



Comprehensive Analysis of an Individualized Immune-Related lncRNA Pair Signature in Gastric Cancer

Chuang Nie¹, Jiabao Zhai¹, Qi Wang¹, Xiaojie Zhu¹, Guanghui Xiang¹, Chang Liu¹, Tianyu Liu¹, Wanyu Wang¹, Yimin Wang², Yashuang Zhao¹, Wenjing Tian^{1*}, Yingwei Xue^{2*} and Haibo Zhou^{1*}

¹Department of Epidemiology, College of Public Health, Harbin Medical University, Harbin, China, ²Department of Gastroenterological Surgery, Harbin Medical University Cancer Hospital, Harbin, China

OPEN ACCESS

Edited by:

Yongqian Shu,
Nanjing Medical University, China

Reviewed by:

Chang Zou,
Jinan University, China
Xin Xie,
The First Affiliated Hospital of Xi'an
Jiaotong University, China

*Correspondence:

Wenjing Tian
tianwenjing@ems.hrbmu.edu.cn
Yingwei Xue
xueyingwei@hrbmu.edu.cn
Haibo Zhou
youdeng23@163.com.com

Specialty section:

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

Received: 30 October 2021

Accepted: 02 February 2022

Published: 22 February 2022

Citation:

Nie C, Zhai J, Wang Q, Zhu X, Xiang G,
Liu C, Liu T, Wang W, Wang Y, Zhao Y,
Tian W, Xue Y and Zhou H (2022)
Comprehensive Analysis of an
Individualized Immune-Related lncRNA
Pair Signature in Gastric Cancer.
Front. Cell Dev. Biol. 10:805623.
doi: 10.3389/fcell.2022.805623

Long noncoding RNAs (lncRNAs) have diverse functions, including immune regulation. Increasing studies have reported immune-related lncRNAs in the prognosis of multiple cancers. In this study, we developed an individualized signature containing 13 immune-related lncRNA pairs (IRLPs) which could predict the overall survival, disease-free survival, progression-free survival, and disease-specific survival of gastric cancer (GC) patients in The Cancer Genome Atlas (TCGA) cohort, and internal and external validations, signature comparisons, and subgroup analyses further confirmed its superiority, stability, and generalizability. Notably, this signature also showed good applicability in discriminating the prognosis of pan-cancer patients. Then, we constructed and validated a nomogram for overall survival based on the signature and clinical factors, which allowed more accurate predictions of GC prognosis. In addition, we revealed that the low survival rate of patients with high-risk scores may be due to their aggressive clinical features, enriched cancer-related signaling pathways, the infiltration of specific immunosuppressive cells, and low tumor mutation burden. We further predicted obviously worse immunotherapeutic responses in the high-risk groups and identified some candidate compounds targeting GC risk group differentiation. This signature based on the IRLPs may be promising for predicting the survival outcomes and immunotherapeutic responses of GC patients in clinical practice.

Keywords: long noncoding RNA, immune, gastric cancer, prognostic signature, immunotherapy

INTRODUCTION

Gastric cancer (GC) is one of the most prevalent malignancies and the fourth common cause of cancer deaths worldwide with 1.09 and 0.77 million new GC cases and deaths estimated in 2020, respectively (Sung et al., 2021). Due to the insidiousness of the onset of GC, patients are often diagnosed at an advanced stage, with a median overall survival (OS) of less than 12 months when becoming metastatic at distant sites (Digkila and Wagner, 2016). The tumor node metastasis (TNM) staging system is currently the gold standard for guiding the clinical treatment of GC (Egner, 2010). However, patients with the same clinicopathological features and similar therapeutic strategies could exhibit large differences in prognosis because of the complex genetic heterogeneity, which indicates

the deficiencies of the TNM staging system in prognostic stratification (Shah and Ajani, 2010). Therefore, novel molecular markers with more clinical utilities are needed to improve prognostic prediction and guide the clinical treatment of GC patients.

Long noncoding RNAs (lncRNAs) are a common type of RNA molecules composed of over 200 nucleotides in length that can regulate gene expression through diverse mechanisms such as chromatin remodeling, genomic imprinting, transcription, and post-transcriptional processing (Rafiee et al., 2018). lncRNAs exert an essential role in the initiation and progression of cancers and have shown great potential as biomarkers in cancer diagnosis, prognosis, and treatment (Zhou et al., 2015; Bhan et al., 2017). More recently, increasing evidence has suggested that lncRNAs participate in tumorigenesis not only by influencing the genome or transcriptome but also by modulating cancer immunity (Denaro et al., 2019). To be specific, lncRNAs could regulate gene expression that is correlated with immune responses, and thus altering immune cell infiltration status in the tumor microenvironment (TME), which could profoundly affect tumor invasiveness, progression, and prognosis (Chen et al., 2017; Lazăr et al., 2018). In recent years, cancer immunotherapy, such as immune checkpoint blockade (ICB), has emerged as a breakthrough for the treatment of many cancers, especially melanoma and non-small-cell lung carcinoma (Postow et al., 2015; Rizvi et al., 2015). Although the clinical trials of ICB in the treatment of patients with advanced GC have achieved some encouraging results, the durable efficacy is limited to a small number of patients, and the objective response rate varies greatly between different studies, which impedes the development of immunotherapies for GC (Smyth et al., 2020; Zeng et al., 2021). Studies have indicated that the immunological components of TME are implicated in the antitumor processes and associated with clinical response to ICB (Lazăr et al., 2018; Zeng et al., 2021). Thus, considering the critical role of lncRNAs in regulating the immune microenvironment, identifying lncRNA signatures that are correlated with immunity may promote our understanding of GC immunobiology and help improve the clinical benefits of immunotherapies.

Using large-scale gene expression data in the public databases, recent studies have developed a variety of immune-related lncRNA based prognostic signatures for various types of cancers, including GC (Shen et al., 2020; Chen et al., 2021; Wang et al., 2021c). However, these studies are not further validated in other independent populations, which may limit their clinical applications in the survival prediction for cancer patients. In addition, these signatures are based on quantifying the expression levels of genes, which need further normalization of model gene expression due to the potential biological heterogeneity among different datasets and inevitable detection biases across different platforms (Leek et al., 2010). Fortunately, two-biomarker combinations based on the relative ranking of gene expression levels could avoid the complexity of data preprocessing and have shown great potential in cancer diagnosis and prognostic prediction (Heinäniemi et al., 2013; Li et al., 2017).

In the present work, we carried out an integrative analysis using public gene expression datasets from various GC cohorts, together with a novel pairing algorithm, to identify and validate an immune-related lncRNA pair signature (IRLPS) for improving prognostic prediction of GC. Then, we explored the underlying biological mechanisms and investigated its correlations with immune cell infiltrations, immunotherapeutic responses, and genomic mutations. In addition, we tried to figure out the potential compounds that might have therapeutic values in GC. Finally, we combined the IRLPS with clinical factors for constructing an individualized nomogram to facilitate clinical applications.

MATERIALS AND METHODS

Data Collection and Preprocessing

The workflow of this study is shown in **Supplementary Figure S1**. Firstly, we downloaded the uniformly reprocessed RNA-seq data from the *recount2* platform using the “TCGAbiolinks” R package, including 416 GC samples and 37 adjacent normal samples from The Cancer Genome Atlas (TCGA) project and 204 normal gastric samples from The Genotype-Tissue Expression (GTEx) project (Collado-Torres et al., 2017b; Mounir et al., 2019). Then, the raw count gene expression matrix and corresponding transcripts per million (TPM) normalized gene expression matrix was obtained using the “scale_counts” and “getTPM” functions of the “recount” R package, respectively (Collado-Torres et al., 2017a). Genes whose expression level was zero in more than 50% of the samples were filtered out. Clinical information of 381 patients was extracted from the UCSC Xena after excluding those with missing survival or clinical data. TCGA dataset was used as the training cohort. We also downloaded the somatic mutation data of GC patients and gene expression and survival data of TCGA pan-cancer cohort for further analysis. In addition, two external validation GC cohorts, namely the Asian Cancer Research Group (ACRG) cohort (GSE62254, N = 300) and the Singapore cohort (Singapore batches A GSE15459 and B GSE34942, N = 248), were downloaded from the Gene Expression Omnibus (GEO) database. Both of the two external expression datasets were derived from the Affymetrix Human Genome U133 Plus 2.0 (HG-U133 Plus 2.0) array and normalized on the log₂ scale. The batch effect in the Singapore cohort was removed using the ComBat algorithm of the “sva” R package as described by Lei et al. (2013) (Leek et al., 2012). All these cohorts have OS and disease-free survival (DFS) data except for the lack of DFS data in the Singapore cohort, while only the TCGA cohort has two other survival indicators, namely progression-free survival (PFS) and disease-specific survival (DSS). The clinical characteristics of each cohort were presented in **Supplementary Table S1**.

Immune-Related lncRNA Profile Mining and Differential Expression Analysis

Significant lncRNA-pathway pairs across 33 cancer types were downloaded from the ImmLnc database (Li et al., 2020). Then,

the extracted 3,044 lncRNAs related to GC immunity were defined as the immune-related lncRNAs in this study. The lncRNA expression profile in RNA-seq data was extracted according to the GENCODE (release 36) GTF file. For consistency, the microarray lncRNA expression profiles were obtained by re-annotation. Briefly, the probe sets of Affymetrix HG-U133 Plus 2.0 array were aligned to the human genome (GRCh38) with GENCODE v36 annotation using the SeqMap tool with no mismatch (Jiang and Wong, 2008). Then, by keeping the probes uniquely mapped to target gene sequences, a total of 4,231 probes covering 3,346 lncRNAs were obtained. When multiple probes matched with the same lncRNA, the average expression value of these probes was used as the expression level of corresponding lncRNA. To identify differentially expressed immune-related lncRNAs, the “limma” R package was applied to the read count matrix of the 416 GC and 241 normal samples from the *recount2* project, and the thresholds were set as $|\log_2 \text{fold change}| (|\log_2 \text{FC}|) > 1.0$ and false discovery rate (FDR) < 0.05 (Ritchie et al., 2015). Since the lncRNA expression profiles in TCGA and GEO databases were different, we focused on the overlapped lncRNAs in the present study.

Definition and Construction of the Immune-Related lncRNA Pairs

In this study, we used the IRLPs to develop a prognostic signature for GC, in which the expression levels of lncRNAs were pairwise compared in each sample to generate a score for each paired combination. Briefly, if the expression level of lncRNA 1 was higher than that of the lncRNA 2, the score of the IRLP was assigned 1, otherwise, the score was 0. In addition, those IRLPs with a constant value (scores of 0 or 1 assigned to more than 80% of the samples in any dataset) were removed because they may represent the platform-dependent bias measurements and do not contain sufficient discriminative information to predict patients' survivals.

Construction and Validation of a Prognostic Signature Based on Immune-Related lncRNA Pairs

Univariate Cox regression analysis was performed in the training cohort based the constructed IRLPs matrix and corresponding clinical data, and prognostic IRLPs with $p < 0.01$ were chosen as the candidates. Then, a Cox regression model that combined with the least absolute shrinkage and selection operator (LASSO) algorithm was further adopted to select a panel of IRLPs to develop the prognostic signature. The optimal tuning parameter was estimated by 10-fold cross-validation at one standard error beyond the minimum partial likelihood deviance. Finally, the selected IRLPs were incorporated into a stepwise based multivariate Cox regression model to construct the IRLPS. The risk score was calculated by adding up each IRLP score, with the weights assigned by the corresponding coefficients generated in the multivariate Cox analysis.

Based on the median risk score in the training cohort, patients in each cohort were classified into the high- and low-risk groups. The Kaplan-Meier method was applied to compare the survival between the two groups. Time-dependent receiver operator characteristic (ROC) curves were adopted to evaluate the predictive performance of the IRLPS. Univariate and multivariate Cox regression analyses were performed to assess the clinical significance of the IRLPS. In addition, stratified analyses were conducted to assess the robustness of the signature in predicting GC prognosis in different clinical subgroups. Furthermore, we compared our IRLPS with four published prognostic signatures that were also constructed using the TCGA data based on both time-dependent ROC and concordance index (C-index) analysis (Wang et al., 2021a; Mao et al., 2021b; Wang et al., 2021c; Xu et al., 2021).

