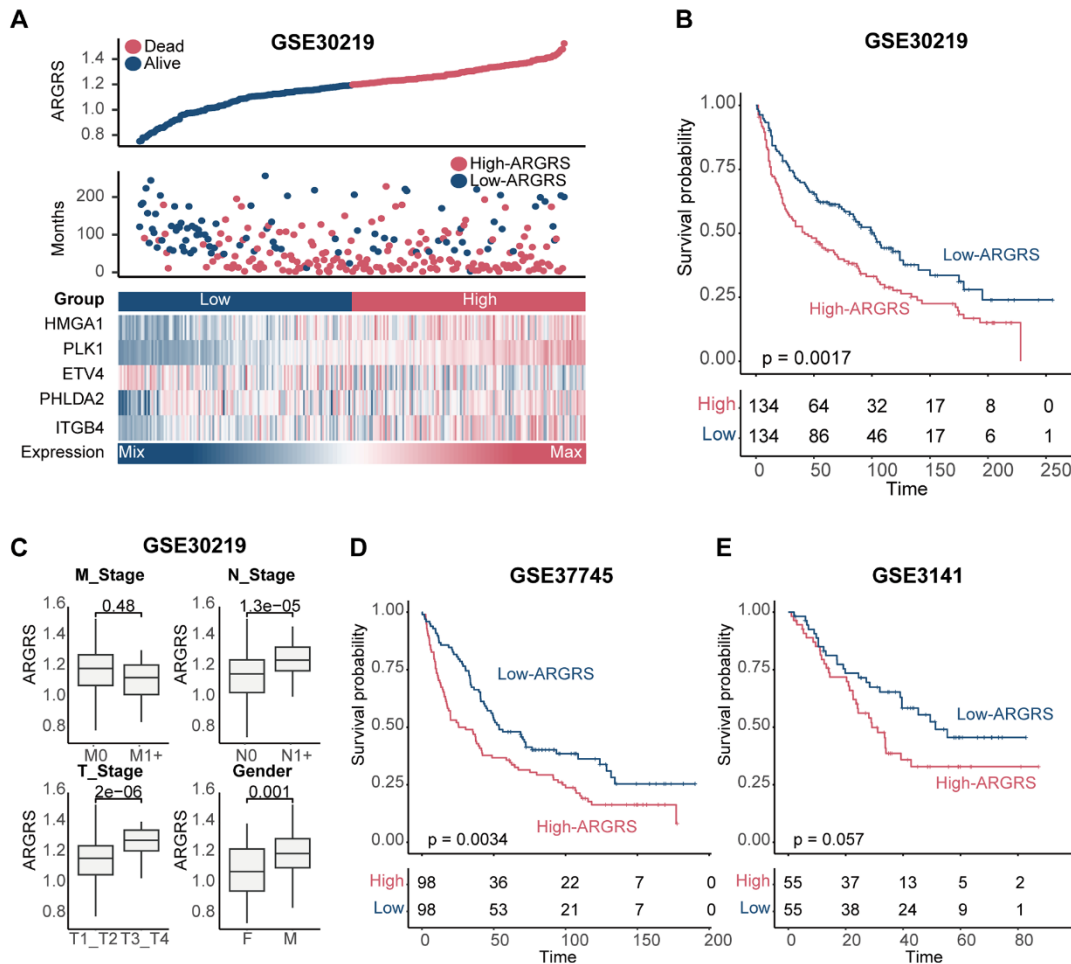

Figure 3. Validation of prognostic model in other independent cohorts. (**A**) Distributions of ARGRS, survival status of NSCLC patients, and expression profiles of the gene signatures in the GSE30219 cohort. (**B**) Survival analysis showing the prognostic value of ARGRS in the GSE30219 cohort. (**C**) Comparison of ARGRS among different clinicopathologic features. (**D**, **E**) Survival analysis showing the prognostic value of ARGRS in the (**D**) GSE37745 and (**E**) GSE3141 cohort.

In another independent cohort, we also found the similar results. ARGRS was negatively correlated with ImmuneScore (R^2 = -0.32, p < 0.0001), but positively with tumor purity (R^2 = 0.36, p < 0.0001, Supplementary Figure 7A–7C). Patients with the ARGRS-low phenotype had higher enrichment scores of immune-related characteristics, such as receptors and MHC molecules (Supplementary Figure 7D). Additionally, patients in the ARGRS- low groups also had higher levels of the immune-related signaling pathways and subpopulations (Supplementary Figure 7E), indicating activate immunological status of patients with
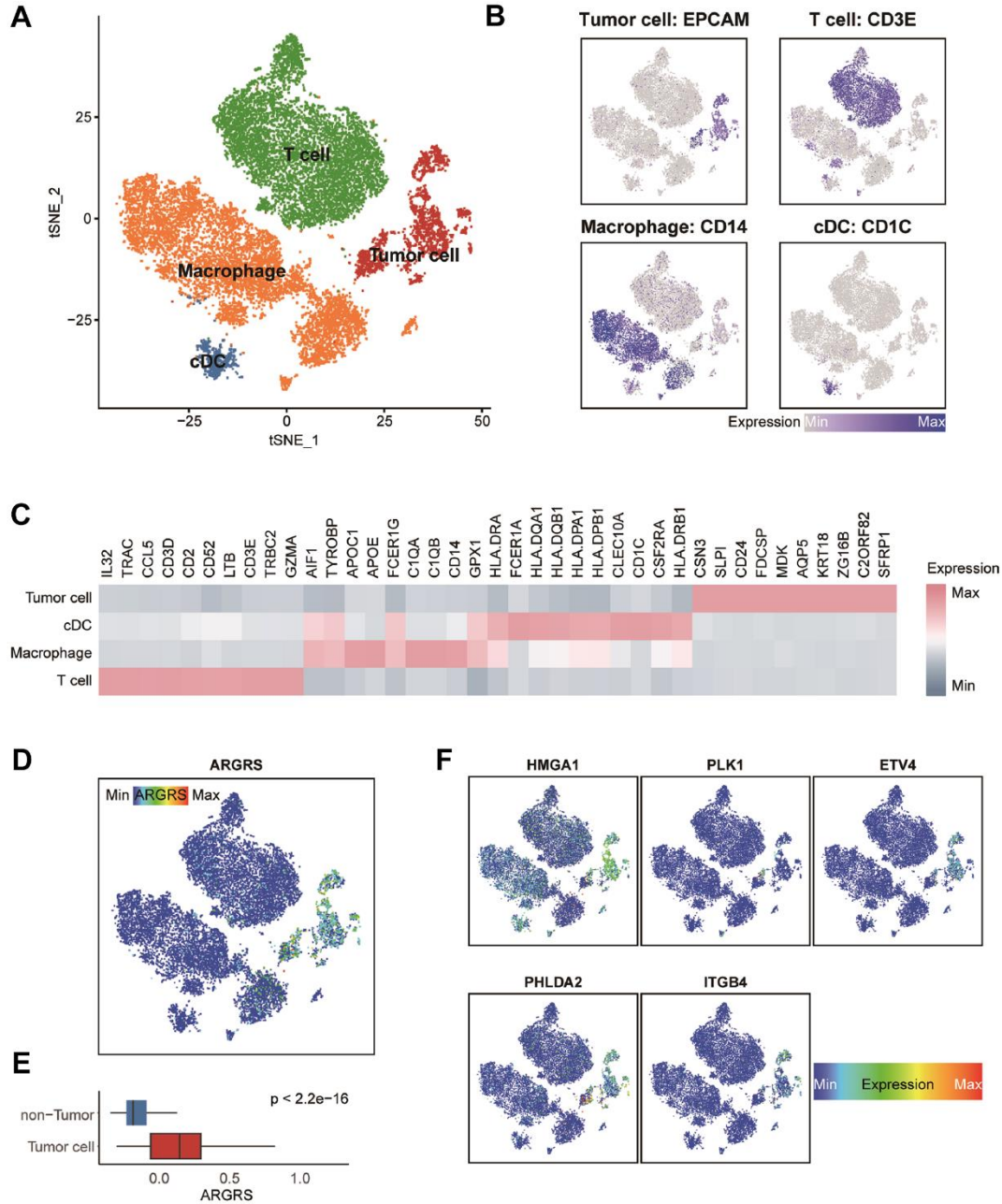

Figure 4. Transcriptomic clustering of NSCLC patients from GSE148071 dataset. (**A**) Marker-based cell type identification analysis allowed the prediction of four broad cell types across all profiled single cells. (**B**) Expression levels of cell type signatures overlaid on the t-SNE representation. EPCAM for tumor cells, CD3E for T cells, CD14 for macrophages and CD1C for cDC. (**C**) Gene expression heatmap of top-10 cell type-specific marker genes as measured by Wilcoxon rank-sum test. (**D**) ARGRS overlaid on the t-SNE representation. (**E**) Boxplot showing the levels of ARGRS between tumor and non-tumor cells. Horizontal lines in the boxplots represent the median, the lower and upper hinges correspond to the first and third quartiles, and the whiskers extend from the hinge up to 1.5 times the interquartile range from the hinge. (**F**) Expression levels of the five genes in the prognostic model overlaid on the t-SNE representation.
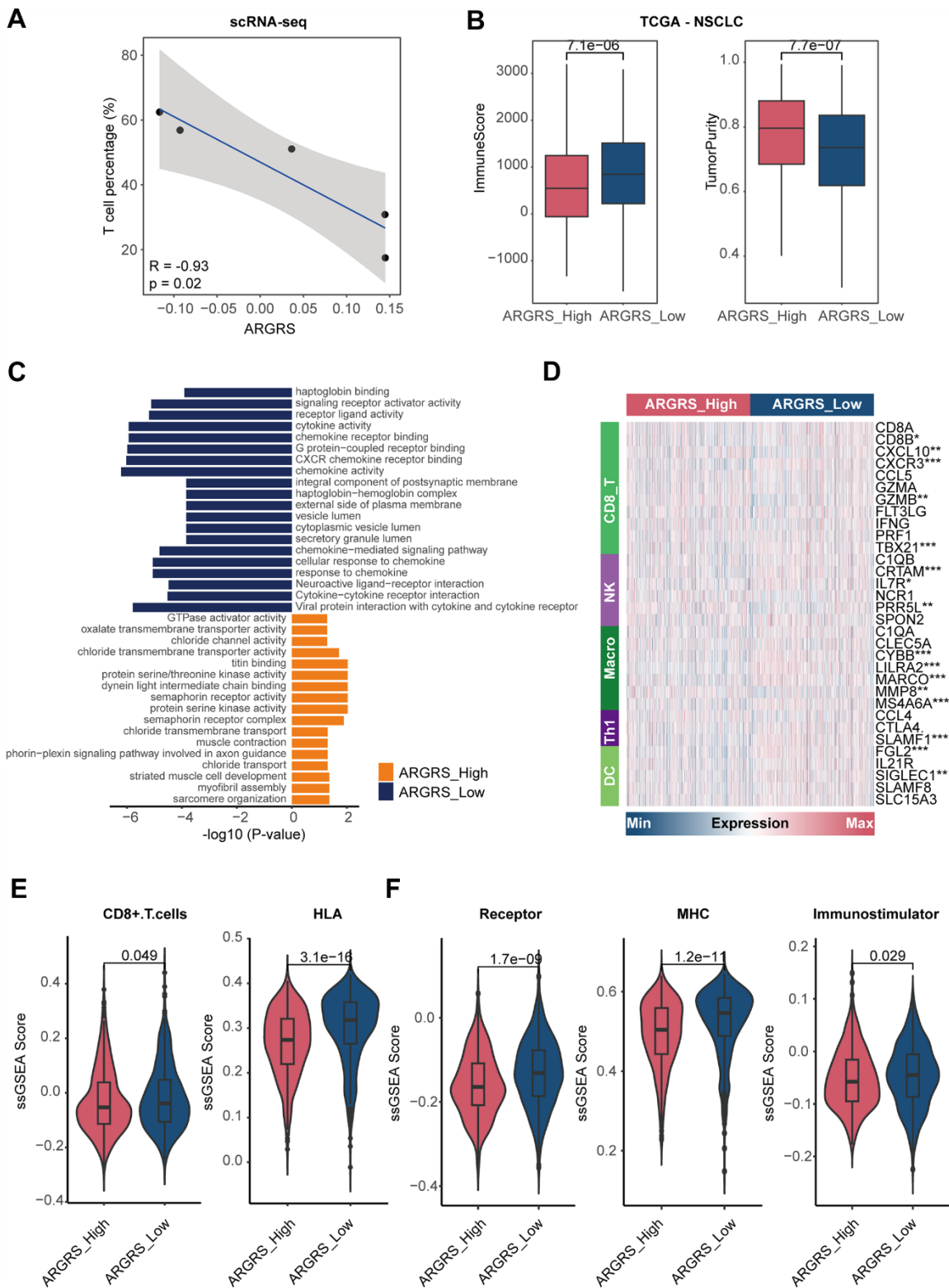
Figure 5. Patients with higher ARGRS exhibit low immune infiltration. (**A**) The correlation between the fraction of T cells and ARGRS levels at the single-cell level. X-axis: the fraction of T cells of each NSCLC patient. Y-axis: the median levels of ARGRS of each NSCLC patient. (**B**) Boxplot showing the levels of ImmuneScore (left) and TumorPurity (right) between the ARGRS-high and low groups in the TCGA NSCLC cohort. (**C**) GO and KEGG analyses of genes up-regulated in the ARGRS-high and low groups in the TCGA NSCLC cohort. (**D**) Heatmap showing the levels of immunological markers between ARGRS-high and low groups in the TCGA NSCLC cohort. (**E**, **F**) Comparison of immunological signatures between ARGRS-high and low groups in the TCGA NSCLC cohort.

