Research Paper

Construction of immune-related gene pairs signature to predict the overall survival of osteosarcoma patients

Long-Qing Li[1,*], Liang-Hao Zhang[2,*], Yan Zhang[1], Xin-Chang Lu[1], Yi Zhang[1], Yong-Kui Liu[1], Manhas Abdul Khader[1], Jia-Wen[1], Tao-Liu[1], Jia-Zhen Li[1]

[1]Department of Orthopaedic Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, PR China
[2]Department of Urology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan, PR China
*Equal contribution

Correspondence to: Jia-Zhen Li; **email:** jzhli6411@zzu.edu.cn

ABSTRACT

The purpose of this study is to establish the prognosis of osteosarcoma patients based on the characteristics of immune-related gene pairs. We used the lasso Cox regression model to construct and verify the signature consisting of 14 immune-related gene pairs. This signature can accurately predict the overall survival of osteosarcoma patients and is an independent prognostic factor for osteosarcoma patients. For this we constructed a signature-based nomogram. The results of the nomogram show that our signature can bring clinical net benefits. We then assessed the abundance of infiltrating immune cells in each sample, and combine the results of the gene set enrichment analysis of a single sample to explore the differences in the immune microenvironment between IRPG signature groups. The result of gene set enrichment analysis shows the strong relationship between signature and immune system. Finally, we evaluated the relationship between signature and immunotherapy efficiency using algorithms such as TIMI and SubMap to explore patients who might benefit from immunotherapy. In conclusion, our signature can predict the overall survival rate of osteosarcoma patients and provide potential guidance for exploring patients who may benefit from immunotherapy.

INTRODUCTION

Osteosarcoma is a malignant bone tumor, which is mainly associated with children and adolescents [1]. Today, newly diagnosed osteosarcoma patients are mainly treated by chemotherapy and surgery, and the five-year survival rate can reach 60-70%. However, some patients who are not sensitive to chemotherapy or have metastases only have a five-year survival rate of 20-30%, and new personalized treatment plans need to be developed to improve the prognosis of these patients [2, 3]. Today, the most important guidelines for risk stratification and clinical decision making for patients with osteosarcoma are still clinical characteristics such as metastasis. However, patients with the same clinical characteristics and the same treatment showed very different clinical results [4]. Therefore, we have a need to consider the new prognostic factors to more accurately stratify patients to develop personalized treatment plans.

Today, the immune system is considered to play a vital role in the initiation and development of tumors [5]. Several immunotherapies including immune checkpoint inhibitors have been developed for this characteristic of tumors. Among them, immune checkpoint inhibitors targeting cytotoxic T lymphocyte associated antigen 4 (CTLA-4) and programmed cell death 1 (PD-1) have shown great potential in the treatment of various tumors [6–8]. Related research is also in full swing in patients

with osteosarcoma, promising to improve the survival of patients with osteosarcoma [9, 10]. However, new treatment methods require us to have a new understanding of tumors. Unfortunately, the molecular mechanism of tumor immunity in osteosarcoma remains undetermined. In addition, it is important to more accurately identify patients who are more likely to benefit from immunotherapy [11].

With the development of high-throughput gene detection technology and the establishment of large-scale gene expression data sets, researchers can more accurately identify key molecular features and combine them with clinical features to more precisely stratify patients to develop individualized treatment plan [12–14]. Previous research based on gene expression characteristics to develop multiple multi-gene signatures can identify high-risk patients. However, there are differences in the measured gene expression levels due to the different platforms for detecting gene expression. This brings certain difficulties to the comprehensive use of these data [15]. Recently, researchers have provided a new way to solve this difficulty, which is to normalize and scale based on the relative ranking of gene expression levels. This approach has produced reliable results in various studies [16–18]. Therefore, the purpose of this study is to research the value of immune-related gene pairs (IRGPs) in osteosarcoma, in predicting patient-survival and to explore their potential in predicting the effectiveness of immunotherapy.

RESULTS

Construction and evaluation of IRGPs signature

A total of 141 patients with osteosarcoma were included in our study. The TCGA cohort has 88 patients and the GSE21257 cohort has 53 patients. Three patients in the TCGA cohort lacked valid clinical information and were therefore not included in survival-related analysis. Table 1 shows the basic characteristics of these patients. A total of 210 immune-related genes are considered to have high variability and used to construct immune-related gene pairs (Supplementary Table 1). The IRGPs with low variation was deleted and a total of 2260 IRGPs were retained for further analysis (Supplementary Tables 2, 3). We selected the TCGA cohort as the training set and identified a total of 32 immune-related gene pairs related to prognosis (P <0.005) (Supplementary Table 4). We used Lasso Cox proportional hazard regression on the training set to define the IRGP signature, and selected the final model consisting of 14 IRGPs. The IRPG signature consists of 21 IRGs as shown in Table 2. We used the time-dependent ROC curve analysis to determine the optimal

cut-off value of the IRGP signature. As shown in Figure 1, the optimal cut-off value of the IRGP signature is -1.66. We divided patients into high-risk group and low-risk group according to cut-off value. As shown in Figure 1, compared to patients in the low-risk group the overall survival of patients in the high-risk group is significantly reduced. Subsequently, we conducted univariate and multivariate Cox regression analysis to explore whether IRGP signatures can be used as independent predictors of prognosis. As shown in Figure 2, after adjusting for variables such as metastasis status, the result showed that the IRGP signature is an independent prognostic factor for the TCGA cohort (HR:6.202, 95%CI: 3.545-10.848, P< 0.001). Figure 3 is the receiver working curve of IRGP signature, metastasis status, age and gender. As shown in Figure 3, the area under the curve (AUC) of the IRGP signature is 0.906, which indicates that our signature has excellent predictive ability.

Verification of IRGP signature

In the GSE21257 data set, patients were divided into high-risk group and low-risk group according to the same signature and the same cut-off value. As shown in Figure 1C, patients in the low-risk group have a longer overall survival. After adjusting for age, gender and other variables, the results of multivariate Cox regression showed that IRGP signature is an independent prognostic factor (HR:1.560, 95%CI: 1.033-2.354, P= 0.034). As shown in Figure 3, IRGP signatures still have excellent prediction capabilities in the verification set (AUC=0.918).

Correlation between IRGP and clinical characteristics

We further analyzed the relationship between IRGP signatures and clinical characteristics. We divided patients into different subgroups based on clinical variables. As shown in Figure 2E, IRGP signatures can predict the overall survival of patients in different subgroups. As shown in Figure 4A, in the TCGA cohort, patients in the metastasis group had higher IRGP values. In addition, based on the patient's metastatic status and IRGP signature, we divided the patients in the TCGA data set into four groups. As shown in Figure 4D, there was no significant difference in overall survival between metastatic patients and non-metastatic patients in the low-risk group. Among patients in the metastasis group, patients in the low-risk group had a longer overall survival than those in the high-risk group. The GSE21257 cohort simultaneously recorded the metastatic status of the patient at the time of the initial diagnosis and the status within 5 years after the diagnosis. Therefore, in the GSE21257 cohort, we investigated the relationship between IRGP signatures

Table 1. Summary of clinical characteristics of Osteosarcoma patient data sets in the study.

Characteristic	Training cohort (TCGA n=88)	Validation cohort (GSE21257 n=53)
Vital status, n (%)		
Alive	57(64.8)	30(62.3)
Dead	29(33.0)	23(37.7)
Unknow	2(0.2)	NA
Age, n (%)		
> = 18	19(21.6)	20(37.7)
<18	69(78.4)	33(62.3)
Gender, n (%)		
Male	51(58.0)	34(64.2)
Female	37(42.0)	19(35.8)
Metastasis, n (%)		
M0	66(75)	19(35.8)
M1	22(25)	34(64.2)
Histological, n (%)		
Osteoblastic	NA	32(60.4)
Other	NA	21(39.6)

NA represents information not available.

Table 2. Information on 14 IRGPs.

Gene pair	Gene1	Gene2	Coefficient
HLA-DQB1|STC2	HLA-DQB1	STC2	-0.2334234
APOD|CCL2	APOD	CCL2	0.38929255
F2R|SEMA3B	F2R	SEMA3B	-0.0684453
F2R|ANGPTL4	F2R	ANGPTL4	-0.1359965
HCK|PLXNB1	HCK	PLXNB1	-0.5042191
HCK|STC1	HCK	STC1	-0.8295186
RAC3|FGFRL1	RAC3	FGFRL1	0.84832068
SEMA3A|GAL	SEMA3A	GAL	-0.6758144
SEMA3B|LTBP4	SEMA3B	LTBP4	0.06249894
SEMA3B|FGFRL1	SEMA3B	FGFRL1	0.44268731
SEMA5A|C5AR1	SEMA5A	C5AR1	0.26480471
EDNRA|STC2	EDNRA	STC2	-0.7028178
PLXNB1|FGFRL1	PLXNB1	FGFRL1	0.14262903
ANGPTL2|SORT1	ANGPTL2	SORT1	-0.1548456

and these two tumor metastasis states. As shown in Figure 4B, 4C, both in the initial diagnosis and within 5 years of diagnosis, patients in the metastasis group had higher IRGP signature values. Finally, we excluded patients who had metastases at the time of diagnosis. Based on the time of sarcoma metastasis, the occurrence of sarcoma metastasis is defined as the outcome. As shown in Figure 4E–4F, patients with high IRGP signatures have a higher risk of metastasis. The results of multivariate analysis indicate that IRGP signature is an independent risk factor.

Construction and evaluation of nomogram based on IRGP signature

We constructed a nomogram based on the clinical variables and IRGP signatures of the TCGA dataset. Subsequently, the nomogram was further verified in the GSE21257 dataset. The results of the calibration chart show that the nomogram performance is the best in predicting the 3-year OS in the two cohorts. In the TCGA cohort, the nomogram C index was 0.903, while in the GSE21257 cohort, the nomogram C index was

0.702. The result of the calibration curve shows that the nomogram has good discrimination ability. The results of the decision curve analysis show that the combination model can bring net benefits in predicting overall survival (Figure 5).

Relationship between IRGP signature and immune cell infiltration

As mentioned above, the abundance of 10 immune-related cells was calculated using the MCP-counter method. As shown in Figure 6A, a significant difference was observed between the two groups of patients. Compared with patients in the high-risk group, the abundance of the 7 cell populations in the patients in the low-risk group was higher (B-cell lineage, CD8+T cells, Cytotoxic lymphocytes, Fibroblasts, Monocytic lineage cells, NK cells, T cells). In addition, as shown in Figure 6B, the degree of immune cell infiltration is negatively correlated with the IRGP value. Similarly, the results of ssGSEA showed that in the high-risk group, most of the 29 immune-related gene sets or immune cells scored lower. We further explored the relationship between these immune-related gene sets or immune cells and the IRGP signature value. As shown in Figure 7, most gene sets or immune cell scores are negatively correlated with IRGP signature values. Subsequently, we used CIBERSORT to infer the relative proportion of 22 infiltrating immune cells in each sample. As shown in Figure 7E, 7F, in the TCGA cohort, the relative proportion of T cells CD8 and T cells CD4 memory activated was higher in the low-risk group, while the relative proportion of T cells CD4 navie was higher in the high-risk group. In the GSE21257 cohort, only the relative abundance of Mast cells activated and Dendritic cells activated differed. Although the radar chart showed that the relative abundance of T cells CD8 was

higher in the low-risk group, it did not reach a statistical difference. It is worth noting that the infiltration ratio of M0 macrophages and M2 macrophages is higher in all patients.