Functional Annotations and Gene Set Enrichment Analysis

The differentially expressed genes (DEGs) between the high- and low-risk groups were analyzed by the “limma” R package, and the DEGs with $|\log_2 \text{FC}| > 1.0$ and $\text{FDR} < 0.05$ were subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses by using the “clusterProfiler” R package (Yu et al., 2012). In addition, we performed a pre-ranked Gene Set Enrichment Analysis (GSEA) to identify the cancer hallmarks and immunologic signatures that are correlated with the IRLPS (Subramanian et al., 2005). The hallmark gene sets (h.all.v7.4. entrez.gmt) and immunologic gene sets (c7. all.v7.4. entrez.gmt) were obtained from the Molecular Signatures Database. Gene sets with nominal $p < 0.05$ and $\text{FDR} < 0.25$ were considered significantly enriched.

Exploration of Immunological Features

Several generally acknowledged methods, including CIBERSORT, EPIC, MCP-counter, xCell, ESTIMATE were applied to analyze the infiltrating characteristics of immune cells and stromal cells based on the gene expression data of tumor tissues (Yoshihara et al., 2013; Newman et al., 2015; Becht et al., 2016; Aran et al., 2017; Racle et al., 2017). Then, the correlations between the IRLPS and immune/stromal cells were calculated by Spearman correlation analysis. In addition, since immune checkpoint gene expression may be correlated with the treatment responses of ICB therapy, we further assessed the associations between the IRLPS and several key immune checkpoint molecules (Goodman et al., 2017).

Prediction of Immunotherapeutic Response

To explore the response of immunotherapy in different risk groups, we applied the Tumor Immune Dysfunction and Exclusion (TIDE) algorithm to generate predicted results about ICB therapy (anti-PD1 and anti-CTLA4) for each GC patient (Jiang et al., 2018; Fu et al., 2020). Given that tumor mutation burden (TMB) has emerged as a promising biomarker for predicting the immunotherapeutic response, we also calculated the TMB of GC patients in the TCGA cohort. TMB was defined as the total number of nonsynonymous mutations

per megabase in the coding area of a tumor genome (Wang et al., 2021b). The “maftools” and “GenVisR” R packages were used to analyze and visualize common mutations (Skidmore et al., 2016; Mayakonda et al., 2018).

Calculation of Stemness Index

Cancer stemness has been reported to be associated with suppressed immune response and significantly poor outcome for most cancers (Miranda et al., 2019). Thus, we investigated whether the established IRLPS was correlated with the stemness phenotype of GC. As previously reported, we utilized the one-class logistic regression (OCLR) algorithm to calculate the RNA expression-based stemness index (mRNAsi) for each tumor sample (Malta et al., 2018). The mRNAsi was mapped to the range of 0–1 through a linear transformation that subtracted the minimum and divided by the maximum.

Connectivity Map Analysis

To explore potential compounds that might be useful for GC treatment, the top 1,000 DEGs (FDR<0.05) between high- and low-risk groups were employed to query the Connectivity Map (CMap) database (Subramanian et al., 2017). Then, compounds with a negative enrichment score and $p < 0.05$ were selected for the subsequent mechanism of action (MoA) analysis.

Construction and Validation of an Individualized Nomogram

To facilitate the clinical utilization of the IRLPS, we established a nomogram in the TCGA cohort by integrating the IRLPS with independent clinical prognostic factors from multivariate Cox regression. Then, the nomogram was applied to the ACRG cohort for external validation (the Singapore cohort was not included due to the lack of corresponding clinical data). C-index and time-dependent ROC analyses were both applied to evaluate the performance of the nomogram. Moreover, calibration curves were used to display the differences between the actual observed and predicted survival possibility. Decision curve analyses (DCA) were used to assess the net benefits of different clinical decisions.

Statistical Analysis

All the statistical analyses were conducted on the R software (version 4.0.0). The R packages used in this study are listed in **Supplementary Table S2**. Internal validation was performed using the bootstrap method. The Wilcoxon rank-sum test and Kruskal-Wallis test were used to evaluate the correlations of IRLPS score with clinical factors, as well as the correlations of other evaluated factors (TIDE score, mRNAsi, and TMB et al.) with IRLPS groups. The chi-square test was used to analyze the differences of specific gene mutations among IRLPS groups. In addition, Spearman correlation analyses were performed to investigate the relationship between IRLPS score and immune cells and immune checkpoints. $p < 0.05$ was considered statistically significant. Bonferroni correction was performed in the subgroup analysis.

RESULTS

Differentially Expressed Immune-Related lncRNAs

A total of 970 differentially expressed immune-related lncRNAs were identified in GC, among which 429 were upregulated and 541 were downregulated (**Supplementary Figure S2A**). Then, we crossed these lncRNAs with 3,346 re-identified lncRNAs in the microarray datasets to obtain 312 shared lncRNAs for the following analysis (**Supplementary Figure S1** and **Supplementary Figure S2B**), and the details were shown in **Supplementary Table S3**.

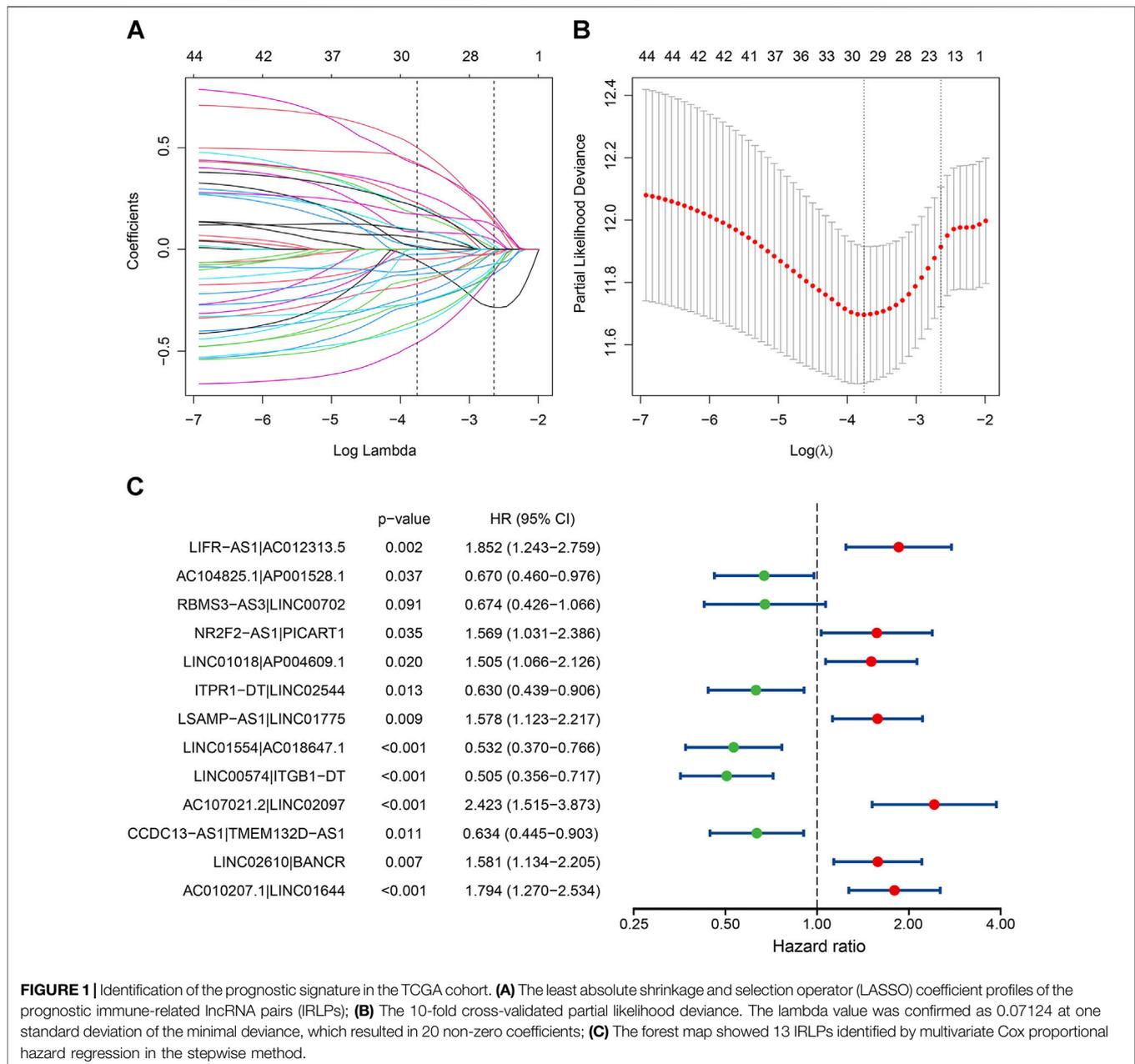
Construction and Internal Evaluation of the Immune-Related lncRNA Pairs

A total of 48,516 IRLPs were built by pairwise comparison using the 312 shared lncRNAs, and 2,186 of them were kept after excluding those with a constant value. With univariate Cox analysis, 44 OS-related IRLPs were identified in the training cohort ($p < 0.01$) (**Supplementary Table S4**). Then, these prognostic IRLPs were subjected to LASSO-Cox regression analysis. The LASSO coefficient profiles and partial likelihood deviance plot were shown in **Figures 1A,B**, illustrating the selected 20 IRLPs. Then, we further applied the multivariate Cox analysis and ultimately obtained 13 IRLPs for constructing the IRLPS and generating the risk scores (**Figure 1C**).

Based on the median cutoff value, patients in the training cohort were divided into the high- and low-risk groups. The Kaplan-Meier survival curves revealed that patients in the high-risk group had significantly poorer OS than those in the low-risk group ($p < 0.0001$) (**Figure 2A**). Time-dependent ROC analysis suggested that our IRLPS had a good performance in predicting OS of GC patients with a 5-years average area under the curve (AUC) of 0.817 (**Figure 2B**). The heatmap of the 13 IRLPs showed that a score of 1 in AC107021.2|LINC02097, NR2F2-AS1|PICART1, AC010207.1|LINC01644, LINC01018|AP004609.1, LINC02610|BANCR, LSAMP-AS1|LINC01775, and LIFR-AS1|AC012313.5 was more distributed in the high-risk group, indicating their harmful roles in GC prognosis, while the remaining IRLPs showed the reverse trend (**Figure 2C**). In addition, the distribution of the risk scores also indicated poor survival in the high-risk group patients (**Figure 2D**). Interestingly, we further found that patients in the high-risk groups had a poorer DFS, PFS, and DSS compared with those in the low-risk groups (all $p < 0.0001$), and the high predictive ability of the IRLPS was revealed by the time-dependent ROC analysis (5-years average AUCs of 0.737, 0.764 and 0.813, respectively) (**Supplementary Figures S3A–S3F**). Moreover, internal validations by bootstrapping verified the performance of the signature in predicting different survival indicators (**Supplementary Table S5**).

External Validation of the Immune-Related lncRNA Pairs

To assess the robustness of the IRLPS in the independent external population, the formula and the cutoff value used in the training



cohort were applied to two GEO datasets. The results showed that patients in the high-risk groups also exhibited obviously worse OS than those in the low-risk groups for both the ACRG ($p = 0.002$) and Singapore cohort ($p = 0.0028$) (Figures 2E,G), and the corresponding 5-years average AUC values were 0.612 and 0.649, respectively (Figures 2F,H). Similar results were observed in the entire GEO validation cohort and the whole cohort (all $p < 0.0001$; 5-years average AUCs of 0.625 and 0.679, respectively) (Supplementary Figures S4A–S4D). In terms of the DFS, the discriminative ability of the IRLPS was also validated in the ACRG cohort ($p = 0.0031$; 5-years average AUC of 0.613) and the combined cohort ($p < 0.0001$; 5-years average AUC of 0.644) (Supplementary Figures S5A–S5D). Notably, we were pleasantly

surprised to find that the IRLPS could also predict all survival indicators of digestive tract cancers and even pan-cancer with corresponding low-risk groups showed survival advantages (all $p < 0.01$) (Supplementary Figures S6A–S6H). Furthermore, compared with other traditional prognostic signatures, the IRLPS showed an obviously higher accuracy for OS prediction in GC patients as demonstrated by the ROC and C-index analyses (Supplementary Figures S7A, S7B).

