A Novel Inflammatory-Related Gene Signature Based Model for Risk Stratification and Prognosis Prediction in Lung Adenocarcinoma

Wen-Yu Zhai[1,2†], Fang-Fang Duan[3†], Si Chen[1,2†], Jun-Ye Wang[1], Yao-Bin Lin[1,2], Yi-Zhi Wang[1,2], Bing-Yu Rao[1,2], Ze-Rui Zhao[1,2*] and Hao Long[1,2*]

[1]State Key Laboratory of Oncology in Southern China, Department of Thoracic Surgery, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, China, [2]Lung Cancer Research Center, Sun Yat-Sen University, Guangzhou, China, [3]State Key Laboratory of Oncology in Southern China, Department of Medical Oncology, Collaborative Innovation Center for Cancer Medicine, Sun Yat-Sen University Cancer Center, Guangzhou, China

Inflammation is an important hallmark of cancer and plays a role in both neogenesis and tumor development. Despite this, inflammatory-related genes (IRGs) remain to be poorly studied in lung adenocarcinoma (LUAD). We aim to explore the prognostic value of IRGs for LUAD and construct an IRG-based prognosis signature. The transcriptomic profiles and clinicopathological information of patients with LUAD were obtained from The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO). Least absolute shrinkage and selection operator (LASSO) analysis and multivariate Cox regression were applied in the TCGA set to generate an IRG risk signature. LUAD cases with from the GSE31210 and GSE30219 datasets were used to validate the predictive ability of the signature. Analysis of the TCGA cohort revealed a five-IRG risk signature consisting of EREG, GPC3, IL7R, LAMP3, and NMUR1. This signature was used to divide patients into two risk groups with different survival rates. Multivariate Cox regression analysis verified that the risk score from the five-IRG signature negatively correlated with patient outcome. A nomogram was developed using the IRG risk signature and stage, with C-index values of 0.687 (95% CI: 0.644–0.730) in the TCGA training cohort, 0.678 (95% CI: 0.586–0.771) in GSE30219 cohort, and 0.656 (95% CI: 0.571–0.740) in GSE30219 cohort. Calibration curves were consistent between the actual and the predicted overall survival. The immune infiltration analysis in the TCGA training cohort and two GEO validation cohorts showed a distinctly differentiated immune cell infiltration landscape between the two risk groups. The IRG risk signature for LUAD can be used to predict patient prognosis and guide individual treatment. This risk signature is also a potential biomarker of immunotherapy.

Keywords: lung adenocarcinoma, inflammatory, gene, prognostic signature, risk stratification

INTRODUCTION

Lung cancer is the second most common malignancy and has the highest mortality rate (Sung et al., 2021). Lung adenocarcinoma (LUAD) is its major histological type, accounting for more than 40% of all lung cancer cases (Lortet-Tieulent et al., 2014; Abe and Tanaka, 2016). Advancements in diagnostic and therapeutic strategies, especially those in molecular targeted therapy and immunotherapy, have done much to improve the outcomes of LUAD (Steven et al., 2016; Vargas and Harris, 2016); however, patient survival rates remain low (Zhao et al., 2018). The existing tumor-node-metastasis (TNM) system accounts for only a few factors and neglects heterogeneity in molecular characteristics (Balachandran et al., 2015). Thus, it is necessary to develop new prognostic biomarkers for LUAD to identify heterogeneous patients who are candidates for individual antineoplastic therapy.

Inflammation is the immune system's response to infection, trauma, and other stresses (Gomes et al., 2014). It is also important in cancer, where it plays multiple roles (Hanahan and Weinberg, 2011). On the one hand, the local inflammatory response promotes the occurrence and development of tumors by releasing important molecules and carcinogens to the tumor microenvironment. These can include survival factors that limit apoptosis, growth factors that sustain the proliferative signaling, and pro-angiogenic factors (Ben-Baruch, 2006; Hanahan and Weinberg, 2011). On the other hand, the lymphocyte-to-monocyte ratio and the neutrophil-to-lymphocyte ratio, which are indicators of the systemic inflammatory response, and have been found to be prognostic factors in patients with LUAD (Takahashi et al., 2016; Minami et al., 2018).

Several studies have reported the prognostic value of different inflammatory-related genes (IRGs). For example, He Z et al. found that BTG Anti-Proliferation Factor 2 (BTG2) expression can suppress the proliferation and metastasis of NSCLC cells (He et al., 2015). Recently, the potential prognostic value of IRGs has also been explored in colorectal cancer, low-grade glioma, and oral cavity squamous cell carcinoma (Bai et al., 2019; Liang et al., 2021; Xiang et al., 2021).

The predictive ability of a multi-gene model is superior to that of a single-gene model (Srivastava and Gopal-Srivastava, 2002). In this study, we used data from The Cancer Genome Atlas (TCGA) to isolate an IRG-based signature associated with the survival rates of LUAD patients. Data from the Gene Expression Omnibus (GEO) was then used to validate the predictive ability of this signature. Finally, a nomogram was developed using the aforementioned IRG-based signature to more precisely predict the outcome of LUAD patients.

MATERIALS AND METHODS

Data Collection and Preparation

RNA-seq data and clinical information of patients with LUAD were downloaded from the TCGA (https://tcga-data.nci.nih.gov/tcga/) and GEO databases (https://www.ncbi.nlm.nih.gov/geo/). Cases with incomplete clinical information and follow-ups of less than 5 days were excluded. Finally, 488 cases from the TCGA database were used as the training cohort while 226 cases from the GSE31210, and 85 cases in GSE30219 datasets were used as the validation cohort. The different gene expression datasets were normalized using the "limma" and "SVA" R packages. The Masked Somatic Mutation data (varscan. Somatic. Maf) was analyzed using the "maftools" R package (Mayakonda and Koeffler, 2016). Finally, IRGs were selected and downloaded from hallmark gene sets in the Molecular Signatures Database (http://www.gseamsigdb.org/gsea/msigdb/cards/HALLMARK_INFLAMMATORY_RESPONSE.html). These are shown in **Supplementary Table S1**.

Construction and Validation of the Prognostic IRG Signature

Differentially expressed genes (DEGs) were defined as having false discovery rates (FDR) < 0.05 and log2 |fold change| > 2. Differential expression analysis between LUAD tumor tissues and para-carcinoma tissues was performed using the "limma" R package. In the training cohort, the differentially expressed IRGs were first subjected to univariate Cox regression analysis ($p < 0.05$). Following this, the least absolute shrinkage and selection operator (LASSO) regression analysis was performed to narrow down the prognostically significant candidate IRGs. Then, multivariate Cox regression analysis was used to determine the best weighting coefficient of each prognostically significant candidate IRG. This IRG signature included all the differentially expressed and prognostically significant IRGs. The risk score for each case was calculated according to normalized expression levels of IRGs and their corresponding regression coefficients following the same kind literature (Huang, et al., 2021). The specific formula for this is as follows: Risk score = sum (expression level of each IRG × corresponding coefficients).

In the TCGA training cohort, all patients were divided into a high-risk and low-risk group according to the cut-off value of risk scores derived from maximally selected log-rank statistics, which was used the R package "Maxstat" with the Horton and Lausen (HL) p value approximation method (Hothorn and Zeileis 2008) The Kaplan-Meier method was used to compute the overall survival (OS), and the log-rank test was used to compare OS between the two groups.

To validate the IRG signature, the risk score of LUAD cases in the GSE31210, and GSE30219 datasets were calculated using the same formula as the TCGA cohort. Cases in the validation set were also divided into two groups according to the cut-off point of risk score obtained from the maximally selected log-rank statistics. Survival curves of the low- and the high-risk groups in the validation cohort were also estimated using the Kaplan-Meier method and were compared via the log-rank test.

Gene Set Enrichment Analyses

To investigate the potential molecular mechanisms of the signature IRGs that were identified, Gene Set Enrichment Analysis (GSEA) was used to find enriched terms between the high-risk and low-risk patients in the TCGA and GEO cohort. After excluded gene with no expression value, all mRNA was

