

Exploring immune-related signatures for predicting immunotherapeutic responsiveness, prognosis, and diagnosis of patients with colon cancer

Lichao Cao^{1,2}, Ying Ba³, Jin Yang^{1,2}, Hezi Zhang³

¹Provincial Key Laboratory of Biotechnology of Shaanxi Province, Northwest University, Xi'an, China

²Key Laboratory of Resource Biology and Biotechnology in Western China, Ministry of Education, School of Life Sciences, Northwest University, Xi'an, China

³Shenzhen Nucleus Gene Technology Co., Ltd., Shenzhen, China

Correspondence to: Jin Yang, Hezi Zhang; **email:** yangjin@nwu.edu.cn; hezizhang2020@163.com; <https://orcid.org/0000-0003-4355-6802>

Keywords: colon cancer, tumor immune microenvironment, prognosis, diagnosis, immunotherapeutic responsiveness

Received: April 4, 2022

Accepted: June 14, 2022

Published: June 20, 2022

Copyright: © 2022 Cao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

The present study focused on identifying the immune-related signatures and exploring their performance in predicting the prognosis, immunotherapeutic responsiveness, and diagnosis of patients with colon cancer. Firstly, the immunotherapeutic response-related differential expressed genes (DEGs) were identified by comparing responders and non-responders from an anti-PD-L1 cohort using the edgeR R package. Then, the immunotherapeutic response related DEGs was intersected with immune-related genes (IRGs) to obtain the immunotherapeutic response and immune-related genes (IRIGs). Then, an immunotherapeutic response and immune-related risk score (IRIRScore) model consisting of 6 IRIGs was constructed using the univariable Cox regression analysis and multivariate Cox regression analysis based on the COAD cohort from the cancer genome atlas (TCGA) database, which was further validated in two independent gene expression omnibus database (GEO) datasets (GSE39582 and GSE17536) and anti-PD-L1 cohort. A nomogram with good accuracy was established based on the immune-related signatures and clinical factors (C-index = 0.75). In the training dataset and GSE39582, higher IRIRScore was significantly associated with higher TMN and advanced pathological stages. Based on the anti-PD-L1 cohort, patients who were sensitive to immunotherapy had significantly lower risk score than non-responders. Furthermore, we explored the immunotherapy-related signatures based on the training dataset. Kaplan-Meier curve revealed a high level of T cells regulatory (Tregs) was significantly related to poor overall survival (OS), while a high level of T cells CD4 memory resting was significantly related to better OS. Besides, the TMB value of patients in the high-risk group was significantly higher than those in a low-risk group. Moreover, patients in the high-risk group had significantly higher expression levels of immune checkpoint inhibitors. In addition, the immune-related signatures were applied to establish prediction models using the random forest algorithm. Among them, *TDGF1* and *NRG1* revealed excellent diagnostic predictive performance (AUC >0.8). In conclusion, the current findings provide new insights into immune-related immunotherapeutic responsiveness, prognosis, and diagnosis of colon cancer.

INTRODUCTION

Colon cancer is one of the most malignant tumors with high mortality worldwide despite advancements in

tumor screening, early diagnosis, and treatment [1]. In recent years, increasing studies explored the functions of tumor immune microenvironment, which may play important roles in tumorigenesis [2]. Moreover,

numerous studies have identified immune-related signatures in the tumor microenvironment for predicting the prognosis for various cancers, including colon cancer [3–5].

Immunotherapy, especially immune checkpoint inhibitors (ICIs), focuses on harnessing the host immune system to combat tumor progression and metastasis and achieve the goal of controlling and eliminating tumors, which has revolutionized the treatment for colon cancer [6]. However, the responsiveness rate to ICIs rarely exceeds 20%, which depends on several key factors, including mutational load, tumor-infiltrating lymphocytes, and regulatory checkpoint receptors [7]. However, the research on molecular markers for predicting the immunotherapeutic responsiveness of colon cancer patients is limited.

Therefore, it is critical to explore effective and robust novel molecular signatures for predicting the immunotherapeutic responsiveness, prognosis, and diagnosis of patients with colon cancer. The present study focused on the immune-related signatures identified in the anti-PD-L1 cohort and systematically explored their performance in predicting the prognosis, immunotherapeutic responsiveness, and diagnosis of patients with colon cancer.

MATERIALS AND METHODS

Acquiring and preprocessing the relevant data

An anti-PD-L1 cohort (IMvigor210) was obtained using the IMvigor210CoreBiologies R package [8], which was treated with atezolizumab (anti-PD-L1 agent); only the samples with immunotherapeutic response information were included. The samples that displayed the complete response (CR) or partial response (PR) were categorized as responders and the samples that displayed stable (SD) or progressive disease (PD) were categorized as non-responders. In addition, the scores (estimate, stromal, and immune) of each sample were calculated using the estimate R package [9]. The differences in the clinical features and the scores between responders and non-responders were statistically compared using the Wilcoxon test, and the detailed information is shown in Table 1 and Supplementary Figure 1A–1C.

A total of 471 COAD samples were downloaded from UCSC Xena platform (<https://xenabrowser.net/datapages/>), including the mRNA expression data, mutation profiling data, and clinical and survival information. Subsequently, the samples ($n = 471$) were divided into the normal group ($n = 39$) and tumor group ($n = 432$);

the tumor group was used as the training dataset, and the detailed information is listed in Supplementary Table 1.

Two independent GEO datasets (GSE39582 and GSE17536), considered testing datasets, were downloaded from the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). The GSE39582 included 556 colon cancer patients, and GSE17536 included 177 colon cancer patients, with available clinical and survival information. The detailed information was shown in Supplementary Tables 2 and 3.

A total of 1509 IRGs were obtained from the ImmPort database (<https://immport.niaid.nih.gov/>).

Identification of immune-related genes from the anti-PD-L1 cohort

The edgeR R package [10] was adopted to compare the gene expression differences between the responders and non-responders. Genes that met the cut-off criteria of false discovery rate (FDR) < 0.05 and $|\log_2$ fold change (FC)| > 1 were considered immunotherapeutic response related DEGs, and then intersected with IRGs to obtain the IRIGs. Subsequently, the Gene ontology (GO) functional, Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed on IRIGs using the clusterProfiler R package [11].

Construction of estimate-stromal-immune scores model

The ESTIMATE algorithm calculated the immune, stromal and estimate scores for each colon cancer sample of the TCGA-COAD cohort [9]. Then, the colon cancer patients were divided into the high-score and low-score groups based on the median levels of the estimate, stromal, and immune scores, respectively. The association between the estimate-stromal-immune scores and OS was analyzed using the Kaplan-Meier curve. The pROC R package was used to plot the receiver operating characteristic (ROC) curves to evaluate the predictive capacities of the estimate-stromal-immune scores model.

Identification and validation of immune-related signatures for predicting immunotherapeutic responsiveness and prognosis of patients with colon cancer

The univariable Cox regression analysis was performed on the IRIGs using the training dataset with $P < 0.05$ as the criteria to identify the prognostic signature. Subsequently, we established the IRIRScore model for

Table 1. The detailed information on Anti-PD-L1 cohort.

	CR/PR (N = 68)	SD/PD (N = 230)	P-value
Sex			
F	11 (16.2%)	54 (23.5%)	0.125
M	57 (83.8%)	176 (76.5%)	
Baseline.ECOG.Score			
0	37 (54.4%)	84 (36.5%)	0.0313
1	27 (39.7%)	138 (60.0%)	
2	4 (5.9%)	8 (3.5%)	
sizeFactor			
Mean (SD)	1.12 (0.419)	1.06 (0.360)	0.393
Median (Min, Max)	1.11 (0.306, 1.95)	1.02 (0.264, 1.97)	
OS			
Alive	63 (92.6%)	46 (20.0%)	0.125
Death	5 (7.4%)	184 (80.0%)	
OS.time			
Mean (SD)	590 (94.1)	271 (194)	<0.001
Median (Min, Max)	618 (278, 734)	214 (5.91, 715)	
FMOne.mutation.burden.per.MB			
Mean (SD)	16.9 (13.3)	8.61 (6.26)	<0.001
Median (Min, Max)	14.0 (1.00, 62.0)	7.00 (0, 44.0)	
Missing	7 (10.3%)	57 (24.8%)	
Neoantigen.burden.per.MB			
Mean (SD)	2.46 (2.34)	1.06 (1.09)	<0.001
Median (Min, Max)	1.80 (0.216, 11.7)	0.725 (0.0392, 6.20)	
Missing	15 (22.1%)	67 (29.1%)	

Abbreviation: OS: overall survival.

predicting the prognosis of colon cancer patients using the multivariate Cox regression analysis and the coefficient of the prognostic indicator was obtained. The IRIRScore for each patient was calculated according to the formula:

$$\text{IRIRScore} = \sum \text{Cox coefficient of gene } \chi^i \times \text{scale expression value of gene } \chi^i$$

Moreover, two independent GEO datasets with accession numbers GSE39582 and GSE17536 and anti-PD-L1 cohort were used to further validate the prognostic signatures of the constructed IRIRScore model. Besides, the Kruskal-Wallis test was used to compare the difference in IRIRScore between responders and non-responders in the anti-PD-L1 cohort.

In addition, a nomogram was designed to visualize the prognostic value of different patients' features, including IRIRScore, StromalScore, ImmuneScore, age,

gender, and tumor stage. The calibration curves were plotted to evaluate the predicted probabilities compared to the ideal predictive line using the rms R package. Moreover, the hazard ratios of the patients' features were illustrated by a forest plot.

Statistical analysis

To verify and access the predictive capabilities of the constructed IRIRScore model, the ROC curves and Kaplan-Meier curves were plotted using the training dataset (tumor samples in TCGA-COAD cohort) and testing dataset (GSE39582, GSE17536, and anti-PD-L1 cohort). The optimal cutoff value for grouping patients were chosen at the points of the ROC curves where the difference between true positive and false positive was the most significant. Additionally, the associations between the risk score level and clinical features, including tumor stages and microsatellite status, were performed using the Kruskal-Wallis test.

Exploring potential immunotherapy related signatures in the tumor microenvironment of colon cancer

Firstly, the training dataset was divided into a high-risk group and a low-risk group based on the constructed IRIScore model. Moreover, the proportions of 22 types of tumor-infiltrating immune cells of each sample in the TCGA-COAD cohort was estimated using the CIBERSORT algorithm [12]. Then, the statistical difference in immune landscape between the high-risk and low-risk score groups was compared using an unpaired *t*-test. The association between the significant differential immune cell types (*P*-value < 0.01) and OS were further accessed using Kaplan-Meier curve. Moreover, the correlations between tumor-infiltrating immune cells and IRIScore were analyzed using the *cor.test* in R software (version 4.0.2, <http://www.R-project.org>).

In addition, the maftools R package was utilized to calculate the TMB value and visualize the mutation profiles of the high- and low-risk groups [13]. Subsequently, the TMB value between high- and low-risk groups was compared by the Kruskal-Wallis test, and the differences in the mRNA levels of immune checkpoints and their ligands between the high- and low-risk groups were statistically compared using the Wilcoxon test.

Preliminary exploration of IRIGs as predictive diagnostic indicators base on TCGA-COAD cohort

To explore the potential application of the IRIGs in the diagnosis prediction of colon cancer, we compared the

expressions of the IRIGs between tumor and normal samples using the Wilcoxon test. Moreover, the IRIGs were employed as independent biomarkers to establish diagnosis prediction models using Random Forest (RF) algorithm. Besides, we tried to use the IRIGs as keywords to search immunohistochemical images in the Human Protein Atlas (HPA) (<https://www.proteinatlas.org/>), which may experimentally verify our findings.