Subgroup Analyses

According to the clinical traits, including age, gender, tumor stage, grade, T stage, N stage, M stage, and Lauren subtype, GC patients in both separate and combined cohorts were divided into

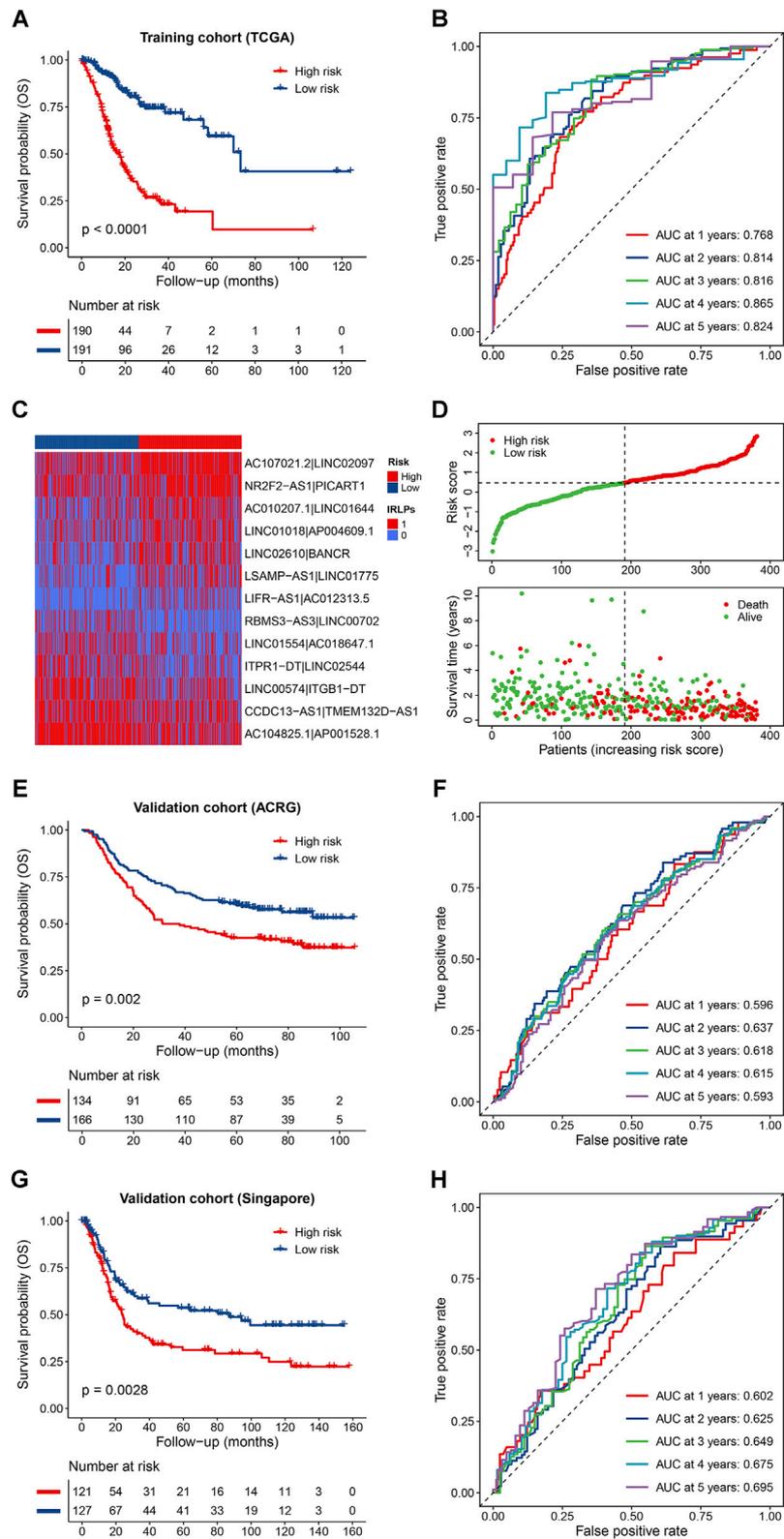


FIGURE 2 | Predictive effects of the prognostic signature. **(A,B)** Kaplan-Meier curves and time-dependent receiver operator characteristic (ROC) curves of the signature for predicting the overall survival (OS) of patients in the TCGA cohort; **(C,D)** Heatmap of the immune-related lncRNA pairs (IRLPs) in the high- and low-risk groups, rank of the risk score, and distribution of the survival status; **(E,F)** Kaplan-Meier curves and time-dependent ROC curves of the signature for predicting the OS of patients in the Asian Cancer Research Group (ACRG) cohort; **(G,H)** Kaplan-Meier curves and time-dependent ROC curves of the signature for predicting the OS of patients in the Singapore cohort.

TABLE 1 | Univariate and multivariate Cox regression analysis of clinicopathologic factors and IRLPS for predicting OS in the training and validation cohorts.

For OS variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	p Value	HR (95% CI)	p Value
TCGA training cohort (N = 381) ^a				
Age (≥60 vs. <60 years)	1.574 (1.084–2.286)	0.017	1.740 (1.186–2.553)	0.005
Gender (Male vs. Female)	1.216 (0.861–1.718)	0.267	—	—
Histological grade (G3 vs. G1-2)	1.436 (1.016–2.030)	0.040	1.345 (0.945–1.916)	0.100
T stage (T3-4 vs. T1-2)	1.748 (1.157–2.643)	0.008	1.362 (0.857–2.163)	0.191
N stage (N1-3 vs. N0)	2.014 (1.342–3.022)	7.21e-04	1.017 (0.609–1.700)	0.948
M stage (M1 vs. M0)	2.213 (1.272–3.849)	0.005	2.242 (1.254–4.007)	0.006
Tumor stage (Stage III-IV vs. Stage I-II)	2.057 (1.456–2.905)	4.25e-05	1.652 (1.027–2.657)	0.039
IRLPS (high- vs. low-risk)	4.541 (3.129–6.591)	1.68e-15	4.549 (3.096–6.685)	1.22e-14
ACRG validation cohort (N = 300)				
Age (≥60 vs. <60 years)	1.239 (0.882–1.741)	0.217	—	—
Gender (Male vs. Female)	0.905 (0.647–1.265)	0.559	—	—
T stage (T3-4 vs. T1-2)	2.396 (1.741–3.297)	8.22e-08	1.336 (0.902–1.981)	0.149
N stage (N1-3 vs. N0)	2.816 (1.434–5.527)	0.003	1.729 (0.847–3.531)	0.133
M stage (M1 vs. M0)	3.840 (2.482–5.942)	1.52e-09	2.989 (1.872–4.772)	4.51e-06
Lauren subtype (Diffused vs. Intestinal)	1.677 (1.198–2.347)	0.003	1.130 (0.789–1.617)	0.505
Lauren subtype (Mixed vs. Intestinal)	2.128 (1.167–3.879)	0.014	1.705 (0.928–3.131)	0.086
Tumor stage (Stage III-IV vs. Stage I-II)	3.408 (2.341–4.960)	1.53e-10	2.092 (1.293–3.386)	0.003
IRLPS (high- vs. low-risk)	1.646 (1.196–2.264)	0.002	1.646 (1.190–2.277)	0.003
Singapore validation cohort (N = 248)				
Age (≥60 vs. <60 years)	0.973 (0.659–1.437)	0.889	—	—
Gender (Male vs. Female)	1.128 (0.775–1.639)	0.530	—	—
Lauren subtype (Diffused vs. Intestinal)	1.308 (0.896–1.911)	0.164	—	—
Lauren subtype (Mixed vs. Intestinal)	0.948 (0.471–1.907)	0.880	—	—
Tumor stage (Stage III-IV vs. Stage I-II)	6.136 (3.725–10.110)	1.05e-12	5.995 (3.633–9.892)	2.40e-12
IRLPS (high- vs. low-risk)	1.726 (1.201–2.479)	0.003	1.578 (1.093–2.278)	0.015

IRLPS: immune-related lncRNA, pair signature; OS: overall survival; HR: hazard ratio; CI: confidence interval; TCGA: The cancer genome atlas, ACRG: Asian cancer research group.

^aPatients with Gx (histological grade cannot be assessed) or Nx (lymph nodes cannot be assessed) or Mx (distant metastasis cannot be assessed) were included in the analysis but were not displayed in this table.

sixteen subgroups. As shown in **Supplementary Table S6**, the IRLPS could significantly discriminate the OS of patients in every subgroup of the TCGA cohort (all $p < 0.025$), and similar results were observed in the validation cohorts especially for the male, advanced stage (stage III-IV), N1-3 stage, M0 stage, and intestinal adenocarcinoma groups (all $p < 0.025$). In addition, we further performed subgroup analyses in the DFS datasets and suggested that the IRLPS remained an indicator of the DFS in the younger (<60 years), male, advanced stage, high grade (G3), N1-3 stage, M0 stage, and intestinal adenocarcinoma groups (all $p < 0.025$) (**Supplementary Table S7**).

Independent Prognostic Role of the Immune-Related lncRNA Pairs

To investigate whether the IRLPS was an independent prognostic factor in GC, we conducted the Cox regression analysis. Univariate Cox regression analysis indicated that the high-risk groups had poorer OS than the low-risk groups in the TCGA (HR = 4.541, 95% CI: 3.129–6.591, $p = 1.68 \times 10^{-15}$), ACRG (HR = 1.646, 95% CI: 1.196–2.264, $p = 0.002$), and Singapore cohorts (HR = 1.726, 95% CI: 1.201–2.479, $p = 0.003$) (**Table 1**). After adjusting for the clinical features in the multivariate Cox analysis, we found that the IRLPS was an independent prognostic factor in the TCGA (HR = 4.549, 95% CI: 3.096–6.685, $p = 1.22 \times 10^{-14}$), ACRG (HR = 1.646, 95% CI: 1.190–2.277, $p = 0.003$), and

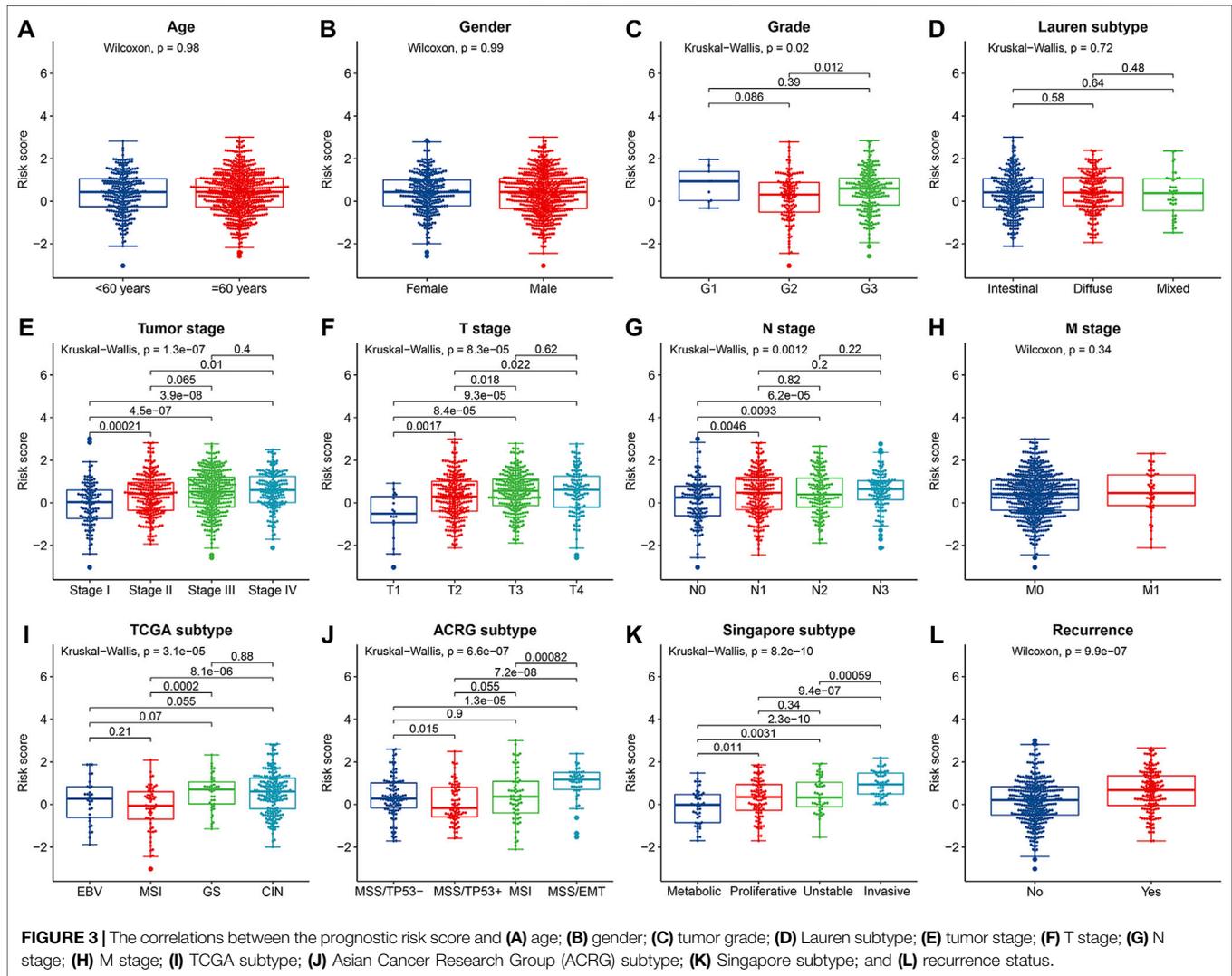
Singapore cohorts (HR = 1.578, 95% CI: 1.093–2.278, $p = 0.015$) (**Table 1**). In addition, the IRLPS was also independently associated with the DFS of GC patients in the TCGA (HR = 3.342, 95% CI: 1.804–6.192, $p = 1.26 \times 10^{-4}$) and ACRG cohorts (HR = 1.600, 95% CI: 1.118–2.290, $p = 0.010$) (**Supplementary Table S8**).