Relationship between IRGP signature and immunotherapy efficiency

Considering that our signature is based on IRGP, we further explored the relationship between IRGP signature and immunotherapy response. First, we tested the predictive ability of IRGP signatures in the GSE78220 data set. We used IRGP signatures to divide patients into high-risk and low-risk groups. Kaplan-Meier curves show that patients in the high-risk group have a poorer prognosis than those in the lower-risk group. In addition, IRGP signature values of patients with poor immunotherapy response showed an increasing trend. Unfortunately, it did not reach statistical significance (Figure 8A, 8B).

Subsequently, we explored the expression of 7 immune checkpoint genes in two groups of patients. As shown in Figure 8C, except for the PDCD1 gene, all other immune checkpoint genes are highly expressed in the low-risk group. We further explored the relationship between IRGP signature and TMIT. Since there is no optimal cutoff value for TMIT classification, we regard the IRGP signature as a continuous variable. Our results show that the IRGP signature value is negatively correlated with TIL and the expression of CD8A gene and CD247 gene. Therefore, patients in the low-risk group are more likely to be classified as TMIT I.

Finally, we used SubMap analysis to further study the relationship between IRGP signature and immunotherapy efficiency. Using subclass mapping, the expression
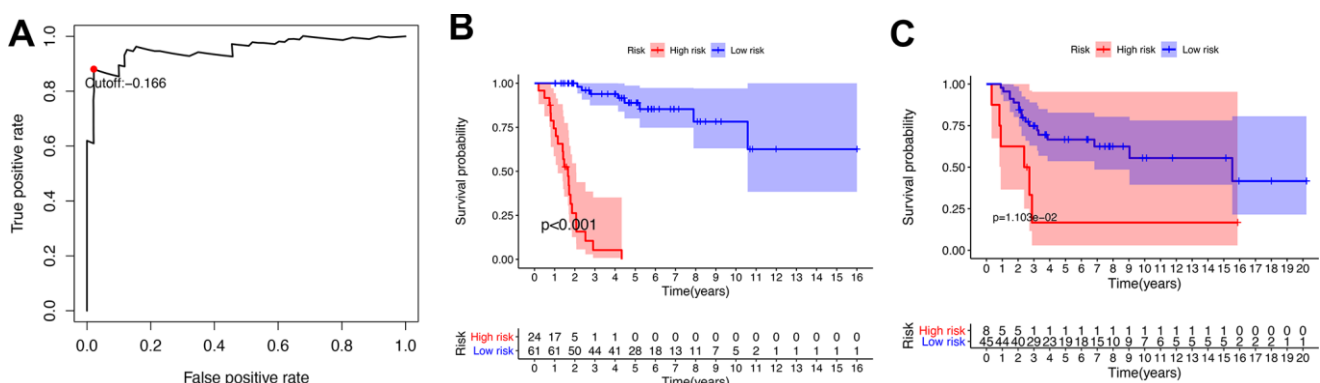

Figure 1. Establishment and verification of IRGP signature. (**A**) Time-dependent ROC curve for IRGP signature in the TCGA cohort. The optimal cut-off value of IRGP signature is -0.166, and patients are divided into high-risk group and low-risk group according to the cut-off value (**B**) Kaplan–Meier curves of overall survival according to IRGP signature groups in the TCGA cohort. (**C**) Kaplan–Meier curves of overall survival according to IRGP signature groups in the GSE21257 cohort.
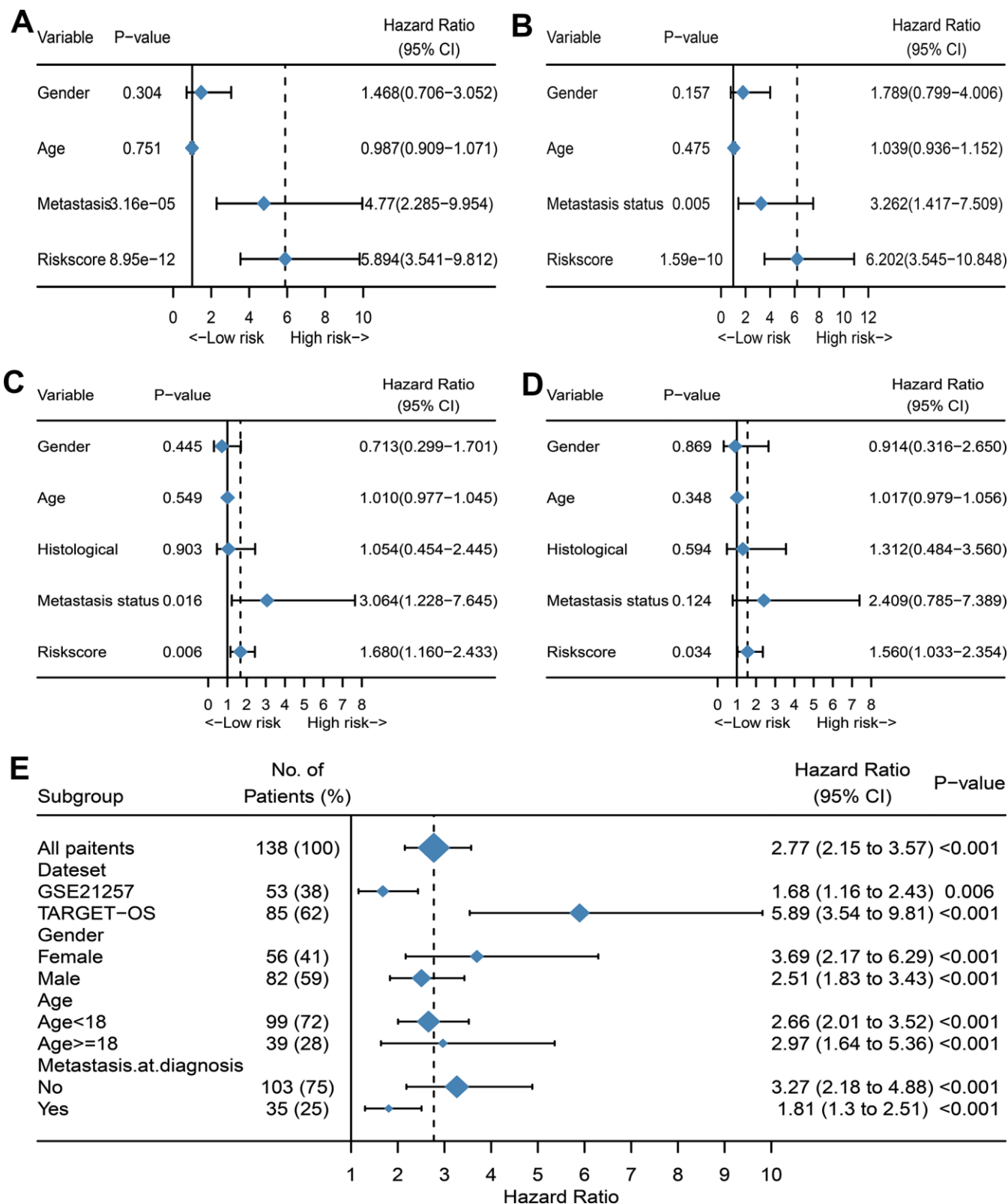
Figure 2. Evaluate whether IRGP signature is an independent prognostic factor. (**A**) Forest plot of univariate Cox regression results of IRPG signature and clinical characteristics of TCGA cohort. (**B**) Forest plot of univariate Cox regression results of IRPG signature and clinical characteristics of GSE21257 cohort. (**C**) Forest plot of multivariate Cox regression results of TCGA cohort. (**D**) Forest plot of multivariate Cox regression results of IRPG signature and clinical characteristics of GSE21257 cohort. (**E**) Forest plots of the associations between IRGP and overall survival in various subgroups.

profiles of the two groups of patients (high-risk group and low-risk group) were compared with a published immunotherapy data set. This data set records the expression data of 47 melanoma patients treated with programmed cell death protein 1 (PD-1) immune checkpoint inhibitors or cytotoxic T lymphocyte associated protein 4 (CTLA-4) immune checkpoint inhibitors. The results showed that the expression profiles of patients in the low-risk group were correlated with those in the PD-L1 response group. This indicates that patients in the low-risk group are more likely to benefit from PD-L1 therapy (Figure 8G).

Gene set enrichment analysis

Gene set enrichment analysis was performed in patients in the high-risk group and the low-risk group. As shown in Figure 9, a large number of immune-related gene sets are enriched in the low-risk group, while the gene sets enriched in the high-risk group are those that are not related to the immune process.