FIGURE 1 | Flow chart of data collection and analysis.
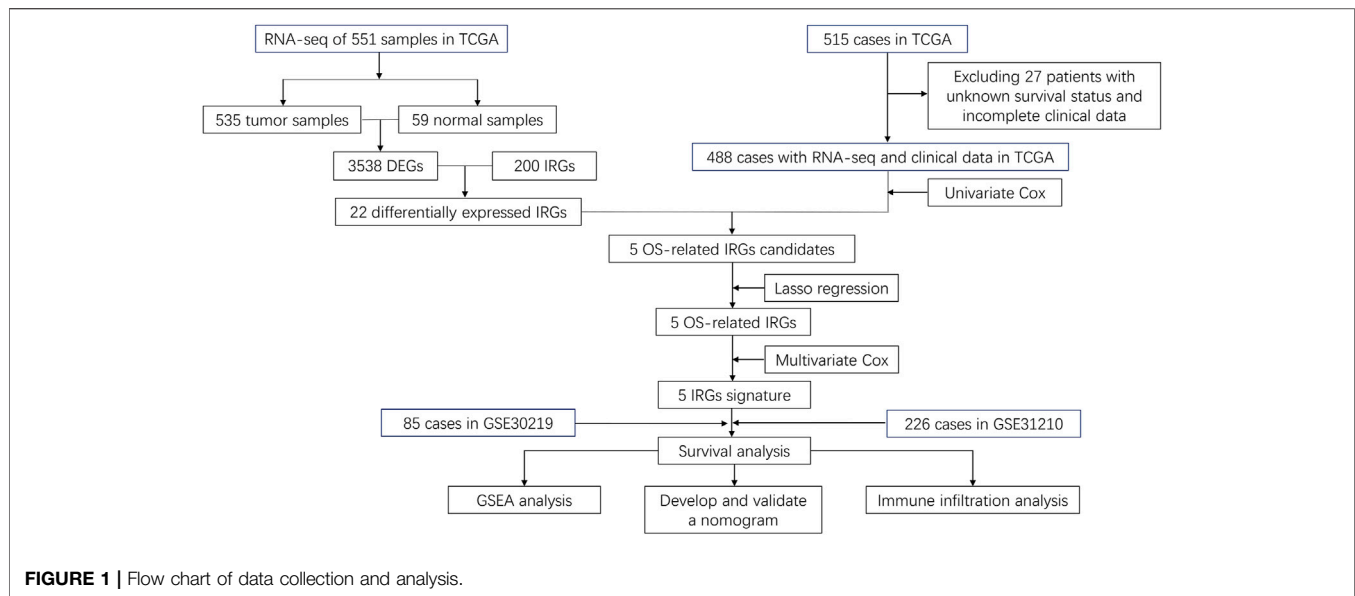
TABLE 1 | Patients' characteristics.

	TCGA training cohort n = 488	GSE31210 cohort n = 226	GSE30219 cohort n = 85
Gender			
Male	223 (45.7)	105 (46.5)	66 (77.6)
Female	265 (54.3)	121 (53.5)	19 (22.4)
Age (year)	65.2 ± 10.0	59.6 ± 7.4	61.5 ± 9.3
Smoking history			
No	68 (13.9)	114 (50.4)	0 (0)
Yes or ever	406 (83.2)	112 (49.6)	85 (100)
Unknown	14 (2.9)	0 (0)	0 (0)
Stage			
I	262 (53.7)	168 (74.3)	81 (95.3)
II	121 (24.8)	58 (25.7)	3 (3.5)
III	80 (16.4)	0 (0)	0 (1.2)
IV	25 (5.1)	0 (0)	0 (0)

entered in GSEA which performed in Java GSEA v. 4.0.1 with the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway in C2, and Gene Ontology (GO) terms in C5. After performing 1,000 permutations, genes with a false discovery rate q < 0.05 were deemed significantly enriched.

Immune Cell Infiltration and Tumor Mutation Burden Analyses

After normalizing the expression data in the TCGA and GEO dataset, we used a single sample GSEA (ssGSEA) to evaluate 28 immune cells using the R package "GSVA" (Tamborero et al., 2018). The results were visualized in a heatmap. With the perm set to 1,000, the CIBERSORT software package was used to perform the CIBERSORT algorithm and evaluate the proportion of 22 types of infiltrating immune cells based on LM22 (Newman et al., 2015). In TCGA cohort, according to the length of the human exon, the TMB calculated for each patient was calculated as the total mutation frequency/35 MB. Dividing the total number of mutations by the size of the coding region of the target is the resulting TMB per megabase. The Mann-Whitney U test was performed to compare the differential expression levels of PDCD1, CD274, PDCD1LG2, CTLA4, CD276, CD80, CD86, VTCN1, and the TMB between the two risk groups.

Establishment of a Predictive Nomogram

Using the TCGA training set, a nomogram integrating the IRG signature and stage to predict individual survival was established. In addition, calibration curves and the area under the curve (AUC) for the OS probability at 1, 3, 5 years were plotted to evaluate the predictive accuracy of this nomogram in the TCGA set and the GEO validation set.

Statistical Analyses

Continuous data are shown as the mean ± SD and were compared using Student's t-test. Categorical variables were analyzed using

FIGURE 2 | Identification of a prognosis-related IRG-based signature in the TCGA training cohort. **(A)** Selection of the optimal candidate genes in the LASSO model. **(B)** LASSO coefficients of prognosis-associated ARGs, each curve represents a gene.
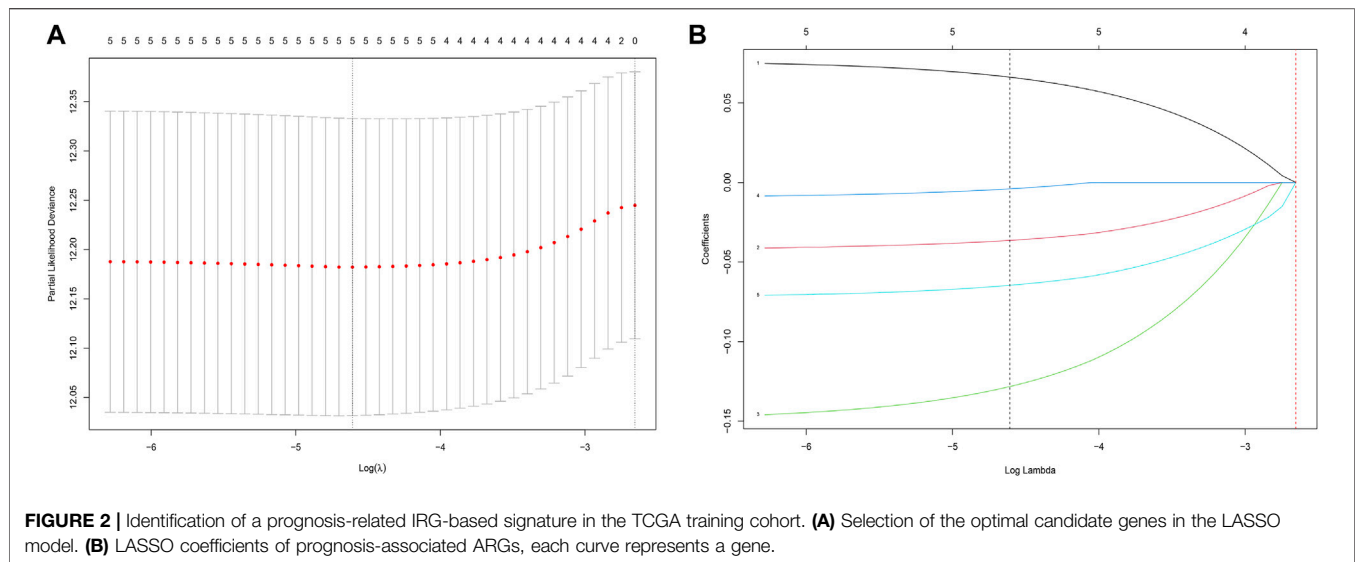
the chi-square (χ^2) test. Cox regression analyses were performed to determine the independent prognostic factors for OS. A prognostic nomogram model was established using the "rms" R package, while its predictive accuracy was assessed via the creation of calibration curves. Statistical analysis was performed using SPSS (version 22.0) and R software (version 4.0.1). The threshold of statistical significance was set at a p-value < 0.05.

RESULTS

Identification of a Prognosis-Related IRG Signature

After differential expression analysis in the TCGA dataset between 535 tumor tissues and 59 normal tissues, we discovered 2,849 upregulated and 689 downregulated DEGs, including 22 differentially expressed IRGs. After excluding 27 cases with unsatisfied follow-up or those lacking important clinical information, 488 cases from the TCGA training set were included to identify prognosis-related IRGs and to construct an IRG-based signature. In addition, 226 cases from the GSE31210, and 85 cases in GSE30219 datasets were used to validate the IRG-based signature (**Figure 1**). The clinicopathological factors of the three datasets are shown in **Table 1**.

After univariate Cox analysis using the gene expression profiles of each differentially expressed IRG, five OS-related IRGs from the TCGA training cohort were found (**Supplementary Figure S1**). These five IRGs were also found to be significant in the LASSO regression analysis (**Figures 2A,B**), which became the IRG signature candidates. Multivariate Cox regression analysis determined the corresponding regression coefficients of each candidate in this ARG risk signature. Finally, a five-IRG risk signature was constructed according to the 488 LUAD cases in the TCGA cohort, whose risk scores were specifically calculated based on a linear combination of gene expression levels and their corresponding regression coefficients. The specific formula for this is as follows: Risk score = EREG ×

0.076861—GPC3 × 0.042023—IL7R × 0.150135—LAMP3 × 0.009471—NMUR1 × 0.072508.

Prognostic Value of the IRG Signature in the Training Cohort

In the TCGA training cohort, the cut-off value of risk scores was determined as -2.0 using the maximally selected log-rank statistics (**Supplementary Figure S2**) after dividing the cases into low-risk and high-risk groups. **Figure 3A** shows the distribution of the risk scores. As shown in **Figure 3B**, the high-risk group has significantly more deaths than the high-risk group. A heatmap showing the differential expression of these 5 IRGs between the low-risk and high-risk groups is depicted in **Figure 3C**. The adjusted FDR and log2 |fold change| of these five IRGs were shown in **Supplementary Table S2**. Additionally, patients in the low-risk group have significantly longer OS time compared with the high-risk group ($p = 1.697e{-}05$) (**Figure 3D**).