the ARGRS- low phenotype. To further investigate the potential factors mediating the formation of inflamed or dessert TME, we deconstructed the interactions among cell populations. Results showed that compared with ARGRS- tumor cells, ARGRS+ tumor cells presented significantly more specific interactions with immune subpopulations (Supplementary Figure 8A). ARGRS+ tumor cells can communicate with T cell via many ligand-receptors (Supplementary Figure 8B), such as TIGIT-NECTIN2 [45] and LGALS9-CD44 [46], which were involved in suppressing tumor T cell infiltration, and promoted the cell state transition of immune cells towards a more immunosuppressive and exhaustive status, indicating the potential mechanism in mediating the dessert TME in the ARGRS-high group.

Patients with higher ARGRS were resistance to immunotherapy

Previous studies reported that the reactivity of immunotherapy was affected by the immunological status of patients. To be specific, hot tumors, featured by T-cell inflammation, were more sensitive to immunotherapies, while cold tumors are resistant to many treatments [9, 10]. Combined with the distinct immunological status between the ARGRS-high and low groups, the two groups may have different response of immunotherapy.

Here, to investigate the immunotherapeutic in the high- and ARGRS-low groups, we downloaded the expression matrix and clinical annotations of NSCLC patients in the cohorts received immunotherapy. In accordance with results above, ARGRS was negatively correlated with almost all immunological characteristics in the GSE126044 dataset (Figure 6A). Also, patients with the ARGRS-high phenotype had higher levels of ImmuneScore, but low tumor purity (Figure 6B, 6C). Besides, patients in the ARGRS-high group were centralized in the NR group (Figure 6D), indicating the immunotherapeutic resistance of them. Consistently, analysis of another immunotherapy cohort also showed the same results. Patients in the ARGRS-high group exhibit lower immune infiltration but higher tumor purity (Figure 6E–6G). Moreover, NSCLC patients in the ARGRS-high group were more likely to recurrence or progress after immunotherapy than the ARGRS-low group (Figure 6H), further supporting the immuno-therapeutic resistance of them.

HMGA1 was the representative biomarker for NSCLC

Next, we chose a representative biomarker from the five genes in the model for further experiments. Firstly, we summarized the expressed fraction of these genes at the single-cell levels. Compared with other

genes in the model, tumor cells had obviously transcriptional levels and the highest expressed fraction of HMGA1 (Figure 7A, 7B, and Supplementary Figure 9A, 9B). Besides, at the high-resolution dataset, compared with other four genes, HMGA1 showed the most significant negative correlation with T cell inflamed and positive with tumor cells (Supplementary Figure 10). A Additionally, HMGA1 expressed was negatively correlated with ImmuneScore, but positively with tumor purity in multiple independent cohorts (Figure 7C–7F), suggesting the guidance of HMGA1 in distinguishing the hot and cold NSCLC tumors.

Next, we verified the expression pattern of HMGA1 in the in-house cohort. We found that HMGA1 expression was significantly higher in tumor tissues than in para-tumor tissues (Figure 8A, 8B), and HMGA1 was specifically expressed in NSCLC (Figure 8C, 8D). We also classified NSCLC into immune-hot and cold tumors based on the CD8 scores. Besides, HMGA1 was significantly reduced in immune-hot tumors (Figure 8E, 8F) and negatively correlated with CD8 scores (Figure 8G). Overall, the negative correlation between HMGA1 expression and T cell infiltration can be verified in the in-house cohort, which greatly improves the confidence of the public cohort results.

DISCUSSION

Nowadays, with the more profound of molecular and microenvironment characteristics of NSCLC, immuno-therapy strategies, such as immune checkpoint inhibitors, has changed first-line treatment of advanced NSCLC, as we all know. Meanwhile, more than 10 indications for immunotherapies in NSCLC clinical practice has been approved by Chinese official organization [47–49].

Even while some patients have benefitted from this new therapy strategy in terms of survival, some individuals are still unable to get lasting relief. Tumor cell apoptosis can result in medication resistance, which in turn affects tumor cell survival in the bloodstream, which is essential for the development of metastasis. Jin et al. found that PLAG1 GDH1 axis promotes apoptotic resistance and tumor metastasis in LKB1-deficient malignancies via CamKK2 AMPK signaling pathway [50]. According to research, CPT1A-mediated fatty acid oxidation can encourage the spread of colorectal cancer cells by impairing the process of cell death [51]. Anoikis has been extensively investigated in relation to tumor growth and metastasis, but little is known about how it alters the tumor immune milieu and hence mediates tumor progression. We concentrate on the part anoikis play in the microenvironment of the tumor and

investigate if anoikis can mediate the regulation of immunization, altering the growth and metastasis of the tumor.

We identified tumor-specific anoikis-related genes in NSCLC patients and used these genes to build a predictive model. Patients with higher ARGRS exhibited worse clinical outcomes, showing that ARGRS plays a role in the malignancy of NSCLC. Furthermore, using single-cell transcriptional datasets, we demonstrated that, when compared to non-tumor cells, tumor cells showed greater levels of ARGRS and signature expression in the prognostic model. Notably, ARGRS was negatively linked with the proportion of T cells at the single-cell level, meaning that ARGRS-high NSCLC patients were more likely to have the TME
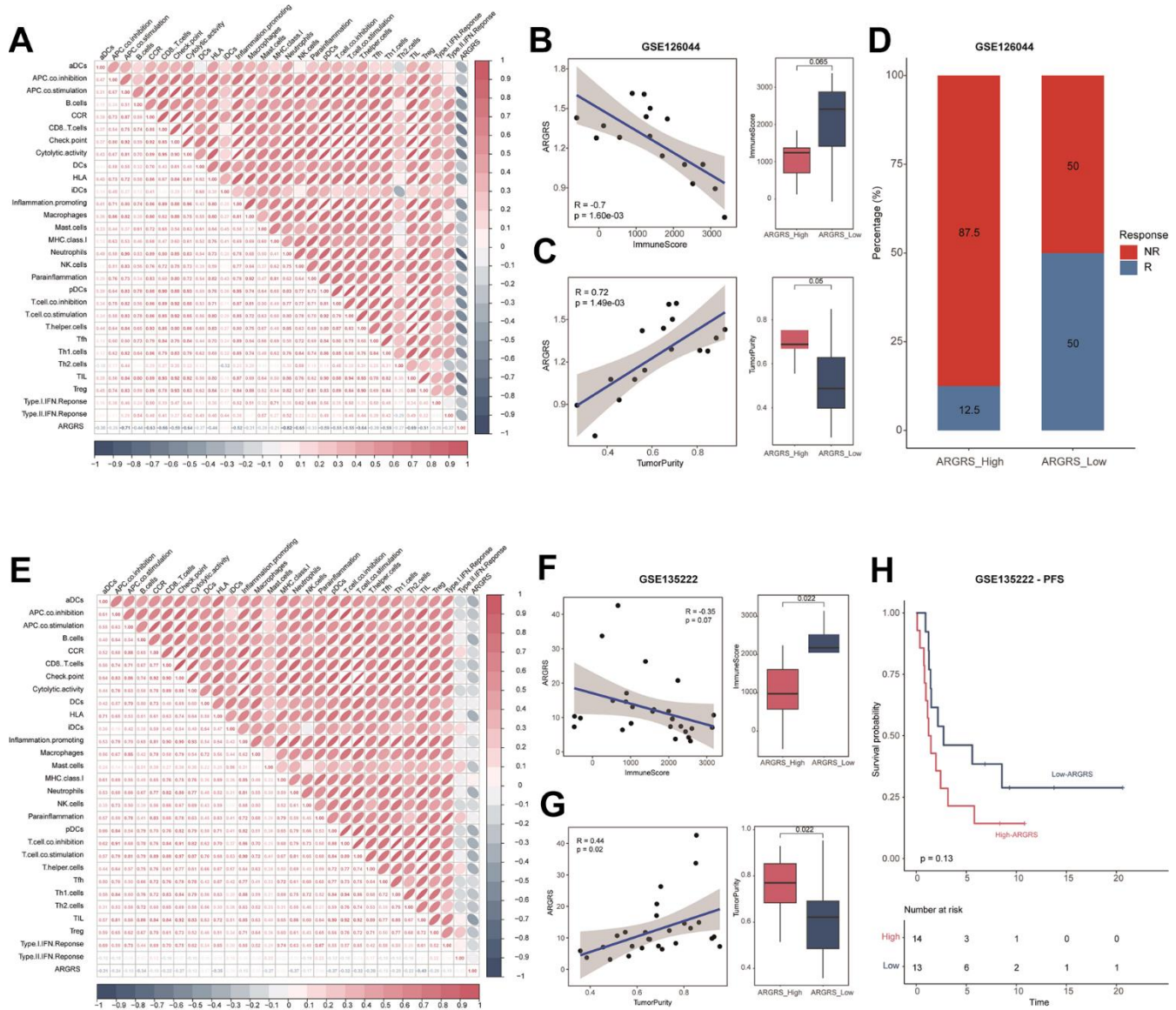

Figure 6. Patients with higher ARGRS were resistance to immunotherapy. (**A**) The correlation among 29 immune cell types and immune-related pathways and ARGRS in the GSE126044 cohort. (**B**) Left: Correlation between ARGRS and ImmuneScore in the GSE126044 cohort. Right: Boxplot showing the levels of ImmuneScore between ARGRS-high and low groups. (**C**) Left: Correlation between ARGRS and TumorPurity in the GSE126044 cohort. Right: Boxplot showing the levels of TumorPurity between ARGRS-high and low groups. (**D**) Barplot showing the percentage of immunotherapy responsive and non-responsive NSCLC patients in ARGRS-high and low groups. (**E**) The correlation among 29 immune cell types and immune-related pathways and ARGRS in the GSE135222 cohort. (**F**) Left: Correlation between ARGRS and ImmuneScore in the GSE135222 cohort. Right: Boxplot showing the levels of ImmuneScore between ARGRS-high and low groups. (**G**) Left: Correlation between ARGRS and TumorPurity in the GSE135222 cohort. Right: Boxplot showing the levels of TumorPurity between ARGRS-high and low groups. (**H**) Survival analysis showing the prognostic value of ARGRS in the GSE135222 cohort.

with minimal immune infiltration. Further investigation revealed that the results from the independent NSCLC cohorts were compatible with the findings at the high dimensionality datasets. Patients in the ARGRS-high group, in particular, showed lower levels of immune features but higher tumor purity.