DISCUSSION

The immune system has been shown to play a key role in the occurrence and development of tumors [19, 20]. Immunotherapy has also shown great potential in a variety of tumors [11, 21]. Therefore, it is of great significance to study the clinical value and potential molecular mechanism of immune-related genes in osteosarcoma. Although many signatures have recently been developed that can predict patient prognosis, these signatures still have some deficiencies in overcoming the batch effects of different platforms. In addition, most of these signatures require pretreatment of gene expression profiles, which affects the widespread use of signatures. In this study, we developed and verified an immune-related gene signature that can predict the overall survival of osteosarcoma patients. The signature of the immune gene pair is derived from the pairwise comparison of gene expression in the same sample, which can be used more widely across different detection platforms [16]. Our signature is composed of 14 immune-related gene pairs, and patients are divided into high-risk group and low-risk group according to the calculated cut-off value. Our results show that the overall survival of patients in the high-risk group is significantly reduced. In addition, the results of univariate and multivariate Cox proportional hazards regression analysis indicate that signature is an independent prognostic factor in predicting the overall survival of osteosarcoma patients. In addition, we found that combining the signature and clinical characteristics to construct a nomogram can more accurately predict
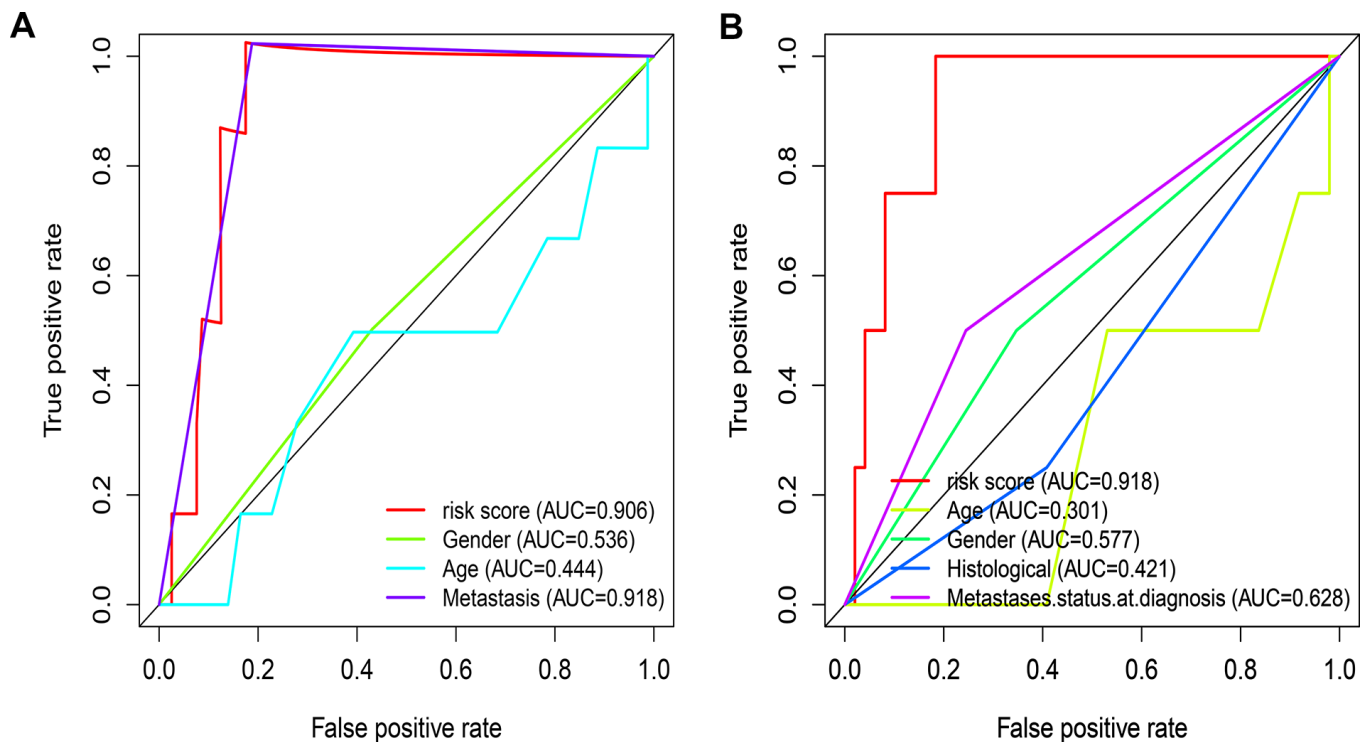

Figure 3. Evaluate the predictive ability of IRGP signature. (**A**) ROC curve of clinical characteristics and IRGP signature in TCGA cohort. (**B**) ROC curve of clinical characteristics and IRGP signature in GSE21257 cohort.

the patient's overall survival and bring net benefits. Finally, we divided the patients in the verification group into a high-risk group and a low-risk group based on the same cutoff value and obtained the same results.

Today, clinical characteristics such as metastatic status are still the most important basis for doctors to stratify risk and make clinical decisions for patients with osteosarcoma [22]. Therefore, we further explored the relationship between signatures and clinical characteristics. Our results show that whether in the TCGA cohort or the GSE21257 cohort, the IRGP signature value of patients in the metastatic group is higher than that in the non-metastatic group. In addition, we divided the patients into 4 groups according to the metastasis status and IRGP signature, and then stratified the patients more accurately. Our results indicate that among the patients in the IRGP signature low-risk group, there was no significant difference in overall survival between patients in the metastatic group and
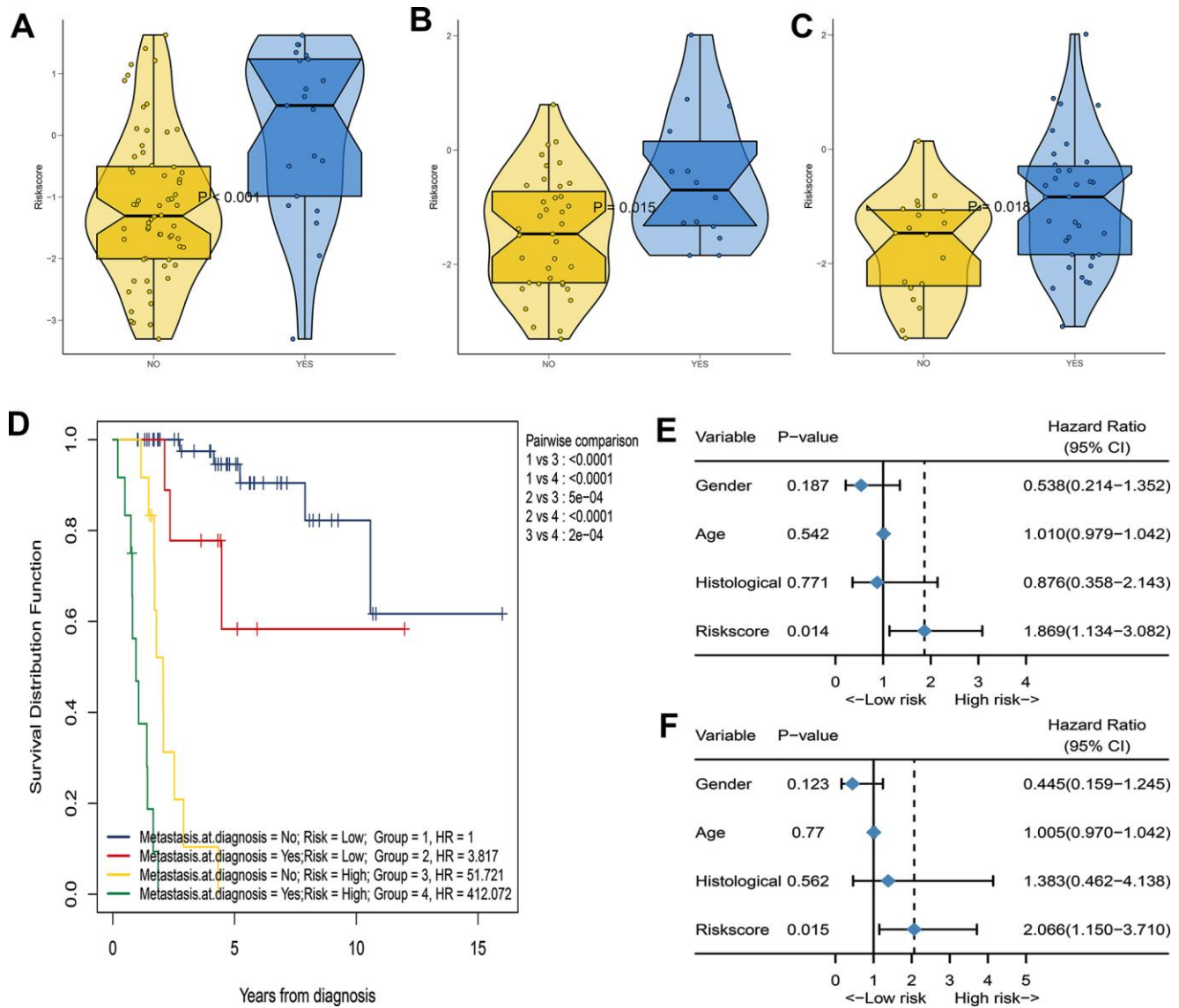

Figure 4. Assess the correlation between IRGP signature and metastasis status. (**A**) Box violin plot of the relationship between the metastasis status at diagnosis and the IRGP signature value in the TCGA cohort. (**B**) Box violin plot of the relationship between the metastasis status at diagnosis and the IRGP signature value in the GSE21257 cohort. (**C**) Box violin plot of the relationship between the metastasis status within 5 years and the IRGP signature value in the GSE21257 cohort. (**C**) Kaplan–Meier curves of overall survival for patients in TCGA cohort stratified by both IRGP signature, and metastasis status. (**D**) Forest plot of univariate Cox regression results between various variables and tumor metastasis in the GSE21257 cohort. (**E**) Forest plot of univariate Cox regression results between various variables and tumor metastasis in the GSE21257 cohort. (**F**) Forest plot of multivariate Cox regression results between various variables and tumor metastasis in the GSE21257 cohort.

those in the non-metastatic group. Among patients in the metastasis group, patients in the low-risk group had better overall survival than those in the high-risk group. This also explains why patients with similar clinical characteristics show completely different clinical outcomes despite undergoing the same treatment. Therefore, combining the IRGP signature with traditional clinical features can more accurately stratify patients and formulate individualized treatment plans. In addition, identifying patients with a high risk of metastasis and enhancing follow-up of these patients

may improve the prognosis of osteosarcoma patients. We further explored the relationship between IRGP signature and tumor metastasis in the GS21257 cohort. Our results indicate that IRGP signature is an independent risk factor for tumor metastasis. Patients with high IRGP signature values have an increased risk of tumor metastasis within five years of diagnosis. Therefore, it is necessary to strengthen the follow-up of patients with high IRGP signature value. It is worth noting that due to the limitation of the sample size, further research is needed to prove our conclusion.
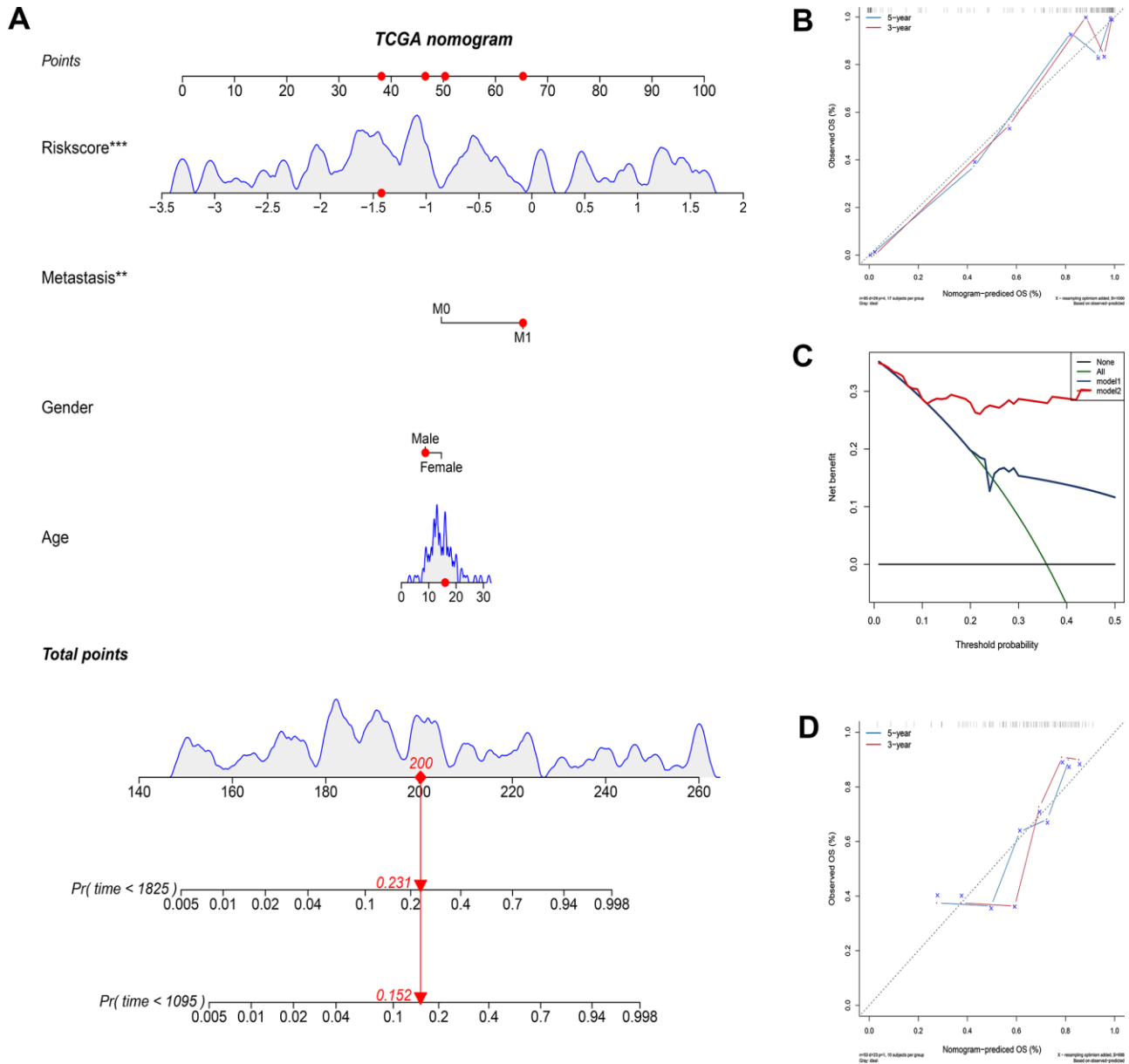

