



# Identification of an Immune-Related Signature Predicting Survival Risk and Immune Microenvironment in Gastric Cancer

Shuang Dai<sup>1†</sup>, Tao Liu<sup>1,2†</sup>, Xiao-Qin Liu<sup>3</sup>, Xiao-Ying Li<sup>4</sup>, Ke Xu<sup>2</sup>, Tao Ren<sup>2\*</sup> and Feng Luo<sup>1\*</sup>

<sup>1</sup> Department of Medical Oncology, Lung Cancer Center, West China Hospital, Sichuan University, Chengdu, China,

<sup>2</sup> Department of Oncology, The First Affiliated Hospital of Chengdu Medical College, Chengdu Medical College, Chengdu, China, <sup>3</sup> Department of Oncology, Jintang First People's Hospital, Chengdu, China, <sup>4</sup> Department of Abdominal Cancer, Cancer Center, West China Hospital, Sichuan University, Chengdu, China

## OPEN ACCESS

### Edited by:

Yongqian Shu,  
Nanjing Medical University, China

### Reviewed by:

Donald Palmer,  
Royal Veterinary College (RVC),  
United Kingdom  
Xianquan Zhan,  
Shandong First Medical University,  
China

### \*Correspondence:

Tao Ren  
584239854@qq.com  
Feng Luo  
hxlufeng@163.com

<sup>†</sup> These authors have contributed  
equally to this work and share first  
authorship

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Cell and Developmental  
Biology

**Received:** 29 March 2021

**Accepted:** 08 October 2021

**Published:** 02 November 2021

### Citation:

Dai S, Liu T, Liu X-Q, Li X-Y, Xu K,  
Ren T and Luo F (2021) Identification  
of an Immune-Related Signature  
Predicting Survival Risk and Immune  
Microenvironment in Gastric Cancer.  
*Front. Cell Dev. Biol.* 9:687473.  
doi: 10.3389/fcell.2021.687473

**Background:** Tumor immune microenvironment plays a vital role in tumorigenesis and progression of gastric cancer (GC), but potent immune biomarkers for predicting the prognosis have not been identified yet.

**Methods:** At first, RNA-sequencing and clinical data from The Cancer Genome Atlas (TCGA) were mined to identify an immune-risk signature using least absolute shrinkage and selection operator (LASSO) regression and multivariate stepwise Cox regression analyses. Furthermore, the risk score of each sample was calculated, and GC patients were divided into high-risk group and low-risk group based on their risk scores. Subsequently, the performance of this signature, including the correlation with overall survival (OS), clinical features, immune cell infiltration, and immune response, has been tested in GC data from TCGA database and Gene Expression Omnibus (GSE84437), respectively.

**Results:** An immune signature composed of four genes (MAGED1, ACKR3, FZD2, and CTLA4) was constructed. The single sample gene set enrichment analysis (ssGSEA) indicated that activated CD4<sup>+</sup>/CD8<sup>+</sup> T cell, activated dendritic cell, and effector memory CD8<sup>+</sup> T cell prominently increased in the low-risk group, showing relatively high immune scores and low stromal scores. Further GSEA analysis indicated that TGF- $\beta$ , Ras, and Rap1 pathways were activated in the high-risk group, while Th17/Th1/Th2 differentiation, T cell receptor and PD-1/PD-L1 checkpoint pathways were activated in the low-risk group. Low-risk patients presented higher tumor mutation burden (TMB) and expression of HLA-related genes. The immune-associated signature showed an excellent predictive ability for 2-, 3-, and 5-year OS in GC.

**Conclusion:** The immune-related prognosis model contributes to predicting the prognosis of GC patients and providing valuable information about their response to immunotherapy using integrated bioinformatics methods.

**Keywords:** gastric cancer, immune signature, tumor mutation burden, immune infiltration, survival

## INTRODUCTION

Gastric cancer (GC) is a common digestive tract tumor and the third leading cause of cancer-related deaths (Bray et al., 2018). Although there have been advances in early screening and current therapies, the treatment and survival for GC remain unsatisfactory. The 5-year survival rate for GC is still lower than 30%, and survival for metastatic GC remains below 2 years (Ferlay et al., 2015). Further biomedical research is urgently recommended for screening novel diagnostic biomarkers and therapeutic targets.

Tumor immune microenvironment (TIM) is closely related to the tumorigenesis and tumor progression, as well as resistance to immunotherapy (Bindea et al., 2013; Quail and Joyce, 2013; Fridman et al., 2017). Accumulating studies have highlighted that heterogeneity exists in the proportion of intratumoral immune cell populations, and researchers try to identify those patients with specific immune response (Zeng et al., 2019; Zhang et al., 2019). The oncologists focus on activating immune responses, screening highly specific immune biomarkers, and exploring immune regulatory mechanisms to restore dysregulated immune microenvironment and improve survival outcomes (Eso and Seno, 2020). For example, the development of immune checkpoint inhibitors (ICIs) (Tang et al., 2018) has proposed the predictive biomarkers including microsatellite-instability (MSI) status, PD-1/PD-L1 expression, and tumor mutation burden (TMB) (Lee et al., 2016; Samstein et al., 2019). However, there is still no effective biomarker to predict sensitivity or resistance to ICIs. It is far from enough to fully understand the roles of tumor immune landscape in GC.

Bioinformatics analysis has been widely employed to screen prognostic signatures (Rooney et al., 2015; Jiang et al., 2018). With the help of ssGSEA and CIBERSORT algorithms, some studies estimated the immune cell subsets and immune scores in order to gain a better understanding of the immune infiltration status and treatment response of immunotherapy in GC (Bindea et al., 2013; Newman et al., 2015). However, there has been no appropriate study that constructed a robust signature to predict immune status and the survival of GC based on immune-related genes, or could systematically evaluate the relationship between immune genomic characteristics and TIM (Shi et al., 2018; Zeng et al., 2019). Therefore, our study took advantage of TCGA and GEO database to determine key immune-associated genes and build a reliable signature using Lasso and multivariate Cox regression analyses. In addition, this signature further examined the predictive effect on the response of ICIs.

**Abbreviations:** GC, gastric cancer; STAD, stomach cancer; ICIs, immune checkpoint inhibitors; MSI, microsatellite-instability; IPS, Immunophenoscore; TMB, tumor mutation burden; TIM, tumor immune microenvironment; MDSC, myeloid-derived suppressor cells; Lasso regression analysis, the least absolute shrinkage and selection operator regression analysis; AUCs, area under the ROC curves; ssGSEA, single sample gene set enrichment analysis; SMGs, significantly mutated genes; GSEA, gene set enrichment analysis; KEGG, Kyoto Encyclopedia of Genes and Genomes; IRGs, immune-related genes; HLA, human leukocyte antigen; MHC, major histocompatibility complex.