Given of the important role of immunological status of TME in immunotherapeutic response [9, 10, 52], we further investigated the responses of the ARGRS-high and low groups in immunotherapy cohorts. Approximately 87.5% of patients in the ARGRS-high group did not achieve remission after immunotherapy, and were more likely to recurrence or progress, suggesting the immunotherapeutic resistance of these patients. There is now evidence that NSCLC patients can benefit from immunotherapy. However, limited by the inadequate research of characteristics to distinguish
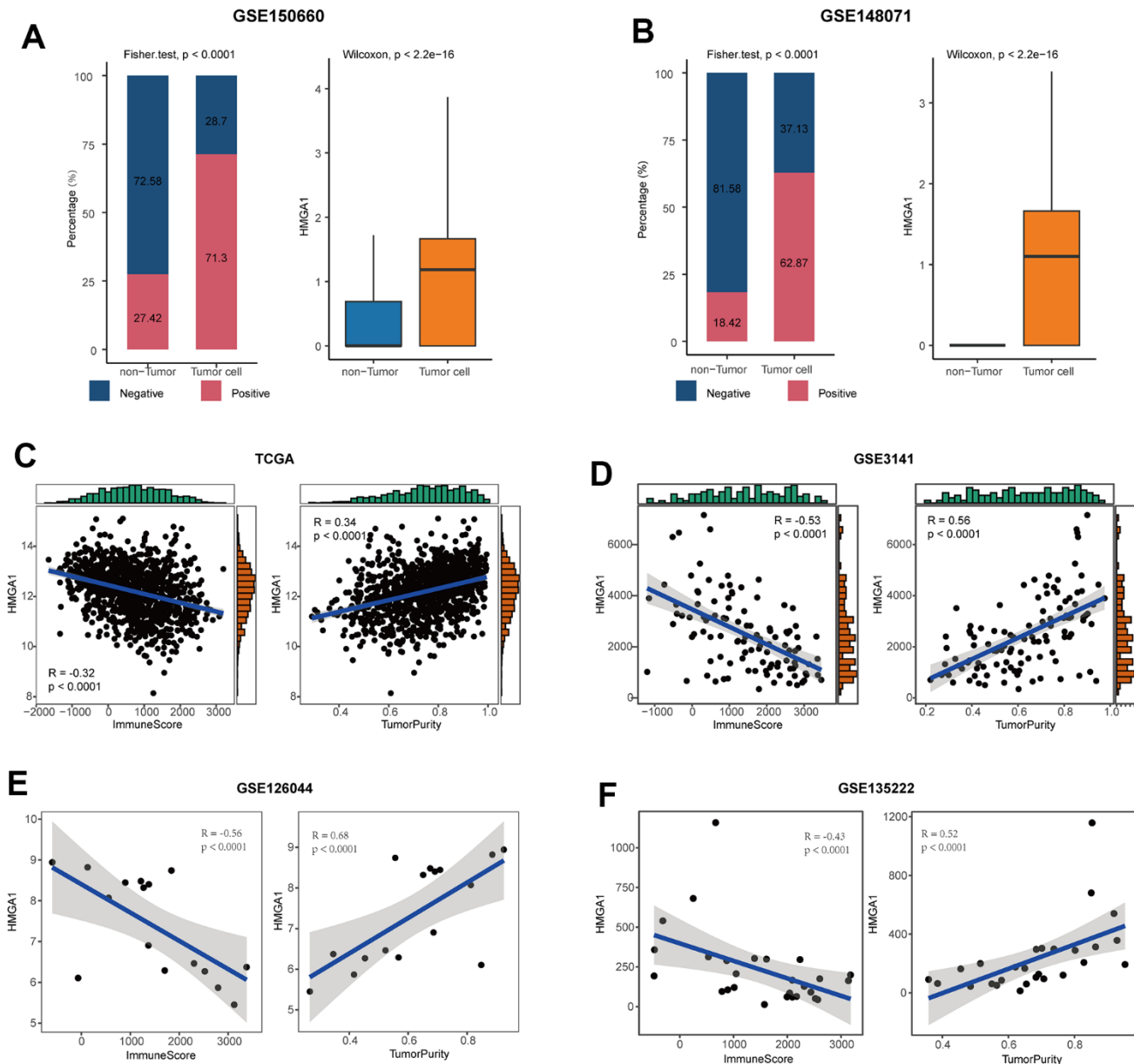

Figure 7. HMGA1 was the representative biomarker. (**A**) Right: bar plot showing the expressed fraction of HMGA1 between the non-tumor and tumor cells in the GSE150660 cohort. Left: comparison of the expression levels of HMGA1 between the non-tumor and tumor cells in the GSE150660 cohort. (**B**) Right: bar plot showing the expressed fraction of HMGA1 between the non-tumor and tumor cells in the GSE148071 cohort. Left: comparison of the expression levels of HMGA1 between the non-tumor and tumor cells in the GSE148071cohort. (**C**–**F**) Correlations between HMGA1 expression and ImmuneScore and tumor purity in (**C**) the TCGA, (**D**) the GSE3141, (**E**) the GSE126044, and (**F**) the GSE135222 cohorts.

the inflamed and desert TME, numerous patients received immunotherapy without obtaining effective results. Therefore, we performed a comprehensive study on explore the ARGs and delved into the prognosis and immune microenvironment characteristics of NSCLC. Additionally, given the regulating upstream feature of ARGRS, targeting it can activate the patient's immune system coincident with tumor cell killing.

In addition, we further validated the expression and immuno-correlations of HMGA1 in NSCLC in the in-house cohort. We found that HMGA1 was highly expressed in tumor tissues and enriched in immuno-cold tumors, indicating that HMGA1 could shape immuno-cold TME and promote immuno-suppression. In fact, reversal of HMGA1-mediated immunosuppression could improve hepatocellular carcinoma therapy [53]. In the term of molecular mechanisms, HMGA1 acted as a crucial regulator of tumor-promoting macrophage recruitment by activating NF-κB-CCL2 signaling and also regulated PD-L1 expression to promote immunosuppression [54, 55].

CONCLUSIONS

In this study, we investigated NSCLC-specific anoikis-related genes and constructed a prognostic model. Patients with higher ARGRS had worse clinical outcomes. Besides, we proved that the ARGRS and signatures in the models were highly expressed on tumor cells, compared with the non-tumor cells at the single-cell level. Notably, we found that the ARGRS was negatively correlated with the fraction of T cells. And other independent cohorts also showed the same results. To be specific, the ARGRS-high group had low immune infiltration, while the ARGRS-low group showed more active immunological status. Furthermore, in the immunotherapy cohort, patients who did not achieve remission or had tumor progression after immunotherapy were enriched in the ARGRS-high group, suggesting that patients with the ARGRS-high phenotype were more likely to be resistant to immunotherapy. To sum up, we constructed an anoikis-related gene model, and explained the relationship between ARGRS and immunological status, which can help to develop more personalized and precise treatment strategies in clinical practice.
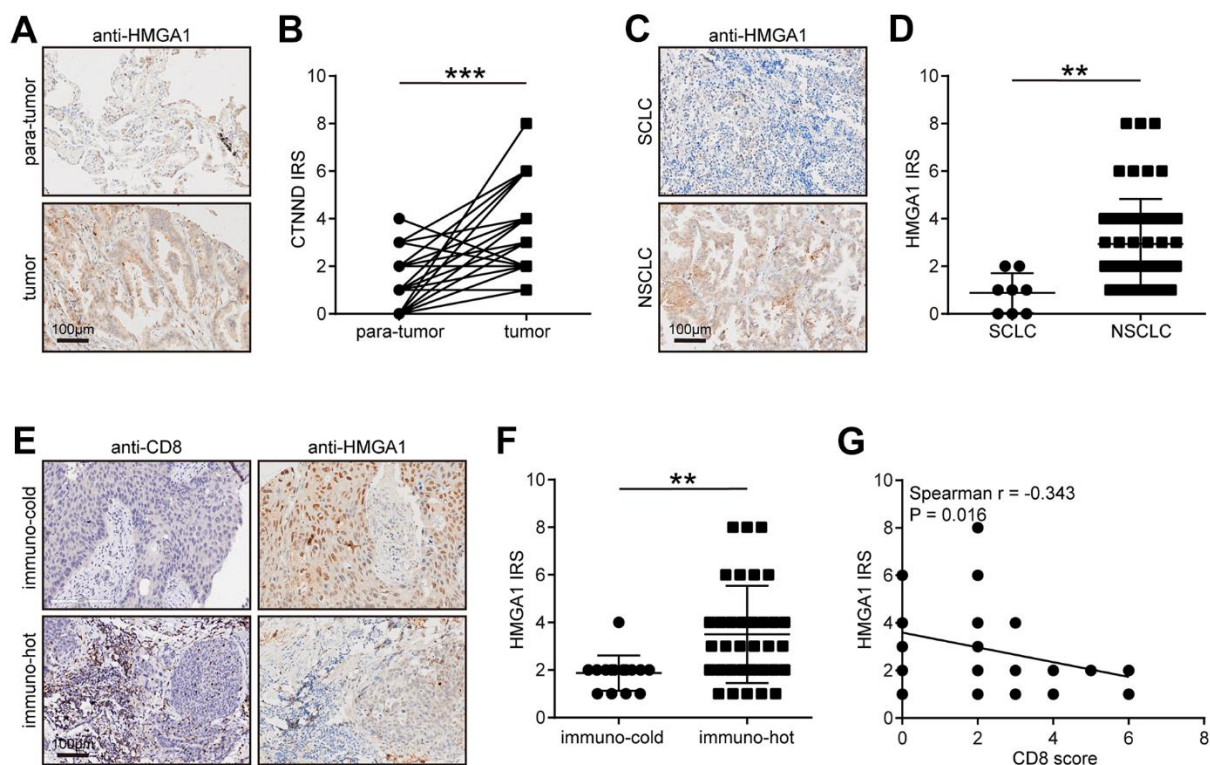

Figure 8. Validation of CTNND1 expression and its correlations with the TME features in the in-house cohort. (**A**, **B**) Representative images uncovering the expression of HMGA1 in para-tumor and tumor tissues and quantitative analysis. (**C**, **D**) Representative images uncovering the expression of HMGA1 in NSCLC and SCLC tumor tissues and quantitative analysis. (**E**, **F**) Representative images uncovering the expression of HMGA1 in immuno-cold and immuno-hot tumor tissues and quantitative analysis. (**G**) Correlation between HMGA1 expression and CD8 score.

Abbreviations

NSCLC: non-small cell lung cancer; LUAD: lung adenocarcinoma; LUSC: lung squamous cell carcinoma; TME: tumor microenvironment; ARG: anoikis-related gene; GEO: Gene Expression Omnibus; OS: overall survival; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; tSNE: t-Distributed Stochastic Neighbor Embedding; SNN: shared nearest neighbor; PC: principal component; PCA: principal component analysis; IHC: immuno-histochemistry; TMA: tissue microarray; FDA: Food and Drug Administration; PD-1: programmed cell death protein-1; ICI: immune checkpoint inhibitor.