Figure 5. Construct and evaluate nomograms in TCGA and GSE21257 cohort. (**A**) Nomogram to predict the probability of TCGA patients mortality based on IRGP and clinical variables. (**B**) The calibration plot for internal validation of the nomogram. (**C**) Decision curve analyses of the nomograms based on IRGP signature for 3-year overall survival. (**D**) The calibration plot of the nomogram in the GSE21257 data set is used for external verification. *P<0.05, **P<0.01, ***P<0.001.

Recently, immune checkpoint inhibitor therapy has shown promising clinical benefits [23, 24]. Historically, sarcomas were the first tumor model for which immunotherapy was suggested as a relevant therapeutic strategy [25]. Unfortunately, recent studies have shown that PD-1 inhibitors have limited activity in osteosarcoma. It is important to identify patients who may benefit from this treatment strategy. [9, 11, 26] Therefore, we used three methods to evaluate the relationship between IRGP signature and immunotherapy efficiency. First, we found that IRGP signatures can predict the outcome of patients with metastatic melanoma who received anti-PD-L1 treatment and that patients with poor treatment response showed an increasing trend in IRGP signature values. However, due to sample size limitations, this result must be interpreted with caution. Subsequently, we further explored the relationship between IRGP signature and TMIT. Similarly, patients with low IRGP signature values are more likely to be classified as TIMT type I due to high PD-L1 gene expression and high CD8A gene expression, and therefore are more likely to benefit from anti-PD-1 therapy. Interestingly, a recent study showed that the expression of PD-L1 gene is related to the poor prognosis of osteosarcoma patients, which is contrary to our conclusion [27]. Therefore, we used univariate Cox regression to explore the relationship between PD-L1 gene expression and the overall survival of osteosarcoma

patients. The results show that PD-L1 gene expression is a protective factor for the overall survival of osteosarcoma patients (Supplementary Figure 1). Therefore, further research is needed to verify the relationship between PD-L1 gene expression and the prognosis of osteosarcoma patients. Finally, we compared the expression profiles of TCGA patients with melanoma patients treated with immune checkpoint inhibitors by submap analysis. There is a certain correlation between the expression profiles of the low-risk group and the PD-L1 response group. In summary, we speculate that patients in the low-risk group may benefit from anti-PD-L1 therapy. However, it is necessary to further study the relationship between IRGP signature and immunotherapy efficiency in osteosarcoma patients.

Our results show differences in biological processes and immune infiltration between the two groups. Among them, the results of GSEA show that a large number of immune-related pathways are enriched in the low-risk group, while immune-unrelated pathways are enriched in the high-risk group. From this point of view, patients in the low-risk group are also more likely to benefit from immunotherapy. The results of MCP-counter and CIBERSORT showed that there was also a difference in immune cell infiltration between the two groups. Among them, CD8T cells were more infiltrated in the low-risk group. Consistent with our conclusion,
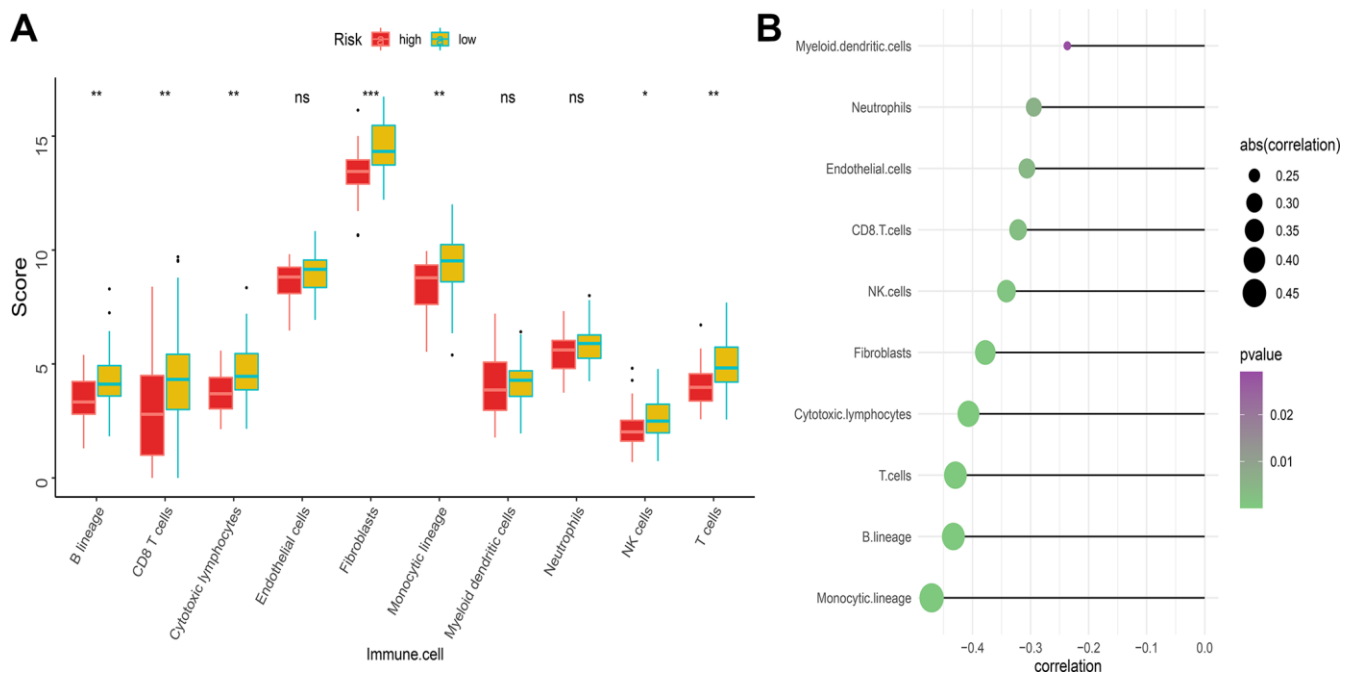

Figure 6. Difference of immune infiltration among patients in IRGP signature group. (**A**) Box plot showing the absolute abundance scores of 10 immune cell and stromal cell populations in two groups of patients. (**B**) Correlation matrix between absolute abundance scores of immune cells and stromal cells and IRGP values. The size of the bubble represents the degree of correlation and the color of the bubble represents the p-value of the correlation. (ns represents no significance, *P < 0.05, **P < 0.01, ***P < 0.001).
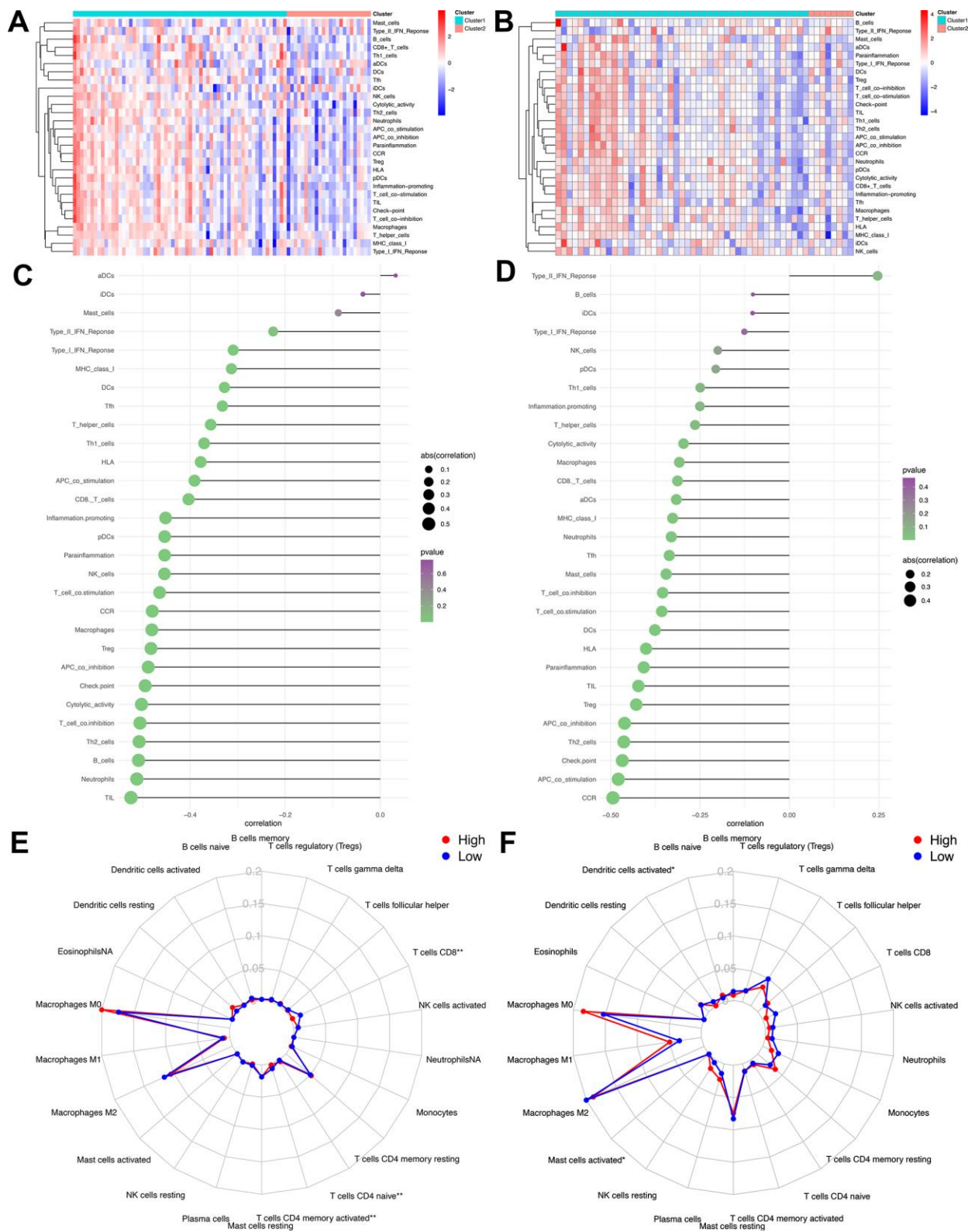
Figure 7. Assess the relationship between immune microenvironment and IRGP signature. (**A**) Heat map of results of single sample gene set enrichment analysis in TCGA cohort. (**B**) Heat map of results of single sample gene set enrichment analysis in GSE21257 cohort. (**C**) Relationship between 29 immune-related gene sets and IRGP signature values in TCGA cohort. (**D**) Relationship between 29 immune-related gene sets and IRGP signature values in GSE21257 cohort. (**E**) Radar chart of the relationship between 22 immune cell infiltration and IRGP signature grouping in TCGA cohort. (**F**) Radar chart of the relationship between 22 immune cell infiltration and IRGP signature grouping in GSE21257 cohort.