Prognostic Value of the IRG Signature in the Validation Cohort

According to the risk score based on the maximally selected log-rank statistics, 226 cases were divided into the high- ($N = 186$) and low-risk ($N = 40$) groups in the GSE31210 validation cohort. The distribution of risk scores is presented in **Figure 4A**. Similar to the training cohort, more patients died in the high-risk group compared with the low-risk group (**Figure 4B**). As shown in **Figure 4C**, the expression profiles of the five IRGs between the low-risk and high-risk groups were plotted in the heatmap and the adjusted FDR and log2 | fold change | of these five IRGs were shown in **Supplementary Table S2**. The survival curves showed that patients in the low-risk group had a better OS than patients in the high-risk group ($p = 1.359e{-}02$) (**Figure 4D**).
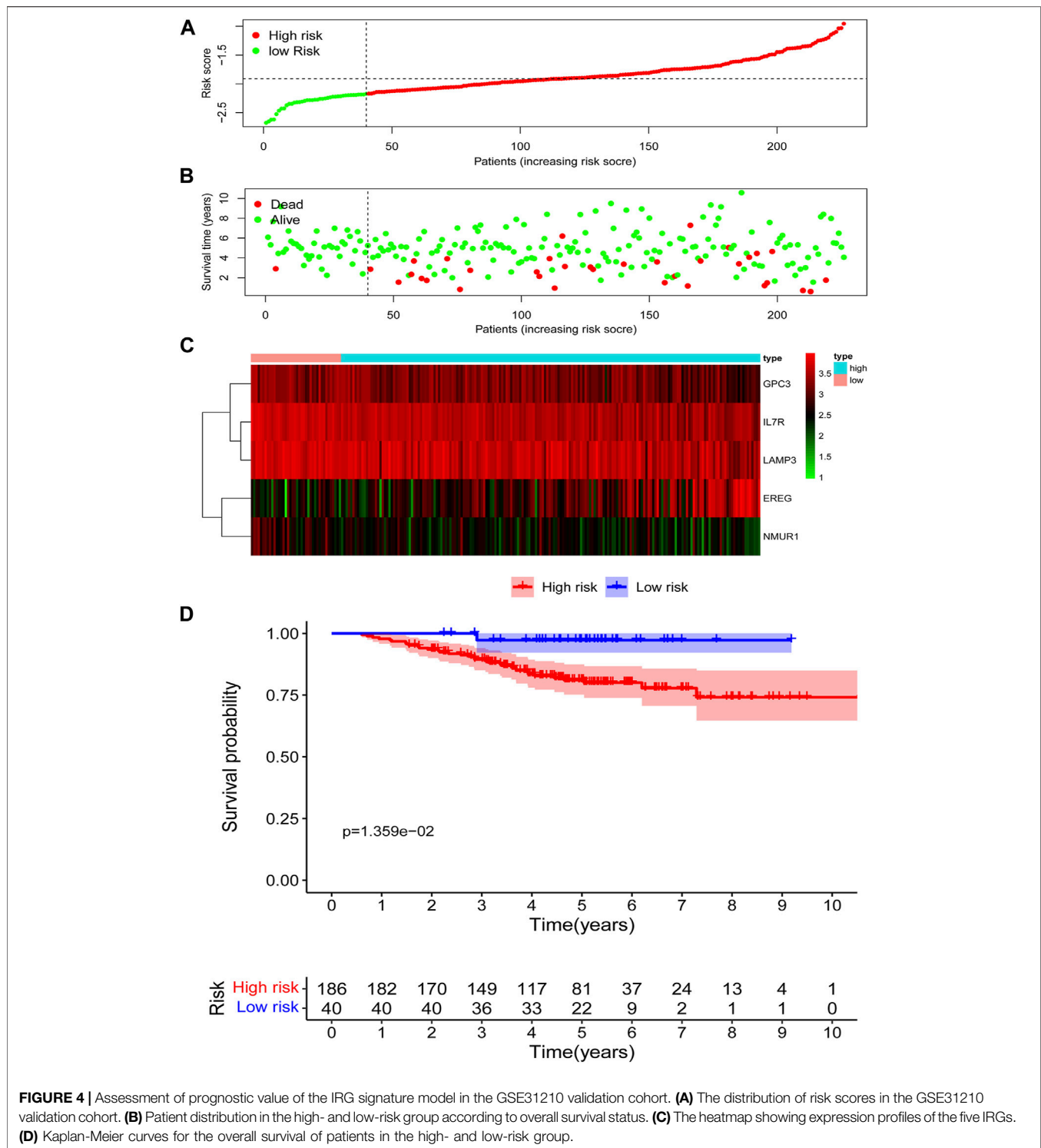
The maximally selected log-rank statistics divided 85 cases in GSE30219 validation cohort into high- ($N = 36$) and low-risk
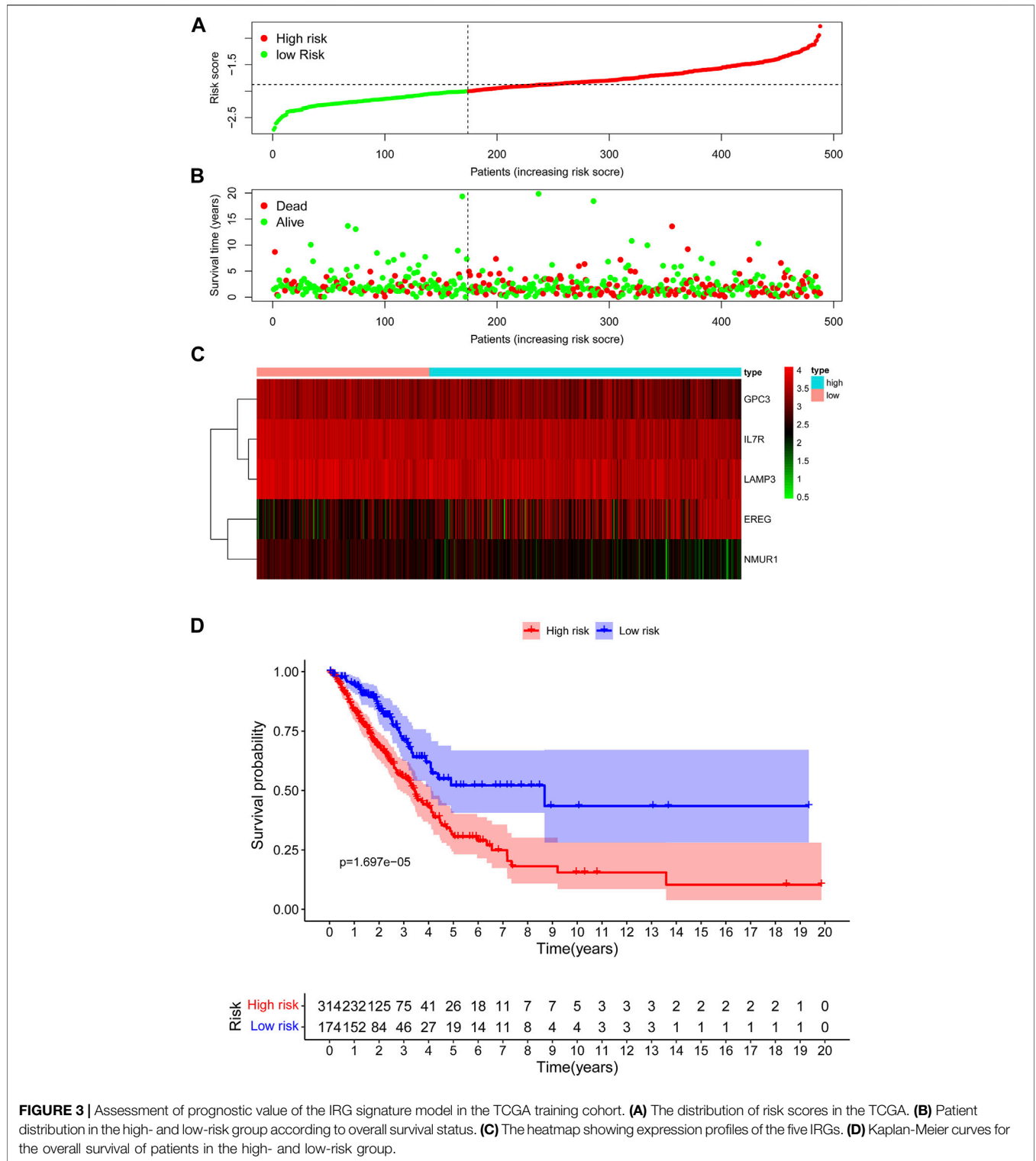
FIGURE 3 | Assessment of prognostic value of the IRG signature model in the TCGA training cohort. **(A)** The distribution of risk scores in the TCGA. **(B)** Patient distribution in the high- and low-risk group according to overall survival status. **(C)** The heatmap showing expression profiles of the five IRGs. **(D)** Kaplan-Meier curves for the overall survival of patients in the high- and low-risk group.