AUTHOR CONTRIBUTIONS

FW and LC designed the study and reviewed the data. YW, ZZ, QQ, and SX performed experiments and drafted the manuscript. YW, ZZ, and QQ performed bioinformatics analyses. All the authors have read, revised, and approved the manuscript.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ETHICAL STATEMENT

Due to that the lung cancer tissue microarray was provided by Outdo BioTech Co. Ltd. (Shanghai, China, https://www.superchip.com.cn/), the Clinical Research Ethics Committee of Outdo BioTech approved the research of the TMAs (approval No. (YB-M-05-02).

FUNDING

REFERENCES

1. Alberg AJ, Brock MV, Ford JG, Samet JM, Spivack SD. Epidemiology of lung cancer: Diagnosis and management of lung cancer, 3rd ed: American College of Chest Physicians evidence-based clinical practice guidelines. Chest. 2013; 143:e1S–29S.
https://doi.org/10.1378/chest.12-2345
PMID:23649439

2. Relli V, Trerotola M, Guerra E, Alberti S. Abandoning the Notion of Non-Small Cell Lung Cancer. Trends Mol Med. 2019; 25:585–94.
https://doi.org/10.1016/j.molmed.2019.04.012
PMID:31155338

3. Mitchell PL, John T. Lung cancer in 2016: immunotherapy comes of age. Lancet Respir Med. 2016; 4:947–9.
https://doi.org/10.1016/S2213-2600(16)30379-4
PMID:27890503

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

5. Yu S, Li Y, Ren H, Zhou H, Ning Q, Chen X, Hu T, Yang L. PDK4 promotes tumorigenesis and cisplatin resistance in lung adenocarcinoma via transcriptional regulation of EPAS1. Cancer Chemother Pharmacol. 2021; 87:207–15.
https://doi.org/10.1007/s00280-020-04188-9
PMID:33221963

6. Kennedy LB, Salama AK. A review of cancer immunotherapy toxicity. CA Cancer J Clin. 2020; 70:86–104.
https://doi.org/10.3322/caac.21596 PMID:31944278

7. Jhunjhunwala S, Hammer C, Delamarre L. Antigen presentation in cancer: insights into tumour immunogenicity and immune evasion. Nat Rev Cancer. 2021; 21:298–312.
https://doi.org/10.1038/s41568-021-00339-z
PMID:33750922

8. Xu J, Gao Z, Liu K, Fan Y, Zhang Z, Xue H, Guo X, Zhang P, Deng L, Wang S, Wang H, Wang Q, Zhao R, Li G. The Non-N6-Methyladenosine Epitranscriptome Patterns and Characteristics of Tumor Microenvironment Infiltration and Mesenchymal Transition in Glioblastoma. Front Immunol. 2022; 12:809808.
https://doi.org/10.3389/fimmu.2021.809808
PMID:35154083

9. Gajewski TF, Corrales L, Williams J, Horton B, Sivan A, Spranger S. Cancer Immunotherapy Targets Based on Understanding the T Cell-Inflamed Versus Non-T Cell-Inflamed Tumor Microenvironment. Adv Exp Med Biol. 2017; 1036:19–31.
https://doi.org/10.1007/978-3-319-67577-0_2
PMID:29275462

10. Hu R, Han Q, Zhang J. STAT3: A key signaling molecule for converting cold to hot tumors. Cancer Lett. 2020; 489:29–40.
https://doi.org/10.1016/j.canlet.2020.05.035
PMID:32522692

11. Mao W, Cai Y, Chen D, Jiang G, Xu Y, Chen R, Wang F, Wang X, Zheng M, Zhao X, Mei J. Statin shapes

inflamed tumor microenvironment and enhances immune checkpoint blockade in non-small cell lung cancer. JCI Insight. 2022; 7:e161940.
https://doi.org/10.1172/jci.insight.161940
PMID:35943796

12. Mei J, Cai Y, Wang H, Xu R, Zhou J, Lu J, Yang X, Pan J, Liu C, Xu J, Zhu Y. Formin protein DIAPH1 positively regulates PD-L1 expression and predicts the therapeutic response to anti-PD-1/PD-L1 immunotherapy. Clin Immunol. 2023; 246:109204.
https://doi.org/10.1016/j.clim.2022.109204
PMID:36503156

13. Taddei ML, Giannoni E, Fiaschi T, Chiarugi P. Anoikis: an emerging hallmark in health and diseases. J Pathol. 2012; 226:380–93.
https://doi.org/10.1002/path.3000 PMID:21953325

14. Cai Q, Yan L, Xu Y. Anoikis resistance is a critical feature of highly aggressive ovarian cancer cells. Oncogene. 2015; 34:3315–24.
https://doi.org/10.1038/onc.2014.264 PMID:25132267

15. Sakamoto S, Kyprianou N. Targeting anoikis resistance in prostate cancer metastasis. Mol Aspects Med. 2010; 31:205–14.
https://doi.org/10.1016/j.mam.2010.02.001
PMID:20153362

16. Yu Y, Song Y, Cheng L, Chen L, Liu B, Lu D, Li X, Li Y, Lv F, Xing Y. CircCEMIP promotes anoikis-resistance by enhancing protective autophagy in prostate cancer cells. J Exp Clin Cancer Res. 2022; 41:188.
https://doi.org/10.1186/s13046-022-02381-7
PMID:35655258

17. Wang LN, Zhang ZT, Wang L, Wei HX, Zhang T, Zhang LM, Lin H, Zhang H, Wang SQ. TGF-β1/SH2B3 axis regulates anoikis resistance and EMT of lung cancer cells by modulating JAK2/STAT3 and SHP2/Grb2 signaling pathways. Cell Death Dis. 2022; 13:472.
https://doi.org/10.1038/s41419-022-04890-x
PMID:35589677

18. Ahn HM, Ryu J, Song JM, Lee Y, Kim HJ, Ko D, Choi I, Kim SJ, Lee JW, Kim S. Anti-cancer Activity of Novel TM4SF5-Targeting Antibodies through TM4SF5 Neutralization and Immune Cell-Mediated Cytotoxicity. Theranostics. 2017; 7:594–613.
https://doi.org/10.7150/thno.15629 PMID:28255353

19. Rousseaux S, Debernardi A, Jacquiau B, Vitte AL, Vesin A, Nagy-Mignotte H, Moro-Sibilot D, Brichon PY, Lantuejoul S, Hainaut P, Laffaire J, de Reyniès A, Beer DG, et al. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. Sci Transl Med. 2013; 5:186ra66.
https://doi.org/10.1126/scitranslmed.3005723
PMID:23698379

20. Botling J, Edlund K, Lohr M, Hellwig B, Holmberg L, Lambe M, Berglund A, Ekman S, Bergqvist M, Pontén F, König A, Fernandes O, Karlsson M, et al. Biomarker discovery in non-small cell lung cancer: integrating gene expression profiling, meta-analysis, and tissue microarray validation. Clin Cancer Res. 2013; 19:194–204.
https://doi.org/10.1158/1078-0432.CCR-12-1139
PMID:23032747

21. Bild AH, Yao G, Chang JT, Wang Q, Potti A, Chasse D, Joshi MB, Harpole D, Lancaster JM, Berchuck A, Olson JA Jr, Marks JR, Dressman HK, et al. Oncogenic pathway signatures in human cancers as a guide to targeted therapies. Nature. 2006; 439:353–7.
https://doi.org/10.1038/nature04296 PMID:16273092

22. Cho JW, Hong MH, Ha SJ, Kim YJ, Cho BC, Lee I, Kim HR. Genome-wide identification of differentially methylated promoters and enhancers associated with response to anti-PD-1 therapy in non-small cell lung cancer. Exp Mol Med. 2020; 52:1550–63.
https://doi.org/10.1038/s12276-020-00493-8
PMID:32879421

23. Jung H, Kim HS, Kim JY, Sun JM, Ahn JS, Ahn MJ, Park K, Esteller M, Lee SH, Choi JK. DNA methylation loss promotes immune evasion of tumours with high mutation and copy number load. Nat Commun. 2019; 10:4278.
https://doi.org/10.1038/s41467-019-12159-9
PMID:31537801

24. Rebhan M, Chalifa-Caspi V, Prilusky J, Lancet D. GeneCards: integrating information about genes, proteins and diseases. Trends Genet. 1997; 13:163.
https://doi.org/10.1016/s0168-9525(97)01103-7
PMID:9097728

25. Rouillard AD, Gundersen GW, Fernandez NF, Wang Z, Monteiro CD, McDermott MG, Ma'ayan A. The harmonizome: a collection of processed datasets gathered to serve and mine knowledge about genes and proteins. Database (Oxford). 2016; 2016:baw100.
https://doi.org/10.1093/database/baw100
PMID:27374120

26. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16:284–7.
https://doi.org/10.1089/omi.2011.0118
PMID:22455463

27. Gene Ontology Consortium. Gene Ontology Consortium: going forward. Nucleic Acids Res. 2015; 43:D1049–56.
https://doi.org/10.1093/nar/gku1179 PMID:25428369

28. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. 2000;

28:27–30.
https://doi.org/10.1093/nar/28.1.27
PMID:10592173

29. Chi Y, Remsik J, Kiseliovas V, Derderian C, Sener U, Alghader M, Saadeh F, Nikishina K, Bale T, Iacobuzio-Donahue C, Thomas T, Pe'er D, Mazutis L, Boire A. Cancer cells deploy lipocalin-2 to collect limiting iron in leptomeningeal metastasis. Science. 2020; 369:276–82.
https://doi.org/10.1126/science.aaz2193
PMID:32675368

30. Wu F, Fan J, He Y, Xiong A, Yu J, Li Y, Zhang Y, Zhao W, Zhou F, Li W, Zhang J, Zhang X, Qiao M, et al. Single-cell profiling of tumor heterogeneity and the microenvironment in advanced non-small cell lung cancer. Nat Commun. 2021; 12:2540.
https://doi.org/10.1038/s41467-021-22801-0
PMID:33953163

31. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018; 36:411–20.
https://doi.org/10.1038/nbt.4096
PMID:29608179

32. Korsunsky I, Millard N, Fan J, Slowikowski K, Zhang F, Wei K, Baglaenko Y, Brenner M, Loh PR, Raychaudhuri S. Fast, sensitive and accurate integration of single-cell data with Harmony. Nat Methods. 2019; 16:1289–96.
https://doi.org/10.1038/s41592-019-0619-0
PMID:31740819

33. Li Q, Wang R, Yang Z, Li W, Yang J, Wang Z, Bai H, Cui Y, Tian Y, Wu Z, Guo Y, Xu J, Wen L, et al. Molecular profiling of human non-small cell lung cancer by single-cell RNA-seq. Genome Med. 2022; 14:87.
https://doi.org/10.1186/s13073-022-01089-9
PMID:35962452

34. Mei J, Cai Y, Chen L, Wu Y, Liu J, Qian Z, Jiang Y, Zhang P, Xia T, Pan X, Zhang Y. The heterogeneity of tumour immune microenvironment revealing the CRABP2/CD69 signature discriminates distinct clinical outcomes in breast cancer. Br J Cancer. 2023; 129:1645–57.
https://doi.org/10.1038/s41416-023-02432-6
PMID:37715025

35. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc. 2020; 15:1484–506.
https://doi.org/10.1038/s41596-020-0292-x
PMID:32103204

36. Cai Y, Ji W, Sun C, Xu R, Chen X, Deng Y, Pan J, Yang J, Zhu H, Mei J. Interferon-Induced Transmembrane Protein 3 Shapes an Inflamed Tumor Microenvironment and Identifies Immuno-Hot Tumors. Front Immunol. 2021; 12:704965.
https://doi.org/10.3389/fimmu.2021.704965
PMID:34456915