CD8 + T lymphocyte infiltration has a positive effect on the prognosis of various tumors, such as hepatocellular carcinoma and breast cancer [28, 29]. In addition, CD8 + T lymphocyte infiltration has been found to improve the prognosis of patients with osteosarcoma [30]. At the same time, a high proportion of M0 and M2 macrophage infiltration existed in both groups. Interestingly, despite many recent studies on tumor-associated macrophages,
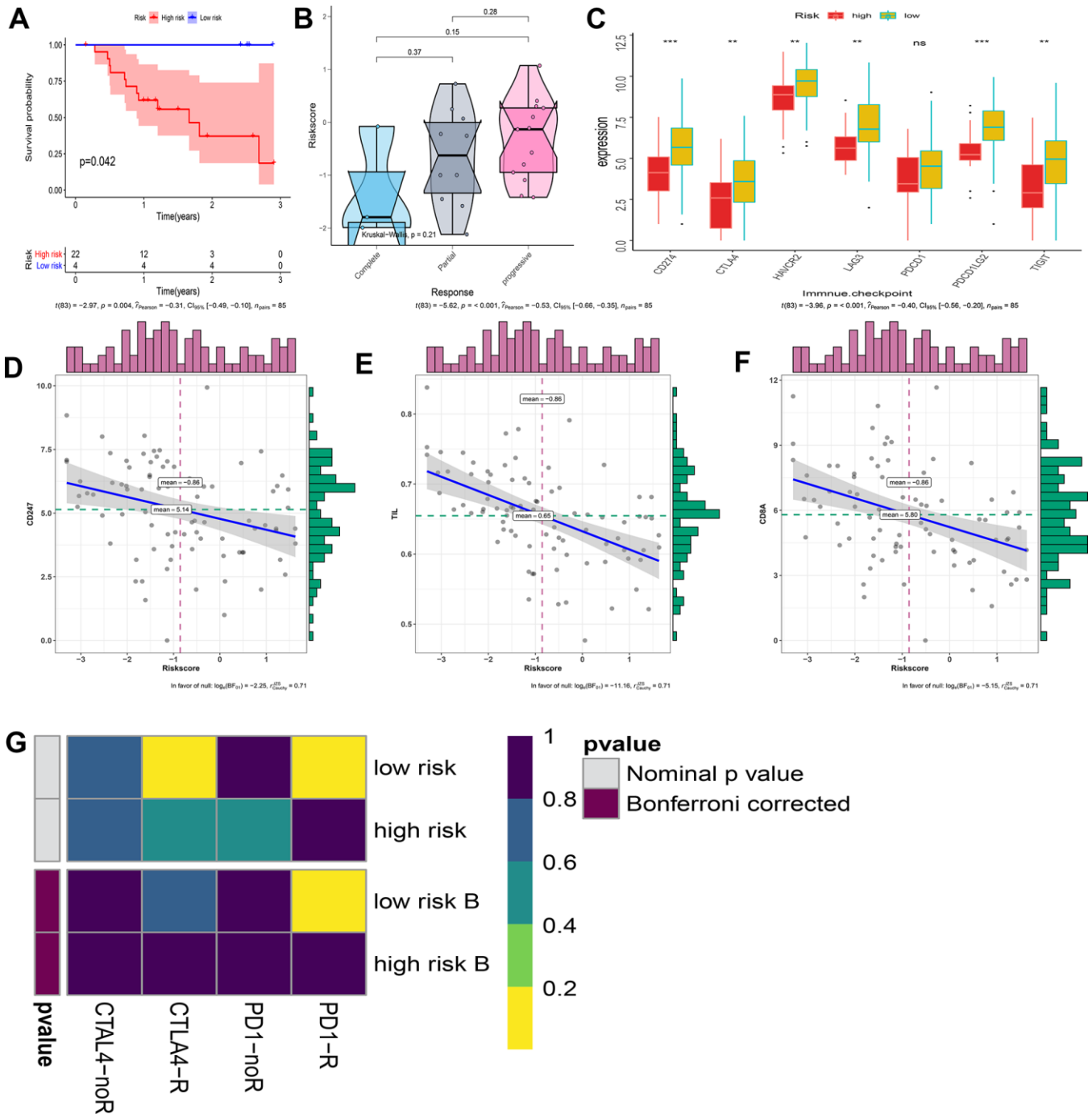

Figure 8. Explore patients who may hope to benefit from immunotherapy. (**A**) Kaplan–Meier curves of overall survival according to IRGP groups in GSE78220. (**B**) Box violin plot of the relationship between the immunotherapy response and the IRGP signature value in the GSE78220 cohort. (**C**) Box plot showing the expression of 7 immune checkpoint genes in two groups of patients. (**D**) Correlation between IRGP value and CD247 gene expression. (**E**) Correlation between IRGP value and TIL. (**F**) Correlation between IRGP value and CD8A gene expression. (**G**) Heatmap of correlation between expression profiles of patients in the IRGP group and patients receiving immunotherapy. The color of the grid represents the correlation P-value.

the conclusions do not seem to be consistent. Many studies have shown that macrophages are usually associated with immunosuppressive microenvironments. In addition, a high number of M2 macrophages have been shown to be associated with poor prognosis in various tumors [31–33]. A recent study also showed that macrophages reduce the sensitivity of osteosarcoma to neoadjuvant chemotherapy drugs by secreting interleukin1β [34]. Interestingly, the results of Buddingh et al. indicate that in the case of osteosarcoma, the direct or indirect antitumor activity of macrophages exceeds its possible tumor supporting effect [35]. In addition, the
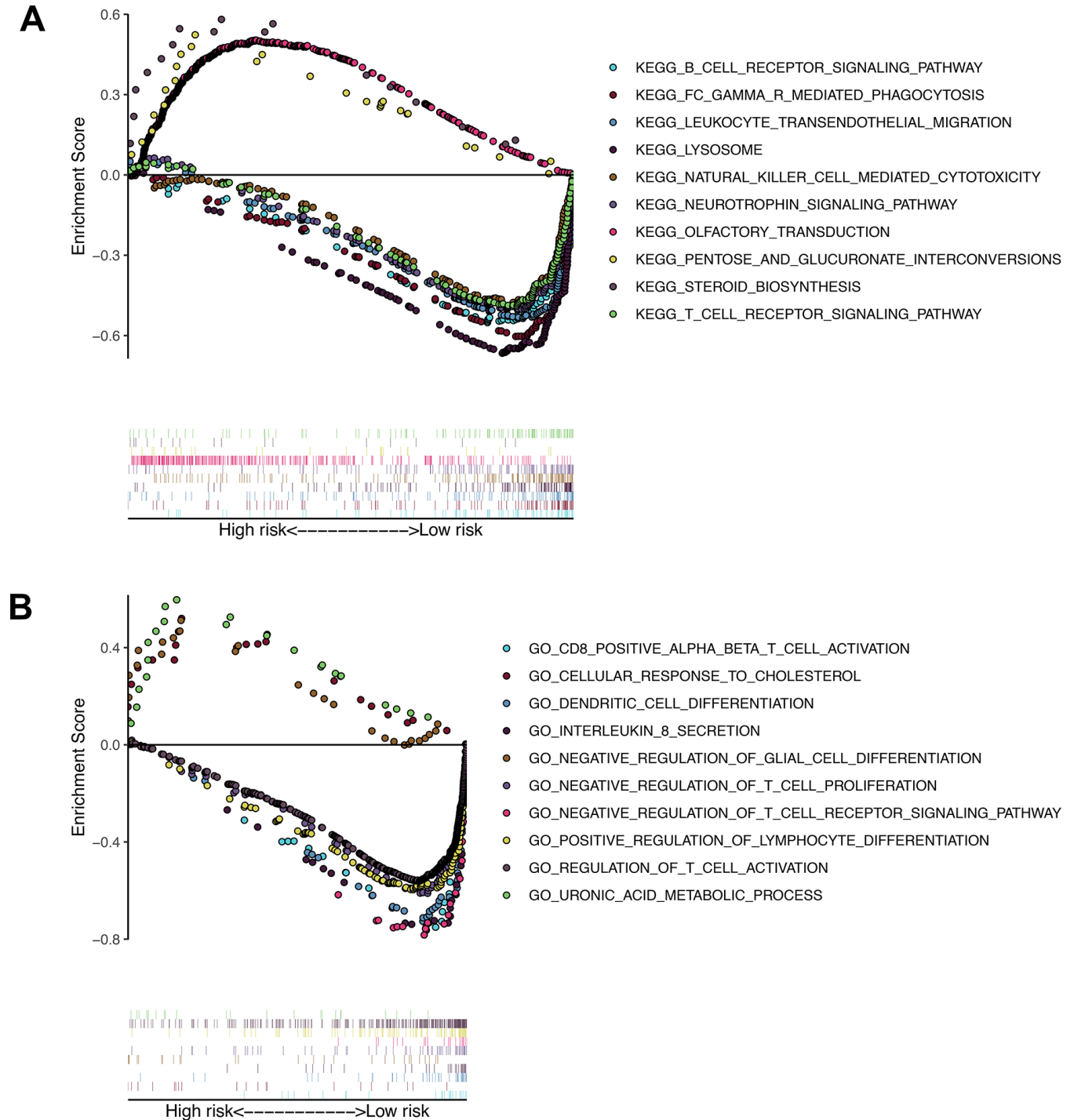

Figure 9. Results of gene set enrichment analysis in the TCGA cohort. (**A**) The significantly enriched KEGG pathways in TCGA cohort by GSEA. (**B**) The significantly enriched GO terms in TCGA cohort by GSEA.

results of Anne Gomez-Brouchet et al. showed that, contrary to the results in other solid tumors, the presence of CD163-positive M2-polarized macrophages is essential for inhibiting osteosarcoma progression [36]. Therefore, it is necessary to further study the role of macrophages in osteosarcoma.