Clinical Correlation Analysis for the Immune-Related lncRNA Pairs

The correlations between the IRLPS and the clinicopathological features were analyzed in the whole cohort. The results showed that the IRLPS was not associated with age, gender, M stage, and Lauren subtype. However, we found significant correlations between the IRLPS and tumor grade ($p = 0.020$), tumor stage ($p = 1.30 \times 10^{-7}$), T stage ($p = 8.30 \times 10^{-5}$), N stage ($p = 0.0012$), recurrence status ($p = 9.90 \times 10^{-7}$), and the molecular subtypes of TCGA ($p = 3.10 \times 10^{-5}$), ACRG ($p = 6.60 \times 10^{-7}$), and Singapore cohorts ($p = 8.20 \times 10^{-10}$) (**Figures 3A–3L**), suggesting that the IRLPS may be involved in the progression of GC.

Functional Enrichment Analysis

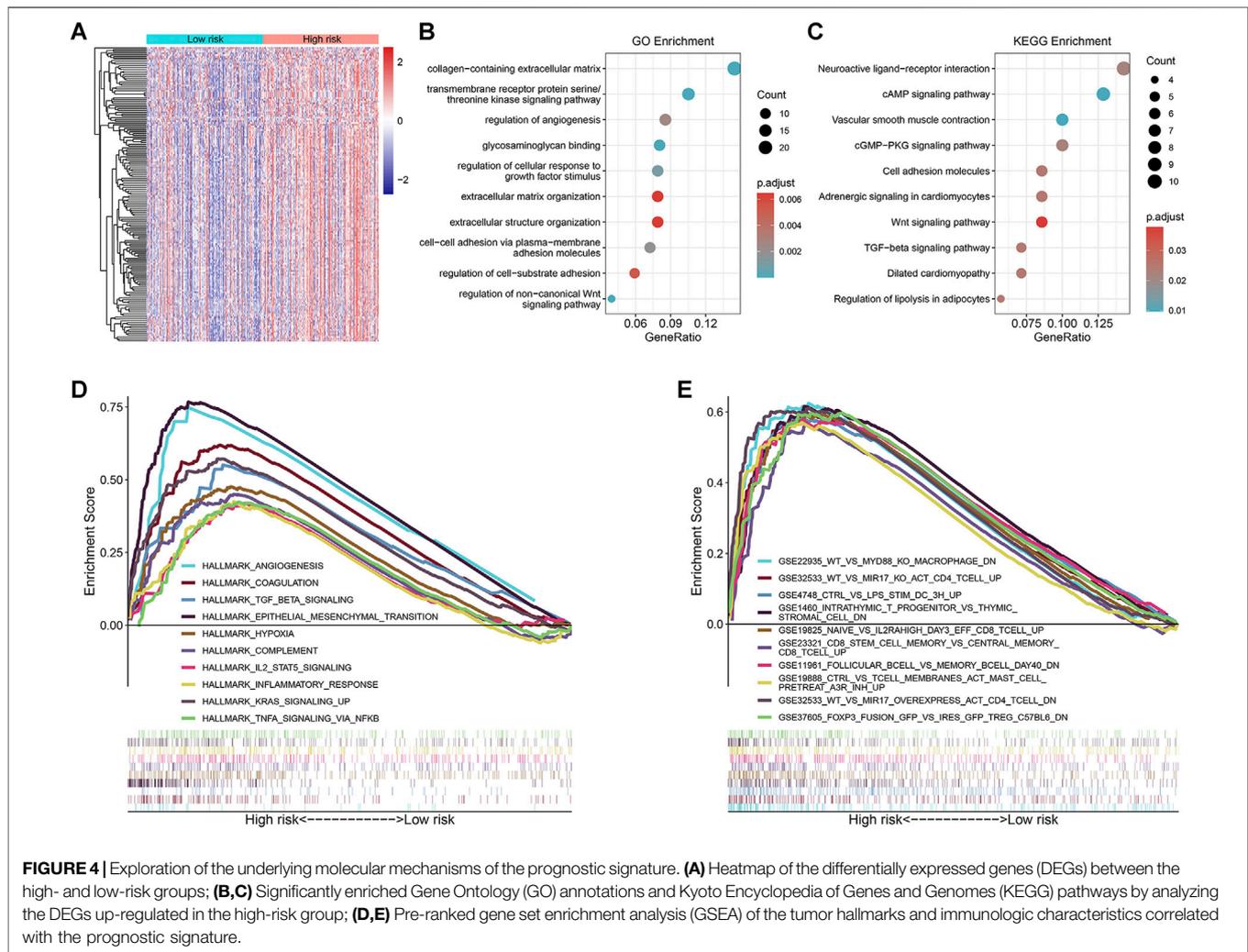
Using gene expression data of the TCGA cohort, 165 significant DEGs were identified between the high- and low-risk groups (**Figure 4A**). Among them, 163 genes were up-regulated in the high-risk group and only 2 genes were



down-regulated (Supplementary Table S9). Therefore, we conducted GO annotation and KEGG enrichment analysis using the up-regulated DEGs and found that these genes were enriched in GO terms such as collagen-containing extracellular matrix (ECM), regulation of angiogenesis, and regulation of cellular response to growth factor stimulus (Figure 4B and Supplementary Table S10), and KEGG pathways such as cAMP, cGMP-PKG, and TGF- β signaling (Figure 4C and Supplementary Table S11). In addition, subsequent pre-ranked GSEA of tumor hallmarks revealed that the high-risk group also exhibited an obvious enrichment of common pathways including angiogenesis, epithelial-mesenchymal transition (EMT), and TGF- β signaling (Figure 4D and Supplementary Table S12). Moreover, we observed the enrichment of numerous immunologic signatures such as follicular B cells vs. memory B cells down and naive vs. effector CD8 T cells up in the high-risk group (Figure 4E and Supplementary Table S13), which implied the immune-related modulation of the signature.

Estimation of the Immune Cell Infiltration and Immunosuppressive Molecules With Immune-Related lncRNA Pairs

Since the IRLPS were initially connected with immune pathways, we consequently explored whether this signature was correlated with tumor immune microenvironment based on the gene expression data of the three cohorts. We found that activated tumor-infiltrating immune cells such as CD4⁺ T cells and T helper cells were more negatively correlated with this signature, while monocytes, macrophages, mast cells, endothelial cells, and cancer-associated fibroblasts showed positive correlations in the TCGA cohort (all $p < 0.05$) (Figure 5A). Similar results were observed in the ACRG and Singapore cohorts (Supplementary Figure S8A and Supplementary Figure S9A). In addition, the stromal score from the ESTIMATE algorithm was positively associated with the IRLPS, while the immune score showed no correlation (Figure 5A, Supplementary Figure S8A and Supplementary Figure S9A). The above results are detailed in Supplementary Table S14. Furthermore, we found that the IRLPS score was generally



negatively correlated with the expression of *CLTA-4*, *LGALS9*, and *HVEM* (all $p < 0.05$) and positively correlated with *PD-L2* and *CD276* in the three cohorts (all $p < 0.05$), but no significant associations were observed between *PD-1* and *PD-L1* expression and the signature (Figure 5B, Supplementary Figure S8B and Supplementary Figure S9B).

Prediction of the Response to Immunotherapy in Different Risk Groups

Using the TIDE algorithm, we found that the high-risk group had a higher TIDE score than the low-risk group ($p = 3.2 \times 10^{-7}$) (Figure 5C), suggesting that the high-risk patients were less likely to benefit from ICB therapy, which was in line with the lower predicted proportion of responders in the high-risk group (28.95 vs. 51.83%, $p = 8.71 \times 10^{-6}$) (Figure 5G). In addition, we found that the high-risk group had higher T cell exclusion and dysfunction scores but lower microsatellite instability scores (all $p < 0.05$) (Figures 5D–F). Consistent prediction results were observed in both the ACRG and Singapore cohorts regarding the immunotherapy response rates, TIDE scores, and T cell

exclusion scores (all $p < 0.05$) (Supplementary Figures S8C–S8G and Supplementary Figures S9C–S9G). Besides, we observed obvious differences in the distribution of predicted responders across the ACRG ($p = 2.32 \times 10^{-8}$) and Singapore molecular subtypes ($p = 7.75 \times 10^{-8}$), with the EMT (2.17%) and invasive subtypes (8.33%) exhibited the worst response rates, respectively (Supplementary Figure S8G and Supplementary Figure S9G). No significant difference was found regarding the TCGA subtypes ($p = 0.080$) (Figure 5G). Interestingly, we observed higher mRNasi in the low-risk-groups compared with the high-risk groups (all $p < 0.01$) (Figure 5H, Supplementary Figure S8H and Supplementary Figure S9H), implying the better prognosis of patients with higher mRNasi as previously reported (Mao et al., 2021a). Furthermore, patients in the low-risk group had obviously higher TMB than those in the high-risk group ($p = 0.00095$) (Figure 5I), which further supported that the low-risk patients may have enhanced responses in ICB therapy.

Somatic Mutation Analysis

Next, we performed somatic mutation analysis to investigate the genomic differences between the two risk groups of the TCGA cohort.