37. Mei J, Fu Z, Cai Y, Song C, Zhou J, Zhu Y, Mao W, Xu J, Yin Y. SECTM1 is upregulated in immuno-hot tumors and predicts immunotherapeutic efficacy in multiple cancers. iScience. 2023; 26:106027.
https://doi.org/10.1016/j.isci.2023.106027
PMID:36818292

38. Mei J, Cai Y, Xu R, Zhu Y, Zhao X, Zhang Y, Mao W, Xu J, Yin Y. Protocol to identify novel immunotherapy biomarkers based on transcriptomic data in human cancers. STAR Protoc. 2023; 4:102258.
https://doi.org/10.1016/j.xpro.2023.102258
PMID:37119142

39. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14:7.
https://doi.org/10.1186/1471-2105-14-7
PMID:23323831

40. Mei J, Liu Y, Yu X, Hao L, Ma T, Zhan Q, Zhang Y, Zhu Y. YWHAZ interacts with DAAM1 to promote cell migration in breast cancer. Cell Death Discov. 2021; 7:221.
https://doi.org/10.1038/s41420-021-00609-7
PMID:34453038

41. Cancer Genome Atlas Network. Genomic Classification of Cutaneous Melanoma. Cell. 2015; 161:1681–96.
https://doi.org/10.1016/j.cell.2015.05.044
PMID:26091043

42. Mezquita L, Jové M, Nadal E, Kfoury M, Morán T, Ricordel C, Dhooge M, Tlemsani C, Léna H, Teulé A, Álvarez JV, Raimbourg J, Hiret S, et al. High Prevalence of Somatic Oncogenic Driver Alterations in Patients With NSCLC and Li-Fraumeni Syndrome. J Thorac Oncol. 2020; 15:1232–9.
https://doi.org/10.1016/j.jtho.2020.03.005
PMID:32179180

43. Offin M, Chan JM, Tenet M, Rizvi HA, Shen R, Riely GJ, Rekhtman N, Daneshbod Y, Quintanal-Villalonga A, Penson A, Hellmann MD, Arcila ME, Ladanyi M, et al. Concurrent RB1 and TP53 Alterations Define a Subset of EGFR-Mutant Lung Cancers at risk for Histologic Transformation and Inferior Clinical Outcomes. J Thorac Oncol. 2019; 14:1784–93.
https://doi.org/10.1016/j.jtho.2019.06.002
PMID:31228622

44. Jamme P, Fernandes M, Copin MC, Descarpentries C, Escande F, Morabito A, Grégoire V, Jamme M, Baldacci S, Tulasne D, Kherrouche Z, Cortot AB. Alterations in

the PI3K Pathway Drive Resistance to MET Inhibitors in NSCLC Harboring MET Exon 14 Skipping Mutations. J Thorac Oncol. 2020; 15:741–51.
https://doi.org/10.1016/j.jtho.2020.01.027
PMID:32169477

45. Ho DW, Tsui YM, Chan LK, Sze KM, Zhang X, Cheu JW, Chiu YT, Lee JM, Chan AC, Cheung ET, Yau DT, Chia NH, Lo IL, et al. Single-cell RNA sequencing shows the immunosuppressive landscape and tumor heterogeneity of HBV-associated hepatocellular carcinoma. Nat Commun. 2021; 12:3684.
https://doi.org/10.1038/s41467-021-24010-1
PMID:34140495

46. Yang R, Sun L, Li CF, Wang YH, Yao J, Li H, Yan M, Chang WC, Hsu JM, Cha JH, Hsu JL, Chou CW, Sun X, et al. Galectin-9 interacts with PD-1 and TIM-3 to regulate T cell death and is a target for cancer immunotherapy. Nat Commun. 2021; 12:832.
https://doi.org/10.1038/s41467-021-21099-2
PMID:33547304

47. Liu L, Bai H, Wang C, Seery S, Wang Z, Duan J, Li S, Xue P, Wang G, Sun Y, Du X, Zhang X, Ma Z, Wang J. Efficacy and Safety of First-Line Immunotherapy Combinations for Advanced NSCLC: A Systematic Review and Network Meta-Analysis. J Thorac Oncol. 2021; 16:1099–117.
https://doi.org/10.1016/j.jtho.2021.03.016
PMID:33839365

48. Vaddepally RK, Kharel P, Pandey R, Garje R, Chandra AB. Review of Indications of FDA-Approved Immune Checkpoint Inhibitors per NCCN Guidelines with the Level of Evidence. Cancers (Basel). 2020; 12:738.
https://doi.org/10.3390/cancers12030738
PMID:32245016

49. Wang Y, Liu JJ, Dransfield PJ, Zhu L, Wang Z, Du X, Jiao X, Su Y, Li AR, Brown SP, Kasparian A, Vimolratana M, Yu M, et al. Discovery and Optimization of Potent GPR40 Full Agonists Containing Tricyclic Spirocycles. ACS Med Chem Lett. 2013; 4:551–5.
https://doi.org/10.1021/ml300427u
PMID:24900707

50. Jin L, Chun J, Pan C, Kumar A, Zhang G, Ha Y, Li D, Alesi GN, Kang Y, Zhou L, Yu WM, Magliocca KR, Khuri FR, et al. The PLAG1-GDH1 Axis Promotes Anoikis Resistance

and Tumor Metastasis through CamKK2-AMPK Signaling in LKB1-Deficient Lung Cancer. Mol Cell. 2018; 69:87–99.e7.
https://doi.org/10.1016/j.molcel.2017.11.025
PMID:29249655

51. Wang YN, Zeng ZL, Lu J, Wang Y, Liu ZX, He MM, Zhao Q, Wang ZX, Li T, Lu YX, Wu QN, Yu K, Wang F, et al. CPT1A-mediated fatty acid oxidation promotes colorectal cancer cell metastasis by inhibiting anoikis. Oncogene. 2018; 37:6025–40.
https://doi.org/10.1038/s41388-018-0384-z
PMID:29995871

52. Mei J, Cai Y, Xu R, Yu X, Han X, Weng M, Chen L, Ma T, Gao T, Gao F, Xia T, Zhu Y, Zhang Y. Angiotensin-converting enzyme 2 identifies immuno-hot tumors suggesting angiotensin-(1-7) as a sensitizer for chemotherapy and immunotherapy in breast cancer. Biol Proced Online. 2022; 24:15.
https://doi.org/10.1186/s12575-022-00177-9
PMID:36284262

53. Yan B, Liu C, Li H, Wen N, Jiao W, Wang S, Zhang Y, Zhang T, Zhang H, Lv Y, Fan H, Liu X. Reversal of HMGA1-Mediated Immunosuppression Synergizes with Immunogenic Magnetothermodynamic for Improved Hepatocellular Carcinoma Therapy. ACS Nano. 2023; 17:9209–23.
https://doi.org/10.1021/acsnano.3c00004
PMID:37162457

54. Chang X, Liu J, Yang Q, Gao Y, Ding X, Zhao J, Li Y, Liu Z, Li Z, Wu Y, Zuo D. Targeting HMGA1 contributes to immunotherapy in aggressive breast cancer while suppressing EMT. Biochem Pharmacol. 2023; 212:115582.
https://doi.org/10.1016/j.bcp.2023.115582
PMID:37146833

55. Chen J, Ji K, Gu L, Fang Y, Pan M, Tian S. HMGA1 Promotes Macrophage Recruitment via Activation of NF-κB-CCL2 Signaling in Hepatocellular Carcinoma. J Immunol Res. 2022; 2022:4727198.
https://doi.org/10.1155/2022/4727198
PMID:35785026

SUPPLEMENTARY MATERIALS

Supplementary Figures
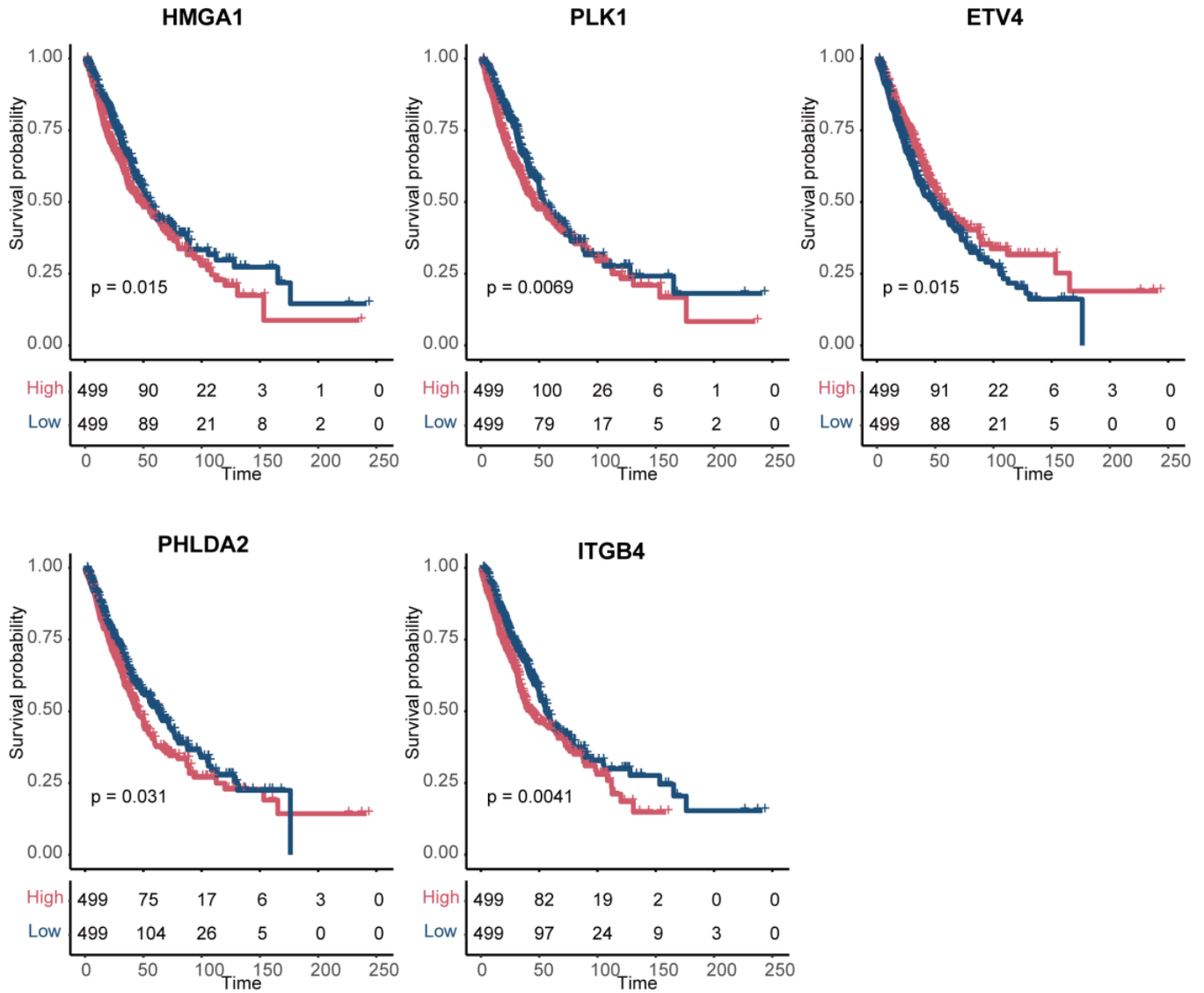
Supplementary Figure 1. Survival analysis showing the prognostic value of the five genes in the prognostic model in the TCGA cohort.
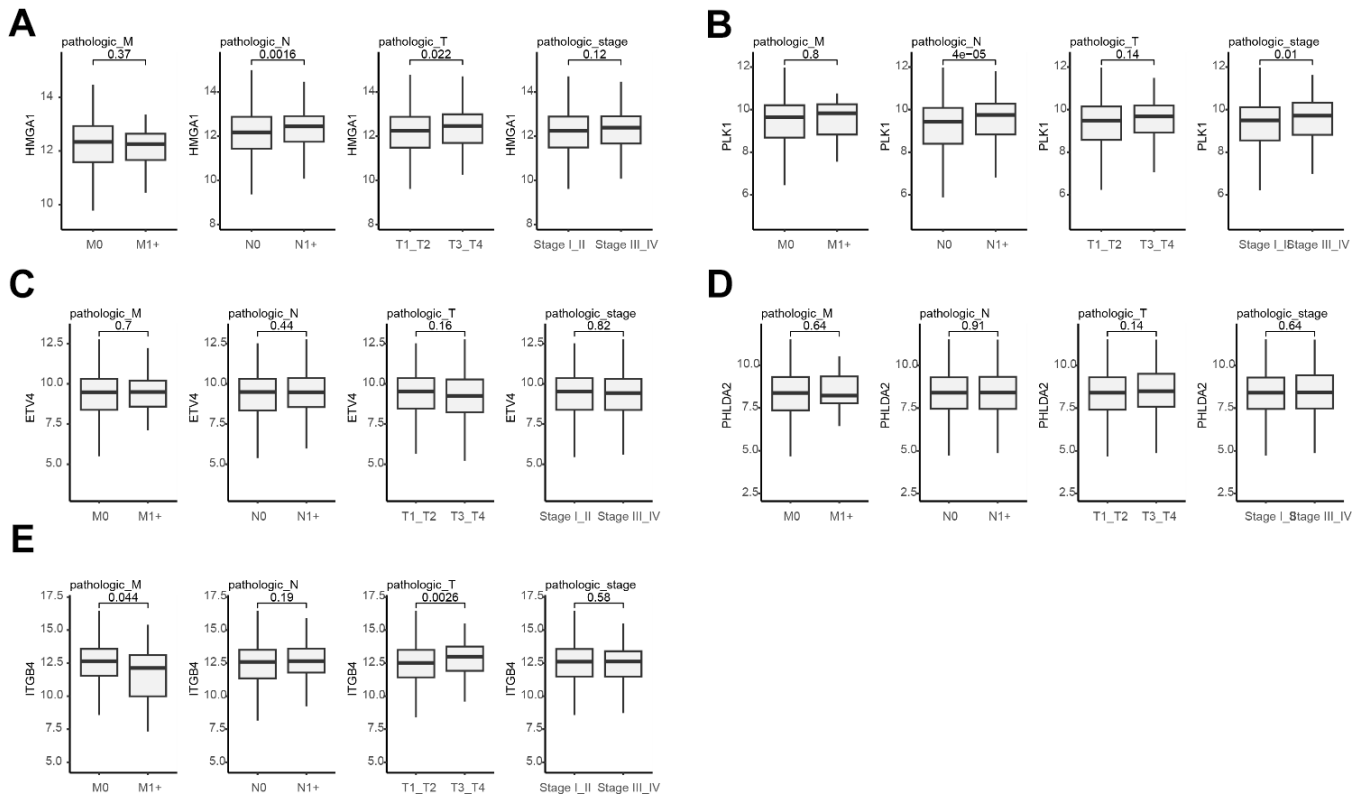
Supplementary Figure 2. Comparing the gene expression of signatures in the prognostic model among patients with different pathological characteristics. (**A**–**E**) Boxplot showing the expression levels of HMGA1 (**A**), PLK1 (**B**), ETV4 (**C**), PHLDA1 (**D**), and ITGB4 (**E**) among patients with different pathological characteristics.