## MATERIALS AND METHODS

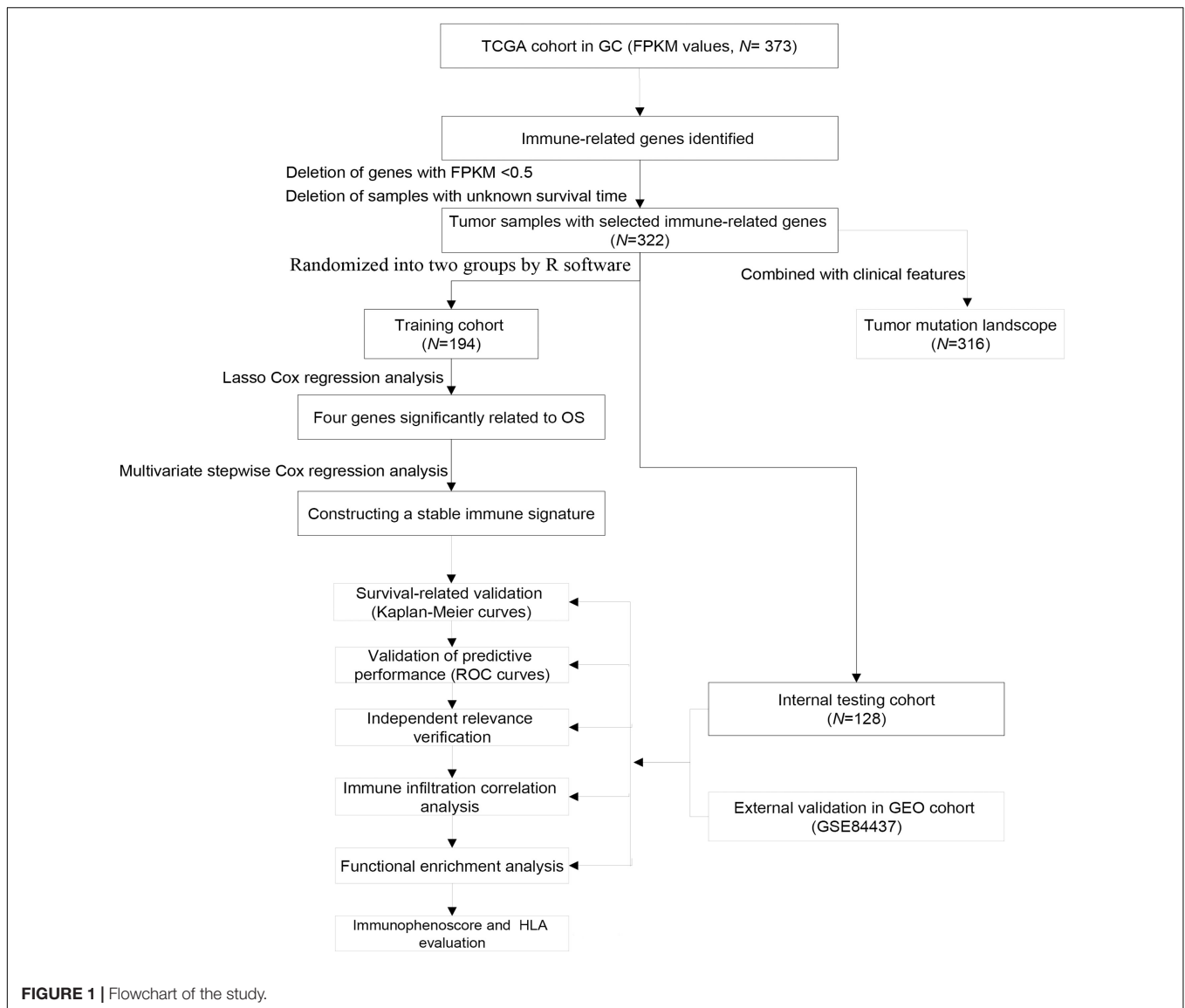
### Data Download and Preprocessing

The FPKM values of gene expression profiles ( $n = 373$ ) and corresponding clinical information were derived from the TCGA database (December, 2020), including 343 GC samples and 30 adjacent non-tumor samples. Tumor samples with the corresponding clinical data were randomized into training group and internal testing group, and then the data with unknown survival time and 0-month survival time were deleted. The immune-related gene profile is derived from the ImmPort database (Bhattacharya et al., 2014) and previous publications. GSE84437 cohort including 433 samples was also downloaded from the GEO database as an independent external validation group. The study did not need the approval from the ethics committees because all data were open-access in the TCGA or GEO database. The detailed flowchart of analysis steps is shown in **Figure 1**.

### Identification of Immune-Related Genes and Development of Immune-Related Risk Signature

Based on the ImmPort Database<sup>1</sup> and the previous studies (Li et al., 2017; Gao et al., 2020; Sun et al., 2020), a total of 3,265 immune-related genes (IRGs) were collected. Then, after the deletion of low-expression genes (FPKM < 0.5), these IRGs were included into the Lasso analysis to integrate multiple variables. Lasso Cox regression analysis, a penalty regression tool, can estimate the regression coefficients by maximizing the log-likelihood function (or sum of squared residuals). Through Lasso regression analysis, the high-dimensional data were resampled 1,000 times, which can automatically delete unnecessary covariates and screen out the most stable prognostic factors. We used multivariate stepwise Cox regression analysis (forward and backward) to evaluate the capabilities of immune-related signatures in predicting OS of GC patients. Eventually, we calculated the total risk score of each patient ( $\beta$  genes  $\times$  expression level of genes) in the training and validation cohorts by integrating the regression coefficients and gene expression levels, respectively. Based on the risk score, patients' survival status, risk score levels, and heat maps of the immune signature were also presented. Subsequently, the samples were separated into high- and low-risk groups based on the median risk score of the training set. The predictive capacity of immune-risk signature was tested by Kaplan–Meier survival analysis using the *R* package “Survminer.” Time-dependent receiver operating characteristic (ROC) curves with AUC values were used to evaluate the ability of prognosis classification. Moreover, the independent predictive ability of the immune-risk signature also needed to be further assessed by multivariate analyses compared with other clinical factors including age, grade, and stage treated as continuous variables.

<sup>1</sup><https://immunport.niaid.nih.gov>



## ssGSEA Validation of Risk Signature Related to Immune Cell Infiltration

To further elucidate the biological understanding of immune-related riskScore, ssGSEA was applied to calculate the proportion of immune cells in the TIM of GC based on the comparison of GC expression profiles with the gene sets (Subramanian et al., 2005; Bindea et al., 2013). The gene sets comprising 782 genes were utilized to identify 28 types of immune cells and predict the abundance of corresponding immune cells. These types of immune cells were defined as three major immunocyte types including anti-tumor immunocytes (activated CD4<sup>+</sup> T cells, activated CD8<sup>+</sup> T cells, activated dendritic cells, CD56 bright natural killer cells, central memory CD4<sup>+</sup> T cells, central memory CD8<sup>+</sup> T cells, effector memory CD4<sup>+</sup> T cells, effector memory CD8<sup>+</sup> T cells, natural killer cells, natural killer T cells, type-1 T helper cells, and type-17 T helper cells), pro-tumor immunocytes (CD56 dim natural killer cells, myeloid-derived suppressor cells

[MDSC], immature dendritic cells, neutrophils, plasmacytoid dendritic cells, macrophages, regulatory T cells, and type-2 T helper cells), and intermediate immunocytes (activated B cells, eosinophils, gamma delta T cells, immature B cells, mast cells, memory B cells, monocytes, and follicular helper T cells). We also compared the proportion differences of tumor immune cell infiltration between low- and high-risk immune groups by using Mann-Whitney-Wilcoxon tests.

## Correlation Analyses of Immune Response

Human leukocyte antigen (HLA) genes have a critical function in immune surveillance and response. The human MHC encodes a glycoprotein known as HLA, which plays an important role in T-cell antigen presentation (Mosaad, 2015; Dendrou et al., 2018). The MHC is divided into three subclasses: the class I region includes the classical gene types (HLA-A, HLA-B,

and HLA-C genes), as well as the non-classical types known as MHC Ib (HLA-E, HLA-F, and HLA-G genes); the class II region includes the HLA-DPA1, HLA-DPB1, HLA-DQA1, HLA-DQA2, HLA-DQB1, HLA-DQB2, HLA-DRA, HLA-DRB1, HLA-DRB2, HLA-DRB3, HLA-DRB4, and HLA-DRB5 genes as well as less variable genes involved in antigen processing and presentation; and the class III region does not encode any molecules for antigen delivery and peptide binding, but contains genes only involving in inflammatory responses, leukocyte maturation and the complement cascade (Mosaad, 2015). We compared the expression of HLA-related genes among patients in different immune risk groups. In addition, Immunophenoscore (IPS), as a scoring tool for evaluating the tumor immunogenicity, was applied to analyze the correlation between the new immune signature and intratumor immune response. ESTIMATE algorithm was also utilized to calculate the Tumor Purity rate, ESTIMATE Score, Immune Score, and Stromal Score within the tumor microenvironment of GC in the TCGA and GEO groups, respectively (Yoshihara et al., 2013). The difference of immune scores and stromal scores among patients in different immune risk groups was expressed by “ggplot2” and “ggsignif” (R package).

### Functional Enrichment Analysis

The R package “limma” was used to analyze differential expression between high-risk and low-risk groups, and all genes were ranked by foldchange values. Gene Set Enrichment Analysis of KEGG was performed to clarify the significant annotated pathways ( $P < 0.05$ ) through R package “gseKEGG of clusterProfiler” (Subramanian et al., 2005).