(N = 49) groups. The distribution of risk scores is presented in **Figure 5A**. Similar to the training cohort, more patients died in the high-risk group compared with the low-risk group (**Figure 5B**). As shown in **Figure 5C**, the expression profiles of the five IRGs between the low-risk and high-risk groups were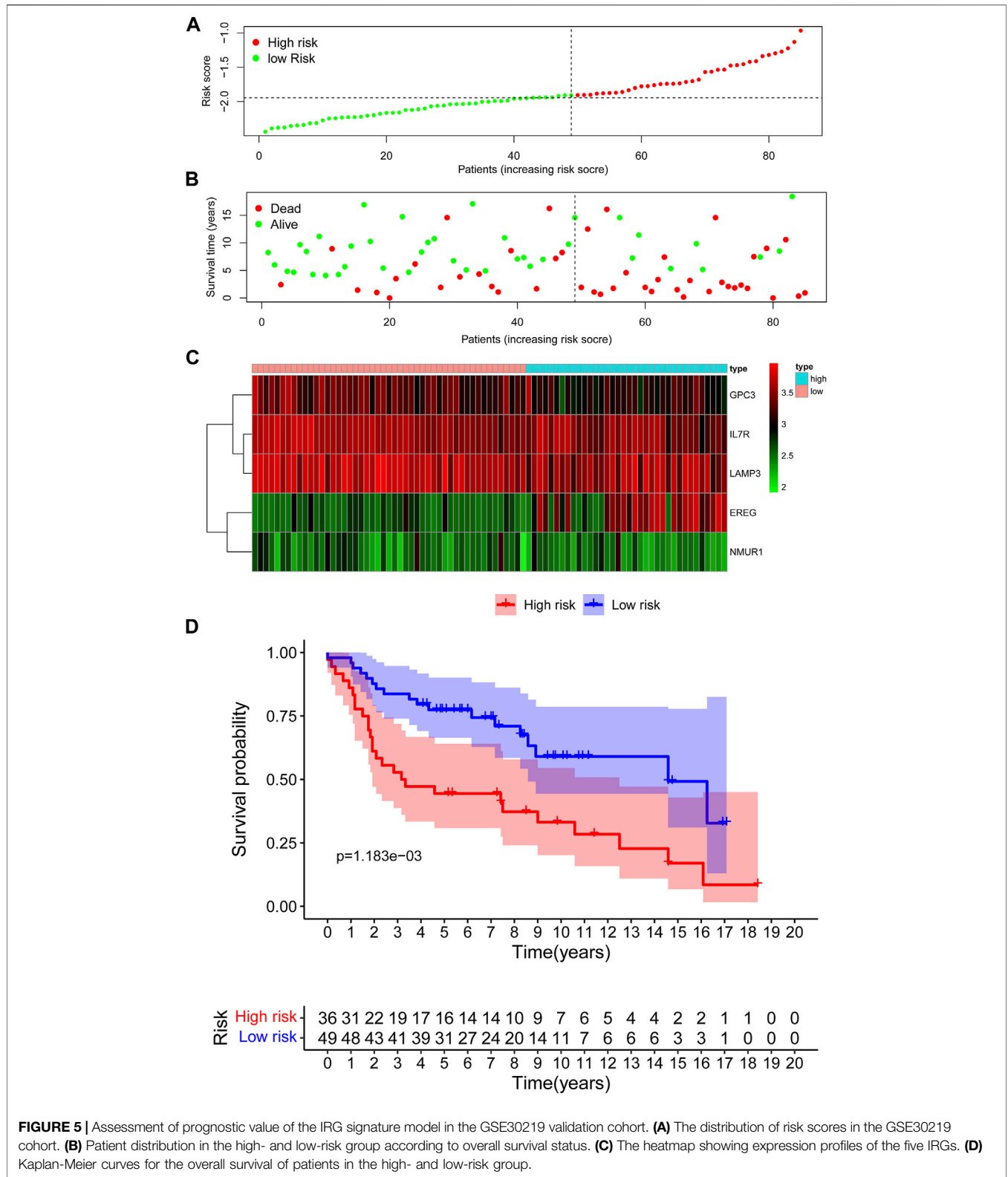 plotted in the heatmap and the adjusted FDR and log2 | fold change | of these five IRGs were shown in **Supplementary Table S2**. The survival curves showed that patients in the low-risk group had a better OS than patients in the high-risk group (p = 1.183e−03) (**Figure 5D**).

FIGURE 4 | Assessment of prognostic value of the IRG signature model in the GSE31210 validation cohort. **(A)** The distribution of risk scores in the GSE31210 validation cohort. **(B)** Patient distribution in the high- and low-risk group according to overall survival status. **(C)** The heatmap showing expression profiles of the five IRGs. **(D)** Kaplan-Meier curves for the overall survival of patients in the high- and low-risk group.

Gene Set Enrichment Analysis for Important Pathways

To investigate the underlying functional mechanisms associated with these five IRGs in patients with LUAD, GSEA through GO and KEGG pathway enrichment analysis was performed between the high-risk and low-risk groups in the TCGA training set and two GEO validation set. In the high-risk group, KEGG enrichment analysis found that genes were primarily enriched in base excision repair, nucleotide excision repair, oxidative phosphorylation, pyrimidine metabolism, and RNA degradation in all three datasets. GO enrichment analysis

FIGURE 5 | Assessment of prognostic value of the IRG signature model in the GSE30219 validation cohort. (A) The distribution of risk scores in the GSE30219 cohort. (B) Patient distribution in the high- and low-risk group according to overall survival status. (C) The heatmap showing expression profiles of the five IRGs. (D) Kaplan-Meier curves for the overall survival of patients in the high- and low-risk group.

found that genes were primarily enriched in metabolism and function of ncRNA, such as ncRNA metabolic process, ncRNA processing and ncRNA transcription; metabolism and function of scRNA, such as snRNA metabolic process and snRNA binding. In patients with low-risk, KEGG enrichment analysis found that genes were primarily enriched in B cell receptor signaling

FIGURE 6 | Gene set enrichment analysis between the low- and high-risk subgroups in TCGA training cohort and two GEO validation cohorts. **(A)** Enriched KEGG terms between high- and low-risk groups in TCGA cohort. **(B)** Enriched GO pathways between high- and low-risk groups in TCGA cohort. **(C)** Enriched KEGG terms between high- and low-risk groups in GSE31210 cohort. **(D)** Enriched GO pathways between high- and low-risk groups in GSE31210 cohort. **(E)** Enriched KEGG terms between high- and low-risk groups in GSE30219 cohort. **(F)** Enriched GO pathways between high- and low-risk groups in GSE30219 cohort.
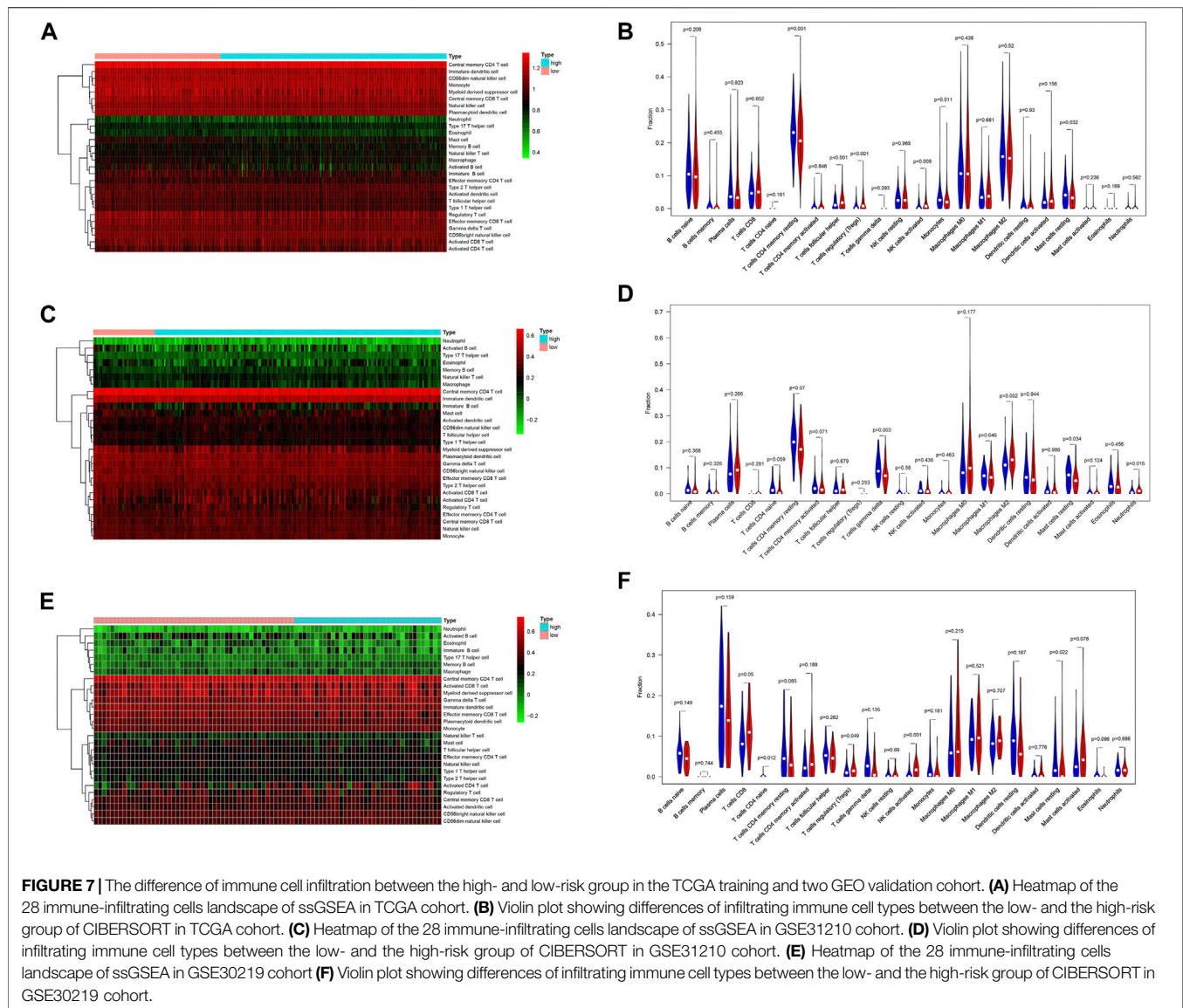
pathway, cell adhesion molecules cams, T cell receptor signaling pathway, and WNT signaling pathway. GO enrichment analysis found that genes were primarily enriched in metabolism and function of calcium ion, such as calcium-mediated signaling, calcium channel complex, calcium dependent protein kinase activity; T cell selection; cargo receptor activity; and scavenger receptor activity (**Figures 6A–F**).