Data availability statement

The immunotherapeutic cohort (IMvigor210) was available according to the guideline on <http://research-pub.gene.com/IMvigor210CoreBiologies> using the IMvigor210CoreBiologies R package. The TCGA-COAD cohort was available in the UCSC Xena (<https://xenabrowser.net/datapages/>). GSE39582 and GSE17536 were downloaded from NCBI-GEO database (<https://www.ncbi.nlm.nih.gov/geo/>).

RESULTS

Identification of IRIGs

Figure 1 shows the whole analysis flow of this study. 1323 immunotherapeutic response-related DEGs (634 upregulated and 689 downregulated) were identified by comparing responders and non-responders in the anti-PD-L1 cohort (Figure 2A, Supplementary Table 4). After the intersection with 1509 IRGs, 108 IRIGs was obtained (Figure 2B, Supplementary Table 5). KEGG analysis results revealed that IRIGs were remarkably

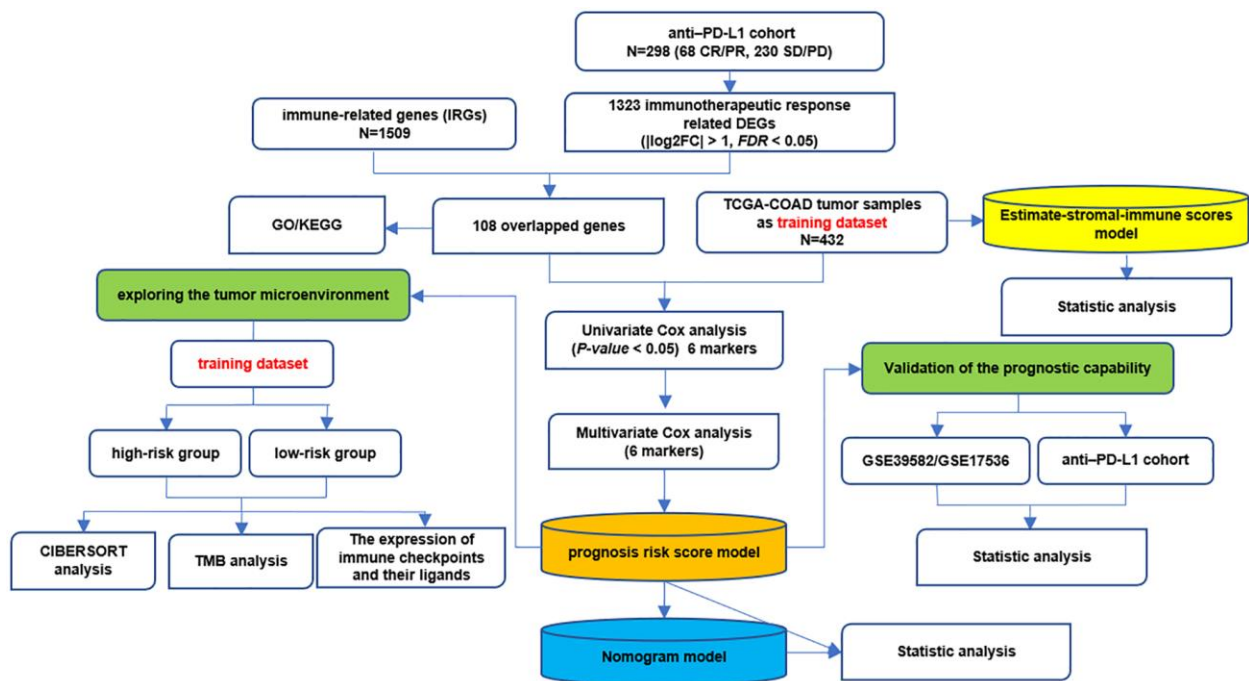


Figure 1. The whole analysis flow of this study.

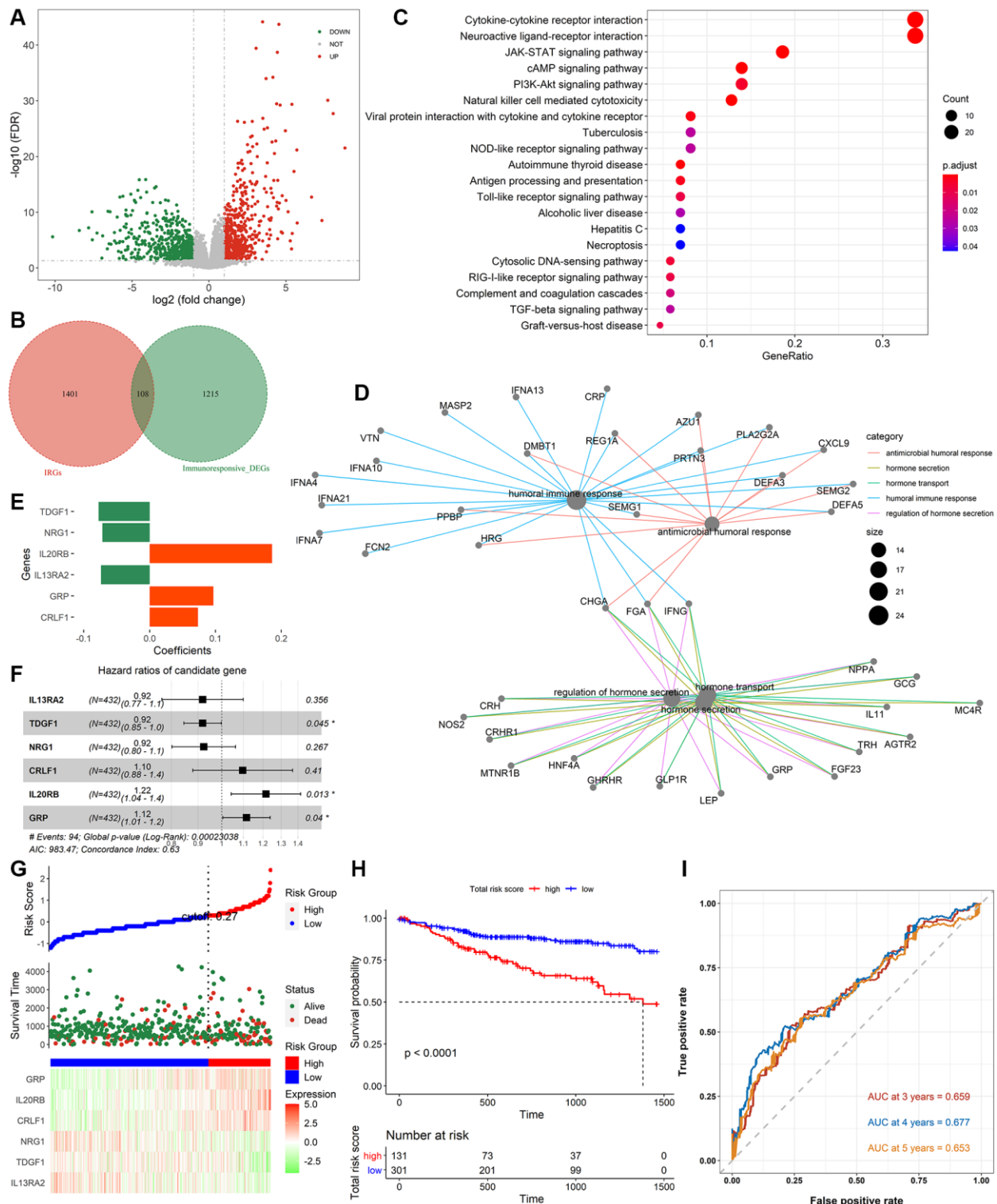


Figure 2. Identification of immune-related signatures and construction of IRIScore prognostic model. (A) Volcano plot on immunotherapeutic response related DEGs between the responders and non-responders in anti-PD-L1 cohort. The green dots represent downregulated genes, while the red dots represent upregulated genes. (B) Venn diagram of the intersection between the immunotherapeutic response related DEGs and IRGs. (C) KEGG pathway enrichment analysis of IRIGs. (D) GO analysis of IRIGs. (E) The coefficients of the 6 IRIGs related to prognosis of colon cancer. (F) Forest plot showing the HR of each IRIG. (G) The distribution of samples in the high- and low risk score groups and their relationship with OS, and the expression pattern of 6 prognostic signatures in high- and low risk score groups. (H) Kaplan-Meier curve exhibited that OS of patients in the low-risk score group was significantly higher than those in the high-risk score group. (I) Time-dependent ROC curve analysis of the IRIScore model.

Table 2. The detailed information on the identified prognostic IRIGs.

Gene Name	HR	HR.95L	HR.95H	p-value
IL13RA2	0.859707917	0.753442337	0.980961205	0.024735341
TDGF1	0.916176837	0.845736901	0.992483591	0.031968386
NRG1	0.878737983	0.786729291	0.981507172	0.021978491
CRLF1	1.242539007	1.029878446	1.499112045	0.023368342
IL20RB	1.261588541	1.097635605	1.450030992	0.001069679
GRP	1.126701281	1.019440489	1.245247555	0.019429694

Abbreviation: HR: hazard ratio.

enriched in terms associated with Cytokine-cytokine receptor interaction, Neuroactive ligand-receptor interaction, JAK-STAT signaling pathway, cAMP signaling pathway, PI3K-Akt signaling pathway, and so on (Figure 2C). Furthermore, the GO enrichment analysis indicated that IRIGs mapped to hormone-related terms, including humoral immune response, antimicrobial humoral response, regulation of hormone secretion, and hormone transport (Figure 2D).

Construction of immune-related prognostic models

The AUC of estimate-stromal-immune scores model was less than 0.6 (ESTIMATEScore: 0.5; StromalScore: 0.513; ImmuneScore: 0.519, Supplementary Figure 1D), indicating its poor performance for predicting the prognosis of colon cancer. Additionally, the Kaplan-Meier curves revealed that the association between the scores (estimate, stromal, or immune) and OS were not significant (ESTIMATEScore: $P = 0.61$, StromalScore: $P = 0.47$, ImmuneScore: $P = 0.94$, Supplementary Figure 1E–1G).

Therefore, effective prognostic signatures for patients with colon cancer needed to be explored. Based on the training dataset, the univariate and multivariate Cox regression analyses were performed on the IRIGs, and 6 genes were found to be significantly related to OS status (Table 2). The coefficient and hazard ratio (HR) of each prognostic indicator is shown in Figure 2E and 2F, respectively. *TDGF1*, *IL13RA2*, and *NRG1* are protective factors (HR <1), while *CRLF1*, *GRP*, and *IL20RB* are risk factors (HR >1). Then, the IRIRScore of each patient was calculated according to the formula described in the Materials and methods section. The scatter diagram in Figure 2G displayed that the survival time of patients in the low-risk group was longer than that in a high-risk group, which was also significantly supported by Figure 2H ($P < 0.0001$). The heatmap in Figure 2G reveals that the expression of risk factors was low in the low-risk group and high in the high-risk group, while the trend of protective factors was the opposite. The AUC of IRIRScore model at 3, 4, and

5 years of OS are 0.659, 0.677, and 0.653, respectively (Figure 2I). Patients with larger tumor sizes yielded higher IRIRScore than those with small tumor size (Kruskal-Wallis test, $P = 3.9e^{-6}$, Figure 3A). Patients with metastases yielded higher IRIRScore than those without metastases (Kruskal-Wallis test, $P = 0.012$, Figure 3B). Patients with more distant or more lymph nodes are involved in patients with higher IRIRScore (Kruskal-Wallis test, $P = 3.2e^{-5}$, Figure 3C). Patients with advanced stage yielded higher IRIRScore than those with early stage (Kruskal-Wallis test, $P = 0.00016$, Figure 3D). Figure 3E reveals that the patients with higher IRIRScore yielded more instability of the microsatellite status (Kruskal-Wallis test, $P = 0.015$, Figure 3E).