### Tumor Mutation Burden Analysis

The somatic mutational annotation file of GC in the TCGA was downloaded from Genome Data Commons. TMB, defined as non-synonymous somatic mutations per Mb, was counted by the total number of mutations in six mutation categories, including all GC splice site mutations, missense and non-sense mutations, in-frame insertions, frameshift mutations, and deletions. The R package “GenVisR” was conducted to investigate the relationship between the immune signature and TMB. Meanwhile, we performed multivariate analyses again to assess the predictive performance of the immune-risk signature by comparing with other factors including age, grade, stage, TMB, and IPS.

### qRT-PCR Analysis

A tissue microarray containing 13 independent STAD tissue samples and paired adjacent non-cancerous tissue samples was purchased from Shanghai Outdo Biotech Co., Ltd. (cat. no. cDNA-HStmA030CS01, Shanghai, China), and used to investigate the expression of four genes (MAGED1, ACKR3, FZD2, and CTLA4). The forward primer sequences were as follows: ACKR3-F: 5'- GACGCTTTTGGTGGGCATGT-3' and the reverse primer sequence was ACKR3-R: 5'- ATT TGATTGCCCGCCTCAGA-3' and the product length was 149 bp; FZD2-F: 5'- CTCCGTCCTCGGAGTGGTTC-3' and the reverse primer sequence was FZD2-R: 5'- GCGAAGC CCTCATGAACAAG-3' and the product length was 117 bp;

MAGED1-F: 5'- AGGTCTGCATAAGCAAGGCG-3' and the reverse primer sequence was MAGED1-R: 5'- TGCTGC CTTCTTCGTCAAGC-3' and the product length was 143 bp; CTLA4-F: 5'- GCAGCAGTTAGTTCGGGGTT-3' and the reverse primer sequence was CTLA4-R: 5'- CTCTGTTGGGGG CATTTCAC-3' and the product length was 123 bp. The primers of  $\beta$ -actin (reference gene): forward: 5'-GAAGAGCTACG AGCTGCCTGA-3' and reverse: 5'- CAGACAGCACTGTGTTG GCG-3'. All RT-PCR operations were performed according to the manufacturer's instructions. In short, a 20- $\mu$ l mixture containing 10  $\mu$ l of TB Green Premix Ex Taq (Tli RNaseH Plus) (TaKaRa Bio Inc., Japan), 1  $\mu$ l of 100  $\mu$ M primer mix (Tsingke Biotechnology Co., Ltd., Beijing, China) and 9  $\mu$ l of RNase/DNase-free sterile water were added to each cDNA sample, and then the cDNA arrays were sealed with a transparent sealing film, and placed on ice for 15 min. The qRT-PCR reaction was carried out in the LightCycler 96 Real-Time PCR System (Roche Diagnostics, Indianapolis, Ind.). The reaction mixture was activated at 50°C for 2 min, pre-denatured at 95°C for 10 min, followed by 40 cycles of amplification reaction at 95°C for 15 s and 60°C for 1 min. Finally, LightCycler 96 software (Version 1.1.0.1320, Roche) was used for the collection and analysis of qRT-PCR data. The relative mRNA expression level was calculated by the  $2^{-\Delta\Delta Ct}$  method taking  $\beta$ -actin as the reference gene.

### Statistical Analysis

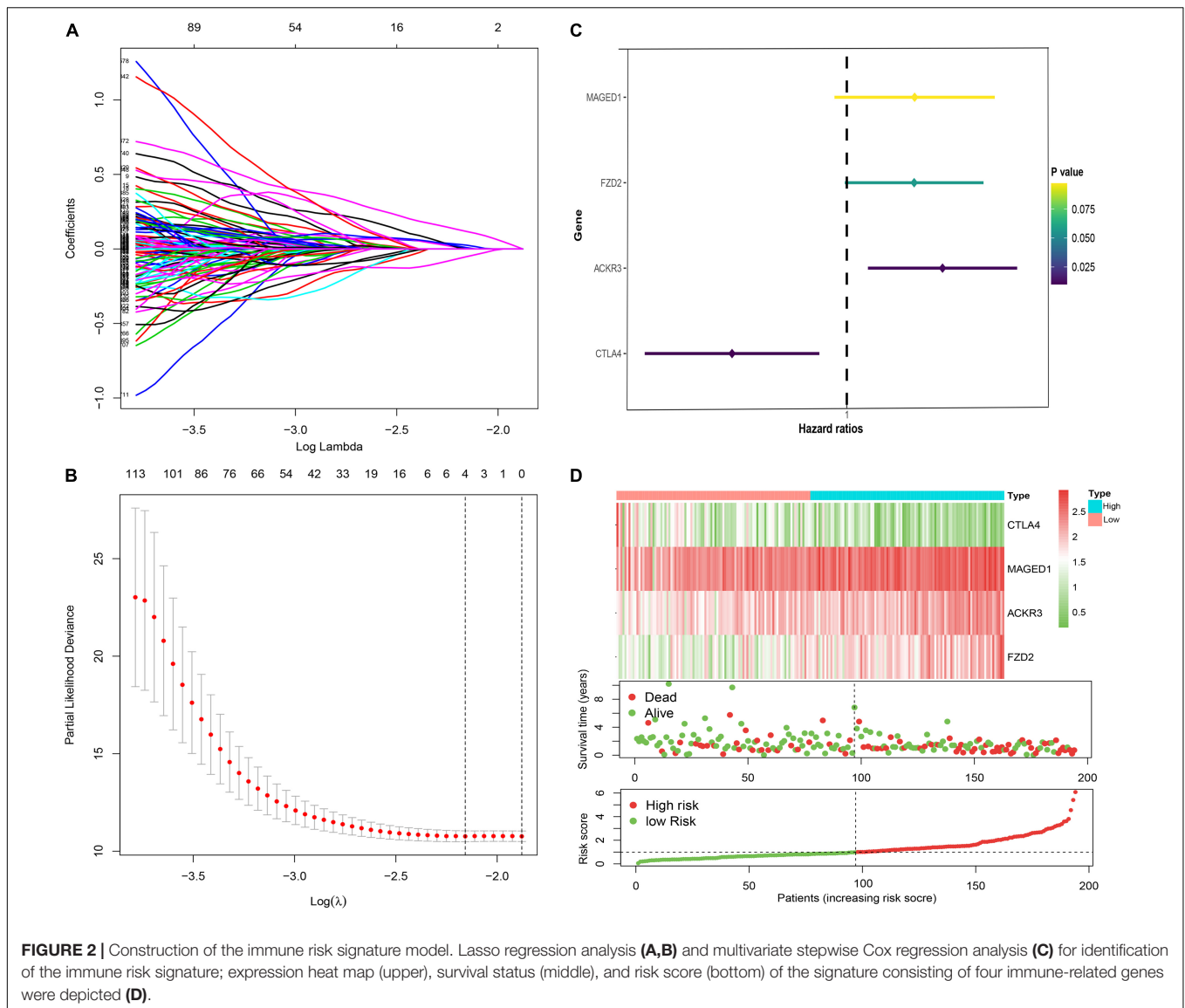
R statistics software (version 3.6.1) was used for all statistical analyses. The chi-square test for qualitative variables was used to estimate the differences between groups. Differences between two or more than two cohorts were examined by non-parametric tests including Mann-Whitney-Wilcoxon test and Kruskal-Wallis  $H$  test in quantitative data. Chi-square test, Mann-Whitney-Wilcoxon test, and Kruskal-Wallis  $H$  test were performed with the R functions “chisq.test,” “wilcox.test,” and “kruskal.test,” respectively.  $P$  values less than 0.05 were considered to be statistically significant.