| Gene	low-ARGRS	high-ARGRS	P−value
TP53	271	381	***
SYNE1	72	145	***
STK11	61	14	***
TTN	253	334	***
RASA1	7	33	***
SCN1A	27	65	***
PTEN	13	43	***
KRAS	103	55	***
EGFR	56	22	***
PLCH1	17	47	***
RFX6	10	33	***
PIK3CA	26	57	***
ORC5	0	11	***
KCNS3	12	0	***
PTPRO	12	35	***
GUCY2C	3	18	***
IGKV1−12	3	18	***
MAPK8IP3	3	18	***
TMEM176A	0	10	***
SLC22A6	14	1	***
DCAF8L2	10	30	**
CEP192	12	33	**
SPICE1	3	17	**
CHAT	9	28	**
CDH10	71	109	**
RADIL	4	19	**
KCNA1	22	5	**
SSRP1	1	12	**
C6	22	47	**
LIG4	7	24	**
ZNF665	7	24	**
COL22A1	57	92	**
RBM11	0	9	**
HECTD3	10	0	**
SPAST	2	14	**
UHRF1BP1	4	18	**
WWC1	3	16	**
DYSF	31	58	**
GRM3	12	31	**
RTN1	12	31	**
SEMA5B	14	34	**
PLEC	36	64	**
INVS	1	11	**
PRMT7	1	11	**
TBC1D15	1	11	**
UGGT2	6	21	**
CCDC67	12	1	**
BEST3	5	19	**
MTMR12	5	19	**
C7orf66	0	8	**
CD2	0	8	**
COLGALT1	0	8	**
DYX1C1	0	8	**
ELMO3	0	8	**
SP3	0	8	**
SUCLG1	0	8	**
THEMIS2	0	8	**
FOXC2	2	13	**
CCDC89	9	0	**
CATSPERB	3	15	**
SP4	3	15	**
NOS1	21	43	**
BAZ2B	14	33	**
DTNA	12	30	**
RALGAPA1	7	22	**
THSD7A	38	65	**
PRLR	9	25	**
ROBO3	15	34	**
BRAF	38	17	**
SLC9A2	11	28	**
MYBPC1	13	31	**
CMYA5	23	45	**
KIAA1217	5	18	**
ARHGEF18	1	10	**
INHA	1	10	**
KRT20	1	10	**
RSBN1L	1	10	**
KCNN2	8	23	**
ATP10A	21	42	**
NCOR2	14	32	**
C14orf37	4	16	**
LRTM2	4	16	**
SUV420H1	4	16	**
ELTD1	19	39	**
OR5L1	19	39	**
CECR1	2	12	**
EEA1	2	12	**
EXT2	2	12	**
IKBIP	2	12	**
PCCA	2	12	**
SLC6A14	2	12	**
VCL	2	12	**
ZNF253	2	12	**
EPN1	3	14	**
IGHV1−45	3	14	**
PCDHGA12	7	21	**
CD163L1	17	36	**
SATB2	17	36	**
APH1A	0	7	**
CARD17	0	7	**
CCDC113	0	7	**
CSRNP1	0	7	**
HSDL1	0	7	**
KCMF1	0	7	**
LACRT	0	7	**
LTV1	0	7	**
PAQR3	0	7	**
RALBP1	0	7	**
RBMX2	0	7	**
RINL	0	7	**
RORC	0	7	**
SSR3	0	7	**
ST3GAL2	0	7	**
UBXN1	0	7	**
ZNF17	0	7	**
OR10G8	15	33	**
CSNK1G1	8	0	**
SPAM1	11	27	**
COL4A4	18	37	**
PAX4	6	19	**
CTNNB1	20	6	**
SETBP1	46	24	**
KMT2D	59	88	**
CYFIP2	5	17	**
LY75−CD302	12	28	**
MPDZ	10	25	**