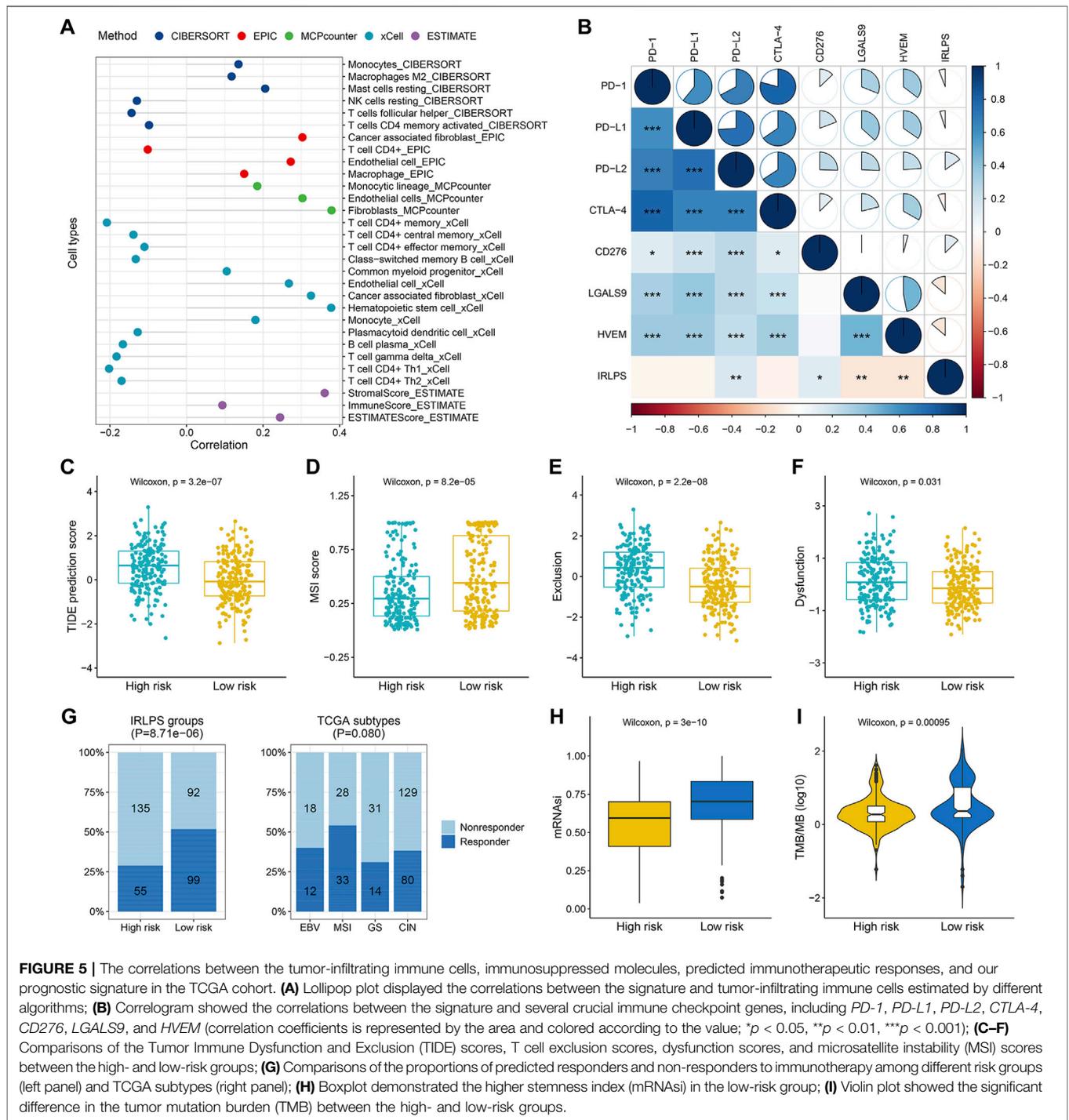
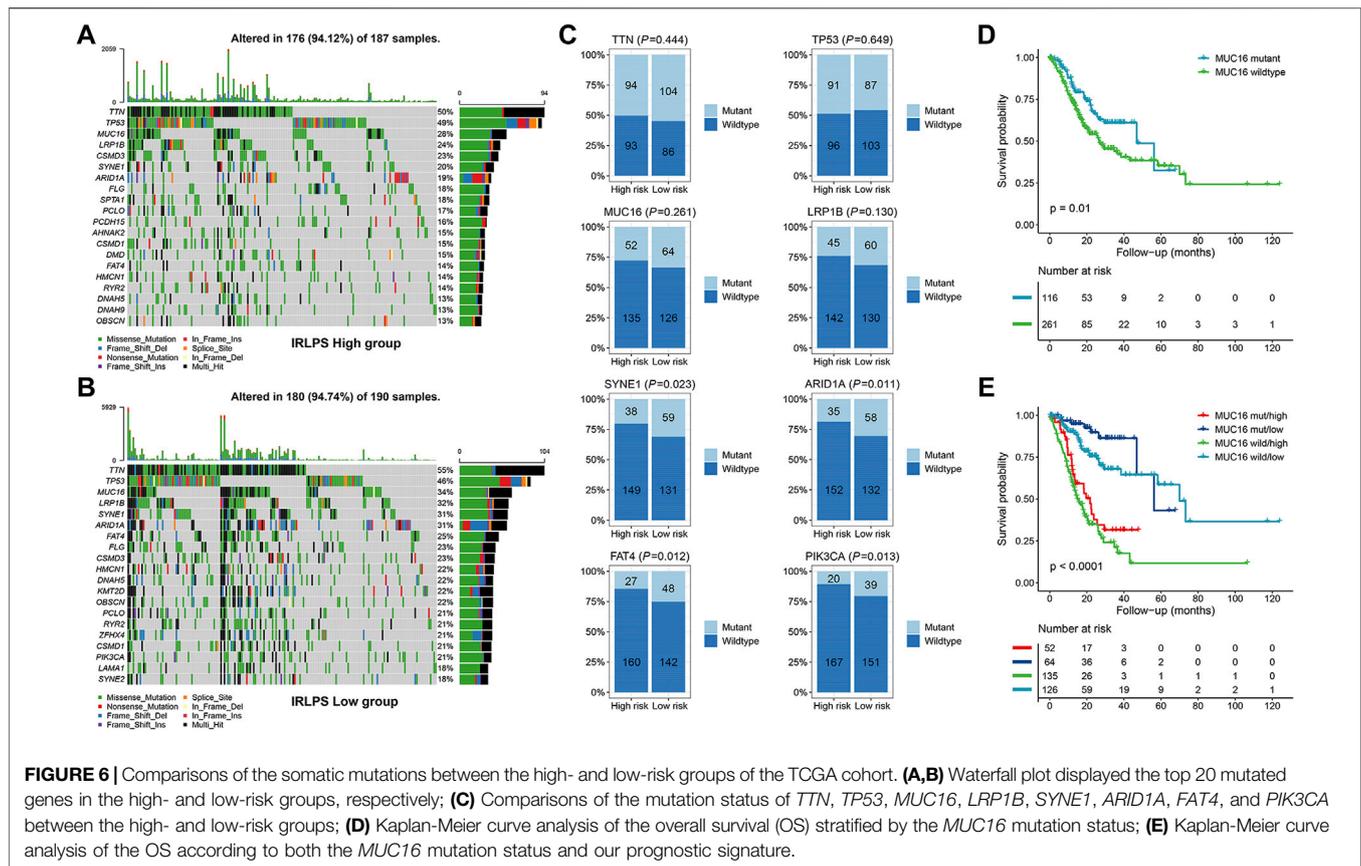


FIGURE 5 | The correlations between the tumor-infiltrating immune cells, immunosuppressed molecules, predicted immunotherapeutic responses, and our prognostic signature in the TCGA cohort. **(A)** Lollipop plot displayed the correlations between the signature and tumor-infiltrating immune cells estimated by different algorithms; **(B)** Correlation heatmap showed the correlations between the signature and several crucial immune checkpoint genes, including *PD-1*, *PD-L1*, *PD-L2*, *CTLA-4*, *CD276*, *LGALS9*, and *HVEM* (correlation coefficients is represented by the area and colored according to the value; **p* < 0.05, ***p* < 0.01, ****p* < 0.001); **(C–F)** Comparisons of the Tumor Immune Dysfunction and Exclusion (TIDE) scores, T cell exclusion scores, dysfunction scores, and microsatellite instability (MSI) scores between the high- and low-risk groups; **(G)** Comparisons of the proportions of predicted responders and non-responders to immunotherapy among different risk groups (left panel) and TCGA subtypes (right panel); **(H)** Boxplot demonstrated the higher stemness index (mRNAs) in the low-risk group; **(I)** Violin plot showed the significant difference in the tumor mutation burden (TMB) between the high- and low-risk groups.

The waterfall plot presented the top 20 genes with the highest mutation frequency in the high- and low-risk groups (Figures 6A,B). Therein, *TTN*, *TP53*, *MUC16*, and *LRP1B* were the most mutated genes in both the groups but no significant differences were found for their mutation frequencies (all *p* > 0.05) (Figure 6C). However, other common mutated genes in GC such as *SYNE1* (*p* = 0.023), *ARID1A* (*p* = 0.011), *FAT4* (*p* = 0.012), and *PIK3CA* (*p* = 0.013) exhibited significantly higher mutation frequencies in the low-

risk patients (Figure 6C). *MUC16* mutations were correlated with favorable survival outcomes in GC patients (Figure 6D) and may be applicable for prognostic prediction and immunotherapeutic guidance for GC (Li et al., 2018). Hence, we explored whether *MUC16* mutation combined with the IRLPS could yield different prognoses in GC patients. The results showed that the risk groups determined by the IRLPS showed significant survival differences in both the *MUC16* mutation (*MUC16* mut/high vs. *MUC16* mut/low,



$p < 0.0001$) and wildtype subgroups (*MUC16* wild/high vs. *MUC16* wild/low, $p < 0.0001$) (Figure 6E), indicating that the predictive ability of the IRLPS was not affected by *MUC16* mutation status. Intriguingly, patients in the *MUC16* mut/high group had a worse OS than patients in the *MUC16* wild/low group ($p < 0.0001$). Moreover, the *MUC16* mut/low group had the best OS rate, and the *MUC16* wild/high group had the worst OS rate (Figure 6E).

Identification of Potential Compounds Targeting the Immune-Related lncRNA Pairs

According to the MoA analysis, there were 29 corresponding molecular pathways targeted by 41 compounds (Figure 7 and Supplementary Table S15). Among them, adrenergic receptor antagonist was the most critical MOA which was shared by six compounds including carteolol, nadolol, pindolol, terazosin, timolol, and vincamine. Then, three compounds (diphenhydramine, thioperamide, and trimethobenzamide) shared the same MoA as histamine receptor antagonist, two compounds (indoprofen and naproxen) shared the same MoA as cyclooxygenase inhibitor, and two compounds (lisuride and quinpirole) shared the same MoA as dopamine receptor agonist. Further studies are warranted to assess the potential therapeutic values of these compounds in GC.

Development and Evaluation of the Prognostic Nomogram

Multivariate Cox analysis showed that age, M stage, tumor stage, and IRLPS were independent prognostic factors for the OS in the TCGA cohort. Thus, we developed an individualized nomogram using these variables to further improve the prediction of the OS for GC patients (Figure 8A). The C-index of the nomogram was 0.760 (95% CI: 0.719–0.801) in the TCGA cohort, and similar results were observed when using bootstrapping for the internal validation (C-index: 0.762, 95% CI: 0.722–0.800) (Supplementary Table S16). The nomogram was further validated in the independent ACRG cohort (C-index: 0.653, 95% CI: 0.608–0.698), and it had higher C-index than any other individual included factors such as the tumor stage (Supplementary Table S16). In addition, the ROC curves were also displayed to illustrate the high predictive accuracy of the nomogram (Figures 8B,D). Subsequently, we draw the calibration curves for the OS prediction at 1, 3, and 5 years and revealed the high agreement between the predictions by nomogram and actual observations in both the TCGA and ACRG cohorts (Figures 8C,E). Furthermore, the DCA curves showed that the net benefits of the nomogram were significantly higher than the limit curves (Supplementary Figures S10A, S10B), indicating the good clinical applicability of the nomogram.