## RESULTS

### Construction of Immune Risk Signature

A total of four immune genes (MAGED1, FZD2, ACKR3, and CTLA4) were screened out through Lasso regression analysis (Figures 2A,B). In order to strictly examine their prognostic significance and establish an immune-risk signature, these four genes were further included into multivariate stepwise Cox regression analysis, and a prognostic Cox regression model was established (Figure 2C). Based on regression coefficients and expression levels, the formula of risk score was as follows:  $(0.2615 \times \text{expression level of MAGED1}) + (-0.3691 \times \text{expression level of FZD2}) + (0.2597 \times \text{expression level of ACKR3}) + (-0.4425 \times \text{expression level of CTLA4})$ . The patients were divided into high- and low-risk cohorts according to the median value of the training group. Based on the risk score, patients' survival state, risk score levels, and heatmaps of this signature in the training group are shown in Figure 2D





in detail. Validations in the TCGA and GEO sets are shown in **Supplementary Figure 1**. Taken together, these genes are most likely to participate in GC.

## Validation of Prognostic Value of Risk Signature

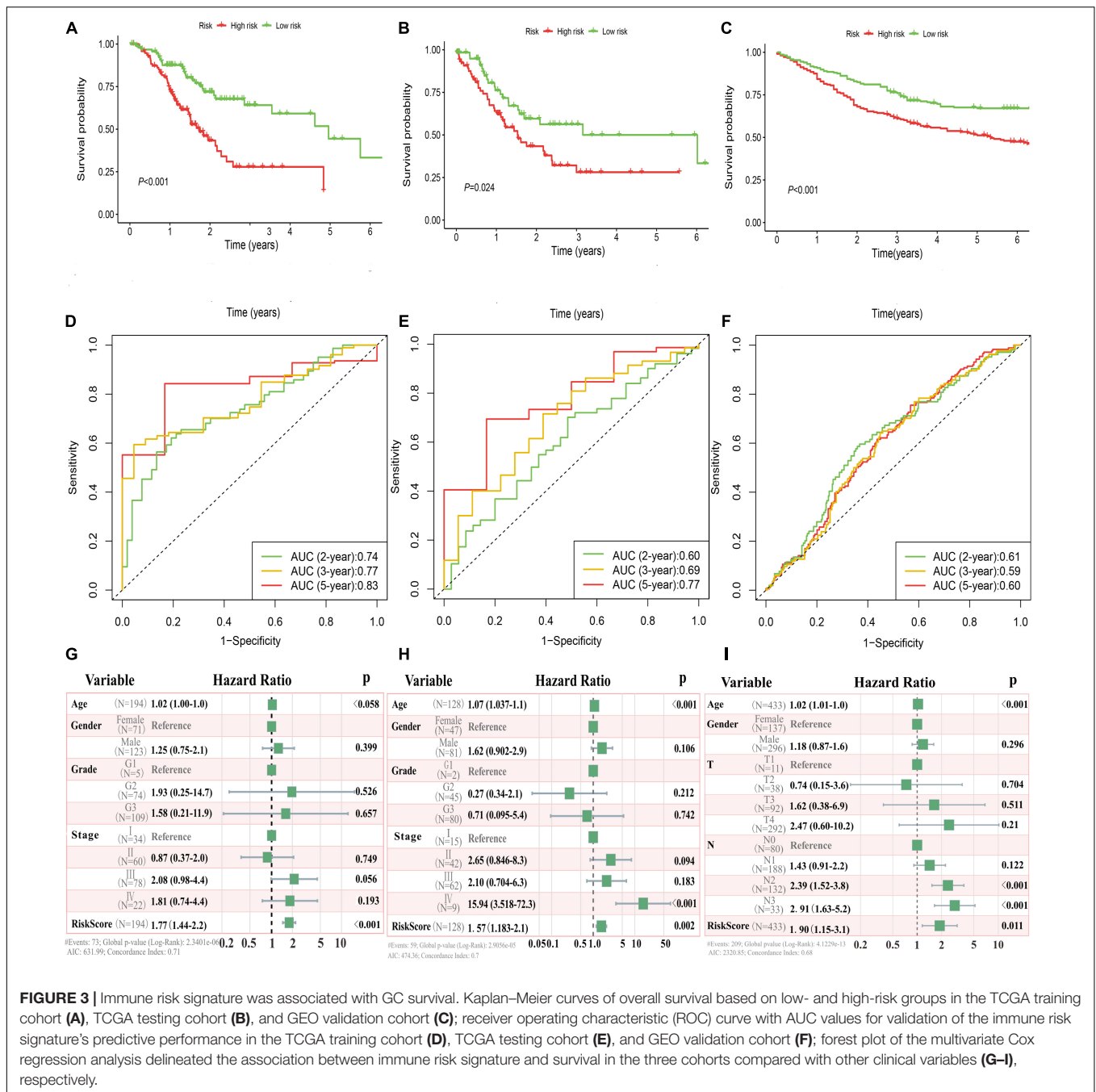
In the TCGA and GEO cohorts, we observed significant differences in prognosis among patients in different immune risk groups *via* Kaplan–Meier curves (**Figures 3A–C**). Moreover, the AUC values of 2-, 3-, and 5-years were 0.74, 0.77, and 0.83 in the training cohort, 0.60, 0.69, and 0.77 in the TCGA testing cohort, and 0.61, 0.59, and 0.60 in the GEO cohort, respectively (**Figures 3D–F**).

Multivariate analyses were conducted to validate whether this immune signature was the risk factor of GC, not concerned with clinical features (age, gender, grade classification, and tumor stage). The results showed that the risk score was an independent

variable related to the survival of GC patients in all three cohorts ( $P < 0.01$ , **Figures 3G–I**).

## Correlation Assessment Between Immune Risk Signature and Immune Cell Infiltration

We would like to further evaluate the relationship between the immune signature and TIM characteristics. Hence, we draw a heatmap by ssGSEA to visualize the relative abundance of 28 immune infiltrating cell subsets from the TCGA training dataset (**Figure 4**). Obviously, the cell subset of anti-tumor lymphocytes including activated dendritic cells ( $P = 0.009$ ), activated  $CD4^+/CD8^+$  T cells ( $P < 0.001$ ), and effector memory CD8 T cells ( $P < 0.001$ ) were enriched in the low-risk signature group. Protumor immunocyte subtypes like MDSC ( $P = 0.002$ ), regulatory T cells ( $P = 0.042$ ), and type 2 T helper cell ( $P = 0.007$ ) and intermediate immunocyte