Tumor Immunity Landscape and TMB in LUAD

To explore the relation between the IRG risk signature and the tumor immunity landscape, the CIBERSORT algorithm and ssGSEA were utilized to evaluate immunity infiltration between low- and high-risk groups in the TCGA training set and 2 GEO validation set. The heatmaps show the immune cell

infiltration landscape of the 28 immune cells in three cohorts, as obtained from the results of the ssGSEA analysis (**Figures 7A,C,E**). The barplot show the immune cell infiltration landscape of the 22 immune cells as obtained from the results of the CIBERSORT algorithm (**Supplementary Figure S3A,C,E**). As shown in **Figures 7B,D,F**, patients in low-risk had significantly higher proportions of infiltrating resting mast cells in all three cohorts. Patients in low-risk also had higher proportions of infiltrating resting CD4$^+$ memory T cells and monocytes in TCGA cohort, and gamma delta T cell in GSE31210 cohorts. Patients in high-risk group had higher proportions of infiltrating of Treg cells, activated NK cells in TCGA and GSE30219 cohort, while those of follicular helper T cells were significantly higher in TCGA cohort. High-risk patients in GSE31210 had higher proportions infiltrating of M2 macrophage and neutrophils. The proportions of infiltrating

FIGURE 7 | The difference of immune cell infiltration between the high- and low-risk group in the TCGA training and two GEO validation cohort. (A) Heatmap of the 28 immune-infiltrating cells landscape of ssGSEA in TCGA cohort. (B) Violin plot showing differences of infiltrating immune cell types between the low- and the high-risk group of CIBERSORT in TCGA cohort. (C) Heatmap of the 28 immune-infiltrating cells landscape of ssGSEA in GSE31210 cohort. (D) Violin plot showing differences of infiltrating immune cell types between the low- and the high-risk group of CIBERSORT in GSE31210 cohort. (E) Heatmap of the 28 immune-infiltrating cells landscape of ssGSEA in GSE30219 cohort (F) Violin plot showing differences of infiltrating immune cell types between the low- and the high-risk group of CIBERSORT in GSE30219 cohort.

28 immune cells in ssGSEA were shown in **Supplementary Figure S3B,D,F**. The proportions of infiltrating eosinophils and mast cells were higher in low-risk group among three cohort. In the cohort with large samples, TCGA and GSE31210, low-risk group had higher proportions infiltrating of activated CD8 T cells, effector memory CD8 T cell, central memory CD4 T cell, type 1 helper T cell and activated B cell, which play an important role in anti-tumor immunity. Somatic mutation analysis revealed that patients in the low-risk group had a lower TMB. (**Supplementary Figure S4**). We also investigated the expression levels of eight immune checkpoint genes between the low- and high-risk groups. There was no difference in the expression level of CD274, CD276, and VTCN1 (**Figure 8**) between the two risk groups. However, patients in the low-risk group had a higher expression level of PDCD1, PDCD1LG2, CD80, CD86, and CTLA4 (**Figure 8**).

Nomogram Based on the IRG Signature for LUAD

After adjusting for gender, age and smoking history, multivariate Cox analysis demonstrated that the stage (HR = 1.60, 95%CI = 1.39–1.85, and $p < 0.001$) and risk score was a negative prognostic factor of OS in the training cohort (HR = 2.39, 95%CI = 1.49–3.83, and $p < 0.001$) (**Figure 9A**). Based on the IRG risk scores and stage, a visualized predictive nomogram for individual OS probability at 1, 3, and 5 years was developed using the data of the training cohort (**Figure 9B**). Bootstrap validation was performed in this nomogram. The C-index of the training cohort was 0.687 (95% CI: 0.644–0.730), while the C-index of the GSE31210 and GSE30219 cohort was 0.678 (95% CI: 0.586–0.771), and 0.656 (95% CI: 0.571–0.740), respectively, which suggested its good performance in predicting OS for LUAD. Calibration curves were drawn in the training and two
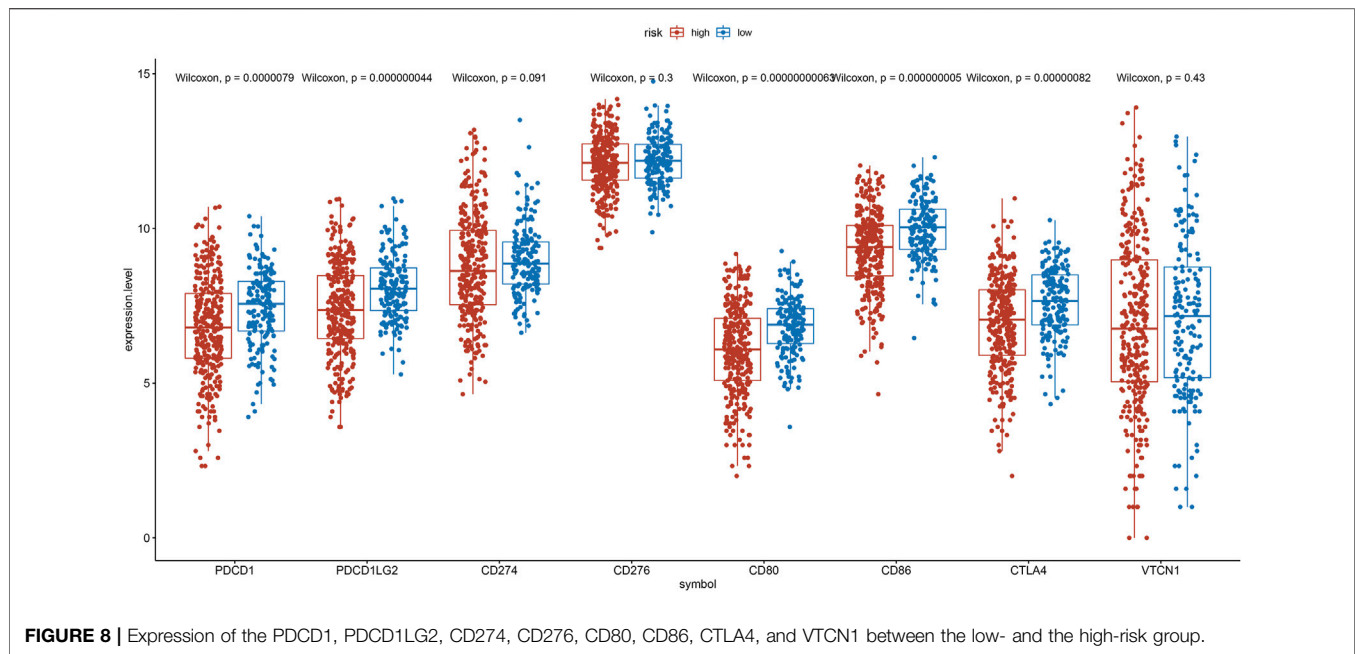
FIGURE 8 | Expression of the PDCD1, PDCD1LG2, CD274, CD276, CD80, CD86, CTLA4, and VTCN1 between the low- and the high-risk group.