Our signature consists of 21 immune-related genes, and previous studies have described the value of some genes in osteosarcoma. Both ANGPTL2 and ANGPTL4 belong to the Angiopoietin-like protein family. Previous studies have shown that ANGPTL2, as a chronic inflammatory mediator, is overexpressed in various tumors and is associated with poor prognosis [37, 38]. In addition, studies have shown that ANGPTL2 is highly expressed in osteosarcoma cells induced by hypoxia / HIF-1α and promotes cell proliferation, invasion, migration and G1 phase arrest. ANGPTL2 also enhances the expression of VEGFA, Ang II and HK2 in mice to enhance angiogenesis and glycolysis [39]. Similarly, ANGPTL4 is highly expressed in hypoxic-induced osteosarcoma cells and promotes osteosarcoma cell proliferation and migration as well as osteoclast formation and bone resorption activity [40]. Previous studies have shown that the CCL2 gene affects the proliferation of osteosarcoma cells through the RANKL signaling pathway. In addition, the expression of CCL2 gene in high-grade osteosarcoma cells increased and promoted the proliferation and invasion of osteosarcoma cells [41]. SEMA3A can inhibit the ability of osteosarcoma cells to stimulate osteoclast production [42]. The signatures based on these gene pairs respond well to the patient's immune status.

It should be admitted that our research still has some limitations. First, although the incidence of osteosarcoma is relatively low, the study involved a small sample size. Second, the conclusion about the efficacy of immunotherapy cannot be verified in osteosarcoma patients. Further research is needed to verify our results. Finally, this study is a retrospective study, and further prospective studies are needed to verify our results.

In conclusion, IRGP signatures can accurately predict the overall survival of osteosarcoma patients and combining signatures with clinical characteristics can bring net benefits. In addition, signatures may identify patients who are more likely to benefit from immunotherapy. Further research is needed to verify our conclusion.

MATERIALS AND METHODS

Clinical samples and data acquisition

We downloaded the level three RNA-Seq expression data of 88 osteosarcoma patients from TCGA database (https://portal.gdc.cancer.gov) (FPKM). Subsequently, the latest clinical data of osteosarcoma patients were downloaded from the TARGET database (https://ocg.cancer.gov/programs/target). Finally, 85 patients with both RNA-seq expression data and valid clinical information were identified for analysis in this study. We retrieved the osteosarcoma dataset with clinical information such as overall survival on the Gene Expression Comprehensive Library (GEO; http://www.ncbi.nlm.nih.gov/geo/). We then downloaded the GSE21257 dataset on the GPL10295 platform (Illumina human-6 v2.0 expression beadchip) [35]. This data set is the only one in the osteosarcoma data set that has the overall survival of the patient and was used to verify our results. Finally, for the analysis of immunotherapy efficiency, the GSE78220 data set (GPL11154 platform (Illumina HiSeq 2000 (Homo sapiens)) was downloaded. GSE78220 dataset records transcriptome data of 28 patients with metastatic melanoma treated with anti-PD-1 agents (pem-brolizumab or nivolumab). Supplementary Table 5 provides detailed clinical information for the GSE78220 data set. If a target gene corresponds to multiple probes, the average expression value of the probes is used to represent the expression level of the gene. At the same time, delete genes whose expression is 0 in all samples.

Identification of prognostic-related IRGPs in patients with osteosarcoma

We downloaded the Immunity Related Gene List (IRG) from the Immunology Database and the Analysis Portal (ImmPort) database [43]. The list contained a total of 1811 unique IRGs, which are related to T cell receptor and B cell antigen receptor signaling pathways, cytotoxicity of natural killer cells, antigen processing and presentation pathways. First, immune-related genes with high variability were identified. Specifically, a particular gene is considered to have high variability if it has a high median absolute deviation (MAD>0.5) value in each dataset. We then used the gene expression levels of these genes in each sample for pairwise comparison to construct IRGP. In a specific sample, if the expression value of the first IRG is greater than that of the second IRG, the score of this IRGPs in the sample is 1, otherwise it is 0. The score of each IRGP in all samples was calculated and the IRGPs with low variation were removed (IRGP with a score of 1 or 0 in more than 80% of the sample in any data set). Finally, IRGPs with higher variability were identified for further analysis. Use TCGA cohort as training set. Univariate Cox regression analysis was performed on these IRGPs in the TCGA cohort, and IRGPs with p <0.005 were considered as prognostic-related IRGPs and used for subsequent analysis.

Construction and evaluation of signatures based on IRGPs

Lasso Cox proportional hazards regression analysis was performed on the above-mentioned prognostic-related IRGPs, and finally an optimal model composed of 14 gene pairs was determined. Subsequently, the optimal-model based IRGP signature of each patient was calculated. In the 3-year overall survival TCGA cohort, time-dependent receptor operating characteristic (ROC) curve analysis was used to determine the optimal cut-off value for IRGP signature [44]. According to the cut-off value of IRGP signature, patients were divided into high-risk group and low-risk group. The log-rank test was used to evaluate the overall survival difference between the low-risk group and the high-risk group and the KM survival curve was drawn. Receiver operating characteristic (ROC) curve analysis was used to evaluate the sensitivity and specificity of IRGPs. A ROC curve including clinical characteristics was drawn, and the area under the curve (AUC) was calculated. Finally, univariate and multivariate Cox regression analysis was used to investigate whether the prognostic value of IRGP was affected by other clinical characteristics.

Verification of signatures based on IRGPs

In order to verify the IRGPs signatures, we calculated the risk score of each patient in the GSE21257 data set using the above method and divided the patients into a high-risk group and a low-risk group according to the cutoff value above. The log-rank test was then used to assess the overall survival difference between the two groups and to plot the KM survival curve. We used univariate and multivariate Cox regression to evaluate the independent prognostic value of signatures. Finally, the receiver operating characteristic (ROC) curve was drawn and the area under the curve (AUC) was calculated.

Evaluation of the relationship between signatures and clinical characteristic

We evaluated the relationship between IRGPs signature and clinical characteristics. First, we evaluated the relationship between the IRGP signature value and clinical variables. Subsequently, we divided patients with osteosarcoma into different subgroups based on clinical variables and explored the prognostic value of IRGP signatures among different subgroups. In addition, we divided patients in the TARGT-OS cohort into four groups based on their metastatic status and IRGP group. The differences in overall survival between the four groups of patients were evaluated. Finally, the relationship between IRGP signature and tumor metastasis was further evaluated in the GSE21257 cohort.

Construction and evaluation of nomograms

We combined the clinical characteristics of the TCGA data set with the IRGP signature to construct a nomogram and verified the nomogram using the GSE21257 data set as external verification We used the C index to evaluate the discriminative power of the nomogram and drew a calibration chart to evaluate the accuracy of the nomogram. We then compared the decision curve analysis between the clinical characteristics model and the combined model including gene signature and clinical characteristics.

Estimation of immune infiltration

First, the immune infiltration assessment was performed using the "microenvironment cell population count (MCP-counter)" method [45]. Using the normalized FPKM expression matrix converted by log2 as input, the absolute abundance scores of ten immune cell and stromal cell populations are generated through the "MCP-counter" package. Research shows that immune cell infiltration assessed by MCP-counter algorithm performs well when comparing between samples [46]. Subsequently, CIBERSORT was used to infer the relative proportion of 22 infiltrating immune cells in each sample for supplementation. In addition, the single sample GSEA was used to evaluate the enrichment of 29 immune-related gene sets in each sample.

Relationship between IRGP and immunotherapy efficiency

First, verify the IRGP signature in the GSE78220 cohort. Considering the significant heterogeneity between different tumors, we recalculated the cut-off value using time-dependent ROC curve analysis. According to the cut-off value, divide the patients into high-risk group and low-risk group and draw KM survival curve. Subsequently, 'Tumor microenvironment immune type (TMIT)' was used to speculate the efficacy of anti-PD-1 / PD-L1 treatment [47, 48]. TMIT divided patients into four types based on PD-L1 and CD8A mRNA expression, which has been shown to predict patients' response to immune checkpoint inhibitors in pan-cancer analysis. At the same time, we also explored the differences in gene expression of the other six immune checkpoints between the two groups of patients. In addition, we used SubMap analysis (Gene Pattern) to compare gene expression profiles of osteosarcoma patients with melanoma patients treated with immunotherapy to indirectly predict the efficacy of immunotherapy in osteosarcoma patients [49, 50].

Gene set enrichment analyses

GSEA software (version 4.0.1) was used to perform gene set enrichment analysis between high-risk and low-risk groups. Recognized the enriched terms in gene ontology (GO) and KEGG in high-risk group and low-risk group respectively. P <0.05 and False discovery rate (FDR) <0.05 are considered statistically significant.

Statistical analysis

Except for gene set enrichment analysis, all statistical analyses involved in this research were conducted through R software (version 3.6.3, R Foundation for Statistical Computing, Vienna, Austria). If no special instructions, p <.05 is considered statistically significant.

Data availability statement

RNA-seq data of the TCGA cohort can be obtained from the TCGA database (https://portal.gdc.cancer.gov). Clinical data of these patients can be obtained from the TARGET database (https://ocg.cancer.gov/programs/target). The data of GSE21257 can be obtained from the GEO database (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE21257).

Abbreviations

OS: osteosarcoma; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; TARGET: Therapeutically Applicable Research To Generate Effective Treatments; MAD: Median absolute deviation; MCP-counter: microenvironment cell population count; CIBERSORT: Cell type Identification By Estimating Relative Subsets Of known RNA Transcripts; ROC: receiver operating characteristic; AUC: area under the curve; GSEA: gene set enrichment analysis; TIL: Tumor infiltrating lymphocytes; PD-L1: Programmed Cell Death-Ligand 1.