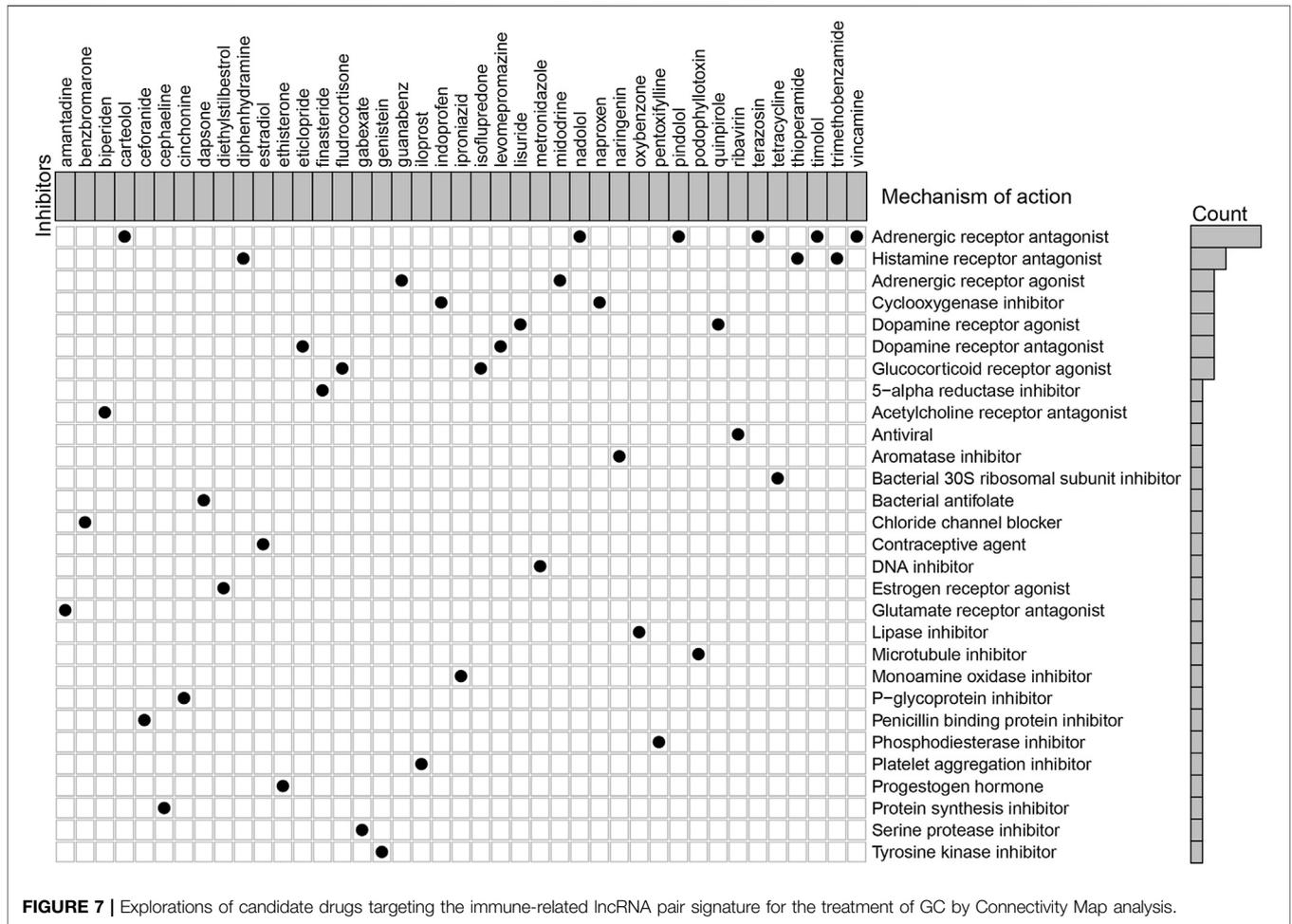


FIGURE 7 | Explorations of candidate drugs targeting the immune-related lncRNA pair signature for the treatment of GC by Connectivity Map analysis.

DISCUSSION

GC is a common malignant tumor originating in the digestive tract, of which the prognostic outcome is undesirable despite improved therapeutic strategies (Smyth et al., 2020). Due to the complex tumor heterogeneity and oncogenic mechanisms, developing an individualized prognostic evaluation system for GC that can guide precise clinical treatment remains a challenge. Growing evidence has revealed that tumor progression depends not only on its intrinsic malignant characteristics but also on the TME (Wang et al., 2019). The immune responses in the TME are crucial determinants of tumorigenesis, antitumor immunity, and prognosis (Zeng et al., 2019). As important immune regulators, lncRNAs could impact the prognosis of cancer patients and serve as potential biomarkers for cancer therapies (Zhou et al., 2015; Denaro et al., 2019; Li et al., 2020). Therefore, we focused on the lncRNAs that are correlated with immune modulation to establish the prognostic signature for GC.

Previous studies have published many genetic prognostic models using traditional method that focused on the exact expression levels of coding genes or noncoding RNAs (Wang et al., 2021a; Mao et al., 2021b; Wang et al., 2021c; Xu et al., 2021). However, these signatures may have limited applicability for the

evaluation of the prognosis of GC because clinicians have to quantify the gene expression to a standardized unit. Here, we adopted a novel pairing method to transform the lncRNA expression profile to 0-or-1 IRLPs matrix and developed a reasonable prognostic signature named IRLPS in GC (Li et al., 2017). Since the IRLPs were generated entirely based on the relative ranking of their expression in the same patients, there is no need to conduct scaling and normalization to make gene expression levels comparable among patients or datasets. In this study, we enrolled three independent relatively large GC cohorts intending to verify the effectiveness of the method in different populations and measurement platforms. In particular, the lncRNA expression profiles in the GEO microarray datasets were mined by the re-annotation method. Previous studies have shown the feasibility of the profiling of lncRNAs through microarray probe re-annotation (Zhang et al., 2012; Zhou et al., 2015; Wang et al., 2018). Considering the differences of the lncRNA expression profiles extracted from the TCGA and GEO database, we focused on the shared lncRNAs between the two data sources to ensure that the analyzed lncRNAs are general and universal across the GC samples.

A total of 13 IRLPs comprising 26 lncRNAs were selected to construct the prognostic model through the combination of

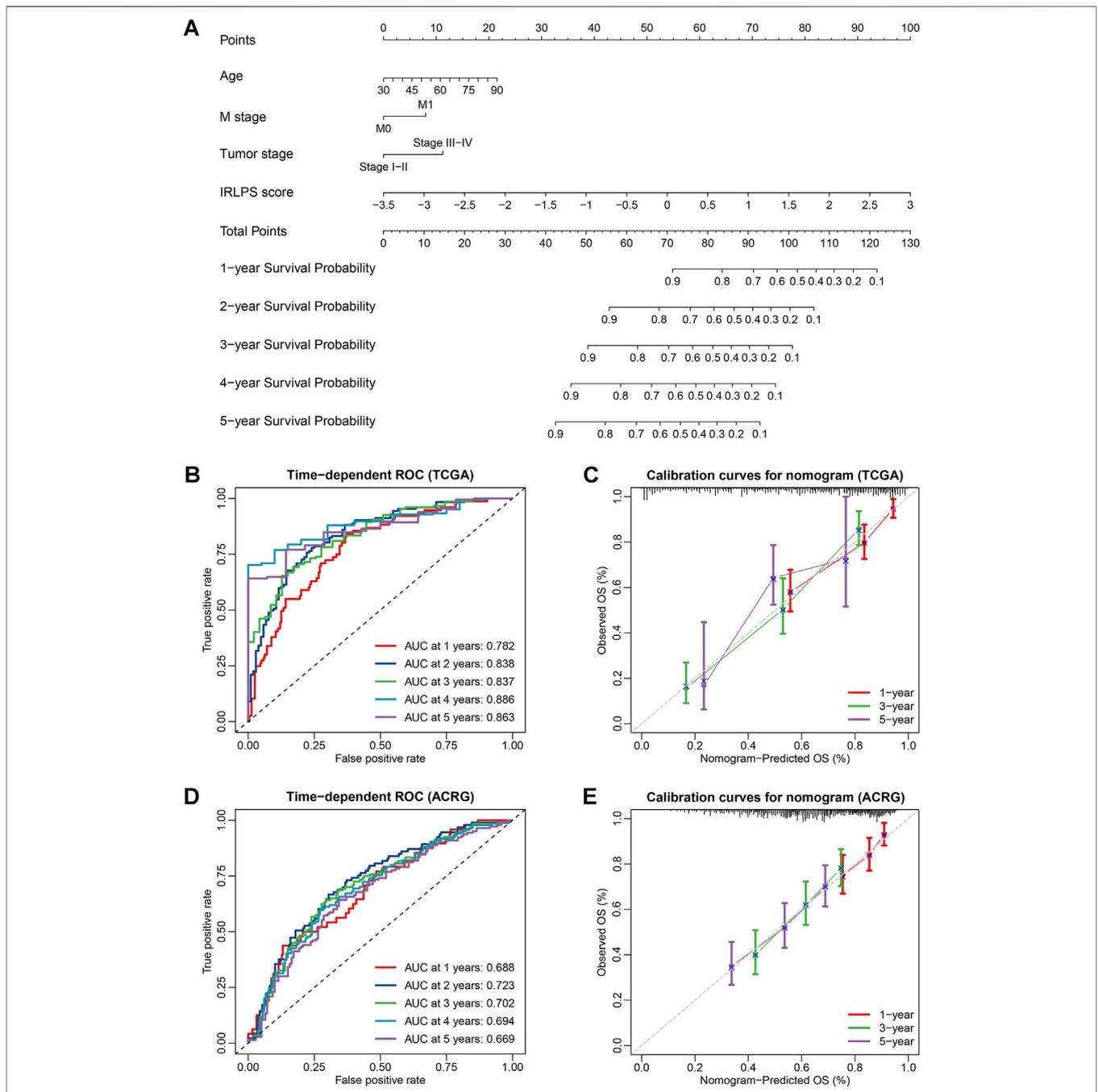


FIGURE 8 | Construction and validation of an individualized nomogram for the overall survival (OS) prediction based on independent prognostic factors. **(A)** A nomogram for the prediction of OS integrating our prognostic signature with independent clinical variables (age, M stage, and tumor stage) was constructed in the TCGA cohort; **(B,D)** Time-dependent receiver operator characteristic (ROC) curves for OS prediction of the nomogram in the TCGA and Asian Cancer Research Group (ACRG) cohorts; **(C,E)** Calibration curves of the nomogram on 1-, 3-, and 5-years survival probability in the TCGA and ACRG cohorts.

univariate Cox regression, LASSO-Cox regression, and multivariate Cox regression analyses. In the TCGA training cohort, the IRLPS showed good predictive ability for not only the OS but also other survival indicators (DFS, DSS, and PFS) of GC patients, which was effectively validated by bootstrapping technique. Then, subgroup analyses and independent external

validations further demonstrated the stability and generalizability of the IRLPS in predicting the OS and DFS of GC patients. Notably, pan-cancer survival analyses also demonstrated the universal applicability of the IRLPS. In the comparisons with the other four published traditional prognostic signatures, we clarified the superiority of the established signature that was

based on the paring algorithm. Moreover, the IRLPS was found to be an independent predictive indicator for both the OS and DFS of GC patients. In particular, considering the complementary value between the IRLPS and clinical prognostic characteristics, we further constructed and validated an individualized nomogram which could provide a more accurate prediction of OS for GC and is more convenient for oncologists to use.

Several lncRNAs in the model have been previously confirmed to play important roles in the occurrence and development of GC. For instances, Pan et al. reported that *LIFR-AS1* could contribute to the proliferation and invasion of GC cells *via* miR-29a-3p/*COL1A2* axis (Pan et al., 2021). Li et al. (2019) found that *PICART1* functioned as a tumor suppressor by regulating the PI3K/AKT and ERK/MAPK signaling pathways. Zhang et al. (2015) demonstrated that *BANCR* was involved in the growth and apoptosis of GC cells *via* the regulation of NF- κ B1, and overexpression of *BANCR* was correlated with unfavorable prognosis of GC patients (Li et al., 2015). In addition, many of the modeling lncRNAs can also contribute to the malignant phenotypes of multiple other cancer types, such as *RBMS3-AS3* in the prostate cancer, *NR2F2-AS1* in the non-small cell lung cancer, and *LINC01554* in the hepatocellular carcinoma (Zhang et al., 2019; Zheng et al., 2019; Jiang et al., 2020). Hence, we believed that the established signature may provide novel biomarkers for the precise treatment of GC.