Validation and assessment of the identified immune-related signatures in predicting prognosis and immunotherapeutic responsiveness

To verify whether the identified immune-related signatures were suitable for different cohorts, The IRIRScore of each patient in two independent GEO datasets (GSE39582 and GSE17536), and the anti-PD-L1 cohort was calculated based on the same method used in the training dataset. Based on the testing datasets, the OS rate of patients in the high-risk group was significantly lower than that in low-risk group ($P < 0.0001$, Supplementary Figure 2A, Supplementary Figure 3A, and 3C), and the areas under the curves (AUCs) of IRIRScore model were calculated (GSE39582: 0.623 at 3-years, 0.607 at 4-years, and 0.563 at 5-years, Supplementary Figure 2B; GSE17536: 0.656 at 3-years, 0.733 at 4-years, and 0.657 at 5-years, Supplementary Figure 3B; Anti-PD-L1 cohort: 0.645, Supplementary Figure 3D).

Based on the testing data from GSE39582, the Kruskal-Wallis test revealed that higher IRIRScore was associated with higher T stage ($P = 0.00038$), metastasis ($P = 0.011$), N stages ($P = 0.0015$), and advanced pathological stage ($P = 0.021$) (Supplementary Figure 2C–2F).

In addition, based on the anti-PD-L1 cohort, we found that patients sensitive to immunotherapy had significantly lower risk scores than non-responders ($P = 7.1e^{-5}$, Supplementary Figure 3E), indicating that responders will have better OS than non-responders.

Construction and validation of the nomogram

A prognostic nomogram was constructed to improve the accuracy of the performance of the IRIRScore model and to provide a quantitative and visual method for

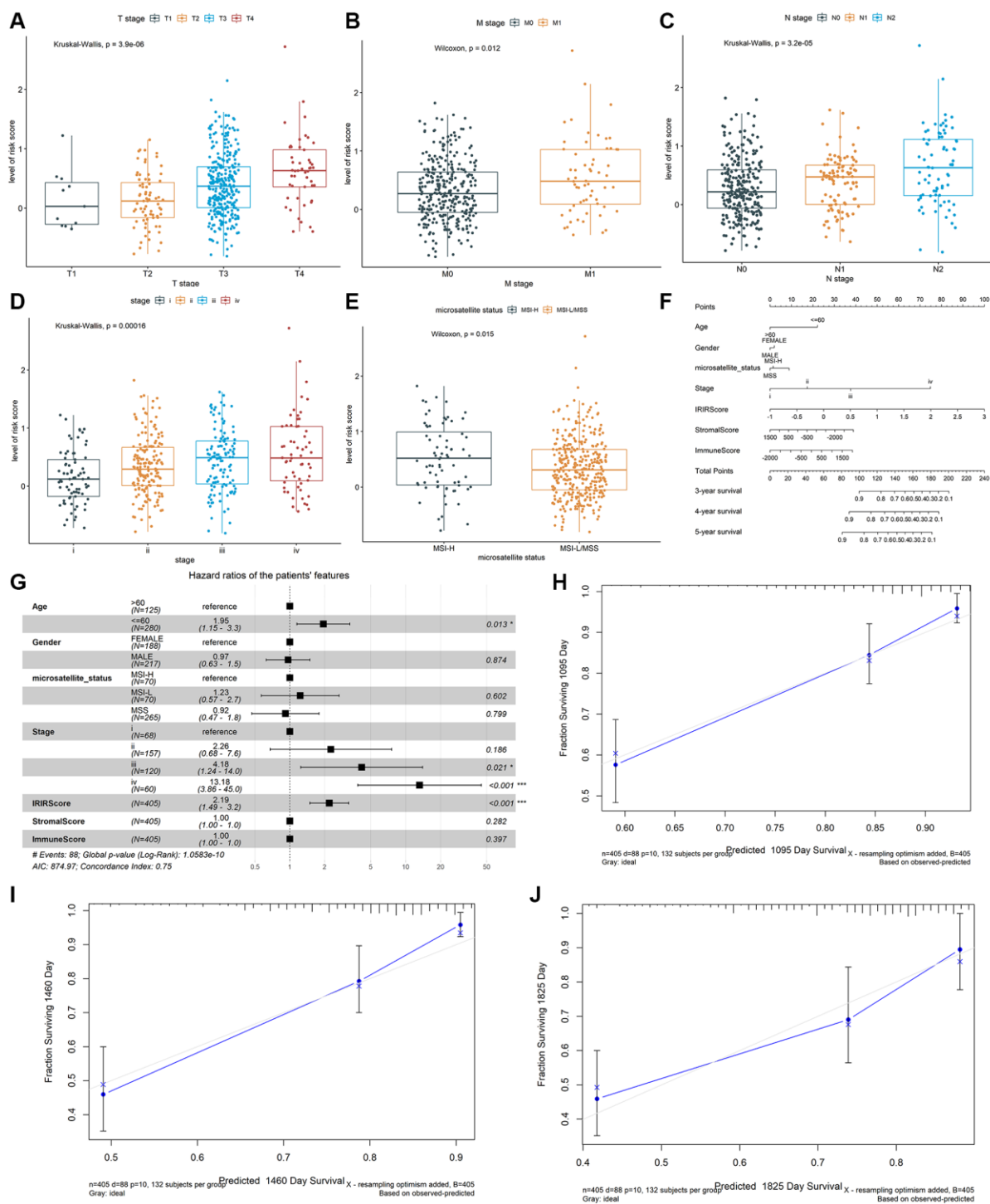


Figure 3. The association between the IRIRScore and the clinical-pathological characteristics, including (A) T stages, (B) M stages, (C) N stages, (D) advanced pathological stages, and (E) microsatellite status Construction and validation of a nomogram. (F) Nomogram to predict the probability of OS in 3, 4 and 5 years for colon cancer. (G) Forest plots showing the associations between patient's characteristics and OS. (H-J) Calibration plot of the nomogram to predict the probability of OS at 3, 4 and 5 years.

predicting 3-, 4-, and 5-years OS probability of patients with colon cancer. In the nomogram, it is easy to estimate the score for each variable on the point scale and calculate the probability of survival at 3-, 4-, 5-years (Figure 3F). Compared with the other variables, IRIRScore had the maximum score points. The forest plot shows that the variables, including the age (>60), advanced pathological stage (III and IV) and IRIRScore are significantly associated with OS ($P < 0.05$, Figure 3G), but the microsatellite status, StromalScore, and ImmuneScore are not ($P > 0.05$).

To verify the predictive performance of the nomogram, we plotted the calibration curves and observed that the predictive curves were close to the ideal curve (Figure 3H–3J). Moreover, the predictive accuracy (C-index) of the nomogram has been dramatically improved from 0.63 to 0.75, indicating good functioning.

Exploring potential immunotherapy related signatures

We estimated the proportions of 22 types of immune cells in patients with colon cancer in the training dataset using the CIBERSORT method. The profiles of high-risk group and low-risk group were shown in Supplementary Figure 3F and 3G. Subsequently, the composition of immune cell types between the high-risk and low-risk groups was compared using the Kruskal-Wallis test. Moreover, remarkable differences were found in T cells CD8, T cells CD4 memory resting, T cells CD4 memory activated, T cells follicular helper, T cells regulatory (Tregs), and Eosinophils ($P < 0.01$, Figure 4A). The Kaplan-Meier curve revealed that high level of T cells regulatory (Tregs) was significantly related to poor OS ($P = 0.029$, Figure 4C), while a high level of T cells CD4 memory resting was significantly related to better OS ($P = 0.0071$, Figure 4B). 10 of the 22 types of tumor-infiltrating immune cells were significantly related to IRIRScore ($P < 0.05$, Supplementary Figure 4). Among them, T cells regulatory (Tregs), Eosinophils, Macrophages M1, T cells follicular helper, T cells CD8, and Macrophages M2 were positively related to IRIRScore, while Dendritic cells activated, Plasma cells, T cells CD4 memory activated, and T cells CD4 memory resting were negatively related to IRIRScore.

The mutation profiles of the high-risk group and low-risk group were plotted and compared using maftools R package. The top 20 significantly mutated genes were *APC*, *TP53*, *TTN*, *KRAS*, *SYNE1*, *MUC16*, *PIK3CA*, *FAT4*, *RYR2*, *OBSCN*, *ZFH4*, *DNAH5*, *PCLO*, *CSMD3*, *ABCA13*, *DNAH11*, *LRP1B*, *FAT3*, *USH2A* and *CSMD1* (Figure 4D). Among them, the mutation rate of *APC* was significantly higher in the low-risk group ($P = 0.000716$, Figure 4E), while the mutation

rate of *MUC16* was significantly higher in the high-risk group ($P = 0.0154$, Figure 4E). However, there was no significant association between the mutation status of *APC* or *MUC16* and patients' OS (Figure 4F). Besides, the TMB value of each sample was calculated and visualized (median value of high-risk group: 2.66/MB; median value of low-risk group: 2/MB, Figure 4G). As shown in Figure 4H, the TMB value of patients in high-risk group was significantly higher than those in low-risk group ($P = 0.00031$). Moreover, the Wilcoxon test was applied to statistically compare the expression of immune checkpoint inhibitor targets between the two groups. The results revealed that the high-risk group had a significantly higher expression level ($P < 0.001$, Figure 5).

The identified IRIGs as potential diagnosis biomarkers

5 of 6 IRIGs were significantly differentially expressed between tumor and normal samples ($P < 0.05$, Figure 6A). Among them, *TDGFI*, and *GRP* were up-regulated in tumor samples, while *NRG1*, *CRLF1*, and *IL20RB* were down-regulated in tumor samples. The IRIGs were used to establish the prediction models with *RF* algorithm, and the AUC of each IRIG was calculated and plotted using the ggroc R package (*IL13RA2*: 0.579; *TDGFI*: 0.813; *NRG1*: 0.933; *CRLF1*: 0.795; *IL20RB*: 0.625; *GRP*: 0.724, Figure 6B). Among them, *TDGFI* and *NRG1* revealed excellent diagnostic predictive performance (AUC > 0.8). The distribution and expression of *CRLF1*, and *NRG1* at the protein level are shown in Figure 6C–6F, whereas *IL13RA2*, *TDGFI*, *IL20RB*, and *GRP* remained inaccessible in HPA.