AUTHOR CONTRIBUTIONS

Long-Qing Li conceived, designed, analyzed the data. Long-Qing Li and Liang-Hao Zhang wrote the manuscript. Manhas Abdul Khader, Xin-Chang Lu and Yi-Zhang revised the manuscript. Yong-Kui Liu, Jia Wen and Tao Liu produced the figures and tables. Jia-Zhen Li and Yan-Zhang supervised the study. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

FUNDING

REFERENCES

1. Bielack SS, Kempf-Bielack B, Delling G, Exner GU, Flege S, Helmke K, Kotz R, Salzer-Kuntschik M, Werner M, Winkelmann W, Zoubek A, Jürgens H, Winkler K. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. J Clin Oncol. 2002; 20:776–90. https://doi.org/10.1200/JCO.2002.20.3.776 PMID:11821461

2. Khanna C, Fan TM, Gorlick R, Helman LJ, Kleinerman ES, Adamson PC, Houghton PJ, Tap WD, Welch DR, Steeg PS, Merlino G, Sorensen PH, Meltzer P, et al. Toward a drug development path that targets metastatic progression in osteosarcoma. Clin Cancer Res. 2014; 20:4200–09. https://doi.org/10.1158/1078-0432.CCR-13-2574 PMID:24803583

3. Yang Y, Han L, He Z, Li X, Yang S, Yang J, Zhang Y, Li D, Yang Y, Yang Z. Advances in limb salvage treatment of osteosarcoma. J Bone Oncol. 2017; 10:36–40. https://doi.org/10.1016/j.jbo.2017.11.005 PMID:29296558

4. Isakoff MS, Bielack SS, Meltzer P, Gorlick R. Osteosarcoma: current treatment and a collaborative pathway to success. J Clin Oncol. 2015; 33:3029–35. https://doi.org/10.1200/JCO.2014.59.4895 PMID:26304877

5. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, Diehn M, West RB, Plevritis SK, Alizadeh AA. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med. 2015; 21:938–45. https://doi.org/10.1038/nm.3909 PMID:26193342

6. El-Khoueiry AB, Sangro B, Yau T, Crocenzi TS, Kudo M, Hsu C, Kim TY, Choo SP, Trojan J, Welling TH 3rd, Meyer T, Kang YK, Yeo W, et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): an open-label, non-comparative, phase 1/2 dose escalation and expansion trial. Lancet. 2017; 389:2492–502.

https://doi.org/10.1016/S0140-6736(17)31046-2
PMID:28434648

7. Siu LL, Even C, Mesía R, Remenar E, Daste A, Delord JP, Krauss J, Saba NF, Nabell L, Ready NE, Braña I, Kotecki N, Zandberg DP, et al. Safety and efficacy of durvalumab with or without tremelimumab in patients with PD-L1-Low/negative recurrent or metastatic HNSCC: the phase 2 CONDOR randomized clinical trial. JAMA Oncol. 2019; 5:195–203.
https://doi.org/10.1001/jamaoncol.2018.4628
PMID:30383184

8. D'Angelo SP, Mahoney MR, Van Tine BA, Atkins J, Milhem MM, Jahagirdar BN, Antonescu CR, Horvath E, Tap WD, Schwartz GK, Streicher H. Nivolumab with or without ipilimumab treatment for metastatic sarcoma (alliance A091401): two open-label, non-comparative, randomised, phase 2 trials. Lancet Oncol. 2018; 19:416–26.
https://doi.org/10.1016/S1470-2045(18)30006-8
PMID:29370992

9. Davis KL, Fox E, Merchant MS, Reid JM, Kudgus RA, Liu X, Minard CG, Voss S, Berg SL, Weigel BJ, Mackall CL. Nivolumab in children and young adults with relapsed or refractory solid tumours or lymphoma (ADVL1412): a multicentre, open-label, single-arm, phase 1-2 trial. Lancet Oncol. 2020; 21:541–50.
https://doi.org/10.1016/S1470-2045(20)30023-1
PMID:32192573

10. Groisberg R, Hong DS, Behrang A, Hess K, Janku F, Piha-Paul S, Naing A, Fu S, Benjamin R, Patel S, Somaiah N, Conley A, Meric-Bernstam F, Subbiah V. Characteristics and outcomes of patients with advanced sarcoma enrolled in early phase immunotherapy trials. J Immunother Cancer. 2017; 5:100.
https://doi.org/10.1186/s40425-017-0301-y
PMID:29254498

11. Le Cesne A, Marec-Berard P, Blay JY, Gaspar N, Bertucci F, Penel N, Bompas E, Cousin S, Toulmonde M, Bessede A, Fridman WH, Sautes-Fridman C, Kind M, et al. Programmed cell death 1 (PD-1) targeting in patients with advanced osteosarcomas: results from the PEMBROSARC study. Eur J Cancer. 2019; 119:151–57.
https://doi.org/10.1016/j.ejca.2019.07.018
PMID:31442817

12. Itzel T, Spang R, Maass T, Munker S, Roessler S, Ebert MP, Schlitt HJ, Herr W, Evert M, Teufel A. Random gene sets in predicting survival of patients with hepatocellular carcinoma. J Mol Med (Berl). 2019; 97:879–88.
https://doi.org/10.1007/s00109-019-01764-2
PMID:31001651

13. Fan CN, Ma L, Liu N. Systematic analysis of lncRNA-miRNA-mRNA competing endogenous RNA network identifies four-lncRNA signature as a prognostic biomarker for breast cancer. J Transl Med. 2018; 16:264.
https://doi.org/10.1186/s12967-018-1640-2
PMID:30261893

14. Zhou R, Zeng D, Zhang J, Sun H, Wu J, Li N, Liang L, Shi M, Bin J, Liao Y, Huang N, Liao W. A robust panel based on tumour microenvironment genes for prognostic prediction and tailoring therapies in stage I-III colon cancer. EBioMedicine. 2019; 42:420–30.
https://doi.org/10.1016/j.ebiom.2019.03.043
PMID:30917936

15. Leek JT, Scharpf RB, Bravo HC, Simcha D, Langmead B, Johnson WE, Geman D, Baggerly K, Irizarry RA. Tackling the widespread and critical impact of batch effects in high-throughput data. Nat Rev Genet. 2010; 11:733–39.
https://doi.org/10.1038/nrg2825 PMID:20838408

16. Li B, Cui Y, Diehn M, Li R. Development and validation of an individualized immune prognostic signature in early-stage nonsquamous non-small cell lung cancer. JAMA Oncol. 2017; 3:1529–37.
https://doi.org/10.1001/jamaoncol.2017.1609
PMID:28687838

17. Heinäniemi M, Nykter M, Kramer R, Wienecke-Baldacchino A, Sinkkonen L, Zhou JX, Kreisberg R, Kauffman SA, Huang S, Shmulevich I. Gene-pair expression signatures reveal lineage control. Nat Methods. 2013; 10:577–83.
https://doi.org/10.1038/nmeth.2445 PMID:23603899

18. Wu J, Zhao Y, Zhang J, Wu Q, Wang W. Development and validation of an immune-related gene pairs signature in colorectal cancer. Oncoimmunology. 2019; 8:1596715.
https://doi.org/10.1080/2162402X.2019.1596715
PMID:31143520

19. Schreiber RD, Old LJ, Smyth MJ. Cancer immunoediting: integrating immunity's roles in cancer suppression and promotion. Science. 2011; 331:1565–70.
https://doi.org/10.1126/science.1203486
PMID:21436444

20. Galon J, Bruni D. Tumor immunology and tumor evolution: intertwined histories. Immunity. 2020; 52:55–81.
https://doi.org/10.1016/j.immuni.2019.12.018
PMID:31940273

21. Brenner MJ, Cho JH, Wong NM, Wong WW. Synthetic biology: immunotherapy by design. Annu Rev Biomed Eng. 2018; 20:95–118.
https://doi.org/10.1146/annurev-bioeng-062117-121147 PMID:29345976

22. Whelan JS, Davis LE. Osteosarcoma, chondrosarcoma, and chordoma. J Clin Oncol. 2018; 36:188–93. https://doi.org/10.1200/JCO.2017.75.1743 PMID:29220289

23. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, Gonzalez R, Robert C, Schadendorf D, Hassel JC, Akerley W, van den Eertwegh AJ, Lutzky J, et al. Improved survival with ipilimumab in patients with metastatic melanoma. N Engl J Med. 2010; 363:711–23. https://doi.org/10.1056/NEJMoa1003466 PMID:20525992

24. Garon EB, Rizvi NA, Hui R, Leighl N, Balmanoukian AS, Eder JP, Patnaik A, Aggarwal C, Gubens M, Horn L, Carcereny E, Ahn MJ, Felip E, et al, and KEYNOTE-001 Investigators. Pembrolizumab for the treatment of non-small-cell lung cancer. N Engl J Med. 2015; 372:2018–28. https://doi.org/10.1056/NEJMoa1501824 PMID:25891174

25. Coley WB. II. Contribution to the knowledge of sarcoma. Ann Surg. 1891; 14:199–220. https://doi.org/10.1097/00000658-189112000-00015 PMID:17859590

26. Xie L, Xu J, Sun X, Guo W, Gu J, Liu K, Zheng B, Ren T, Huang Y, Tang X, Yan T, Yang R, Sun K, et al. Apatinib plus camrelizumab (anti-PD1 therapy, SHR-1210) for advanced osteosarcoma (APFAO) progressing after chemotherapy: a single-arm, open-label, phase 2 trial. J Immunother Cancer. 2020; 8:e000798. https://doi.org/10.1136/jitc-2020-000798 PMID:32376724

27. Orth MF, Buecklein VL, Kampmann E, Subklewe M, Noessner E, Cidre-Aranaz F, Romero-Pérez L, Wehweck FS, Lindner L, Issels R, Kirchner T, Altendorf-Hofmann A, Grünewald TG, Knösel T. A comparative view on the expression patterns of PD-L1 and PD-1 in soft tissue sarcomas. Cancer Immunol Immunother. 2020; 69:1353–62. https://doi.org/10.1007/s00262-020-02552-5 PMID:32222780

28. Itoh S, Yoshizumi T, Yugawa K, Imai D, Yoshiya S, Takeishi K, Toshima T, Harada N, Ikegami T, Soejima Y, Kohashi K, Oda Y, Mori M. Impact of immune response on outcomes in hepatocellular carcinoma: association with vascular formation. Hepatology. 2020. [Epub ahead of print]. https://doi.org/10.1002/hep.31206 PMID:32112577

29. Byrne A, Savas P, Sant S, Li R, Virassamy B, Luen SJ, Beavis PA, Mackay LK, Neeson PJ, Loi S. Tissue-resident memory T cells in breast cancer control and immunotherapy responses. Nat Rev Clin Oncol. 2020; 17:341–48. https://doi.org/10.1038/s41571-020-0333-y PMID:32112054

30. Fritzsching B, Fellenberg J, Moskovszky L, Sápi Z, Krenacs T, Machado I, Poeschl J, Lehner B, Szendrõi M, Bosch AL, Bernd L, Csóka M, Mechtersheimer G, et al. CD8+/FOXP3+-ratio in osteosarcoma microenvironment separates survivors from non-survivors: a multicenter validated retrospective study. Oncoimmunology. 2015; 4:e990800. https://doi.org/10.4161/2162402X.2014.990800 PMID:25949908