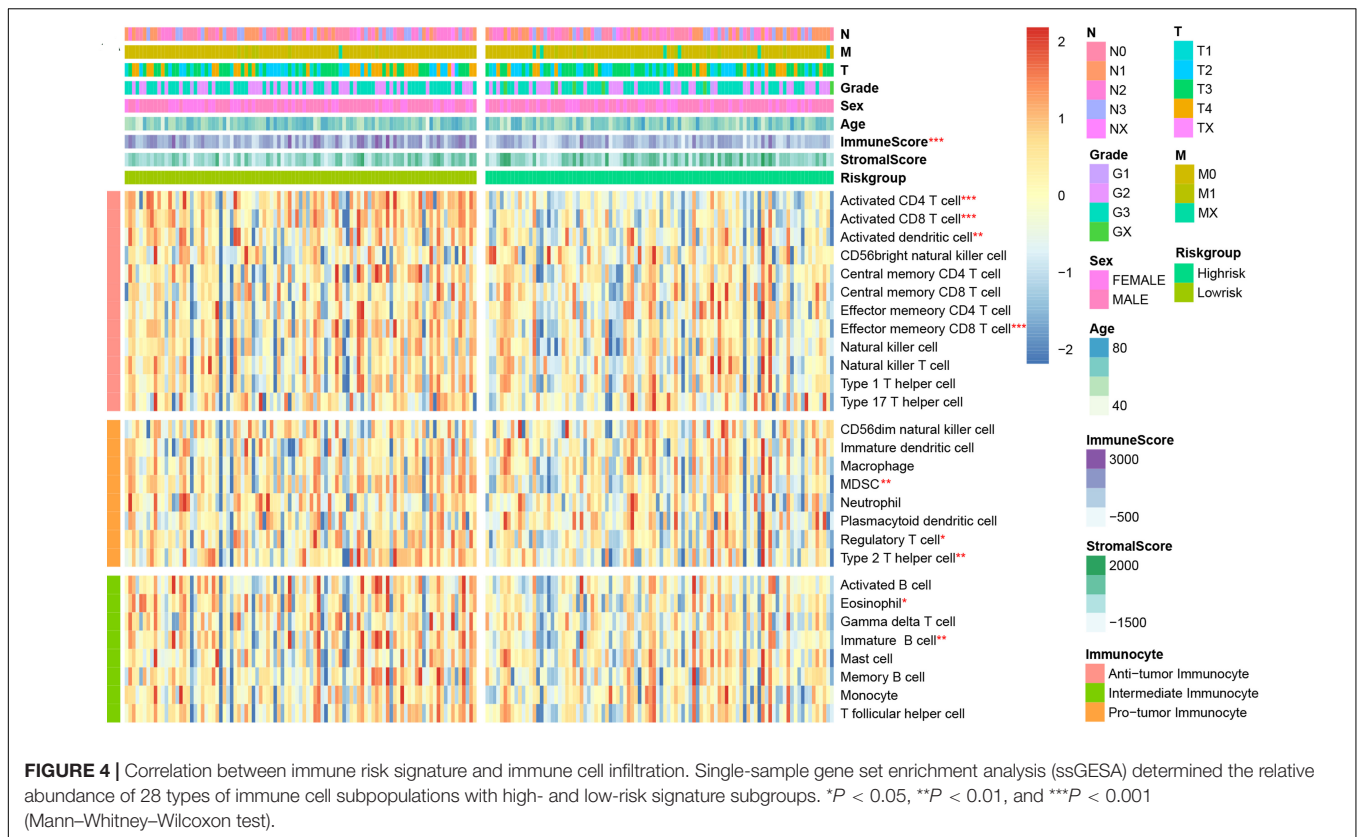


subtypes like eosinophil ( $P = 0.014$ ) and immature B cell ( $P = 0.008$ ) were also elevated in the low-risk signature group. Subsequently, the conclusion was validated again in TCGA and GEO validation datasets (Supplementary Figure 2), and a similar tendency was observed in the two-risk stratification groups. Moreover, immune and stromal scores were counted by the ESTIMATE algorithm. Analysis of differences between groups showed that the immune score of the low-risk group was higher ( $P < 0.001$ ) than that of the high-risk group, but the stromal score was relatively low (Supplementary Figure 3), which was basically in line with the results

observed within the entire TCGA and GEO validation groups ( $P < 0.05$ ).

### Functional Enrichment Analysis

We assessed the newly identified signature’s roles in regulating the gene enrichment pathways *via* the related biological signaling pathway analysis based on GSEA analysis (Figure 5A). The TGF $\beta$ , Ras, and Rap1 pathways involved in tumor proliferation, migration, and invasion were obviously enriched in the immune high-risk cohort. In the immune low-risk signature, Th17/Th1/Th2 differentiation signaling pathways, T cell receptor



as well as PD-L1 expression and PD-1 checkpoint pathways were mainly activated.

## Immune Risk Signature Evaluated Immune Response

We assessed the difference in expression of 24 HLA-related genes in all three datasets. We found that HLA-B, HLA-DMA, HLA-DMB, HLA-DOA, HLA-DOB, HLA-DPA1, HLA-DPB1, HLA-DPB2, HLA-DQA1, HLA-DQB1, HLA-DQB2, HLA-DRA, HLA-DRB1, HLA-DRB5, HLA-E, HLA-F, and HLA-H were elevated ( $P < 0.05$ ) in the immune low-risk signature group, which may indicate a good immune response (Figure 5B). Validation within the entire TCGA and GEO groups was shown in Supplementary Figure 4. According to the results of IPS, the low-risk group had an obviously higher score compared with the high-risk group in the TCGA dataset ( $P = 0.007$ , Figure 6A).

## Immune Risk Signature Determined the Mutation Landscape

To improve the understanding of the correlation between the newly constructed immune signature and mutational landscape, TMB was counted and summed. Patients in the low-risk group tended to exhibit higher TMB than those in the high-risk group in the TCGA dataset ( $P = 0.002$ , Figure 6B). Besides, the study further explored the association between survival and various variables such as clinical factors, IPS, TMB, and newly constructed immune signature, which indicated that the

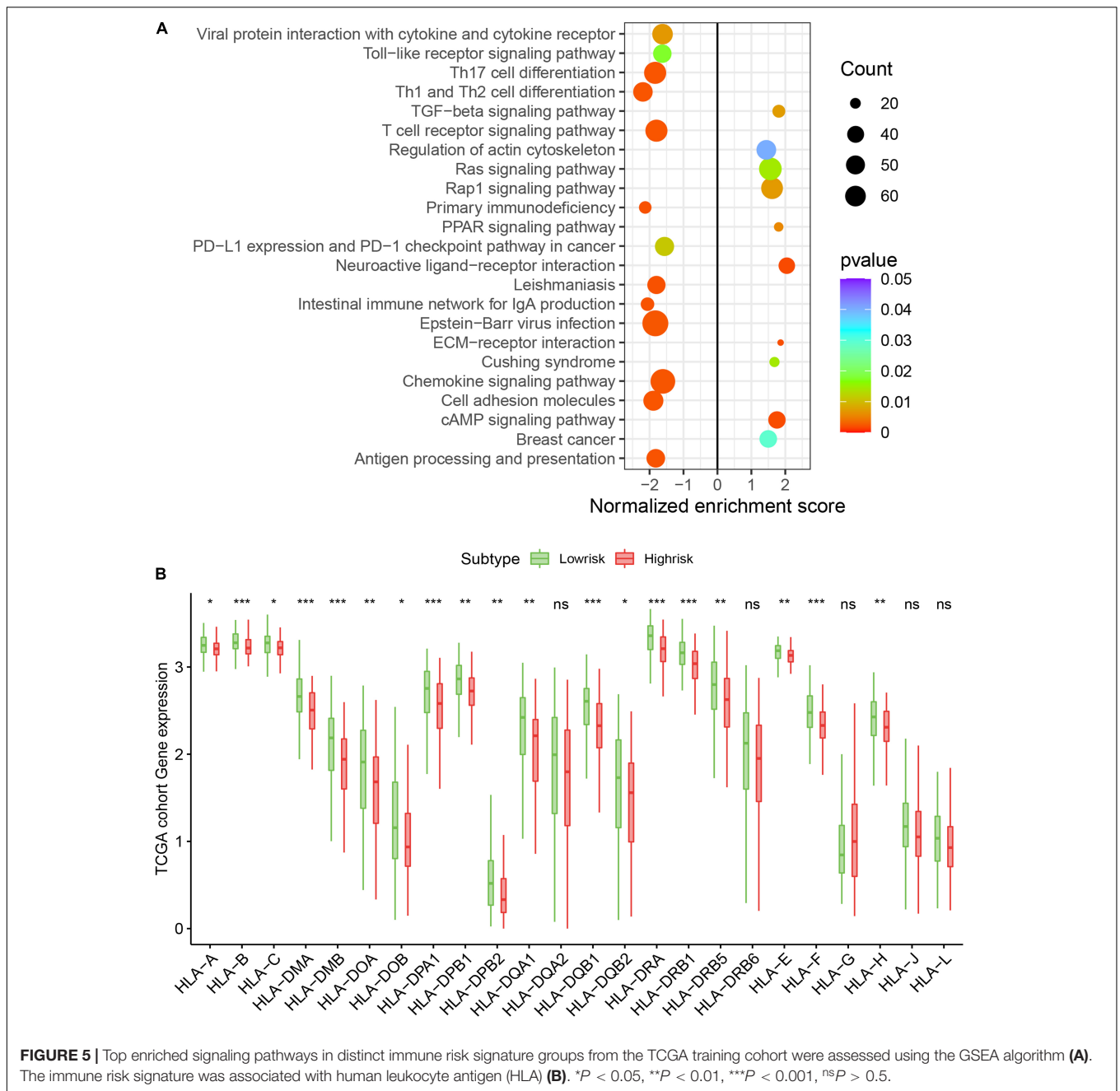
immune signature was still an independent survival-related factor ( $P < 0.001$ , Figure 6C). Significantly mutated genes (SMGs) analysis for GC samples (high-risk vs. low-risk subgroups) was demonstrated in Figure 6D. The SMGs including ARID1A [2 of 164 (1.22%) vs. 12 of 143 (8%),  $P = 0.008$ ], PIK3CA [12 of 164 (7.32%) vs. 33 of 143 (22%),  $P < 0.001$ ], OBSCN [15 of 164 (9.15%) vs. 29 of 143 (19.33%),  $P = 0.015$ ], PLEC [11 of 164 (6.71%) vs. 25 of 143 (16.67%),  $P = 0.01$ ], DMD [14 of 164 (8.54%) vs. 28 of 143 (18.67%),  $P = 0.013$ ], KMT2D [13 of 164 (7.93%) vs. 33 of 143 (22%),  $P < 0.001$ ], and XIRP2 [10 of 164 (6.1%) vs. 20 of 143 (13.33%),  $P = 0.046$ ] were marked in Figure 6D.