Supplementary Figure 3. Comparison of mutant alterations between the high- and low-ARGRS groups. Genes with p-value < 0.01 were displayed.
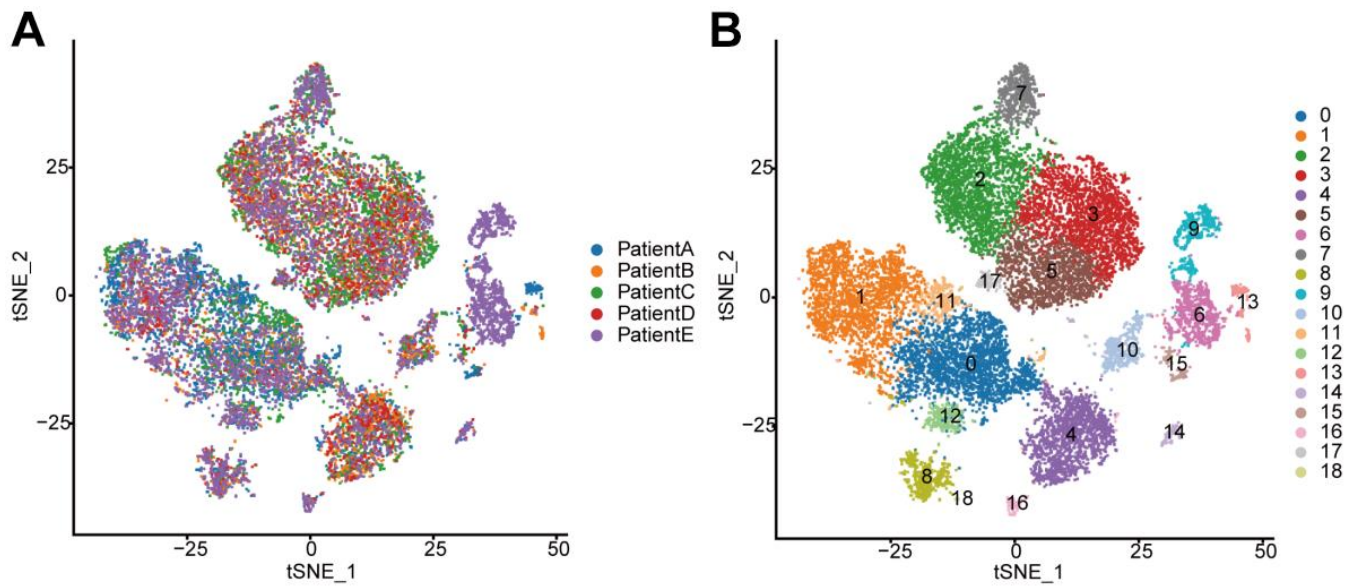
Supplementary Figure 4. Unsupervised clustering of five NSCLC patients from the GSE150660 dataset at the single-cell level.
(**A**) The distribution of NSCLC patients across all profiled single cells. (**B**) Unsupervised clustering of all profiled single cells.
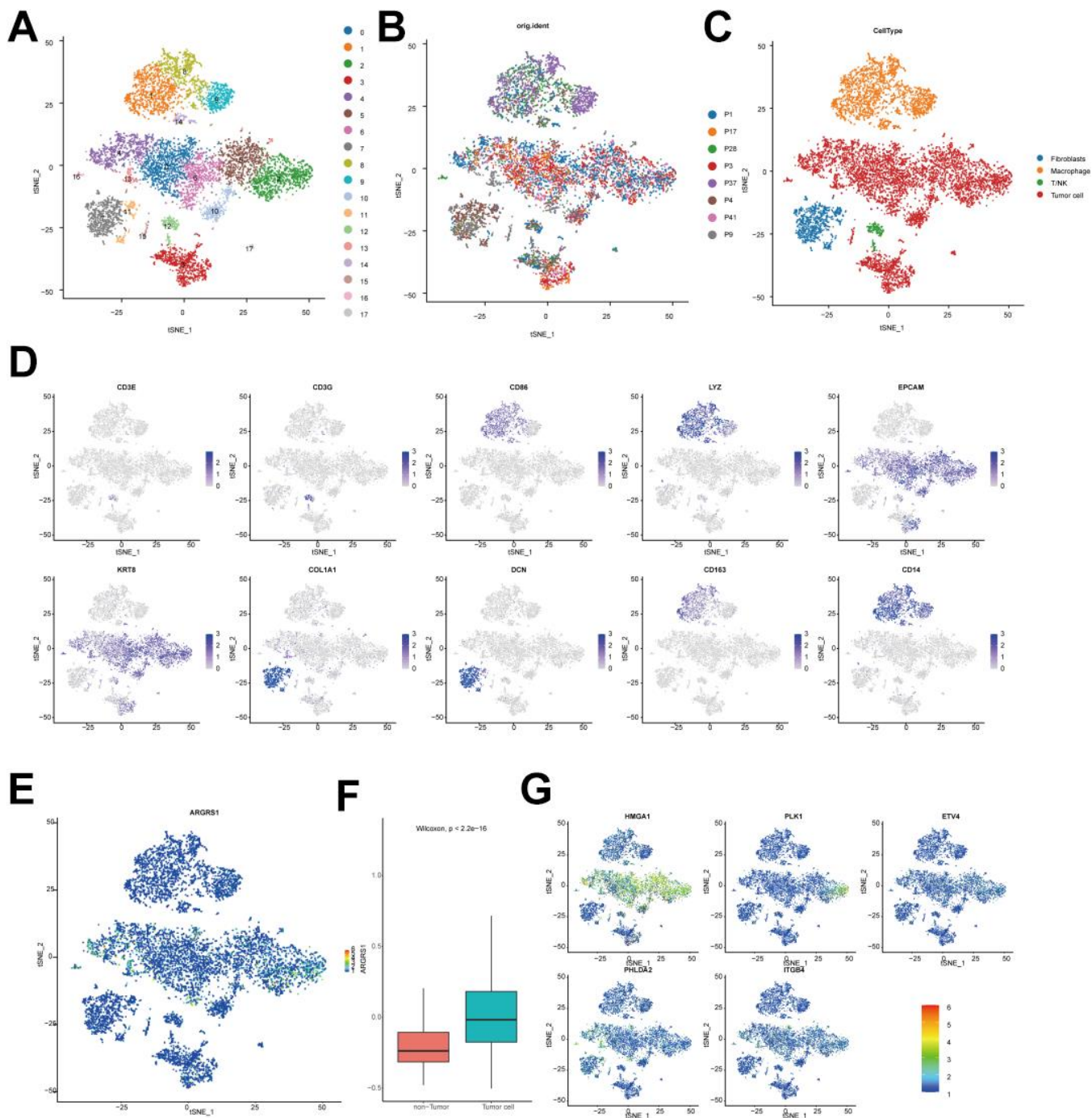
Supplementary Figure 5. Transcriptomic clustering of NSCLC patients in the GSE148071 cohort. (**A**)Unsupervised clustering of 9,563 cells. (**B**) The distribution of NSCLC patients across all profiled single cells. (**C**) Marker-based cell type identification analysis allowed the prediction of four broad cell types across all profiled single cells. (**D**) Expression levels of cell type signatures overlaid on the t-SNE representation. EPCAM and KRT8 for tumor cells, CD3E and CD3G for T cells, CD86, LYZ, CD163 and CD14 for macrophages, COL1A1 and DCN for fibroblasts. (**E**) ARGRS overlaid on the t-SNE representation. (**F**) Boxplot showing the levels of ARGRS between tumor and non-tumor cells. Horizontal lines in the boxplots represent the median, the lower and upper hinges correspond to the first and third quartiles, and the whiskers extend from the hinge up to 1.5 times the interquartile range from the hinge. (**G**) Expression levels of the five genes in the prognostic model overlaid on the t-SNE representation.
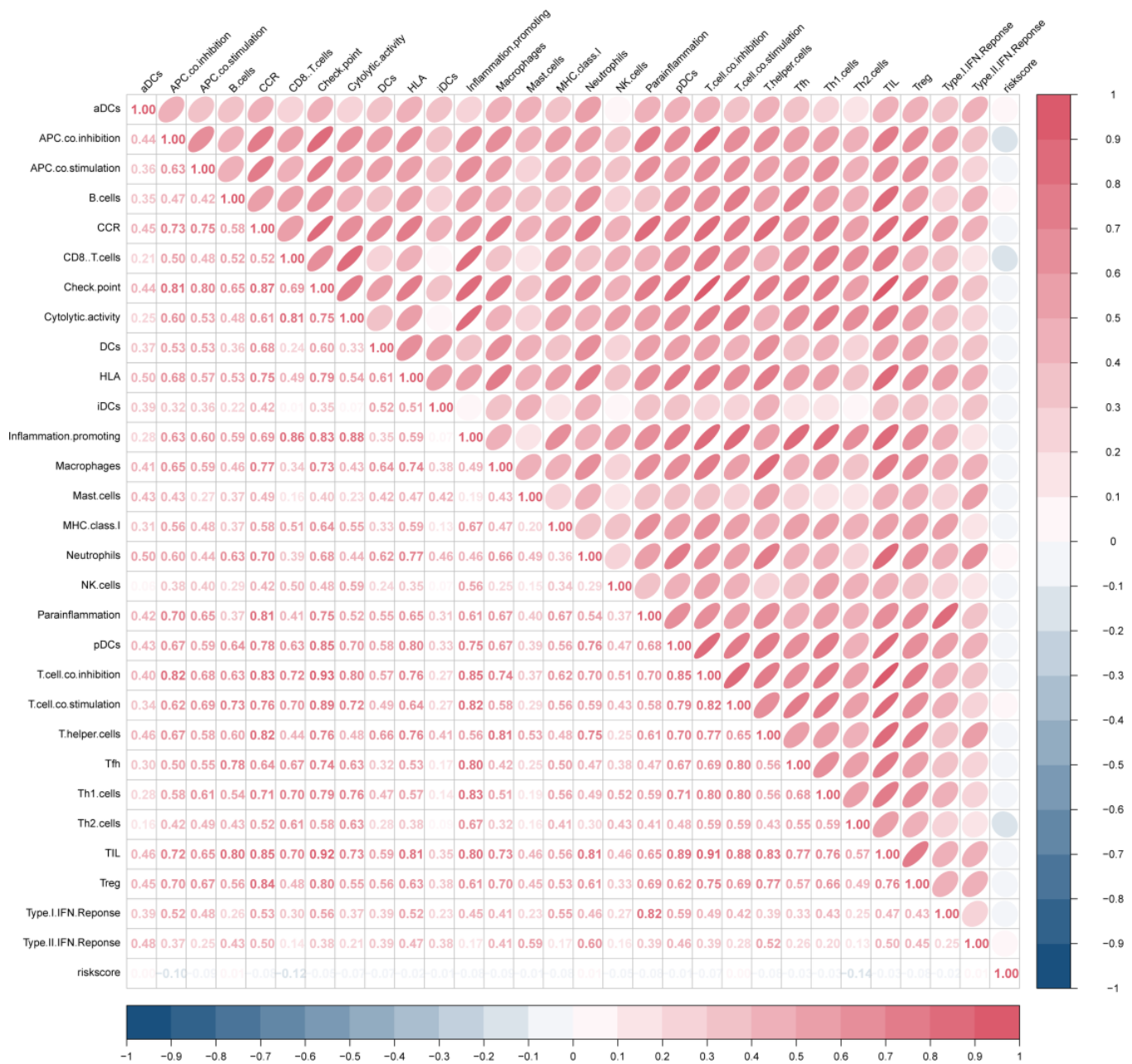
Supplementary Figure 6. The correlation among 29 immune cell types and immune-related pathways and ARGRS in the TCGA cohort.
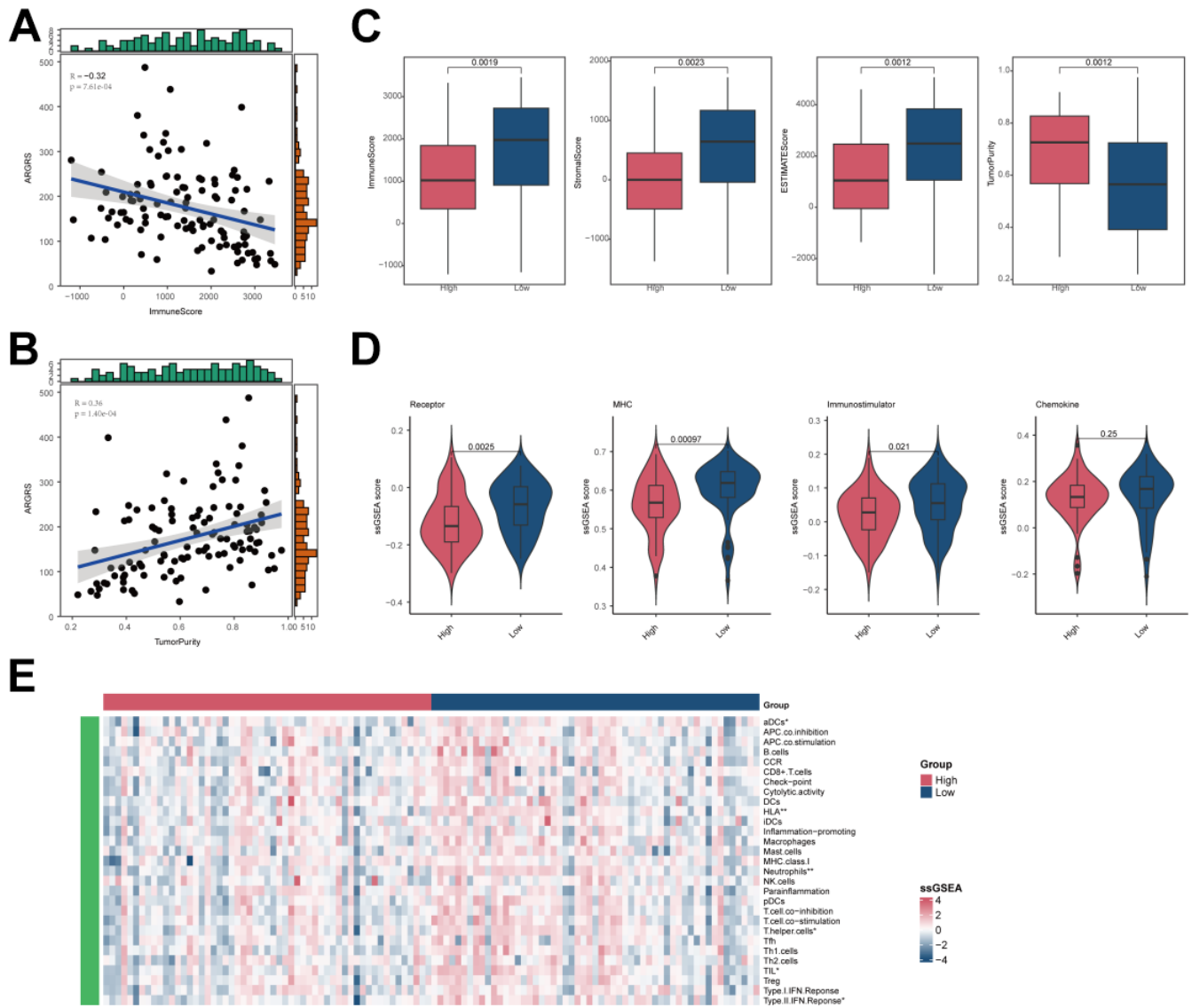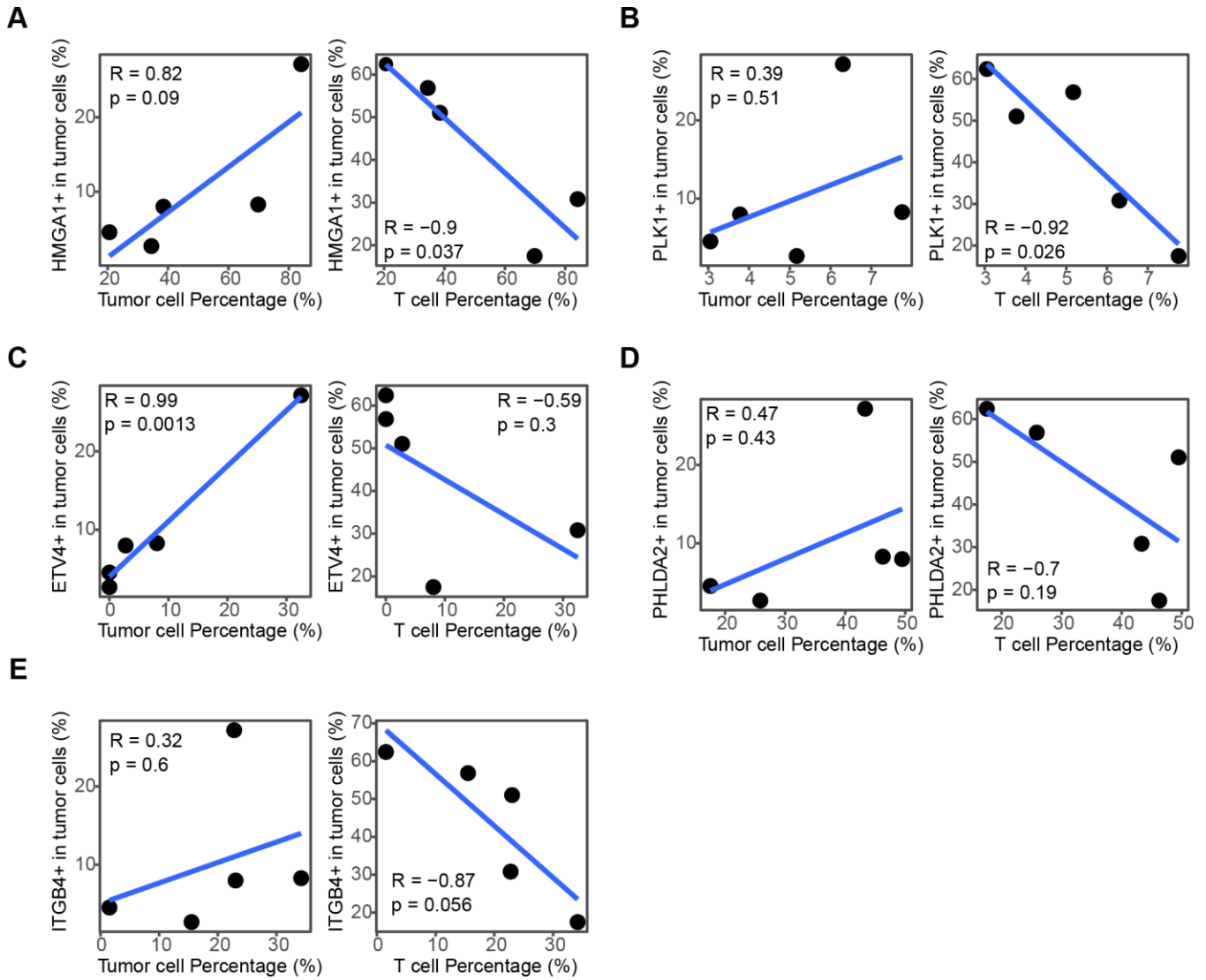