DISCUSSION

Extensive studies have been conducted to identify novel diagnosis, prognosis, and therapeutic targets for colon cancer [14–17]. However, more effective signatures needed to be explored. In this study, we tried to identify immune-related signatures for predicting the immunotherapeutic responsiveness, prognosis, and diagnosis of colon cancer. Firstly, immunotherapeutic responsiveness-related DEGs were identified by comparing the responders and non-responders in the anti-PD-L1 cohort, intersecting with IRGs, then obtained IRIGs. Then, 6 of these IRIGs (*TDGFI*, *IL13RA2*, *NRG1*, *CRLF1*, *GRP*, and *IL20RB*) were selected to establish an IRIRScore prognostic model. Among them, *TDGF*, *NRG1*, and *GRP* have been demonstrated as prognostic indicators for patients with colon cancer [18–20]. *TDGFI*, a member of the epidermal growth factor-cripto *FRL1* cryptic protein family, is involved in activating several different

signaling pathways during embryonic development and cellular transformation, and could be a predictive marker for metachronous metastasis in patients with

colorectal cancer [21]. Previous studies demonstrated that *IL13RA2* was overexpressed in various cancers, such as malignant gliomas [22], and thyroid Carcinoma

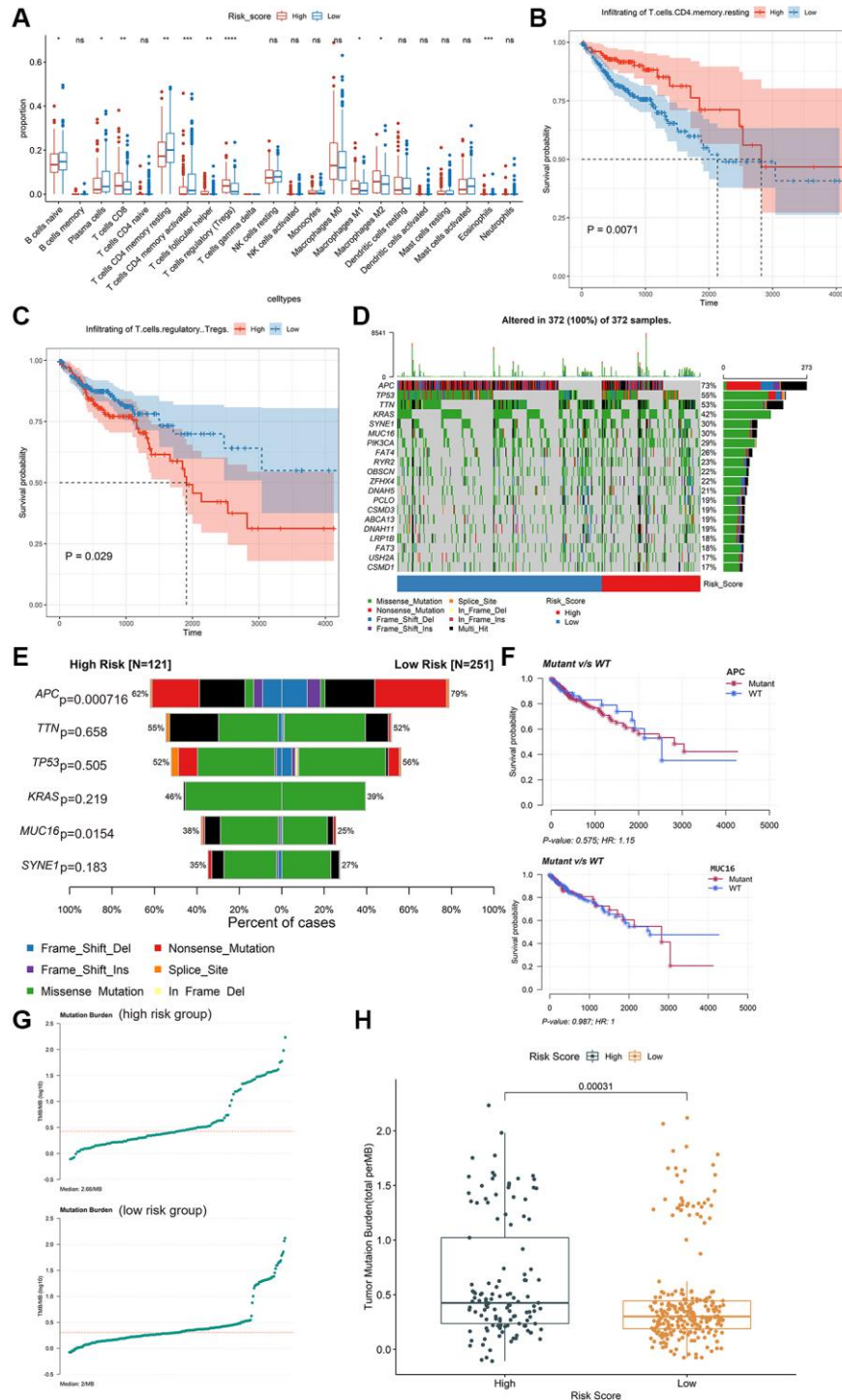


Figure 4. Exploring potential immunotherapy-related signatures. (A) Comparison of the differences in the proportions of immune cells between low-risk group and high-risk group using the Kruskal-Wallis test. The values of P were labelled above each boxplot with asterisks. ($*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$). (B and C) The Kaplan-Meier analysis of the associations between the level of T cells regulatory (Tregs) and T cells CD4 memory resting with patients' OS. (D) The mutation profiles of the high-risk and low-risk groups. (E) Comparison of the mutation rate between high-risk group and low-risk group. (F) The association between the mutation status of *APC* and *MUC16* and patients' OS. (G) The TMB profiles of the low-risk group and high-risk group. (H) Comparison of the difference in TMB between high-risk and low-risk groups.

[23]. In our study, the trend of the expression level of *IL13RA2* was also up-regulated in tumor samples compared with normal samples, though it was not significant ($P = 0.0918$, Figure 6A). *NRG1* can activate *HER3* to promote resistance in tumor cells [24]. *CRLF1*, as the target for miR-3065-3p, could promote the stemness and metastasis of colorectal cancer [25].

Besides, *CRLF1* regulates immune, inflammation, hematopoiesis, cell growth, and differentiation [26], which may relate to tumorigenesis. Moreover, the expression level of *CRLF1* was down-regulated in tumor samples compared with normal samples ($P = 4.1e^{-10}$, Figure 6A, 6E and 6F). Gastrin Releasing Peptide (*GRP*), acting synergistically to promote cell

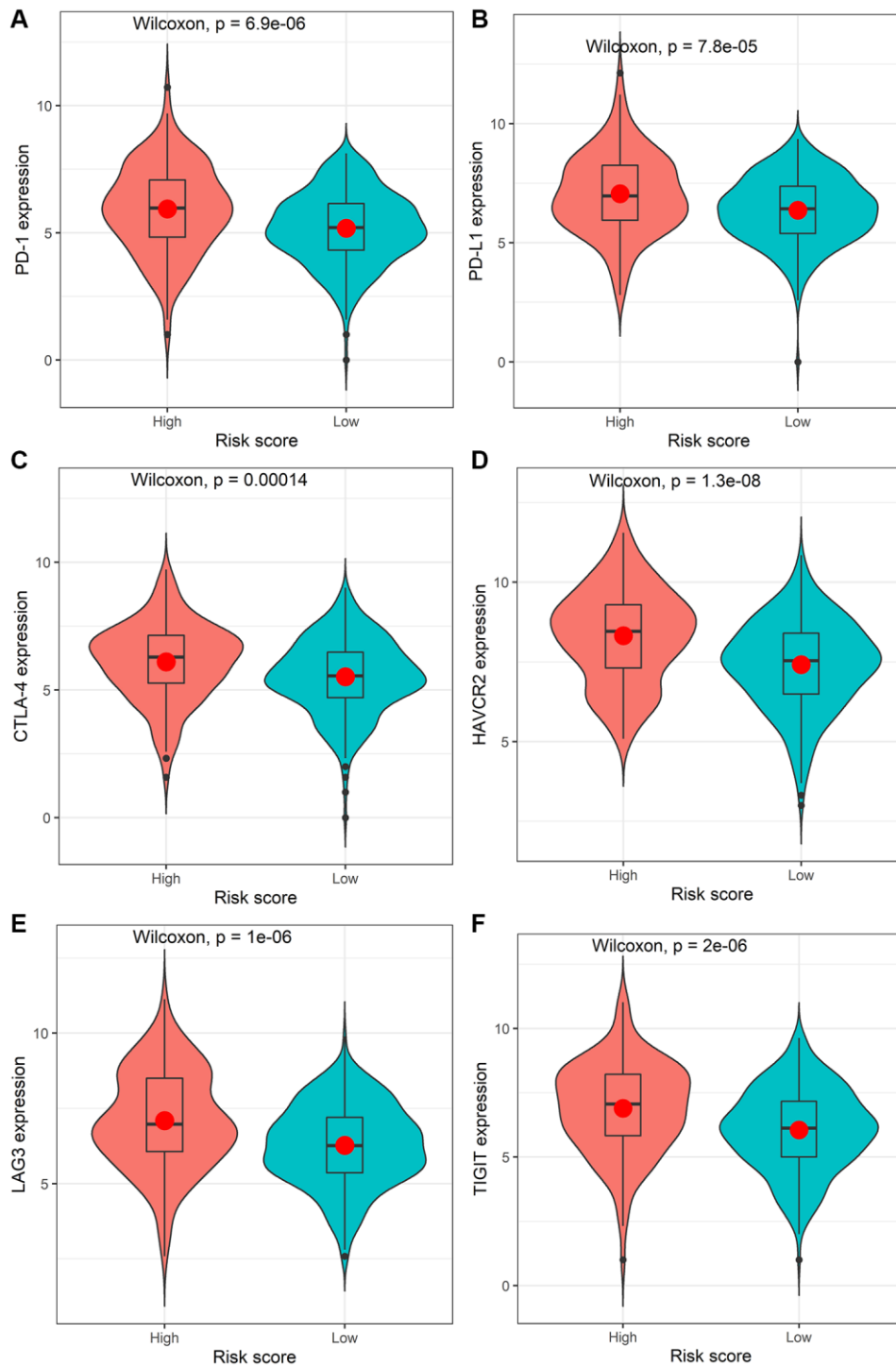


Figure 5. Comparison of the expression levels of the immune checkpoints and their ligands between the high-risk score group and low-risk score group. The expression of *PD-1* (A), *PD-L1* (B), *CTLA-4* (C), *HAVCR2* (D), *LAG3* (E), or *TIGIT* (F).

proliferation, play an essential role in cancer development and are frequently over-expressed in various tumors [27]. According to the results of our study, *GRP* was also up-regulated in tumor samples ($P = 1.9e^{-6}$, Figure 6A). Previous studies have demonstrated that *IL20RB* was associated with the prognosis of various cancers, such

as lung adenocarcinoma [28], pancreatic ductal adenocarcinoma [29], and pancreatic cancer [30].