## Expression Verification in STAD by qRT-PCR

The expression levels of these selected genes were tested using qRT-PCR to further prove their reliability. As shown in Figure 7, all the mRNA transcripts were detected and their expressions in STAD were significantly higher than those in adjacent normal tissues ( $p < 0.05$ ).

## DISCUSSION

Tumor immune contexture plays a key role in tumorigenesis, progression, and metastasis (Dunn et al., 2004; Fridman et al., 2012; Remark et al., 2015). ICIs have been regarded as a promising treatment in advanced or metastatic GC. Yet, up to now, only a small number of GC patients have benefited



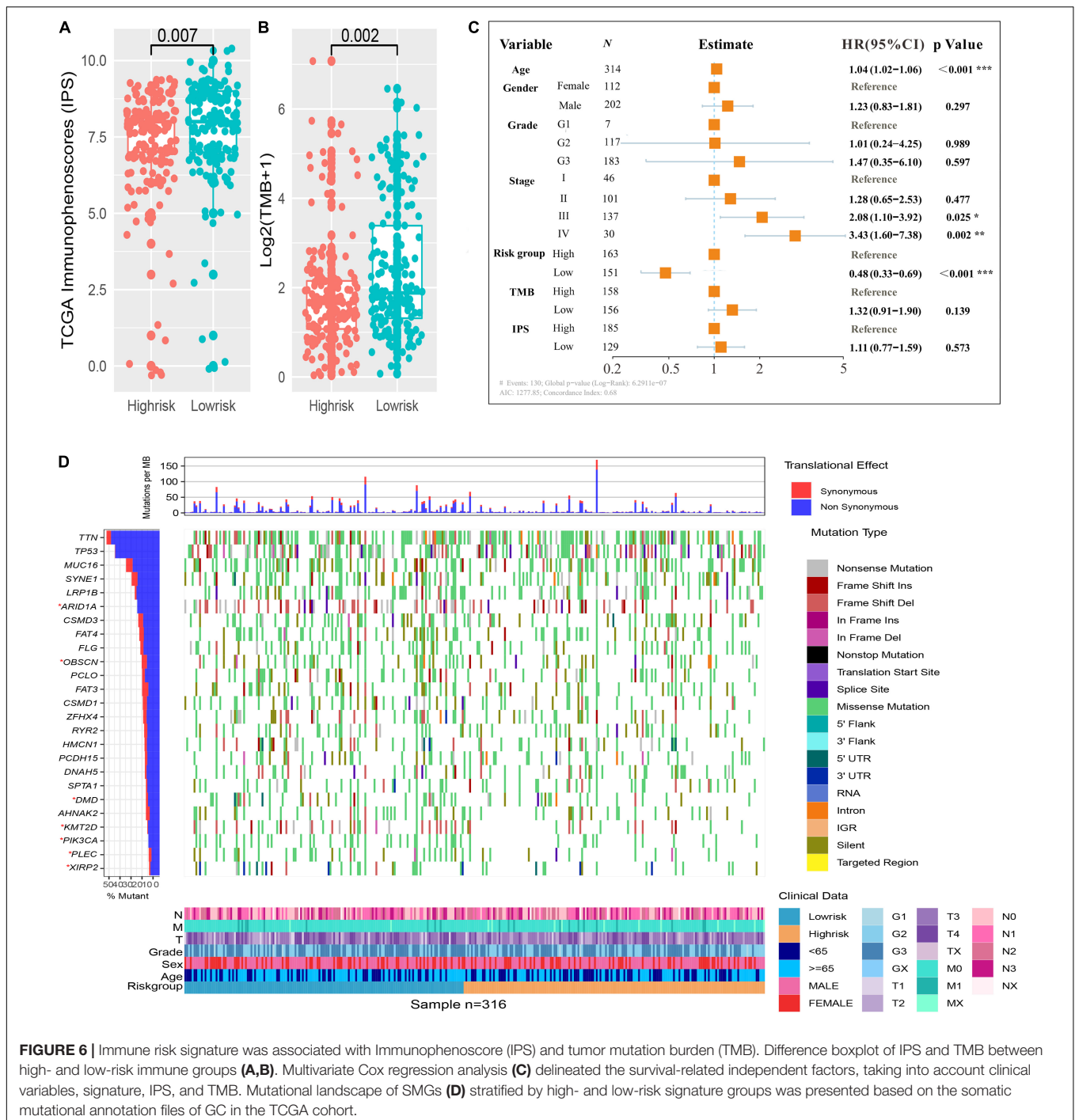
**FIGURE 5 |** Top enriched signaling pathways in distinct immune risk signature groups from the TCGA training cohort were assessed using the GSEA algorithm (A). The immune risk signature was associated with human leukocyte antigen (HLA) (B). \**P* < 0.05, \*\**P* < 0.01, \*\*\**P* < 0.001, ns*P* > 0.5.

from the ICIs treatment. In this research, we have developed an immune risk signature based on four immune-related genes (MAGED1, FZD2, ACKR3, and CTLA4) in a stand-alone dataset, and have validated its performance in TCGA and GEO datasets. This new signature can stratify GC patients into high- and low-risk subgroups and serve as a potential prognostic biomarker of immunotherapy.

Similar to previous investigation (Lee et al., 2008; Kang et al., 2017), our study indicates that many types of immune cells are enriched in low-risk immune groups, including activated CD4<sup>+</sup>/CD8<sup>+</sup> T cell, activated dendritic cell, effector memory CD8<sup>+</sup> T cell, eosinophil, immature B cell, MDSC, regulatory T

cell, and type 2 T helper cell, with better survival. Protumor immunocyte types like MDSC and regulatory T cell, as adverse prognostic factors in GC (Ichihara et al., 2003; Wang B. et al., 2011), similarly increase in the low-risk immune group. This inconsistent result may be due to the actual high ratio of anti-tumor immunocytes compared with protumor immunocyte, presenting a favorable factor in GC survival. In addition, stromal and immune scores calculated by the ESTIMATE algorithm showed that immune scores are significantly elevated and stromal scores depressed in the low-risk immune group, which is consistent with studies that emphasize that these patients may be sensitive to immunotherapy (Mao et al., 2013;



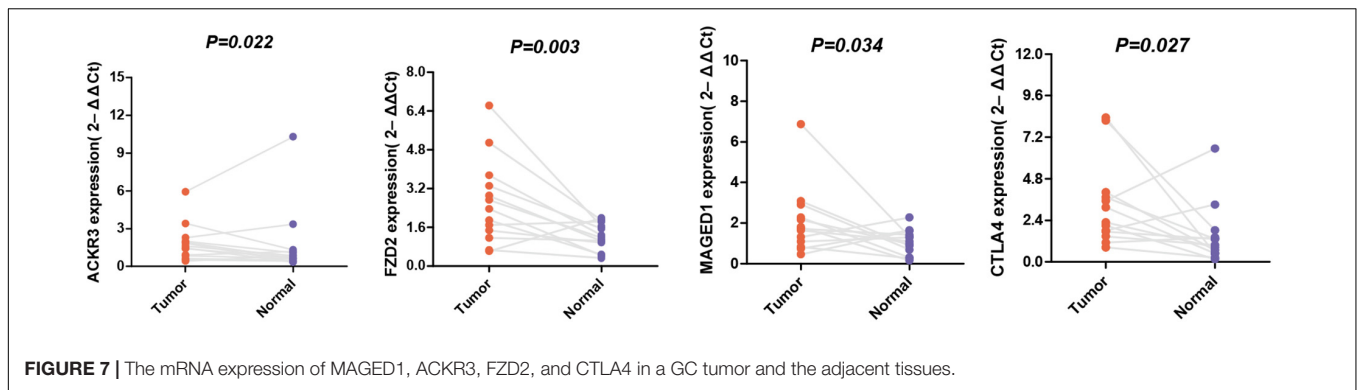


**FIGURE 6 |** Immune risk signature was associated with Immunophenoscore (IPS) and tumor mutation burden (TMB). Difference boxplot of IPS and TMB between high- and low-risk immune groups (A,B). Multivariate Cox regression analysis (C) delineated the survival-related independent factors, taking into account clinical variables, signature, IPS, and TMB. Mutational landscape of SMGs (D) stratified by high- and low-risk signature groups was presented based on the somatic mutational annotation files of GC in the TCGA cohort.