Supplementary Figure 7. Correlation between immunological characteristics and ARGRS in the GSE3141 cohort. (**A**, **B**) Correlation between ARGRS and ImmuneScore (**A**) and tumor purity (**B**). (**C**) Comparison of ImmuneScore, StromalScore, ESTIMATEScore, and tumor purity between ARGRS-high and low groups. (**D**) Comparison of the enrichment of receptors, MHC molecules, Immunostimulator, and Chemokine between ARGRS-high and low groups. (**E**) Heatmap showing the enrichment scores of 29 immunological characteristics.
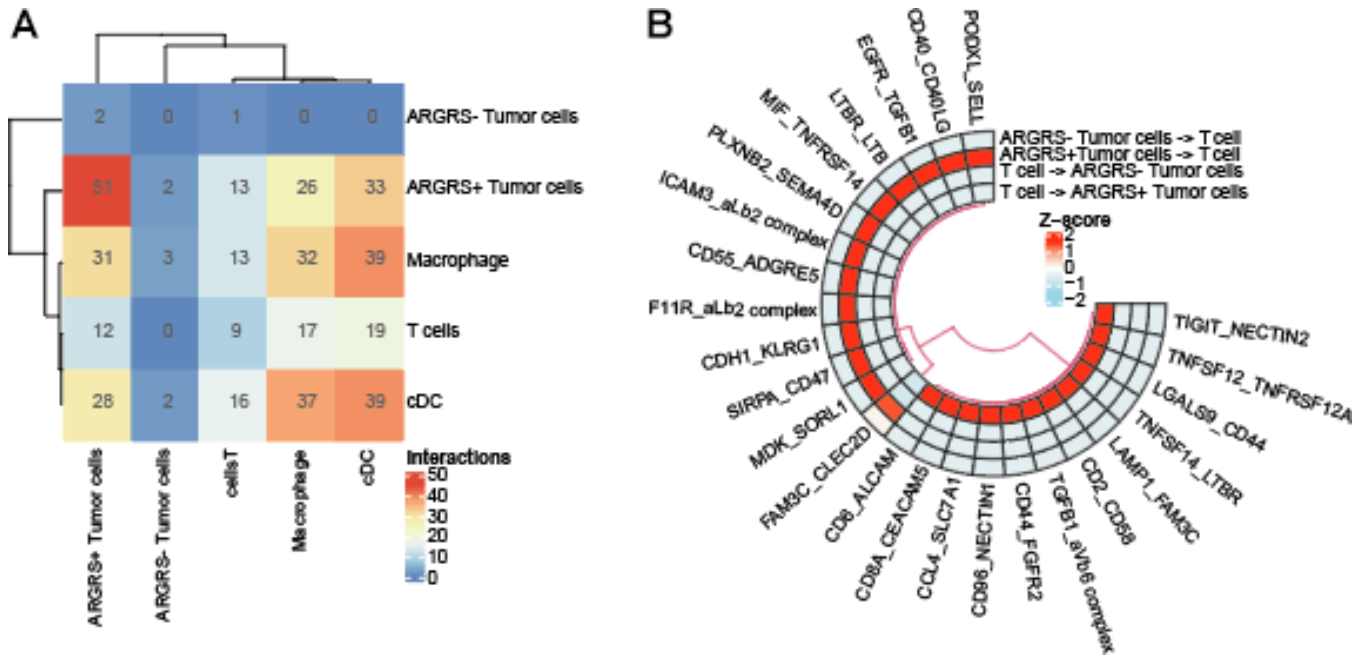
Supplementary Figure 8. Cell-cell communications among cell types in NSCLC (GSE150660). (**A**) Heatmap showing the interaction numbers among cell types. (**B**) Heatmap showing the ligand-receptors between tumor (ARGRS+ and ARGRS-) and T cells.
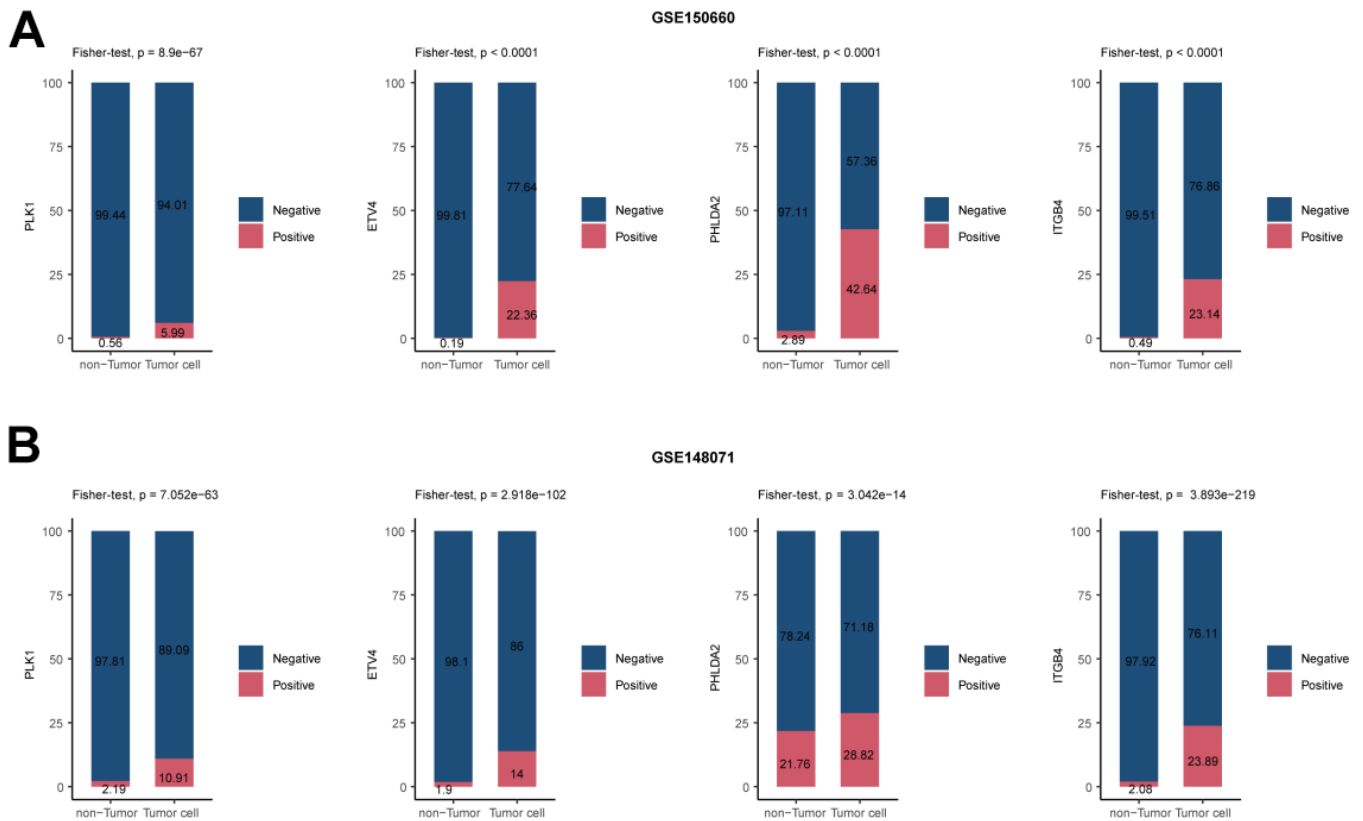
Supplementary Figure 9. Comparison of expressed fraction of genes in the model between non-tumor and tumor cells. (**A**) Bar plot showing the expressed fraction of PLK1, ETV4, PHLDA2, ITGB4 between the non-tumor and tumor cells in the GSE150660 cohort. (**B**) Bar plot showing the expressed fraction of PLK1, ETV4, PHLDA2, ITGB4 between the non-tumor and tumor cells in the GSE148071 cohort.

Supplementary Figure 10. Correlation between expressed fractions of genes in the model and tumor/T cell percentage in GSE150660. (**A**) Correlation between HMGA1+ cells in tumor cells and tumor (left) and T cell percentage (right) in GSE150660 dataset. (**B**) Correlation between PLK1+ cells in tumor cells and tumor (left) and T cell percentage (right) in GSE150660 dataset. (**C**) Correlation between ETV4+ cells in tumor cells and tumor (left) and T cell percentage (right) in GSE150660 dataset. (**D**) Correlation between PHLDA2+ cells in tumor cells and tumor (left) and T cell percentage (right) in GSE150660 dataset. (**E**) Correlation between ITGB4+ cells in tumor cells and tumor (left) and T cell percentage (right) in GSE150660 dataset.