The identified immune-related signatures in predicting prognosis and TNM staging were validated in multiple cohorts, and displayed robust capability. Moreover, a nomogram was established based on the IRIRScore and

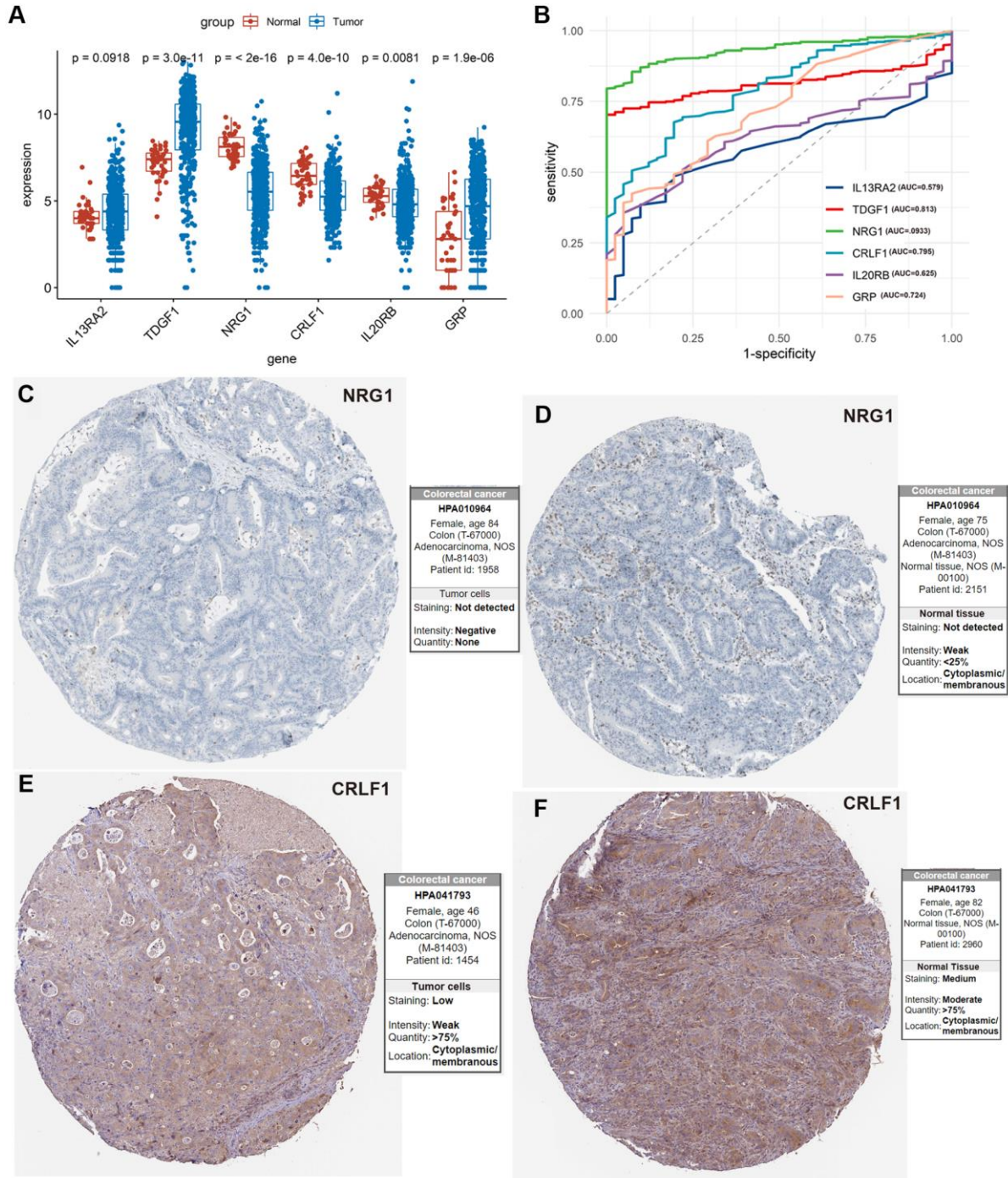


Figure 6. Exploration of the 6 IRIGs as predictive diagnostic indicators based on the TCGA-COAD cohort. (A) Comparison of the expressions of the IRIGs between tumor and normal samples using the Wilcoxon test. **(B)** Establishment of the predictive diagnosis model using the 6 IRIGs. **(C–F)** The comparison of protein expression of *NRG1* and *CRLF1* between tumor and normal tissues.

other clinical characters, and the C-index of the nomogram was remarkably higher than the IRIScore model. These findings may be helpful for the prognosis of patients with colon cancer. Additionally, we found that patients in the high-risk group had significantly higher expression level of immune checkpoint inhibitor targets than those in the low-risk group (Figure 5), suggesting that high-risk patients may respond better to immune checkpoint inhibitors targeting *PD-1*, *PD-L1*, *CTLA-4*, *HAVCR2*, *LAG3*, or *TIGIT*. Besides, higher TMB value (Figure 4H), more instability of the microsatellite status (Figure 3E), or a higher mutation rate of *MUC16* (Figure 4E) are significantly associated with patients in the high-risk group, suggesting that those patients are more likely to benefit from immunotherapy. However, patients in the high-risk group exhibit a low mutation rate of *APC* (Figure 4E), suggesting that *APC* may be a tumor suppressor gene. In addition, patients with sensitive immunotherapeutic responsiveness may have a significantly better OS (Supplementary Figure 3C, 3E).

Furthermore, we investigated the 6 prognostic IRIGs in the diagnostic prediction of colon cancer. Among them, *TDGF1* and *NRG1* could be used as an independent diagnostic indicator and reveal excellent performance (Figure 6B, AUC >0.8).

In conclusion, our research provided new insights into immune-related immunotherapeutic responsiveness, prognosis, and diagnosis of colon cancer. However, there are still some limitations to our study. First, the datasets used in our study are based on public databases, and more experimental data were needed to verify our findings. Moreover, the molecular function and mechanism of the prognostic IRIGs need to be further investigated.

AUTHOR CONTRIBUTIONS

Lichao Cao was involved in the study concept and design, and drafting the manuscript. Hezi Zhang and Ying Ba put forward some kind suggestions. Jin Yang helped in analyzing and interpreting of the data.

CONFLICTS OF INTEREST

Authors Hezi Zhang and Ying Ba were employed by Shenzhen Nucleus Gene Technology Co., Ltd. The remaining authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflicts of interest.

FUNDING

This work was supported by a grant from the Key Science and Technology Program of Shaanxi Province

(2019ZDLSF02-05) and the National Natural Scientific Foundation of China (81974378 and 82003115).

REFERENCES

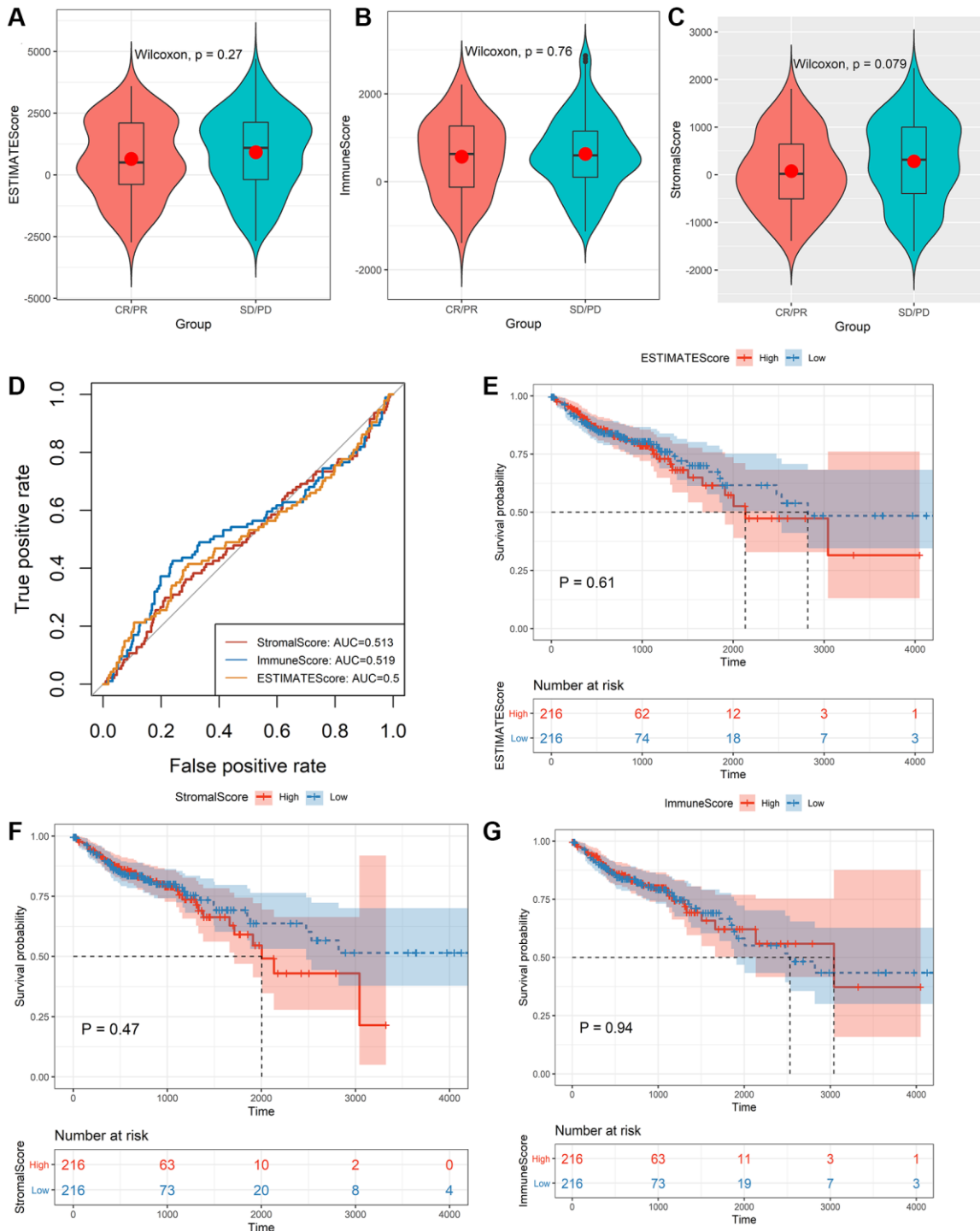
1. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA Cancer J Clin.* 2022; 72:7–33. <https://doi.org/10.3322/caac.21708> PMID:35020204
2. Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, Li J, Li F, Tan HB. Immune cells within the tumor microenvironment: Biological functions and roles in cancer immunotherapy. *Cancer Lett.* 2020; 470:126–33. <https://doi.org/10.1016/j.canlet.2019.11.009> PMID:31730903
3. Cao L, Li T, Ba Y, Chen E, Yang J, Zhang H. Exploring Immune-Related Prognostic Signatures in the Tumor Microenvironment of Colon Cancer. *Front Genet.* 2022; 13:801484. <https://doi.org/10.3389/fgene.2022.801484> PMID:35281839
4. Liu J, Huang X, Liu H, Wei C, Ru H, Qin H, Lai H, Meng Y, Wu G, Xie W, Mo X, Johnson CH, Zhang Y, Tang W. Immune landscape and prognostic immune-related genes in KRAS-mutant colorectal cancer patients. *J Transl Med.* 2021; 19:27. <https://doi.org/10.1186/s12967-020-02638-9> PMID:33413474
5. Li X, Wen D, Li X, Yao C, Chong W, Chen H. Identification of an Immune Signature Predicting Prognosis Risk and Lymphocyte Infiltration in Colon Cancer. *Front Immunol.* 2020; 11:1678. <https://doi.org/10.3389/fimmu.2020.01678> PMID:33013820
6. Franke AJ, Skelton WP, Starr JS, Parekh H, Lee JJ, Overman MJ, Allegra C, George TJ. Immunotherapy for Colorectal Cancer: A Review of Current and Novel Therapeutic Approaches. *J Natl Cancer Inst.* 2019; 111:1131–41. <https://doi.org/10.1093/jnci/djz093> PMID:31322663
7. Lichtenstern CR, Ngu RK, Shalpour S, Karin M. Immunotherapy, Inflammation and Colorectal Cancer. *Cells.* 2020; 9:618. <https://doi.org/10.3390/cells9030618> PMID:32143413
8. Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel EE III, Koeppen H, Astarita JL, Cubas R, Jhunjhunwala S, Banchereau R, Yang Y, et al. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature.* 2018; 554:544–8.