Mariathasan et al., 2018). The GSEA analysis revealed that TGF-β, Ras, and Rap1 signaling pathways are easily activated in high-risk signature and regulate tumorigenesis, proliferation, and invasion, which has been confirmed in the previous studies (Adjei, 2001; Tauriello et al., 2018; Liu et al., 2021). Other pathways activated in low-risk signature like Th17/Th1/Th2 differentiation, T cell receptor signaling as well as PD-L1 expression and PD-1 checkpoint pathways potentially improve the therapeutic efficacy (Thorsson et al., 2018). The new immune

signature helps to screen high-risk population and facilitate the development of precision immunotherapy.

The validation analysis reveals that the immune-related signature performs well in therapeutic and prognostic prediction for GC, and low-risk patients are characterized as higher mutation burden and expression level of HLA genes. HLA, often referred to as the MHC molecule or major histocompatibility complex, varies greatly from person to person. In addition, it is a marker for mutually recognizing the immune cells of



different individuals, which participates in the immune response and has very important biological functions. HLA-I (MHC-I), of great importance, consists of a highly polymorphic  $\alpha$ -heavy chain and a  $\beta$ 2-microglobulin ( $\beta$ 2M) light chain, encoded by the HLA-A, HLA-B, or HLA-C genes, delivers peptides to CD8<sup>+</sup> T cells, and is essential for immune surveillance and cancer immunotherapy (Hazini et al., 2021). Patients with increased expression of HLA-A, HLA-B, or HLA-C genes have a higher efficiency of antigen delivery. HLA-I dysregulation in cancer patients results in poor immune outcomes. Shim et al. (2020) have shown that correcting HLA can improve the immune efficacy of NSCLC. In our study, the immune-related risk score was able to assess patients with different levels of HLA-A, HLA-B, or HLA-C expression, thus guiding the use of immunotherapy to some extent. High TMB is often regarded as a sensitive marker of immune-checkpoint blockade, particularly a high mutation proportion of driver genes (Rizvi et al., 2015; Samstein et al., 2019). The SMGs like ARID1A, PIK3CA, OBSCN, PLEC, DMD, KMT2D, and XIRP2 are enriched in low-risk groups. ARID1A and PIK3CA have significant association with microsatellite instability (Wang K. et al., 2011; Mathur, 2018; Shen et al., 2019), predicting good response for immunotherapy of GC. DMD acts as a favorable gene for survival of gastric adenocarcinomas (Jones et al., 2020). The relationship between other SMGs and immunotherapy have not been reported yet. Moreover, IPS also exhibits a rising tendency in low-risk immune groups. To investigate the survival-related covariates, we included TMB, IPS, and other clinical factors into multivariate analysis, and found that this signature remains an independent prognostic factor. Therefore, it is reasonable to think that people with a low-risk immune systems may probably gain clinical benefits from checkpoint immunotherapy. The new immune signature can distinguish immune status and promote the progress of precision immunotherapy.

There are still several limitations in our study. For one, immune-related features are only constructed based on public datasets, which have not been confirmed in clinical cohorts. Prospective clinical studies involving immunotherapy combined with basic experiments are needed for further validation. For another, as the lack of mutational profile in the GEO database, the mutational landscape is not validated in independent datasets. In addition, the AUC values of the above prognostic signature cannot be very high, and some AUC values are only at fair

levels, especially in the GEO independent validation group. This is because our study only extracted immune-related genes for model construction. So far, studies have identified other gene categories closely linked to tumors and built models accordingly, such as ferroptosis-related genes and m6A-related genes. A mature precision model with strong AUC value level for gastric cancer is bound to be an integrated model through various genomics, proteomics, and even radiomics, which is also bound to be a long process of effort. Despite this, the findings of our study still have noteworthy implications for survival prediction along with the evaluation of precision immunotherapy for GC patients.

In a word, our study has developed a new immune-related signature that could not only predict the prognosis of GC patients but also reveal intratumor immune response, which may guide the development of novel treatment strategies combining chemotherapy, targeted therapy, and immunotherapy.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: Publicly available datasets were analyzed in this study. This data can be found here: TCGA (<https://gdc.cancer.gov/>) and GEO database (<https://www.ncbi.nlm.nih.gov/geo/>).

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

FL and TR: conception and design and manuscript review. TL and SD: methodology development, data collection, and manuscript writing. TL, SD, X-QL, X-YL, and KX: data analysis and interpretation. All authors approved the final manuscript.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2021.687473/full#supplementary-material>

**Supplementary Figure 1** | Expression heat map, survival status, and risk score of the signature in the TCGA testing cohort and GEO validation cohort were depicted.

**Supplementary Figure 2** | Estimation of the relative proportion of 28 types of immune cell subpopulations stratified by high- and low-risk signature groups in the TCGA testing cohort and GEO validation cohort. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$  (Mann–Whitney–Wilcoxon test).

**Supplementary Figure 3** | Difference of immune scores and stromal scores between low- and high-risk immune groups in the TCGA training cohort, TCGA testing cohort, and GEO validation cohort.

**Supplementary Figure 4** | Immune risk signature was associated with human leukocyte antigen (HLA) in the entire TCGA and GEO cohort.