GEO validation cohorts (**Figures 9C,E,G**) to verify the accuracy of the nomogram. At 1, 3, and 5 years, the calibration curves demonstrated a satisfactory consistency between the actual observations and the predicted survival probabilities. The time ROC curve in 1, 3, and 5 years also showed the good performance of this nomogram. The AUC values of the 1-, 3-, and 5-years OS in the TCGA cohort were 0.745, 0.690, and 0.697, respectively (**Figure 9D**). The AUC values of the 1-, 3-, and 5-years OS in the GSE31210 cohort were 0.901, 0.712, and 0.665, respectively (**Figure 9F**). The AUC values of the 1-, 3-, and 5-years OS in the GSE30219 cohort were 0.771, 0.685, and 0.637, respectively (**Figure 9H**).
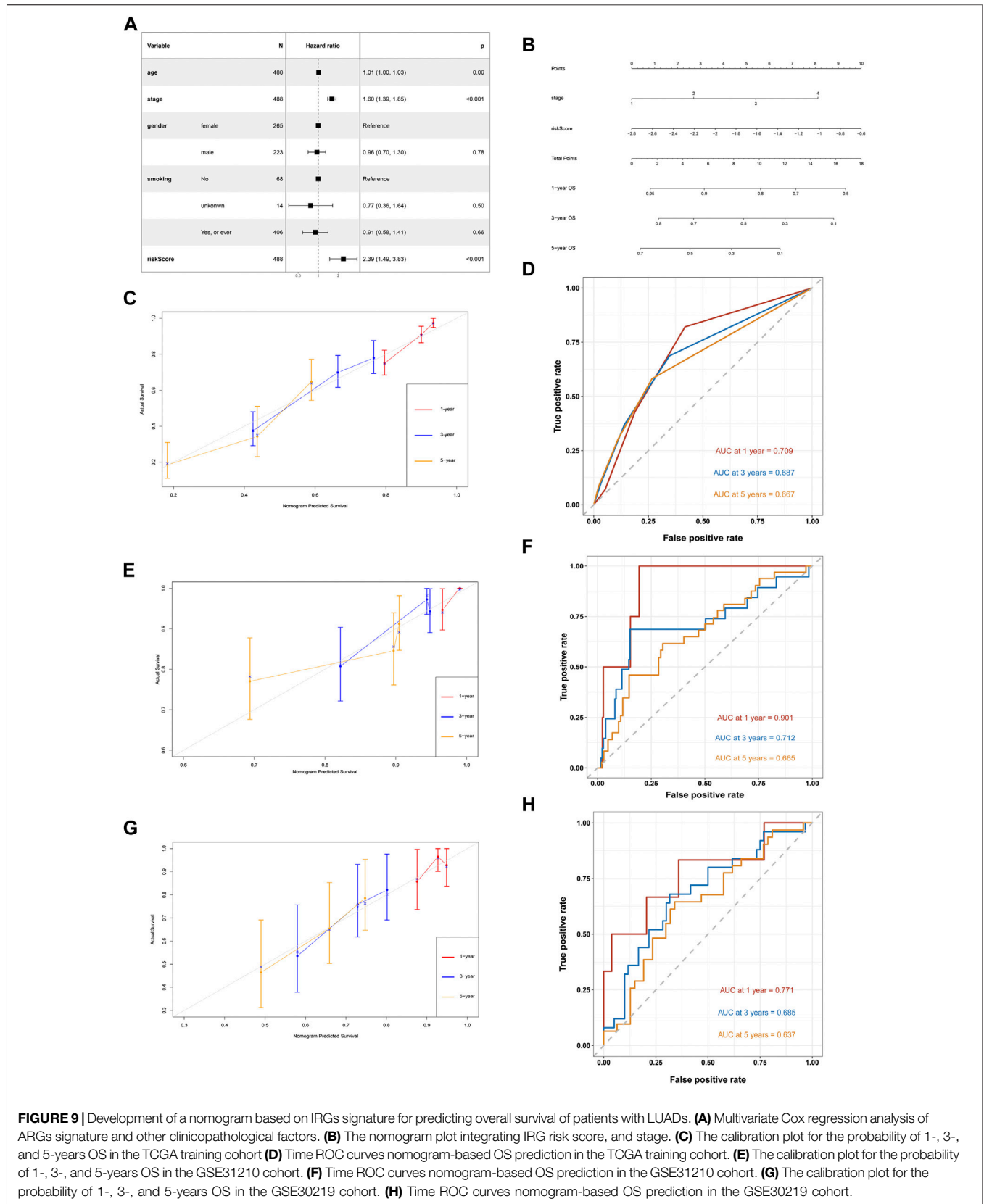
DISCUSSION

In this study, we explored the relationship between the expression level of IRGs and survival from LUAD. We constructed a novel prognostic ARG signature consisting of five IRGs, including EREG, GPC3, ILR7, LAMP3, and NMUR1. In addition, the multivariate Cox analysis confirmed the prognostic value of the IRG signature. Furthermore, we established a nomogram integrating the IRG signature and stage for predicting individual survival and validated its predictive ability in the GSE31210 and GSE30219 datasets. Finally, we also explored the relationship between expression levels of IRGs and immune cell infiltration in LUAD.

With increasing convenience in data collection from open-access public databases, many studies concentrated on the relationship between RNA-seq data of specific gene sets and individual patient outcomes (Wu et al., 2019; Qu et al., 2020; Zhu et al., 2020; Xu and Chen, 2021). These studies were limited to autophagy, aging, and immune infiltration, but most of them lack clinical applications. In addition, the study focused on the

prognostic role of the IRGs in LUAD is lacking. Increasing evidence demonstrated the underlying mechanism of the local inflammatory microenvironment and systemic inflammatory response driving tumorigenesis in many cancers, including lung cancer (Gomes et al., 2014; Lucas et al., 2017; Greten and Grivennikov, 2019; Oya et al., 2020). Several previous studies reported the prognostic value of a single inflammatory-related gene, such as BTG2, TNFRSF10B, and IL1B (Landvik et al., 2012; Schabath et al., 2013; He et al., 2015). However, a comprehensive model with multiple genes and a stronger predictive ability is also necessary. In this study, our nomogram showed good prediction performance in both the training and validation sets.

The IRG risk score formula used in this study indicated that a high gene expression of epiregulin (EREG) was unfavorable for individual survival. EREG is a ligand that belongs to the ERBB family. EREG can bind to the ERBB1 and ERBB4 receptors, activating their intrinsic kinase domain and leading to the phosphorylation of specific tyrosine residues in their receptor's cytoplasmic tail of their receptors (Shelly et al., 1998). The overexpression of EREG is found in LUAD and associated with unfavorable prognoses (Sunaga and Kaira, 2015). Phosphatidylinositol proteoglycan-3 (GPC3) is an extracellular glycoprotein belonging to the heparan sulfate proteoglycan (HSPG) family (Gonzalez et al., 1998). Interestingly, the *GPC3* gene regulates cell proliferation as a tumor suppressor gene in LUAD (Kim et al., 2003) but shows the opposite effect in LUSC (Lin et al., 2012). Interleukin-7 receptor (IL7R) is the receptor of IL-7, and the IL-7/IL-7R interaction seems to have a two-sided effect in lung cancer. On the one hand, IL-7/IL-7R suppresses autophagy by activating the PI3K/Akt/mTOR signaling pathway and promoting lymphangiogenesis (Ming et al., 2009; Jian et al., 2019). On the other hand, IL-7/IL-7R improves chemotherapeutic sensitivity and anti-tumor immunity (Shi et al., 2019; Bednarz-Misa et al., 2021). Lysosome-associated

FIGURE 9 | Development of a nomogram based on IRGs signature for predicting overall survival of patients with LUADs. (A) Multivariate Cox regression analysis of ARGs signature and other clinicopathological factors. (B) The nomogram plot integrating IRG risk score, and stage. (C) The calibration plot for the probability of 1-, 3-, and 5-years OS in the TCGA training cohort (D) Time ROC curves nomogram-based OS prediction in the TCGA training cohort. (E) The calibration plot for the probability of 1-, 3-, and 5-years OS in the GSE31210 cohort. (F) Time ROC curves nomogram-based OS prediction in the GSE31210 cohort. (G) The calibration plot for the probability of 1-, 3-, and 5-years OS in the GSE30219 cohort. (H) Time ROC curves nomogram-based OS prediction in the GSE30219 cohort.

membrane protein 3 (LAMP3) belongs to the LAMP family of proteins, influencing cellular processes such as phagocytosis, aging, and lipid transport. The role of LAMP3 in cancer remains to be elucidated (Alessandrini et al., 2017). Neuromedin U receptor 1 (NMUR1) is one of the receptors of Neuromedin U. A previous study reported that the methylation of the NMUR1 gene was related to poor survival in patients with head and neck squamous cell carcinoma (Misawa et al., 2020). The potential mechanisms and function of NMUR1 in LUAD had not been reported.