To elucidate the potential reasons for the survival differences determined by the IRLPS, we first evaluated the associations of the IRLPS with clinical characteristics and found positive correlations between the signature and tumor differentiation, invasion depth, lymph node metastasis, and clinical stage, indicating the essential role of the IRLPS in GC progression. It was known that molecular characteristics were closely related to the prognosis and treatment responses of GC (Cristescu et al., 2015). Here, we observed that GC patients in different molecular subtypes had obviously different risk scores, suggesting that our signature may reflect tumor heterogeneity and participate in the molecular events contributing to the oncogenesis of GC. In particular, the EMT subtype of the ACRG molecular classifier had the highest risk scores and the lowest proportion of predicted responders, which was consistent with previous reports that this subtype predicts poor survival and immunotherapy response (Kim et al., 2018). Next, functional enrichment analyses were performed on the up-regulated genes to further identify the potential mechanisms involved. We observed that the terms correlated to the ECM were markedly enriched in the high-risk group, indicating the active proliferation and migration of tumor cells because the surrounding ECM could influence cell adhesion and polarity and stimulate growth factor signaling (Walker et al., 2018). In addition, several established cancer-related pathways and hallmarks including cAMP and cGMP-PKG signaling pathway, and angiogenesis and EMT related gene sets were enriched in the high-risk group, indicating the internal regulatory mechanisms of the IRLPS in the invasiveness and poor prognosis of GC. It was also worth noting that both the KEGG enrichment analysis and pre-ranked GSEA identified the TGF- β signaling in the high-risk group. Studies have shown that TGF- β could promote the infiltration of T regulatory cells and inhibit the function of effector T cells and natural killer cells, which induces

immune suppression within the TME, thus causing tumor immune evasion and poor prognosis in the high-risk group (Batlle and Massagué, 2019). Moreover, we found many immunologic gene sets that were up-regulated in the high-risk group, highlighting the role of IRLPS in regulating the immune system.

Accumulating evidence has recently elucidated the effects of the TME components on defining the immunophenotypes of cancers, which could impact cancer tumorigenesis and prognosis (Wu et al., 2021). Thus, understanding the immune landscape of the TME may help find effective ways to improve patients' prognoses and therapeutic benefits. To investigate the relationships between the established signature and immune cells, we adopted several acknowledged methods to estimate the TME characteristics. Our results found that tumors with high IRLPS scores exhibited high infiltration levels of monocytes, macrophages, and mast cells, and low infiltration levels of CD4⁺ T cells and follicular T helper cells, which indicated that we could evaluate the immune response of GC tissue according to the IRLPS model. A previous study revealed that tumor-associated macrophages could secrete anti-inflammatory cytokines that induce an immunosuppressive tumor microenvironment by recruiting T regulatory cells, which suppresses the cell-mediated immune response (Mehta et al., 2021). It was also reported that increased tumor-infiltrating mast cells could foster immune suppression and GC progression through TNF- α -PD-L1 pathway (Lv et al., 2019). These may explain the poor prognosis of the high-risk groups. Except for the tumor-infiltrating immune cells, stromal cells also play important roles in shaping the tumor immunophenotypes (Denton et al., 2018). Fortunately, we observed negative correlations between the IRLPS and stromal scores, and the cancer-associated fibroblasts, as the most prominent components within the tumor stroma that could directly contribute to an immunosuppressive environment, are more abundant in the tumors with higher risk scores (Monteran and Erez, 2019). Accordingly, we obtained the expected results of the worse predicted immunotherapy responses in the patients of the high-risk groups by analyzing the TIDE scores and TMB. The TIDE prediction score was associated with the induction of T cell dysfunction in tumors with high cytotoxic T lymphocyte (CTL) levels and the prevention of T cell infiltration in tumors with low CTL levels and thus represents two primary mechanisms of tumor immune escape (Jiang et al., 2018). In this study, higher TIDE and T cell exclusion scores were found in the high-risk groups of all the investigated cohorts, indicating that their lower ICB response rates might be due to the immune evasion through T cell exclusion. In addition, using the somatic mutation data available in the TCGA cohort, we found obviously lower TMB in the high-risk group. It has been reported that tumors with low TMB generates fewer neoantigens, thus avoiding being attacked by CTLs (Wang et al., 2021b). Taken together, these findings suggested the important role of IRLPS in determining TME cell infiltration and predicting response to ICB therapy.

It should be admitted that this study has some limitations. First, although we validated the effectiveness of the signature in two independent GEO cohorts, the overall accuracy of the signature in the validation cohorts was not as good as the TCGA cohort, which may be due to racial differences between different cohorts or sampling bias caused by genomic intratumor

heterogeneity (Gyanchandani et al., 2016). Second, the clinical value of the signature in predicting immunotherapy response needs to be assessed in prospective clinical trials. Third, the underlying mechanisms of the identified lncRNAs in the immune regulation of GC have not been fully elucidated, and further research on their functions may contribute to their clinical application as novel therapeutic targets.

In summary, this study proposed an innovative signature for predicting the prognosis of GC patients, which may provide new insights into the role of lncRNAs in tumor immunity and facilitate the more effective development of anti-tumor immunotherapy.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary material**.

AUTHOR CONTRIBUTIONS

WT, YX and HZ conceived and designed the study and critically revised the manuscript. CN wrote the original manuscript. CN,

JZ, QW, XZ, and GX analyzed the data and created figures and tables. CL, TL and WW provided technical support. YW and YZ supervised the research. WT provided funding acquisition. All authors approved the final manuscript submitted for publication.

FUNDING

This work was supported by grants from the Heilongjiang Provincial Natural Science Foundation of China (No. LH 2020H020).

ACKNOWLEDGMENTS

The authors gratefully acknowledge the TCGA, GEO, and GTEx programs for their contributions to provide high-quality data to researchers.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcell.2022.805623/full#supplementary-material>

REFERENCES

- Aran, D., Hu, Z., and Butte, A. J. (2017). xCell: Digitally Portraying the Tissue Cellular Heterogeneity Landscape. *Genome Biol.* 18 (1), 220. doi:10.1186/s13059-017-1349-1
- Battle, E., and Massagué, J. (2019). Transforming Growth Factor- β Signaling in Immunity and Cancer. *Immunity* 50 (4), 924–940. doi:10.1016/j.immuni.2019.03.024
- Becht, E., Giraldo, N. A., Lacroix, L., Buttard, B., Elarouci, N., Petitprez, F., et al. (2016). Estimating the Population Abundance of Tissue-Infiltrating Immune and Stromal Cell Populations Using Gene Expression. *Genome Biol.* 17 (1), 218. doi:10.1186/s13059-016-1070-5
- Bhan, A., Soleimani, M., and Mandal, S. S. (2017). Long Noncoding RNA and Cancer: A New Paradigm. *Cancer Res.* 77 (15), 3965–3981. doi:10.1158/0008-5472.Can-16-2634
- Chen, Y. G., Satpathy, A. T., and Chang, H. Y. (2017). Gene Regulation in the Immune System by Long Noncoding RNAs. *Nat. Immunol.* 18 (9), 962–972. doi:10.1038/ni.3771
- Chen, Y., Luo, T. Q., Xu, S. S., Chen, C. Y., Sun, Y., Lin, L., et al. (2021). An Immune-related seven-lncRNA Signature for Head and Neck Squamous Cell Carcinoma. *Cancer Med.* 10 (7), 2268–2285. doi:10.1002/cam4.3756
- Collado-Torres, L., Nellore, A., and Jaffe, A. E. (2017a). Recount Workflow: Accessing over 70,000 Human RNA-Seq Samples with Bioconductor. *F1000Res* 6 (1558), 1558. doi:10.12688/f1000research.12223.1
- Collado-Torres, L., Nellore, A., Kammers, K., Ellis, S. E., Taub, M. A., Hansen, K. D., et al. (2017b). Reproducible RNA-Seq Analysis Using Recount2. *Nat. Biotechnol.* 35 (4), 319–321. doi:10.1038/nbt.3838
- Cristescu, R., Lee, J., Nebozhyn, M., Kim, K.-M., Ting, J. C., Wong, S. S., et al. (2015). Molecular Analysis of Gastric Cancer Identifies Subtypes Associated with Distinct Clinical Outcomes. *Nat. Med.* 21 (5), 449–456. doi:10.1038/nm.3850
- Denaro, N., Merlano, M. C., and Lo Nigro, C. (2019). Long Noncoding RNA S as Regulators of Cancer Immunity. *Mol. Oncol.* 13 (1), 61–73. doi:10.1002/1878-0261.12413
- Denton, A. E., Roberts, E. W., and Fearon, D. T. (2018). Stromal Cells in the Tumor Microenvironment. *Adv. Exp. Med. Biol.* 1060, 99–114. doi:10.1007/978-3-319-78127-3_6
- Digkila, A., and Wagner, A. D. (2016). Advanced Gastric Cancer: Current Treatment Landscape and Future Perspectives. *Wjg* 22 (8), 2403–2414. doi:10.3748/wjg.v22.i8.2403
- Egner, J. R. (2010). AJCC Cancer Staging Manual. *JAMA* 304 (15), 1726–1727. doi:10.1001/jama.2010.1525
- Fu, J., Li, K., Zhang, W., Wan, C., Zhang, J., Jiang, P., et al. (2020). Large-scale Public Data Reuse to Model Immunotherapy Response and Resistance. *Genome Med.* 12 (1), 21. doi:10.1186/s13073-020-0721-z
- Goodman, A., Patel, S. P., and Kurzrock, R. (2017). PD-1-PD-L1 Immune-Checkpoint Blockade in B-Cell Lymphomas. *Nat. Rev. Clin. Oncol.* 14 (4), 203–220. doi:10.1038/nrclinonc.2016.168
- Gyanchandani, R., Lin, Y., Lin, H.-M., Cooper, K., Normolle, D. P., Brufsky, A., et al. (2016). Intratumor Heterogeneity Affects Gene Expression Profile Test Prognostic Risk Stratification in Early Breast Cancer. *Clin. Cancer Res.* 22 (21), 5362–5369. doi:10.1158/1078-0432.CCR-15-2889
- Heinäniemi, M., Nykter, M., Kramer, R., Wienecke-Baldacchino, A., Sinkkonen, L., Zhou, J. X., et al. (2013). Gene-pair Expression Signatures Reveal Lineage Control. *Nat. Methods* 10 (6), 577–583. doi:10.1038/nmeth.2445
- Jiang, H., and Wong, W. H. (2008). SeqMap: Mapping Massive Amount of Oligonucleotides to the Genome. *Bioinformatics* 24 (20), 2395–2396. doi:10.1093/bioinformatics/btn429
- Jiang, P., Gu, S., Pan, D., Fu, J., Sahu, A., Hu, X., et al. (2018). Signatures of T Cell Dysfunction and Exclusion Predict Cancer Immunotherapy Response. *Nat. Med.* 24 (10), 1550–1558. doi:10.1038/s41591-018-0136-1
- Jiang, Z., Zhang, Y., Chen, X., Wu, P., and Chen, D. (2020). Long Noncoding RNA RBMS3-AS3 Acts as a microRNA-4534 Sponge to Inhibit the Progression of Prostate Cancer by Upregulating VASH1. *Gene Ther.* 27 (3–4), 143–156. doi:10.1038/s41434-019-0108-1
- Kim, S. T., Cristescu, R., Bass, A. J., Kim, K.-M., Odegaard, J. I., Kim, K., et al. (2018). Comprehensive Molecular Characterization of Clinical Responses to PD-1 Inhibition in Metastatic Gastric Cancer. *Nat. Med.* 24 (9), 1449–1458. doi:10.1038/s41591-018-0101-z