- <https://doi.org/10.1038/nature25501>
PMID:29443960
9. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun.* 2013; 4:2612.
<https://doi.org/10.1038/ncomms3612>
PMID:24113773
10. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* 2010; 26:139–40.
<https://doi.org/10.1093/bioinformatics/btp616>
PMID:19910308
11. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012; 16:284–7.
<https://doi.org/10.1089/omi.2011.0118>
PMID:22455463
12. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015; 12:453–7.
<https://doi.org/10.1038/nmeth.3337>
PMID:25822800
13. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 2018; 28:1747–56.
<https://doi.org/10.1101/gr.239244.118>
PMID:30341162
14. Cao L, Chen E, Zhang H, Ba Y, Yan B, Li T, Yang J. Construction of a novel methylation-related prognostic model for colorectal cancer based on microsatellite status. *J Cell Biochem.* 2021; 122:1781–90.
<https://doi.org/10.1002/jcb.30131>
PMID:34397105
15. Cai C, Peng Y, Shen E, Wan R, Gao L, Gao Y, Zhou Y, Huang Q, Chen Y, Liu P, Guo C, Feng Z, Zhang X, et al. Identification of tumour immune infiltration-associated snoRNAs (TIIsno) for predicting prognosis and immune landscape in patients with colon cancer via a TIIsno score model. *EBioMedicine.* 2022; 76:103866.
<https://doi.org/10.1016/j.ebiom.2022.103866>
PMID:35144219
16. Cui Z, Sun G, Bhandari R, Lu J, Zhang M, Bhandari R, Sun F, Liu Z, Zhao S. Comprehensive Analysis of Glycolysis-Related Genes for Prognosis, Immune Features, and Candidate Drug Development in Colon Cancer. *Front Cell Dev Biol.* 2021; 9:684322.
<https://doi.org/10.3389/fcell.2021.684322>
PMID:34422808
17. Zhou R, Zhang J, Zeng D, Sun H, Rong X, Shi M, Bin J, Liao Y, Liao W. Immune cell infiltration as a biomarker for the diagnosis and prognosis of stage I-III colon cancer. *Cancer Immunol Immunother.* 2019; 68:433–42.
<https://doi.org/10.1007/s00262-018-2289-7>
PMID:30564892
18. Sun YL, Zhang Y, Guo YC, Yang ZH, Xu YC. A Prognostic Model Based on the Immune-related Genes in Colon Adenocarcinoma. *Int J Med Sci.* 2020; 17:1879–96.
<https://doi.org/10.7150/ijms.45813>
PMID:32788867
19. Xu J, Dai S, Yuan Y, Xiao Q, Ding K. A Prognostic Model for Colon Cancer Patients Based on Eight Signature Autophagy Genes. *Front Cell Dev Biol.* 2020; 8:602174.
<https://doi.org/10.3389/fcell.2020.602174>
PMID:33324651
20. Li C, Shen Z, Zhou Y, Yu W. Independent prognostic genes and mechanism investigation for colon cancer. *Biol Res.* 2018; 51:10.
<https://doi.org/10.1186/s40659-018-0158-7>
PMID:29653552
21. Miyoshi N, Ishii H, Mimori K, Sekimoto M, Doki Y, Mori M. TDGF1 is a novel predictive marker for metachronous metastasis of colorectal cancer. *Int J Oncol.* 2010; 36:563–8.
<https://doi.org/10.3892/ijo.00000530>
PMID:20126975
22. Zeng J, Zhang J, Yang YZ, Wang F, Jiang H, Chen HD, Wu HY, Sai K, Hu WM. IL13RA2 is overexpressed in malignant gliomas and related to clinical outcome of patients. *Am J Transl Res.* 2020; 12:4702–14.
PMID:32913543
23. Chong ST, Tan KM, Kok CYL, Guan SP, Lai SH, Lim C, Hu J, Sturgis C, Eng C, Lam PYP, Ngeow J. IL13RA2 Is Differentially Regulated in Papillary Thyroid Carcinoma vs Follicular Thyroid Carcinoma. *J Clin Endocrinol Metab.* 2019; 104:5573–84.
<https://doi.org/10.1210/jc.2019-00040>
PMID:31290966
24. Zhang Z, Karthaus WR, Lee YS, Gao VR, Wu C, Russo JW, Liu M, Mota JM, Abida W, Linton E, Lee E, Barnes SD, Chen HA, et al. Tumor Microenvironment-Derived NRG1 Promotes Antiandrogen Resistance in Prostate Cancer. *Cancer Cell.* 2020; 38:279–96.e9.
<https://doi.org/10.1016/j.ccell.2020.06.005>
PMID:32679108

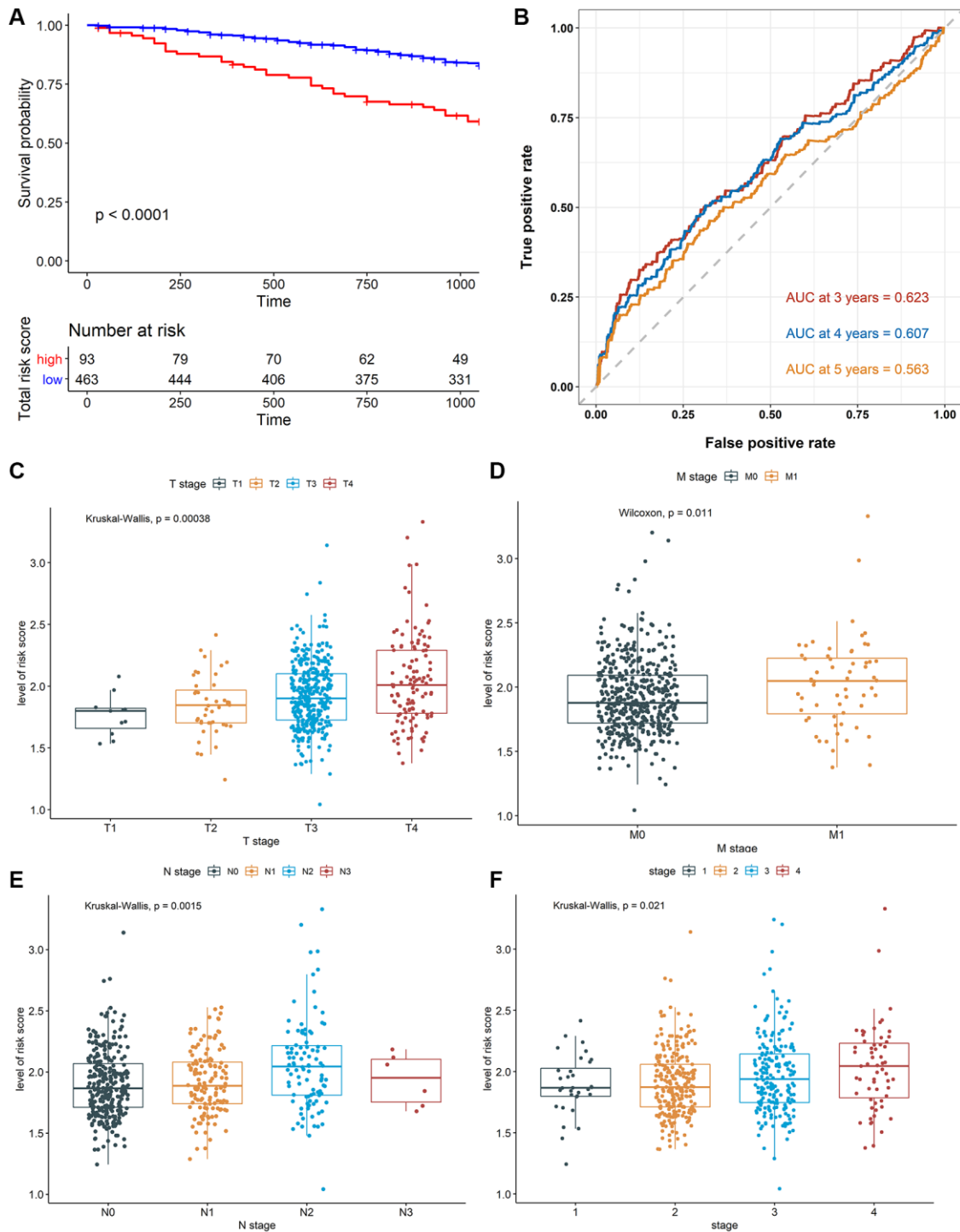
25. Li Y, Xun J, Wang B, Ma Y, Zhang L, Yang L, Gao R, Guan J, Liu T, Gao H, Wang X, Zhang Q. miR-3065-3p promotes stemness and metastasis by targeting CRLF1 in colorectal cancer. *J Transl Med.* 2021; 19:429.
<https://doi.org/10.1186/s12967-021-03102-y>
PMID:[34656128](https://pubmed.ncbi.nlm.nih.gov/34656128/)
26. Crisponi L, Buers I, Rutsch F. CRLF1 and CLCF1 in Development, Health and Disease. *Int J Mol Sci.* 2022; 23:992.
<https://doi.org/10.3390/ijms23020992>
PMID:[35055176](https://pubmed.ncbi.nlm.nih.gov/35055176/)
27. Palmioli A, Nicolini G, Tripodi F, Orsato A, Ceresa C, Donzelli E, Arici M, Coccetti P, Rocchetti M, La Ferla B, Airoldi C. Targeting GRP receptor: Design, synthesis and preliminary biological characterization of new non-peptide antagonists of bombesin. *Bioorg Chem.* 2021; 109:104739.
<https://doi.org/10.1016/j.bioorg.2021.104739>
PMID:[33626451](https://pubmed.ncbi.nlm.nih.gov/33626451/)
28. Zhang M, Zhu K, Pu H, Wang Z, Zhao H, Zhang J, Wang Y. An Immune-Related Signature Predicts Survival in Patients With Lung Adenocarcinoma. *Front Oncol.* 2019; 9:1314.
<https://doi.org/10.3389/fonc.2019.01314>
PMID:[31921619](https://pubmed.ncbi.nlm.nih.gov/31921619/)
29. Wang W, Yan L, Guan X, Dong B, Zhao M, Wu J, Tian X, Hao C. Identification of an Immune-Related Signature for Predicting Prognosis in Patients With Pancreatic Ductal Adenocarcinoma. *Front Oncol.* 2021; 10:618215.
<https://doi.org/10.3389/fonc.2020.618215>
PMID:[33718118](https://pubmed.ncbi.nlm.nih.gov/33718118/)
30. Liu B, Fu T, He P, Du C, Xu K. Construction of a five-gene prognostic model based on immune-related genes for the prediction of survival in pancreatic cancer. *Biosci Rep.* 2021; 41:BSR20204301.
<https://doi.org/10.1042/BSR20204301>
PMID:[34143198](https://pubmed.ncbi.nlm.nih.gov/34143198/)