Tumor inflammation is closely associated with immune cell infiltration in the tumor microenvironment, which contributes to immunotherapy response. However, there is no study exploring the relationship between tumor inflammation and immune cell infiltration in LUAD. The ssGSEA and CIBERSORT algorithms were performed to compare tumor infiltration between the two risk groups in this study. We found that patients in high-risk group had higher proportions of immune cells, which against the anti-tumor immunity, such asregulatory T cells, follicular helper T cells, M2 macrophage, and neutrophils. Previous studies had reported that high proportions of these immune cells promote progression and metastasis in NSCLC (Shi et al., 2014; Marshall et al., 2016; Powell and Huttenlocher, 2016; Pritchard et al., 2020). In patients with low-risk group, a higher proportions infiltrating of resting CD4 memory T cells, activated CD8 T cells, effector memory CD8 T cell, central memory CD4 T cell, type 1 helper T cell, activated B cell and resting mast cells were found, and which contribute to the anti-tumor immunity and are positively associated with prognosis (Chamoto et al., 2006; Lange et al., 2019; Han et al., 2020; Xu et al., 2020). Besides, we investigated the expression level of 8 immune checkpoint genes. The proteins of these 8 genes were costimulatory molecules belong to B7-CD28 family members, which closely associated with anti-tumor immunity (Zhang et al., 2020). Previous study mentioned that patients with high expression level of these 8 genes might had better response of immunotherapy (Li et al., 2019). We found cases in the low-risk group had a higher expression level of PDCD1 (PD-1), PDCD1LG2 (PD-L2), CTLA4, CD80 (B7.1), and CD86 (B7.2), which indicate that patients with a low-risk score may benefit more from immunotherapy.

Some limitations should be considered in this study. First, this IRG prognostic model was established through bioinformatics analyses from data available in the TCGA and GEO databases. Hence, the results of this study need further validation from prospective, multicenter trials or experimental data. In addition, this study preliminarily investigated the potential relationship between the IRG risk signature and immune cells infiltration, so further studies are needed to reveal the underlying mechanisms for this. Finally, although the IRG signature and the stage were integrated into our prognostic nomogram, we cannot identify the contribution of each IRG in this signature.

In conclusion, to determine if the IRG risk signature we deduced was related to OS in patients with LUAD, we constructed and validated a prognostic nomogram for LUAD, including the use of the IRG risk signature and stage, and for predicting individual survival. Patients with a low IRG risk score have a higher expression levels of immune checkpoint gene and this IRG risk signature can be a potential indicator of immunotherapy.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: TCGA databases (https://tcga-data.nci.nih.gov/tcga/) and the GEO databases (https://www.ncbi.nlm.nih.gov/geo/).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethics committee of Sun Yat-sen University Cancer Center. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. No potentially identifiable human images or data are presented in the manuscript.

AUTHOR CONTRIBUTIONS

Conception and design: HL and Z-RZ; Provision of study materials or patients: W-YZ, F-FD, and SC; Collection and assembly of data: J-YW, Y-BL, Y-ZW, and B-YR; Data analysis and interpretation: W-YZ, F-FD, and SC; Manuscript writing and editing: W-YZ and F-FD; Final approval of manuscript: All authors.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.798131/full#supplementary-material

REFERENCES

Abe, Y., and Tanaka, N. (2016). The Hedgehog Signaling Networks in Lung Cancer: The Mechanisms and Roles in Tumor Progression and Implications for Cancer Therapy. *Biomed. Res. Int.* 2016, 1–11. doi:10.1155/2016/7969286

Alessandrini, F., Pezzè, L., and Ciribilli, Y. (2017). LAMPs: Shedding Light on Cancer Biology. *Semin. Oncol.* 44 (4), 239–253. doi:10.1053/j.seminoncol.2017.10.013

Bai, S., Yan, Y.-B., Chen, W., Zhang, P., Zhang, T.-M., Tian, Y.-Y., et al. (2019). Bioinformatic Analysis Reveals an Immune/Inflammatory-Related Risk Signature for Oral Cavity Squamous Cell Carcinoma. *J. Oncol.* 2019, 1–10. doi:10.1155/2019/3865279

Balachandran, V. P., Gonen, M., Smith, J. J., and DeMatteo, R. P. (2015). Nomograms in Oncology: More Than Meets the Eye. *Lancet Oncol.* 16 (4), e173–e180. doi:10.1016/s1470-2045(14)71116-7

Bednarz-Misa, I., Bromke, M. A., and Krzystek-Korpacka, M. (2021). Interleukin (IL)-7 Signaling in the Tumor Microenvironment. *Adv. Exp. Med. Biol.* 1290, 9–49. doi:10.1007/978-3-030-55617-4_2

Benbaruch, A. (2006). Inflammation-associated Immune Suppression in Cancer: the Roles Played by Cytokines, Chemokines and Additional Mediators. *Semin. Cancer Biol.* 16 (1), 38–52. doi:10.1016/j.semcancer.2005.07.006

Chamoto, K., Wakita, D., Narita, Y., Zhang, Y., Noguchi, D., Ohnishi, H., et al. (2006). An Essential Role of Antigen-Presenting cell/T-Helper Type 1 Cell-Cell Interactions in Draining Lymph Node during Complete Eradication of Class II-Negative Tumor Tissue by T-Helper Type 1 Cell Therapy. *Cancer Res.* 66 (3), 1809–1817. doi:10.1158/0008-5472.can-05-2246

Gomes, M., Teixeira, A. L., Coelho, A., Araújo, A., and Medeiros, R. (2014). The Role of Inflammation in Lung Cancer. *Adv. Exp. Med. Biol.* 816, 1–23. doi:10.1007/978-3-0348-0837-8_1

Gonzalez, A. D., Kaya, M., Shi, W., Song, H., Testa, J. R., Penn, L. Z., et al. (1998). OCI-5/GPC3, a Glypican Encoded by a Gene that Is Mutated in the Simpson-Golabi-Behmel Overgrowth Syndrome, Induces Apoptosis in a Cell Line-specific Manner. *J. Cel Biol* 141 (6), 1407–1414. doi:10.1083/jcb.141.6.1407

Greten, F. R., and Grivennikov, S. I. (2019). Inflammation and Cancer: Triggers, Mechanisms, and Consequences. *Immunity* 51 (1), 27–41. doi:10.1016/j.immuni.2019.06.025

Han, J., Khatwani, N., Searles, T. G., Turk, M. J., and Angeles, C. V. (2020). Memory CD8+ T Cell Responses to Cancer. *Semin. Immunol.* 49, 101435. doi:10.1016/j.smim.2020.101435

Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of Cancer: the Next Generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013

He, Z., Liu, Y., Xiao, B., and Qian, X. (2015). miR-25 Modulates NSCLC Cell Radio-Sensitivity through Directly Inhibiting BTG2 Expression. *Biochem. Biophysical Res. Commun.* 457 (3), 235–241. doi:10.1016/j.bbrc.2014.12.094

Hothorn, T., and Zeileis, A. (2008). Generalized Maximally Selected Statistics. *Biometrics* 64 (4), 1263–1269. doi:10.1111/j.1541-0420.2008.00995.x

Huang, H.-y., Wang, Y., Wang, W.-d., Wei, X.-l., Gale, R. P., Li, J.-y., et al. (2021). A Prognostic Survival Model Based on Metabolism-Related Gene Expression in Plasma Cell Myeloma. *Leukemia* 35 (11), 3212–3222. doi:10.1038/s41375-021-01206-4

Jian, M., Yunjia, Z., Zhiying, D., Yanduo, J., and Guocheng, J. (2019). Interleukin 7 Receptor Activates PI3K/Akt/mTOR Signaling Pathway via Downregulation of Beclin-1 in Lung Cancer. *Mol. Carcinogenesis* 58 (3), 358–365. doi:10.1002/mc.22933

Kim, H., Xu, G.-L., Borczuk, A. C., Busch, S., Filmus, J., Capurro, M., et al. (2003). The Heparan Sulfate ProteoglycanGPC3Is a Potential Lung Tumor Suppressor. *Am. J. Respir. Cel Mol Biol* 29 (6), 694–701. doi:10.1165/rcmb.2003-0061OC

Landvik, N. E., Tekpli, X., Anmarkrud, K. H., Haugen, A., and Zienolddiny, S. (2012). Molecular Characterization of a Cancer-Related Single Nucleotide Polymorphism in the Pro-inflammatory interleukin-1B Gene. *Mol. Carcinog.* 51 (Suppl. 1), E168–E175. doi:10.1002/mc.21910

Lange, T., Born, J., and Westermann, J. (2019). Sleep Matters: CD4+ T Cell Memory Formation and the Central Nervous System. *Trends Immunol.* 40 (8), 674–686. doi:10.1016/j.it.2019.06.003

Li, W., Wang, H., Ma, Z., Zhang, J., Ou-Yang, W., Qi, Y., et al. (2019). Multi-omics Analysis of Microenvironment Characteristics and Immune Escape Mechanisms of Hepatocellular Carcinoma. *Front. Oncol.* 9, 1019. doi:10.3389/fonc.2019.01019

Liang, Y., Wu, X., Su, Q., Liu, Y., and Xiao, H. (2021). Identification and Validation of a Novel Inflammatory Response-Related Gene Signature for the Prognosis of Colon Cancer. *J.ir* 14, 3809–3821. doi:10.2147/jir.s321852