## REFERENCES

- Adjei, A. A. (2001). Blocking oncogenic ras signaling for cancer therapy. *J. Natl. Cancer Inst.* 93, 1062–1074. doi: 10.1093/jnci/93.14.1062
- Bhattacharya, S., Andorf, S., Gomes, L., Dunn, P., Schaefer, H., Pontius, J., et al. (2014). ImmPort: disseminating data to the public for the future of immunology. *Immunol. Res.* 58, 234–239. doi: 10.1007/s12026-014-8516-1
- Bindea, G., Mlecnik, B., Tosolini, M., Kirilovsky, A., Waldner, M., Obenaus, A. C., et al. (2013). Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 39, 782–795. doi: 10.1016/j.immuni.2013.10.003
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424. doi: 10.3322/caac.21492
- Dendrou, C. A., Petersen, J., Rossjohn, J., and Fugger, L. (2018). HLA variation and disease. *Nat. Rev. Immunol.* 18, 325–339. doi: 10.1038/nri.2017.143
- Dunn, G. P., Old, L. J., and Schreiber, R. D. (2004). The three Es of cancer immunoeediting. *Annu. Rev. Immunol.* 22, 329–360. doi: 10.1146/annurev.immunol.22.012703.104803
- Eso, Y., and Seno, H. (2020). Current status of treatment with immune checkpoint inhibitors for gastrointestinal, hepatobiliary, and pancreatic cancers. *Therap. Adv. Gastroenterol.* 13:1756284820948773. doi: 10.1177/1756284820948773
- Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136, E359–E386. doi: 10.1002/ijc.29210
- Fridman, W. H., Pagès, F., Sautès-Fridman, C., and Galon, J. (2012). The immune contexture in human tumours: impact on clinical outcome. *Nat. Rev. Cancer.* 12, 298–306. doi: 10.1038/nrc3245
- Fridman, W. H., Zitvogel, L., Sautès-Fridman, C., and Kroemer, G. (2017). The immune contexture in cancer prognosis and treatment. *Nat. Rev. Clin. Oncol.* 14, 717–734. doi: 10.1038/nrclinonc.2017.101
- Gao, Y., Chen, S., Vafaei, S., and Zhong, X. (2020). Tumor-infiltrating immune cell signature predicts the prognosis and chemosensitivity of patients with pancreatic ductal adenocarcinoma. *Front. Oncol.* 10:557638. doi: 10.3389/fonc.2020.557638
- Hazini, A., Fisher, K., and Seymour, L. (2021). Deregulation of HLA-I in cancer and its central importance for immunotherapy. *J. Immunother. Cancer.* 9:e002899. doi: 10.1136/jitc-2021-002899
- Ichihara, F., Kono, K., Takahashi, A., Kawaida, H., Sugai, H., and Fujii, H. (2003). Increased populations of regulatory T cells in peripheral blood and tumor-infiltrating lymphocytes in patients with gastric and esophageal cancers. *Clin. Cancer Res.* 9, 4404–4408.
- Jiang, Y., Zhang, Q., Hu, Y., Li, T., Yu, J., Zhao, L., et al. (2018). ImmunoScore signature: a prognostic and predictive tool in gastric cancer. *Ann. Surg.* 267, 504–513. doi: 10.1097/sla.0000000000002116
- Jones, L., Naidoo, M., Machado, L. R., and Anthony, K. (2020). The Duchenne muscular dystrophy gene and cancer. *Cell. Oncol.* 44, 19–32. doi: 10.1007/s13402-020-00572-y
- Kang, B. W., Kim, J. G., Lee, I. H., Bae, H. I., and Seo, A. N. (2017). Clinical significance of tumor-infiltrating lymphocytes for gastric cancer in the era of immunology. *World J. Gastrointest. Oncol.* 9, 293–299. doi: 10.4251/wjgo.v9.i7.293
- Lee, H. E., Chae, S. W., Lee, Y. J., Kim, M. A., Lee, H. S., Lee, B. L., et al. (2008). Prognostic implications of type and density of tumour-infiltrating lymphocytes in gastric cancer. *Br. J. Cancer.* 99, 1704–1711. doi: 10.1038/sj.bjc.6604738
- Lee, V., Murphy, A., Le, D. T., and Diaz, L. A. Jr. (2016). Mismatch repair deficiency and response to immune checkpoint blockade. *Oncologist.* 21, 1200–1211. doi: 10.1634/theoncologist.2016-0046
- Li, B., Cui, Y., Diehn, M., and Li, R. (2017). Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer. *JAMA Oncol.* 3, 1529–1537. doi: 10.1001/jamaoncol.2017.1609
- Liu, S., Ren, J., and Ten Dijke, P. (2021). Targeting TGFβ signal transduction for cancer therapy. *Signal Transduct. Target. Ther.* 6:8. doi: 10.1038/s41392-020-00436-9
- Mao, Y., Keller, E. T., Garfield, D. H., Shen, K., and Wang, J. (2013). Stromal cells in tumor microenvironment and breast cancer. *Cancer Metastasis Rev.* 32, 303–315. doi: 10.1007/s10555-012-9415-3
- Mariathasan, S., Turley, S. J., Nickles, D., Castiglioni, A., Yuen, K., Wang, Y., et al. (2018). TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 554, 544–548. doi: 10.1038/nature25501
- Mathur, R. (2018). ARID1A loss in cancer: towards a mechanistic understanding. *Pharmacol. Ther.* 190, 15–23. doi: 10.1016/j.pharmthera.2018.05.001
- Mosaad, Y. M. (2015). Clinical role of human leukocyte antigen in health and disease. *Scand. J. Immunol.* 82, 283–306. doi: 10.1111/sji.12329
- 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, 453–457. doi: 10.1038/nmeth.3337
- Quail, D. F., and Joyce, J. A. (2013). Microenvironmental regulation of tumor progression and metastasis. *Nat. Med.* 19, 1423–1437. doi: 10.1038/nm.3394
- Remark, R., Becker, C., Gomez, J. E., Damotte, D., Dieu-Nosjean, M.-C., Sautès-Fridman, C., et al. (2015). The non-small cell lung cancer immune contexture. A major determinant of tumor characteristics and patient outcome. *Am. J. Respir. Crit. Care Med.* 191, 377–390. doi: 10.1164/rccm.201409-1671PP
- Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., et al. (2015). Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science (New York, NY)* 348, 124–128. doi: 10.1126/science.aaa1348
- Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G., and Hacohen, N. (2015). Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 160, 48–61. doi: 10.1016/j.cell.2014.12.033
- Samstein, R. M., Lee, C. H., Shoushtari, A. N., Hellmann, M. D., Shen, R., Janjigian, Y. Y., et al. (2019). Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat. Genet.* 51, 202–206. doi: 10.1038/s41588-018-0312-8
- Shen, X., Zhao, Y., Chen, X., Sun, H., Liu, M., Zhang, W., et al. (2019). Associations of PIK3CA mutations with clinical features and prognosis in gastric cancer. *Future Oncol.* 15, 1873–1894. doi: 10.2217/fon-2018-0335
- Shi, H., Jiang, Y., Cao, H., Zhu, H., Chen, B., and Ji, W. (2018). Nomogram based on systemic immune-inflammation index to predict overall survival in gastric cancer patients. *Dis. Markers* 2018:1787424. doi: 10.1155/2018/1787424
- Shim, J. H., Kim, H. S., Cha, H., Kim, S., Kim, T. M., Anagnostou, V., et al. (2020). HLA-corrected tumor mutation burden and homologous recombination deficiency for the prediction of response to PD-(L)1 blockade in advanced non-small-cell lung cancer patients. *Ann. Oncol.* 31, 902–911. doi: 10.1016/j.annonc.2020.04.004
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc.*

- Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Sun, J., Zhang, Z., Bao, S., Yan, C., Hou, P., Wu, N., et al. (2020). Identification of tumor immune infiltration-associated lncRNAs for improving prognosis and immunotherapy response of patients with non-small cell lung cancer. *J. Immunother. Cancer* 8:e000110. doi: 10.1136/jitc-2019-000110
- Tang, J., Yu, J. X., Hubbard-Lucey, V. M., Neftehinov, S. T., Hodge, J. P., and Lin, Y. (2018). Trial watch: the clinical trial landscape for PD1/PDL1 immune checkpoint inhibitors. *Nat. Rev. Drug Discov.* 17, 854–855. doi: 10.1038/nrd.2018.210
- Tauriello, D. V. F., Palomo-Ponce, S., Stork, D., Berenguer-Llargo, A., Badiarmentol, J., Iglesias, M., et al. (2018). TGF $\beta$  drives immune evasion in genetically reconstituted colon cancer metastasis. *Nature* 554, 538–543. doi: 10.1038/nature25492
- Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48, 812–30.e14. doi: 10.1016/j.immuni.2018.03.023
- Wang, B., Xu, D., Yu, X., Ding, T., Rao, H., Zhan, Y., et al. (2011). Association of intra-tumoral infiltrating macrophages and regulatory T cells is an independent prognostic factor in gastric cancer after radical resection. *Ann. Surg. Oncol.* 18, 2585–2593. doi: 10.1245/s10434-011-1609-3
- Wang, K., Kan, J., Yuen, S. T., Shi, S. T., Chu, K. M., Law, S., et al. (2011). Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat. Genet.* 43, 1219–1223. doi: 10.1038/ng.982
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-García, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4:2612. doi: 10.1038/ncomms3612
- Zeng, D., Li, M., Zhou, R., Zhang, J., Sun, H., Shi, M., et al. (2019). Tumor microenvironment characterization in gastric cancer identifies prognostic and immunotherapeutically relevant gene signatures. *Cancer Immunol. Res.* 7, 737–750. doi: 10.1158/2326-6066.cir-18-0436
- Zhang, D., He, W., Wu, C., Tan, Y., He, Y., Xu, B., et al. (2019). Scoring system for tumor-infiltrating lymphocytes and its prognostic value for gastric cancer. *Front. Immunol.* 10:71. doi: 10.3389/fimmu.2019.00071
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
- Copyright © 2021 Dai, Liu, Liu, Li, Xu, Ren and Luo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.