SUPPLEMENTARY MATERIALS

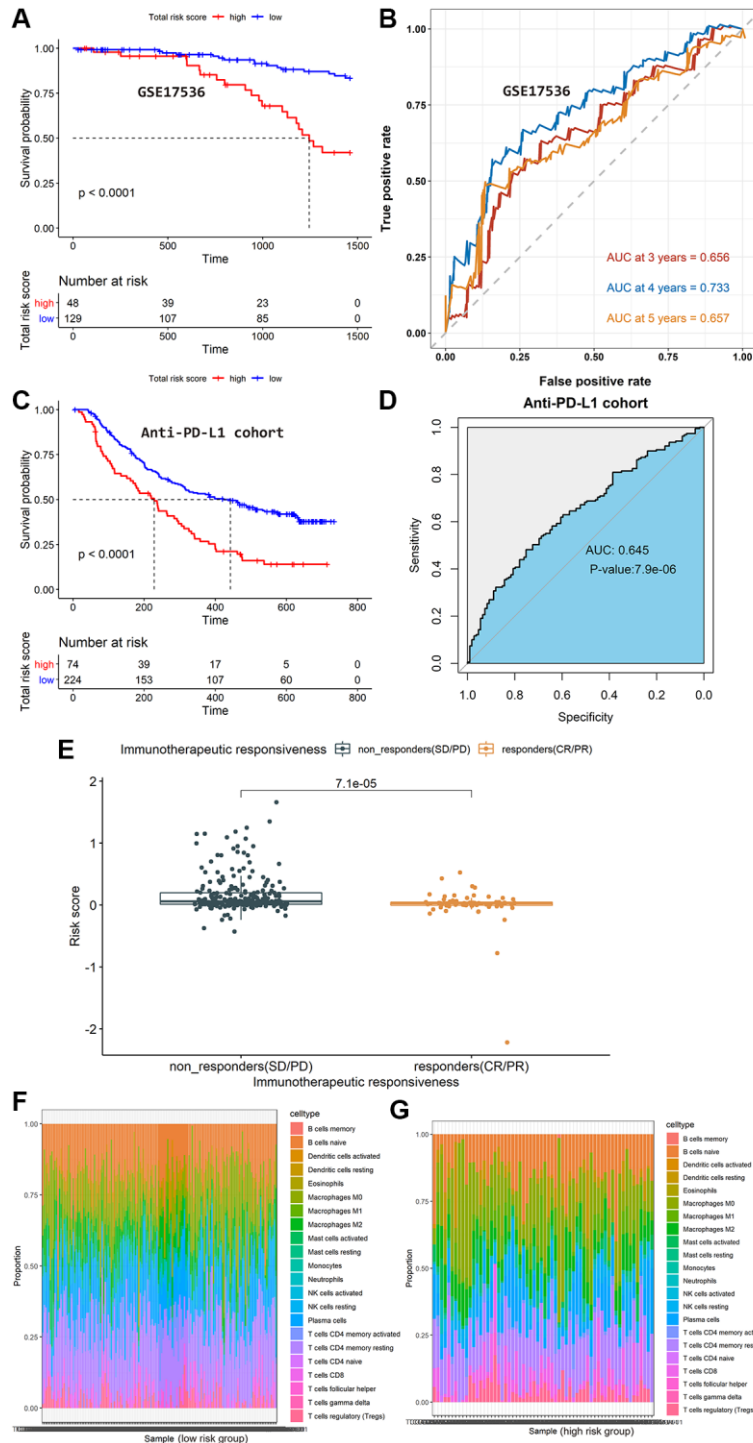
Supplementary Figures



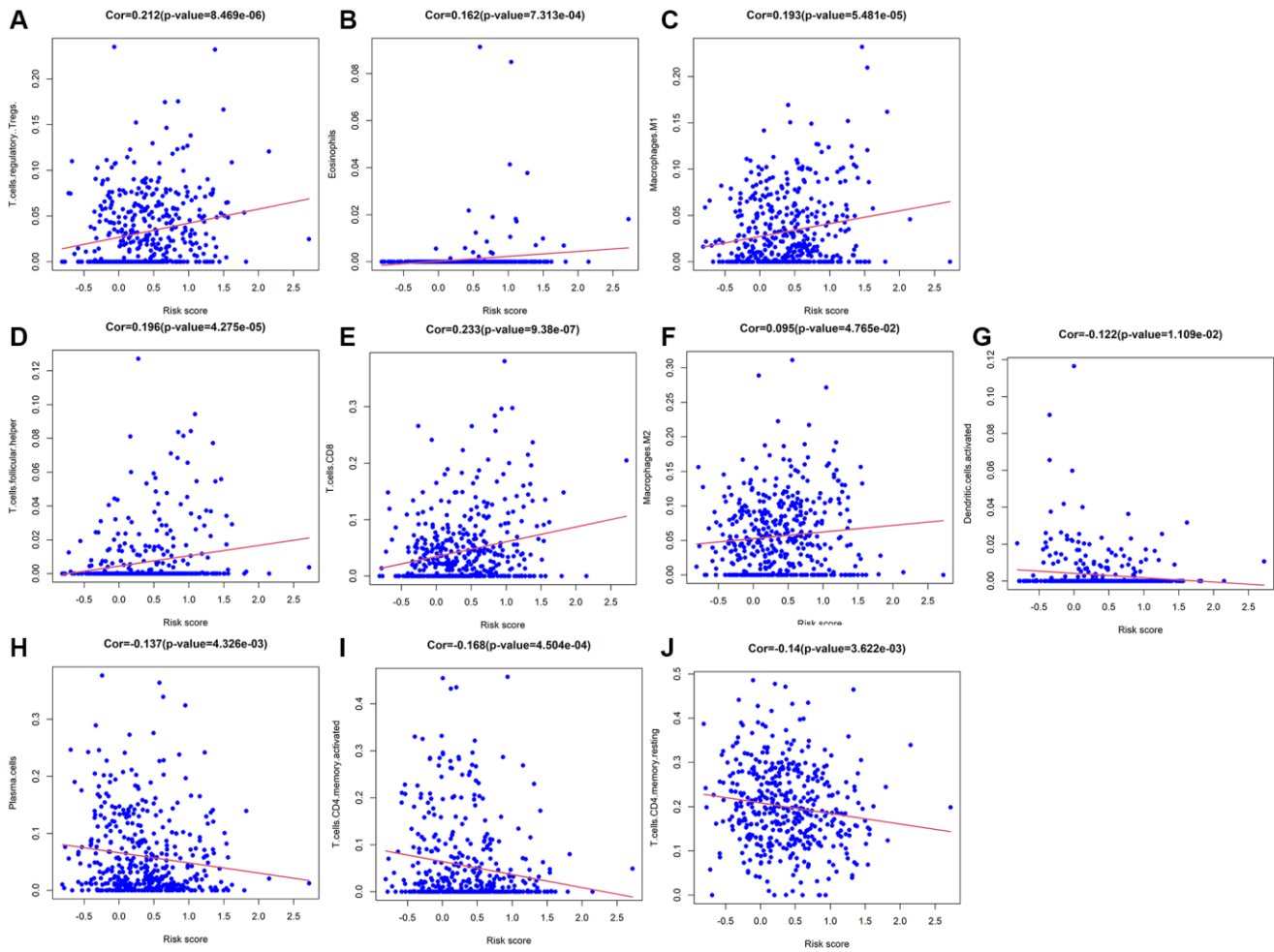
Supplementary Figure 1. Construction and assessment of estimate-stromal-immune scores model. (A–C) Comparison of the differences of the estimate-stromal-immune scores between responders and non-responders using Wilcoxon test. (D) The AUCs of the estimate-stromal-immune scores model. (E–G) Kaplan-Meier curves analyses of the association between the scores (estimate, stromal, or immune) and patients’ OS.



Supplementary Figure 2. Validation of the identified immune-related prognostic signatures using GSE39582. (A) Kaplan-Meier curve analysis. (B) Time-dependent ROC curve analysis. (C–F) Comparison of the association between the IRIRScore of the colon cancer patients and the clinical-pathological characteristics, including (C) T stages, (D) M stages, (E) N stages and (F) advanced pathological stages.



Supplementary Figure 3. Validation of the identified immune-related prognostic signatures using GSE17536 and anti-PD-L1 cohort. (A) Kaplan-Meier curve analysis using GSE17536. (B) Time-dependent ROC curve analysis using 17536. (C) Kaplan-Meier curve analysis using anti-PD-L1 cohort. (D) Time-dependent ROC curve analysis using anti-PD-L1 cohort. (E) Comparing the difference in the risk score between responders and non-responders in anti-PD-L1 cohort. (F–G) The profiles of relative proportions of immune infiltration. (F) Patients in low-risk group, (G) patients in high-risk group.



Supplementary Figure 4. The correlation between IRIScore and the proportions of immune cells infiltrating, including (A) T cells regulatory (Tregs), (B) Eosinophils, (C) Macrophages M1, (D) T cells follicular helper, (E) T cells CD8, (F) Macrophages M2, (G) Dendritic cells activated, (H) Plasma cells, (I) T cells CD4 memory activated, (J) T cells CD4 memory resting.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Table 4.

Supplementary Table 1. The detailed information of TCGA-COAD cohort.

	Tumor (N = 432)	Normal (N = 39)	Overall (N = 471)
Gender			
Female	200 (46.3%)	20 (51.3%)	220 (46.7%)
Male	232 (53.7%)	19 (48.7%)	251 (53.3%)
Age			
≤60	136 (31.5%)	9 (23.1%)	145 (30.8%)
>60	296 (68.5%)	30 (76.9%)	326 (69.2%)
AJCC-Stage			
I	73 (16.9%)	4 (10.3%)	77 (16.3%)
II	166 (38.4%)	21 (53.8%)	187 (39.7%)
III	122 (28.2%)	6 (15.4%)	128 (27.2%)
IV	60 (13.9%)	7 (17.9%)	67 (14.2%)
Missing	11 (2.5%)	1 (2.6%)	12 (2.5%)
AJCC-T			
T1	11 (2.5%)	0 (0%)	11 (2.3%)
T2	75 (17.4%)	5 (12.8%)	80 (17.0%)
T3	296 (68.5%)	28 (71.8%)	324 (68.8%)
T4	49 (11.3%)	6 (15.4%)	55 (11.7%)
Missing	1 (0.2%)	0 (0%)	1 (0.2%)
AJCC-M			
M0	318 (73.6%)	25 (64.1%)	343 (72.8%)
M1	60 (13.9%)	7 (17.9%)	67 (14.2%)
MX	47 (10.9%)	6 (15.4%)	53 (11.3%)
Missing	7 (1.6%)	1 (2.6%)	8 (1.7%)
AJCC-N			
N0	254 (58.8%)	27 (69.2%)	281 (59.7%)
N1	100 (23.1%)	7 (17.9%)	107 (22.7%)
N2	78 (18.1%)	5 (12.8%)	83 (17.6%)

Supplementary Table 2. The detailed information of GSE39582.

	Alive (N = 369)	Death (N = 187)	Overall (N = 556)
Gender			
Female	175 (47.4%)	74 (39.6%)	249 (44.8%)
Male	194 (52.6%)	113 (60.4%)	307 (55.2%)
Age			
≤60	115 (31.2%)	42 (22.5%)	157 (28.2%)
>60	254 (68.8%)	145 (77.5%)	399 (71.8%)
AJCC-Stage			
0	3 (0.8%)	1 (0.5%)	4 (0.7%)
1	27 (7.3%)	5 (2.7%)	32 (5.8%)

2	183 (49.6%)	75 (40.1%)	258 (46.4%)
3	136 (36.9%)	67 (35.8%)	203 (36.5%)
4	20 (5.4%)	39 (20.9%)	59 (10.6%)
AJCC-T			
N/A	8 (2.2%)	12 (6.4%)	20 (3.6%)
T0	1 (0.3%)	0 (0%)	1 (0.2%)
T1	10 (2.7%)	1 (0.5%)	11 (2.0%)
T2	36 (9.8%)	8 (4.3%)	44 (7.9%)
T3	249 (67.5%)	111 (59.4%)	360 (64.7%)
T4	63 (17.1%)	54 (28.9%)	117 (21.0%)
Tis	2 (0.5%)	1 (0.5%)	3 (0.5%)
AJCC-M			
M0	340 (92.1%)	134 (71.7%)	474 (85.3%)
M1	20 (5.4%)	40 (21.4%)	60 (10.8%)
MX	1 (0.3%)	1 (0.5%)	2 (0.4%)
N/A	8 (2.2%)	12 (6.4%)	20 (3.6%)
AJCC-N			
N/A	8 (2.2%)	12 (6.4%)	20 (3.6%)
N+	1 (0.3%)	5 (2.7%)	6 (1.1%)
N0	214 (58.0%)	81 (43.3%)	295 (53.1%)
N1	91 (24.7%)	40 (21.4%)	131 (23.6%)
N2	52 (14.1%)	46 (24.6%)	98 (17.6%)
N3	3 (0.8%)	3 (1.6%)	6 (1.1%)

Abbreviation: AJCC: American Joint Committee on Cancer.