Lin, Q., Xiong, L.-w., Pan, X.-f., Gen, J.-f., Bao, G.-l., Sha, H.-f., et al. (2012). Expression of GPC3 Protein and its Significance in Lung Squamous Cell Carcinoma. *Med. Oncol.* 29 (2), 663–669. doi:10.1007/s12032-011-9973-1

Lortet-Tieulent, J., Soerjomataram, I., Ferlay, J., Rutherford, M., Weiderpass, E., and Bray, F. (2014). International Trends in Lung Cancer Incidence by Histological Subtype: Adenocarcinoma Stabilizing in Men but Still Increasing in Women. *Lung Cancer* 84 (1), 13–22. doi:10.1016/j.lungcan.2014.01.009

Lucas, C., Barnich, N., and Nguyen, H. (2017). Microbiota, Inflammation and Colorectal Cancer. *Ijms* 18 (6), 1310. doi:10.3390/ijms18061310

Marshall, E. A., Ng, K. W., Kung, S. H. Y., Conway, E. M., Martinez, V. D., Halvorsen, E. C., et al. (2016). Emerging Roles of T Helper 17 and Regulatory T Cells in Lung Cancer Progression and Metastasis. *Mol. Cancer* 15 (1), 67. doi:10.1186/s12943-016-0551-1

Mayakonda, A., and Koeffler, H. P. (2016). Maftools: Efficient Analysis, Visualization and Summarization of MAF Files from Large-Scale Cohort Based Cancer Studies. bioRxiv. doi:10.1101/052662

Minami, S., Ihara, S., Kim, S.-H., Yamamoto, S., and Komuta, K. (2018). Lymphocyte to Monocyte Ratio and Modified Glasgow Prognostic Score Predict Prognosis of Lung Adenocarcinoma without Driver Mutation. *World J. Oncol.* 9 (1), 13–20. doi:10.14740/wjon1084w

Ming, J., Zhang, Q., Qiu, X., and Wang, E. (2009). Interleukin 7/interleukin 7 Receptor Induce C-fos/c-jun-dependent Vascular Endothelial Growth Factor-D Up-Regulation: a Mechanism of Lymphangiogenesis in Lung Cancer. *Eur. J. Cancer* 45 (5), 866–873. doi:10.1016/j.ejca.2008.12.006

Misawa, K., Mima, M., Satoshi, Y., Misawa, Y., Imai, A., Mochizuki, D., et al. (2020). Neuropeptide Receptor Genes GHSR and NMUR1 Are Candidate Epigenetic Biomarkers and Predictors for Surgically Treated Patients with Oropharyngeal Cancer. *Sci. Rep.* 10 (1), 1007. doi:10.1038/s41598-020-57920-z

Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust Enumeration of Cell Subsets from Tissue Expression Profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337

Oya, Y., Hayakawa, Y., and Koike, K. (2020). Tumor Microenvironment in Gastric Cancers. *Cancer Sci.* 111 (8), 2696–2707. doi:10.1111/cas.14521

Powell, D. R., and Huttenlocher, A. (2016). Neutrophils in the Tumor Microenvironment. *Trends Immunol.* 37 (1), 41–52. doi:10.1016/j.it.2015.11.008

Pritchard, A., Tousif, S., Wang, Y., Hough, K., Khan, S., Strenkowski, J., et al. (2020). Lung Tumor Cell-Derived Exosomes Promote M2 Macrophage Polarization. *Cells* 9 (5), 1303. doi:10.3390/cells9051303

Qu, Y., Cheng, B., Shao, N., Jia, Y., Song, Q., Tan, B., et al. (2020). Prognostic Value of Immune-Related Genes in the Tumor Microenvironment of Lung Adenocarcinoma and Lung Squamous Cell Carcinoma. *Aging* 12 (6), 4757–4777. doi:10.18632/aging.102871

Schabath, M. B., Giuliano, A. R., Thompson, Z. J., Amankwah, E. K., Gray, J. E., Fenstermacher, D. A., et al. (2013). TNFRSF10B Polymorphisms and Haplotypes Associated with Increased Risk of Death in Non-small Cell Lung Cancer. *Carcinogenesis* 34 (11), 2525–2530. doi:10.1093/carcin/bgt244

Shelly, M., Pinkas-Kramarski, R., Guarino, B. C., Waterman, H., Wang, L.-M., Lyass, L., et al. (1998). Epiregulin Is a Potent Pan-ErbB Ligand that Preferentially Activates Heterodimeric Receptor Complexes. *J. Biol. Chem.* 273 (17), 10496–10505. doi:10.1074/jbc.273.17.10496

Shi, L., Xu, Z., Yang, Q., Huang, Y., Gong, Y., Wang, F., et al. (2019). IL-7-Mediated IL-7R-JAK3/STAT5 Signalling Pathway Contributes to Chemotherapeutic Sensitivity in Non-small-cell Lung Cancer. *Cell Prolif* 52(6), doi:10.1111/cpr.12699

Shi, L., Li, X., Cha, Z., Sun, S., Wang, L., Jiao, S., et al. (2014). Dysregulation of Circulating Follicular Helper T Cells in Nonsmall Cell Lung Cancer. *DNA Cel Biol.* 33 (6), 355–360. doi:10.1089/dna.2013.2332

Srivastava, S., and Gopal-Srivastava, R. (2002). Biomarkers in Cancer Screening: a Public Health Perspective. *J. Nutr.* 132 (8 Suppl. l), 2471s–2475s. doi:10.1093/jn/132.8.2471S

Steven, A., Fisher, S. A., and Robinson, B. W. (2016). Immunotherapy for Lung Cancer. *Respirology* 21 (5), 821–833. doi:10.1111/resp.12789

Sunaga, N., and Kaira, K. (2015). Epiregulin as a Therapeutic Target in Non-small-cell Lung Cancer. *Lctt* 6, 91–98. doi:10.2147/lctt.s60427

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Takahashi, Y., Kawamura, M., Hato, T., Harada, M., Matsutani, N., and Horio, H. (2016). Neutrophil-Lymphocyte Ratio as a Prognostic Marker for Lung Adenocarcinoma after Complete Resection. *World J. Surg.* 40 (2), 365–372. doi:10.1007/s00268-015-3275-2

Tamborero, D., Rubio-Perez, C., Muiños, F., Sabarinathan, R., Piulats, J. M., Muntasell, A., et al. (2018). A Pan-Cancer Landscape of Interactions between Solid Tumors and Infiltrating Immune Cell Populations. *Clin. Cancer Res.* 24 (15), 3717–3728. doi:10.1158/1078-0432.ccr-17-3509

Vargas, A. J., and Harris, C. C. (2016). Biomarker Development in the Precision Medicine Era: Lung Cancer as a Case Study. *Nat. Rev. Cancer* 16 (8), 525–537. doi:10.1038/nrc.2016.56

Wu, P., Wang, Y., Wu, Y., Jia, Z., Song, Y., and Liang, N. (2019). Expression and Prognostic Analyses of ITGA11, ITGB4 and ITGB8 in Human Non-small Cell Lung Cancer. *PeerJ* 7, e8299. doi:10.7717/peerj.8299

Xiang, Z., Chen, X., Lv, Q., and Peng, X. (2021). A Novel Inflammatory lncRNAs Prognostic Signature for Predicting the Prognosis of Low-Grade Glioma Patients. *Front. Genet.* 12, 697819. doi:10.3389/fgene.2021.697819

Xu, J.-Z., Gong, C., Xie, Z.-F., and Zhao, H. (2020). Development of an Oncogenic Driver Alteration Associated Immune-Related Prognostic Model for Stage I-II Lung Adenocarcinoma. *Front. Oncol.* 10, 593022. doi:10.3389/fonc.2020.593022

Xu, Q., and Chen, Y. (2021). An Aging-Related Gene Signature-Based Model for Risk Stratification and Prognosis Prediction in Lung Adenocarcinoma. *Front. Cel. Dev. Biol.* 9, 685379. doi:10.3389/fcell.2021.685379

Zhang, C., Zhang, Z., Sun, N., Zhang, Z., Zhang, G., Wang, F., et al. (2020). Identification of a Costimulatory Molecule-Based Signature for Predicting Prognosis Risk and Immunotherapy Response in Patients with Lung Adenocarcinoma. *Oncoimmunology* 9 (1), 1824641. doi:10.1080/2162402x.2020.1824641

Zhao, Y., Varn, F. S., Cai, G., Xiao, F., Amos, C. I., and Cheng, C. (2018). A P53-Deficiency Gene Signature Predicts Recurrence Risk of Patients with Early-Stage Lung Adenocarcinoma. *Cancer Epidemiol. Biomarkers Prev.* 27 (1), 86–95. doi:10.1158/1055-9965.epi-17-0478

Zhu, J., Wang, M., and Hu, D. (2020). Development of an Autophagy-Related Gene Prognostic Signature in Lung Adenocarcinoma and Lung Squamous Cell Carcinoma. *PeerJ* 8, e8288. doi:10.7717/peerj.8288