- Lazăr, D. C., Avram, M. F., Romoan, I., Cornianu, M., Tăban, S., and Goldi, A. (2018). Prognostic Significance of Tumor Immune Microenvironment and Immunotherapy: Novel Insights and Future Perspectives in Gastric Cancer. *Wjg* 24 (32), 3583–3616. doi:10.3748/wjg.v24.i32.3583
- Leek, J. T., Johnson, W. E., Parker, H. S., Jaffe, A. E., and Storey, J. D. (2012). The Sva Package for Removing Batch Effects and Other Unwanted Variation in High-Throughput Experiments. *Bioinformatics* 28 (6), 882–883. doi:10.1093/bioinformatics/bts034
- Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., et al. (2010). Tackling the Widespread and Critical Impact of Batch Effects in High-Throughput Data. *Nat. Rev. Genet.* 11 (10), 733–739. doi:10.1038/nrg2825
- Lei, Z., Tan, I. B., Das, K., Deng, N., Zouridis, H., Pattison, S., et al. (2013). Identification of Molecular Subtypes of Gastric Cancer with Different Responses to PI3-Kinase Inhibitors and 5-fluorouracil. *Gastroenterology* 145 (3), 554–565. doi:10.1053/j.gastro.2013.05.010
- 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 (11), 1529–1537. doi:10.1001/jamaoncol.2017.1609
- Li, J. F., Li, W. H., Xue, L. L., and Zhang, Y. (2019). Long Non-coding RNA PICART1 Inhibits Cell Proliferation by Regulating the PI3K/AKT and MAPK/ERK Signaling Pathways in Gastric Cancer. *Eur. Rev. Med. Pharmacol. Sci.* 23 (2), 588–597. doi:10.26355/eurrev_201901_16871
- Li, L., Zhang, L., Zhang, Y., and Zhou, F. (2015). Increased Expression of lncRNA BANCRI Is Associated with Clinical Progression and Poor Prognosis in Gastric Cancer. *Biomed. Pharmacother.* 72, 109–112. doi:10.1016/j.biopha.2015.04.007
- Li, X., Pasche, B., Zhang, W., and Chen, K. (2018). Association of MUC16 Mutation with Tumor Mutation Load and Outcomes in Patients with Gastric Cancer. *JAMA Oncol.* 4 (12), 1691–1698. doi:10.1001/jamaoncol.2018.2805
- Li, Y., Jiang, T., Zhou, W., Li, J., Li, X., Wang, Q., et al. (2020). Pan-cancer Characterization of Immune-Related lncRNAs Identifies Potential Oncogenic Biomarkers. *Nat. Commun.* 11 (1), 1000. doi:10.1038/s41467-020-14802-2
- Lv, Y., Zhao, Y., Wang, X., Chen, N., Mao, F., Teng, Y., et al. (2019). Increased Intratumoral Mast Cells Foster Immune Suppression and Gastric Cancer Progression through TNF- α -PD-L1 Pathway. *J. Immunotherapy Cancer* 7 (1), 54. doi:10.1186/s40425-019-0530-3
- Malta, T. M., Sokolov, A., Gentles, A. J., Burzykowski, T., Poisson, L., Weinstein, J. N., et al. (2018). Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell* 173 (2), 338–e15. e315. doi:10.1016/j.cell.2018.03.034
- Mao, D., Zhou, Z., Song, S., Li, D., He, Y., Wei, Z., et al. (2021a). Identification of Stemness Characteristics Associated with the Immune Microenvironment and Prognosis in Gastric Cancer. *Front. Oncol.* 11, 626961. doi:10.3389/fonc.2021.626961
- Mao, R., Liu, K., Zhao, N., Guo, P., Wu, Y., Wang, Z., et al. (2021b). Clinical Significance and Prognostic Role of an Immune-Related Gene Signature in Gastric Adenocarcinoma. *Aging* 13 (13), 17734–17767. doi:10.18632/aging.203266
- Mayakonda, A., Lin, D.-C., Assenov, Y., Plass, C., and Koeffler, H. P. (2018). Maftools: Efficient and Comprehensive Analysis of Somatic Variants in Cancer. *Genome Res.* 28 (11), 1747–1756. doi:10.1101/gr.239244.118
- Mehta, A. K., Kadel, S., Townsend, M. G., Oliwa, M., and Guerriero, J. L. (2021). Macrophage Biology and Mechanisms of Immune Suppression in Breast Cancer. *Front. Immunol.* 12, 643771. doi:10.3389/fimmu.2021.643771
- Miranda, A., Hamilton, P. T., Zhang, A. W., Pattnaik, S., Becht, E., Mezheyeuski, A., et al. (2019). Cancer Stemness, Intratumoral Heterogeneity, and Immune Response across Cancers. *Proc. Natl. Acad. Sci. USA* 116 (18), 9020–9029. doi:10.1073/pnas.1818210116
- Monteran, L., and Erez, N. (2019). The Dark Side of Fibroblasts: Cancer-Associated Fibroblasts as Mediators of Immunosuppression in the Tumor Microenvironment. *Front. Immunol.* 10, 1835. doi:10.3389/fimmu.2019.01835
- Mounir, M., Lucchetta, M., Silva, T. C., Olsen, C., Bontempi, G., Chen, X., et al. (2019). New Functionalities in the TCGAbiolinks Package for the Study and Integration of Cancer Data from GDC and GTEx. *Plos Comput. Biol.* 15 (3), e1006701. doi:10.1371/journal.pcbi.1006701
- 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
- Pan, H., Ding, Y., Jiang, Y., Wang, X., Rao, J., Zhang, X., et al. (2021). lncRNA LIFR-AS1 Promotes Proliferation and Invasion of Gastric Cancer Cell via miR-29a-3p/COL1A2 axis. *Cancer Cell Int* 21 (1), 7. doi:10.1186/s12935-020-01644-7
- Postow, M. A., Callahan, M. K., and Wolchok, J. D. (2015). Immune Checkpoint Blockade in Cancer Therapy. *Jco* 33 (17), 1974–1982. doi:10.1200/jco.2014.59.4358
- Racle, J., de Jonge, K., Baumgaertner, P., Speiser, D. E., and Gfeller, D. (2017). Simultaneous Enumeration of Cancer and Immune Cell Types from Bulk Tumor Gene Expression Data. *Elife* 6. doi:10.7554/eLife.26476
- Rafee, A., Riazi-Rad, F., Havaskary, M., and Nuri, F. (2018). Long Noncoding RNAs: Regulation, Function and Cancer. *Biotechnol. Genet. Eng. Rev.* 34 (2), 153–180. doi:10.1080/02648725.2018.1471566
- Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). Limma powers Differential Expression Analyses for RNA-Sequencing and Microarray Studies. *Nucleic Acids Res.* 43 (7), e47. doi:10.1093/nar/gkv007
- Rizvi, N. A., Hellmann, M. D., Snyder, A., Kvistborg, P., Makarov, V., Havel, J. J., et al. (2015). Mutational Landscape Determines Sensitivity to PD-1 Blockade in Non-small Cell Lung Cancer. *Science* 348 (6230), 124–128. doi:10.1126/science.aaa1348
- Shah, M. A., and Ajani, J. A. (2010). Gastric Cancer—An Enigmatic and Heterogeneous Disease. *JAMA* 303 (17), 1753–1754. doi:10.1001/jama.2010.553
- Shen, Y., Peng, X., and Shen, C. (2020). Identification and Validation of Immune-Related lncRNA Prognostic Signature for Breast Cancer. *Genomics* 112 (3), 2640–2646. doi:10.1016/j.ygeno.2020.02.015
- Skidmore, Z. L., Wagner, A. H., Lesurf, R., Campbell, K. M., Kunisaki, J., Griffith, O. L., et al. (2016). GenVisR: Genomic Visualizations in R. *Bioinformatics* 32 (19), 3012–3014. doi:10.1093/bioinformatics/btw325
- Smyth, E. C., Nilsson, M., Grabsch, H. I., van Grieken, N. C., and Lordick, F. (2020). Gastric Cancer. *The Lancet* 396 (10251), 635–648. doi:10.1016/s0140-6736(20)31288-5
- Subramanian, A., Narayan, R., Corsello, S. M., Peck, D. D., Natoli, T. E., Lu, X., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell* 171 (6), 1437–1452. e1417. doi:10.1016/j.cell.2017.10.049
- 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.* 102 (43), 15545–15550. doi:10.1073/pnas.0506580102
- 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
- Walker, C., Mojares, E., and del Rio Hernández, A. (2018). Role of Extracellular Matrix in Development and Cancer Progression. *Ijms* 19 (10), 3028. doi:10.3390/ijms19103028
- Wang, G., Chen, X., Liang, Y., Wang, W., Fang, Y., and Shen, K. (2018). Long Noncoding RNA Signature and Disease Outcome in Estrogen Receptor-Positive Breast Cancer Patients Treated with Tamoxifen. *J. Breast Cancer* 21 (3), 277–287. doi:10.4048/jbc.2018.21.e39
- Wang, H., Meng, Q., and Ma, B. (2021a). Characterization of the Prognostic m6A-Related lncRNA Signature in Gastric Cancer. *Front. Oncol.* 11, 630260. doi:10.3389/fonc.2021.630260
- Wang, H., Wu, X., and Chen, Y. (2019). Stromal-Immune Score-Based Gene Signature: A Prognosis Stratification Tool in Gastric Cancer. *Front. Oncol.* 9, 1212. doi:10.3389/fonc.2019.01212
- Wang, P., Chen, Y., and Wang, C. (2021b). Beyond Tumor Mutation Burden: Tumor Neoantigen Burden as a Biomarker for Immunotherapy and Other Types of Therapy. *Front. Oncol.* 11, 672677. doi:10.3389/fonc.2021.672677
- Wang, Y., Zou, Y., Zhang, Y., and Li, C. (2021c). Developing a Risk Scoring System Based on Immune-Related lncRNAs for Patients with Gastric Cancer. *Biosci. Rep.* 41 (1), BSR20202203. doi:10.1042/bsr20202203
- Wu, J., Li, L., Zhang, H., Zhao, Y., Zhang, H., Wu, S., et al. (2021). A Risk Model Developed Based on Tumor Microenvironment Predicts Overall Survival and

- Associates with Tumor Immunity of Patients with Lung Adenocarcinoma. *Oncogene* 40 (26), 4413–4424. doi:10.1038/s41388-021-01853-y
- Xu, X., Zhou, E., Zheng, J., Zhang, C., Zou, Y., Lin, J., et al. (2021). Prognostic and Predictive Value of m6A "Eraser" Related Gene Signature in Gastric Cancer. *Front. Oncol.* 11, 631803. doi:10.3389/fonc.2021.631803
- 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
- Yu, G., Wang, L.-G., Han, Y., and He, Q.-Y. (2012). clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A J. Integr. Biol.* 16 (5), 284–287. doi:10.1089/omi.2011.0118
- 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 (5), 737–750. doi:10.1158/2326-6066.Cir-18-0436
- Zeng, D., Wu, J., Luo, H., Li, Y., Xiao, J., Peng, J., et al. (2021). Tumor Microenvironment Evaluation Promotes Precise Checkpoint Immunotherapy of Advanced Gastric Cancer. *J. Immunother. Cancer* 9 (8), e002467. doi:10.1136/jitc-2021-002467
- Zhang, S., Zhang, X., Sun, Q., Zhuang, C., Li, G., Sun, L., et al. (2019). LncRNA NR2F2-AS1 Promotes Tumorigenesis through Modulating BMI1 Expression by Targeting miR-320b in Non-small Cell Lung Cancer. *J. Cel Mol Med* 23 (3), 2001–2011. doi:10.1111/jcmm.14102
- Zhang, X., Sun, S., Pu, J. K. S., Tsang, A. C. O., Lee, D., Man, V. O. Y., et al. (2012). Long Non-coding RNA Expression Profiles Predict Clinical Phenotypes in Glioma. *Neurobiol. Dis.* 48 (1), 1–8. doi:10.1016/j.nbd.2012.06.004
- Zhang, Z.-X., Liu, Z.-Q., Jiang, B., Lu, X.-Y., Ning, X.-F., Yuan, C.-T., et al. (2015). BRAF Activated Non-coding RNA (BANCR) Promoting Gastric Cancer Cells Proliferation via Regulation of NF-Kb1. *Biochem. Biophysical Res. Commun.* 465 (2), 225–231. doi:10.1016/j.bbrc.2015.07.158
- Zheng, Y.-L., Li, L., Jia, Y.-X., Zhang, B.-Z., Li, J.-C., Zhu, Y.-H., et al. (2019). LINC01554-Mediated Glucose Metabolism Reprogramming Suppresses Tumorigenicity in Hepatocellular Carcinoma via Downregulating PKM2 Expression and Inhibiting Akt/mTOR Signaling Pathway. *Theranostics* 9 (3), 796–810. doi:10.7150/thno.28992
- Zhou, M., Guo, M., He, D., Wang, X., Cui, Y., Yang, H., et al. (2015). A Potential Signature of Eight Long Non-coding RNAs Predicts Survival in Patients with Non-small Cell Lung Cancer. *J. Transl Med.* 13, 231. doi:10.1186/s12967-015-0556-3

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 © 2022 Nie, Zhai, Wang, Zhu, Xiang, Liu, Liu, Wang, Wang, Zhao, Tian, Xue and Zhou. 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.