Supplementary Table 3. The detailed information of GSE17536.

	Alive (<i>N</i> = 73)	Death (<i>N</i> = 104)	Overall (<i>N</i> = 177)
Gender			
Female	32 (43.8%)	49 (47.1%)	81 (45.8%)
Male	41 (56.2%)	55 (52.9%)	96 (54.2%)
Age			
≤60	26 (35.6%)	33 (31.7%)	59 (33.3%)
>60	47 (64.4%)	71 (68.3%)	118 (66.7%)
AJCC-stage			
1	4 (5.5%)	20 (19.2%)	24 (13.6%)
2	12 (16.4%)	45 (43.3%)	57 (32.2%)
3	25 (34.2%)	32 (30.8%)	57 (32.2%)
4	32 (43.8%)	7 (6.7%)	39 (22.0%)

Abbreviation: AJCC: American Joint Committee on Cancer.

Supplementary Table 4. The detailed information of the immunotherapeutic response related DEGs.

Supplementary Table 5. The detailed information of IRIGs.

Gene Name	logFC	logCPM	LR	P value	FDR	change
NPY	8.07268	1.07696	122.324	1.96E-28	5.54E-25	UP
SEMG2	4.07083	3.58072	57.4661	3.44E-14	1.39E-11	UP
ALB	-5.7925	8.50241	54.1559	1.85E-13	6.58E-11	DOWN
SCGB3A1	3.0018	0.24711	48.5146	3.28E-12	9.02E-10	UP
IL13RA2	1.69046	1.61056	43.4311	4.39E-11	1.03E-08	UP
HAMP	-2.7598	2.91445	40.748	1.73E-10	3.61E-08	DOWN
ORM1	-5.4108	7.03914	40.1051	2.41E-10	4.89E-08	DOWN
PRL	2.7778	-3.3736	39.2792	3.67E-10	6.96E-08	UP
SEMG1	3.56408	4.18769	38.7377	4.85E-10	8.81E-08	UP
HRG	-5.9652	3.80736	37.9741	7.17E-10	1.29E-07	DOWN
KLRC2	1.30054	1.04129	34.6247	4.00E-09	5.86E-07	UP
SAA2	-3.3183	3.34611	34.1954	4.98E-09	7.11E-07	DOWN
DMBT1	2.09005	5.80462	34.0219	5.45E-09	7.74E-07	UP
IL17RB	1.08385	2.40642	33.9975	5.52E-09	7.80E-07	UP
FGF9	1.40722	-0.0577	33.3619	7.65E-09	1.04E-06	UP
GLP1R	1.66604	-1.0846	33.2311	8.18E-09	1.11E-06	UP
IL1RL1	-1.262	3.07036	32.1206	1.45E-08	1.85E-06	DOWN
SAA1	-2.7309	4.41777	30.667	3.06E-08	3.58E-06	DOWN
SSTR2	1.05883	1.12259	29.3184	6.14E-08	6.65E-06	UP
KLRC3	1.13029	0.34016	28.7809	8.10E-08	8.60E-06	UP
KNG1	-5.0457	3.16261	27.9993	1.21E-07	1.25E-05	DOWN
TDGF1	-3.0484	-2.3637	27.841	1.32E-07	1.34E-05	DOWN
FGA	-5.7725	7.00837	27.3696	1.68E-07	1.63E-05	DOWN
VTN	-2.6408	2.36105	27.0228	2.01E-07	1.89E-05	DOWN
LEP	2.07134	0.97879	26.6754	2.41E-07	2.21E-05	UP
FCN2	-3.8681	-1.2927	25.4887	4.45E-07	3.82E-05	DOWN
INHBC	-2.8824	-1.992	25.196	5.18E-07	4.40E-05	DOWN
RXRG	-2.3664	-1.4136	24.6196	6.98E-07	5.65E-05	DOWN
NRG1	-1.4294	3.85232	23.9393	9.94E-07	7.79E-05	DOWN
PRLR	1.25391	3.20478	23.5388	1.22E-06	9.43E-05	UP
RXFP2	-4.867	-0.0549	22.0335	2.68E-06	0.00018	DOWN
GDF5	-1.7395	-1.5743	21.2493	4.03E-06	0.00025	DOWN
SERPINA3	-1.871	7.13183	21.1975	4.14E-06	0.00026	DOWN
MASP2	-4.8815	-1.1652	20.8945	4.85E-06	0.0003	DOWN
ANGPTL3	-5.3865	1.8288	20.8437	4.98E-06	0.0003	DOWN
CRLF1	-1.2931	1.51951	20.1631	7.11E-06	0.0004	DOWN
CCL16	-1.8697	-2.445	19.9142	8.10E-06	0.00045	DOWN
ADCYAP1R1	1.06672	-0.6269	19.2477	1.15E-05	0.0006	UP
DEFA5	5.28237	-2.572	18.9598	1.34E-05	0.00067	UP
APOH	-6.1389	2.79076	18.8944	1.38E-05	0.00069	DOWN
CXCL9	1.12379	4.63729	18.5748	1.63E-05	0.00079	UP
GCGR	-1.5663	-0.9993	18.5641	1.64E-05	0.00079	DOWN
HTR3E	1.67214	-2.806	18.2836	1.90E-05	0.00088	UP
IL36G	-1.9226	0.04119	18.2516	1.94E-05	0.0009	DOWN
IL5RA	-1.3291	-0.0271	16.9115	3.92E-05	0.00157	DOWN
MTNR1B	1.96453	-0.1287	16.5586	4.72E-05	0.00183	UP
PLA2G2A	-1.4746	6.56759	16.193	5.72E-05	0.00214	DOWN
PPBP	1.22015	1.75033	16.1397	5.88E-05	0.00219	UP
STAB2	-1.4416	2.45814	14.9072	0.00011	0.00362	DOWN
IL11	-1.0718	-0.5759	14.8993	0.00011	0.00363	DOWN
OPRD1	2.0922	-3.7315	14.7374	0.00012	0.00388	UP
NR1I2	-1.3061	-0.7679	14.6061	0.00013	0.00408	DOWN
PCSK2	1.36054	-1.7084	14.5353	0.00014	0.00416	UP
CCL14	-1.0164	0.53176	14.4263	0.00015	0.00436	DOWN

HNF4A	-1.6919	1.32014	13.7062	0.00021	0.00583	DOWN
CGA	-3.3334	1.01221	13.4856	0.00024	0.00638	DOWN
PROC	-1.0469	0.93374	13.4168	0.00025	0.00656	DOWN
NOS2	1.01727	1.82741	13.1488	0.00029	0.00727	UP
ANGPT4	-1.5943	-1.3712	13.0615	0.0003	0.00751	DOWN
TRHR	2.33717	-3.8388	13.0064	0.00031	0.00766	UP
NOS1	-1.4171	-0.5483	12.8839	0.00033	0.008	DOWN
IFNG	1.0106	-1.8537	12.8652	0.00033	0.00807	UP
TMPRSS6	-1.3025	0.39655	12.7854	0.00035	0.00833	DOWN
GFAP	1.55922	-2.0394	12.4555	0.00042	0.00958	UP
CRHR1	1.67197	-3.9325	12.4378	0.00042	0.00963	UP
MC4R	1.16876	-1.6198	12.2442	0.00047	0.01036	UP
AGTR2	-2.3464	-1.3464	12.2296	0.00047	0.01041	DOWN
IL20RB	-1.1126	3.46575	12.0512	0.00052	0.01111	DOWN
PRTN3	-2.4749	-3.6688	11.645	0.00064	0.01288	DOWN
SCTR	-2.445	-0.92	11.6004	0.00066	0.01309	DOWN
CRHR2	-1.6053	-2.3542	11.2189	0.00081	0.01529	DOWN
LECT2	-5.8868	-2.1614	10.7966	0.00102	0.01801	DOWN
KIR3DL1	1.25468	-2.4293	10.6348	0.00111	0.01918	UP
CHGA	-2.3938	1.94464	10.2617	0.00136	0.02219	DOWN
IL36A	-2.0358	-3.1271	10.2502	0.00137	0.02228	DOWN
EPGN	1.09948	1.28689	10.1555	0.00144	0.02301	UP
KIR2DL3	1.14603	-2.5951	10.0633	0.00151	0.02386	UP
CCL21	-1.1071	3.14221	9.19801	0.00242	0.0337	DOWN
NPPA	2.94851	-4.0574	9.18006	0.00245	0.03389	UP
AZU1	-1.1304	-3.6622	9.02722	0.00266	0.03611	DOWN
GRP	-1.1081	-1.1179	8.84953	0.00293	0.0388	DOWN
GDF2	-3.1711	-3.4948	8.11006	0.0044	0.05169	DOWN
IL20	-1.5832	-1.4149	7.75751	0.00535	0.05892	DOWN
REG1A	-2.1373	-1.4859	7.59447	0.00585	0.06245	DOWN
TRH	1.0349	-4.1896	7.14708	0.00751	0.07464	UP
CRP	-2.4574	4.84234	7.14605	0.00751	0.07466	DOWN
LEFTY1	-1.1188	-3.2202	7.10331	0.00769	0.07584	DOWN
IFNA7	-4.6106	-2.5514	6.89712	0.00863	0.08167	DOWN
SLURP1	1.05007	-2.4612	6.73495	0.00945	0.0869	UP
DEFA3	-1.6148	-2.5118	6.56553	0.0104	0.09269	DOWN
IFNA10	-4.477	-2.7021	6.3096	0.01201	0.10241	DOWN
SST	-2.1199	3.65444	6.06229	0.01381	0.11247	DOWN
IFNA13	-2.2656	-4.1362	5.73674	0.01661	0.12666	DOWN
GCG	-6.9615	-0.3735	5.71844	0.01679	0.12747	DOWN
GHRHR	-1.6539	-1.4075	5.62295	0.01773	0.13218	DOWN
IFNA4	-4.3846	-3.2965	5.43798	0.0197	0.14048	DOWN
FGF23	-2.0547	-3.3676	5.33392	0.02091	0.14594	DOWN
CSH2	-2.3671	-3.0927	5.24551	0.022	0.15099	DOWN
NR1I3	-1.1059	-1.8725	5.20664	0.0225	0.15324	DOWN
IFNA21	-4.5568	-2.1071	5.17558	0.02291	0.15507	DOWN
CRH	-1.4468	1.14366	5.0405	0.02476	0.1628	DOWN
HTR3C	-1.3177	-2.9042	4.99308	0.02545	0.16588	DOWN
BMP10	-1.0623	-3.0794	4.75552	0.0292	0.18081	DOWN
KIR2DL1	1.02773	-3.2495	4.4789	0.03432	0.20078	UP
IL17A	1.14292	-2.8353	4.37489	0.03647	0.2088	UP
AMELX	1.3256	-3.9947	4.35373	0.03693	0.21063	UP
NTS	-1.1096	4.24173	4.13124	0.0421	0.22775	DOWN
FSHR	-1.0155	-3.5468	3.86795	0.04922	0.2511	DOWN