31. Lee CH, Espinosa I, Vrijaldenhoven S, Subramanian S, Montgomery KD, Zhu S, Marinelli RJ, Peterse JL, Poulin N, Nielsen TO, West RB, Gilks CB, van de Rijn M. Prognostic significance of macrophage infiltration in leiomyosarcomas. Clin Cancer Res. 2008; 14:1423–30. https://doi.org/10.1158/1078-0432.CCR-07-1712 PMID:18316565

32. Jensen TO, Schmidt H, Møller HJ, Høyer M, Maniecki MB, Sjoegren P, Christensen IJ, Steiniche T. Macrophage markers in serum and tumor have prognostic impact in American joint committee on cancer stage I/II melanoma. J Clin Oncol. 2009; 27:3330–37. https://doi.org/10.1200/JCO.2008.19.9919 PMID:19528371

33. Hagemann T, Wilson J, Burke F, Kulbe H, Li NF, Plüddemann A, Charles K, Gordon S, Balkwill FR. Ovarian cancer cells polarize macrophages toward a tumor-associated phenotype. J Immunol. 2006; 176:5023–32. https://doi.org/10.4049/jimmunol.176.8.5023 PMID:16585599

34. Liang X, Guo W, Ren T, Huang Y, Sun K, Zhang H, Yu Y, Wang W, Niu J. Macrophages reduce the sensitivity of osteosarcoma to neoadjuvant chemotherapy drugs by secreting interleukin-1 beta. Cancer Lett. 2020; 480:4–14. https://doi.org/10.1016/j.canlet.2020.03.019 PMID:32220541

35. Buddingh EP, Kuijjer ML, Duim RA, Bürger H, Agelopoulos K, Myklebost O, Serra M, Mertens F, Hogendoorn PC, Lankester AC, Cleton-Jansen AM. Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. Clin Cancer Res. 2011; 17:2110–19. https://doi.org/10.1158/1078-0432.CCR-10-2047 PMID:21372215

36. Gomez-Brouchet A, Illac C, Gilhodes J, Bouvier C, Aubert S, Guinebretiere JM, Marie B, Larousserie F, Entz-Werlé N, de Pinieux G, Filleron T, Minard V, Minville V, et al. CD163-positive tumor-associated

macrophages and CD8-positive cytotoxic lymphocytes are powerful diagnostic markers for the therapeutic stratification of osteosarcoma patients: an immunohistochemical analysis of the biopsies fromthe French OS2006 phase 3 trial. Oncoimmunology. 2017; 6:e1331193.
https://doi.org/10.1080/2162402X.2017.1331193
PMID:28932633

37. Shimura T, Toiyama Y, Tanaka K, Saigusa S, Kitajima T, Kondo S, Okigami M, Yasuda H, Ohi M, Araki T, Inoue Y, Uchida K, Mohri Y, Kusunoki M. Angiopoietin-like Protein 2 as a Predictor of Early Recurrence in Patients After Curative Surgery for Gastric Cancer. Anticancer Res. 2015; 35:4633–9.
PMID:26254352

38. Toiyama Y, Inoue Y, Shimura T, Fujikawa H, Saigusa S, Hiro J, Kobayashi M, Ohi M, Araki T, Tanaka K, Mohri Y, Kusunoki M. Serum angiopoietin-like protein 2 improves preoperative detection of lymph node metastasis in colorectal cancer. Anticancer Res. 2015; 35:2849–56.
PMID:25964566

39. Wang X, Hu Z, Wang Z, Cui Y, Cui X. Angiopoietin-like protein 2 is an important facilitator of tumor proliferation, metastasis, angiogenesis and glycolysis in osteosarcoma. Am J Transl Res. 2019; 11:6341–55.
PMID:31737187

40. Zhang T, Kastrenopoulou A, Larrouture Q, Athanasou NA, Knowles HJ. Angiopoietin-like 4 promotes osteosarcoma cell proliferation and migration and stimulates osteoclastogenesis. BMC Cancer. 2018; 18:536.
https://doi.org/10.1186/s12885-018-4468-5
PMID:29739381

41. Chen Q, Sun W, Liao Y, Zeng H, Shan L, Yin F, Wang Z, Zhou Z, Hua Y, Cai Z. Monocyte chemotactic protein-1 promotes the proliferation and invasion of osteosarcoma cells and upregulates the expression of AKT. Mol Med Rep. 2015; 12:219–25.
https://doi.org/10.3892/mmr.2015.3375
PMID:25695619

42. de Ridder D, Marino S, Bishop RT, Renema N, Chenu C, Heymann D, Idris AI. Bidirectional regulation of bone formation by exogenous and osteosarcoma-derived Sema3A. Sci Rep. 2018; 8:6877.
https://doi.org/10.1038/s41598-018-25290-2
PMID:29720701

43. Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, Berger P, Desborough V, Smith T, Campbell J, Thomson E, Monteiro R, Guimaraes P, et al. ImmPort: disseminating data to the public for the future of immunology. Immunol Res. 2014; 58:234–39.
https://doi.org/10.1007/s12026-014-8516-1
PMID:24791905

44. Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. Biometrics. 2000; 56:337–44.
https://doi.org/10.1111/j.0006-341x.2000.00337.x
PMID:10877287

45. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, de Reyniès A. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016; 17:218.
https://doi.org/10.1186/s13059-016-1070-5
PMID:27765066

46. Sturm G, Finotello F, Petitprez F, Zhang JD, Baumbach J, Fridman WH, List M, Aneichyk T. Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology. Bioinformatics. 2019; 35:i436–45.
https://doi.org/10.1093/bioinformatics/btz363
PMID:31510660

47. Ock CY, Keam B, Kim S, Lee JS, Kim M, Kim TM, Jeon YK, Kim DW, Chung DH, Heo DS. Pan-cancer immunogenomic perspective on the tumor microenvironment based on PD-L1 and CD8 t-cell infiltration. Clin Cancer Res. 2016; 22:2261–70.
https://doi.org/10.1158/1078-0432.CCR-15-2834
PMID:26819449

48. Chen YP, Zhang Y, Lv JW, Li YQ, Wang YQ, He QM, Yang XJ, Sun Y, Mao YP, Yun JP, Liu N, Ma J. Genomic analysis of tumor microenvironment immune types across 14 solid cancer types: immunotherapeutic implications. Theranostics. 2017; 7:3585–94.
https://doi.org/10.7150/thno.21471
PMID:28912897

49. Hoshida Y, Brunet JP, Tamayo P, Golub TR, Mesirov JP. Subclass mapping: identifying common subtypes in independent disease data sets. PLoS One. 2007; 2:e1195.
https://doi.org/10.1371/journal.pone.0001195
PMID:18030330

50. Roh W, Chen PL, Reuben A, Spencer CN, Prieto PA, Miller JP, Gopalakrishnan V, Wang F, Cooper ZA, Reddy SM, Gumbs C, Little L, Chang Q, et al. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. Sci Transl Med. 2017; 9:eaah3560.
https://doi.org/10.1126/scitranslmed.aah3560
PMID:28251903
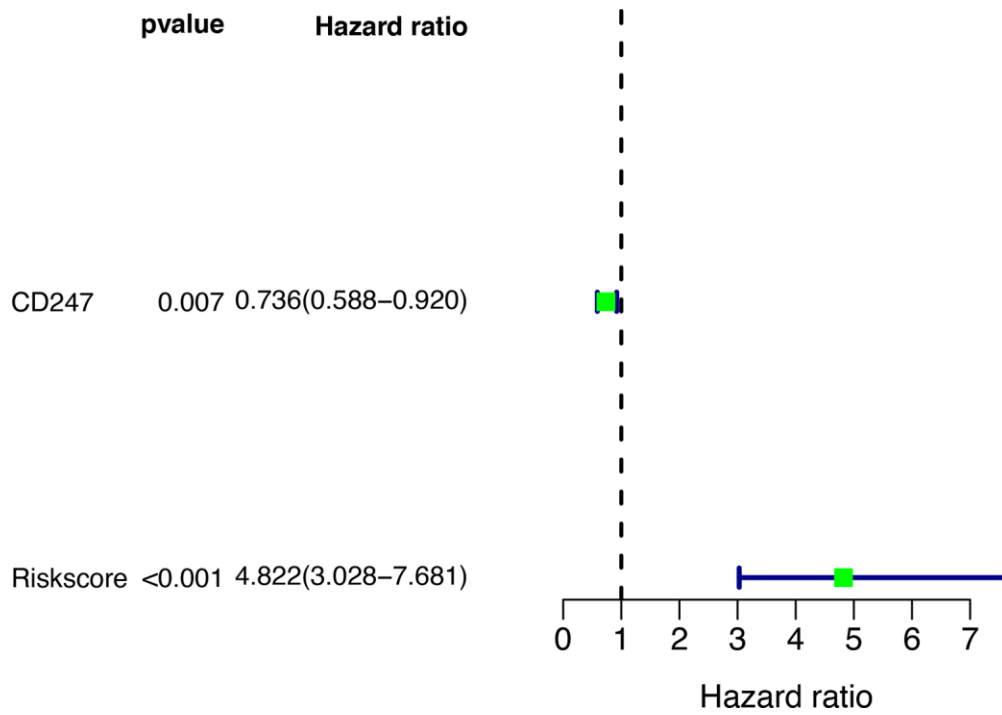
Supplementary Figure 1. Results of univariate COX regression analysis of PD-L1 and Riskscore.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1 to 3.

Supplementary Table 1. IRGs with higher variation for the construction of immune-related gene pairs.

Supplementary Table 2. IRGP with higher variation, which are used for further analysis.

Supplementary Table 3. Scores of 2260 IRGPs with higher variability for each patient.

Supplementary Table 4. Results of Univariate Cox regression of IRGP with higher variation. (P<0.005).

IRGP	HR	HR.95L	HR.95H	coxPvalue
FCER1G|NDRG1	0.27639592	0.12554102	0.60852383	0.00140533
HLA-DQB1|STC2	0.23444675	0.09967148	0.55146446	0.0008881
APOBEC3G|GAL	0.33497266	0.1591734	0.70493365	0.0039641
FABP4|GAL	0.23429033	0.09975861	0.55024781	0.00086442
OASL|GAL	0.33794111	0.1617213	0.70617906	0.00391252
APOD|CCL2	3.03038955	1.43686979	6.39115728	0.00359138
TFRC|ANGPTL2	3.63343365	1.69767116	7.77644128	0.00088986
TFRC|TNFRSF21	3.00990706	1.42652227	6.35078797	0.00382254
F2R|SEMA3B	0.32349419	0.15533762	0.67368412	0.00256725
F2R|ANGPTL4	0.29288792	0.13270808	0.64640626	0.00236392
CCL5|SEMA3B	0.2096095	0.07978568	0.55067705	0.0015214
WNT5A|GAL	0.28471691	0.13684325	0.59238377	0.00077754
HMOX1|FOS	0.32346229	0.15451936	0.67711811	0.00274979
HCK|SEMA5A	0.28401534	0.12572024	0.64162071	0.00246845
HCK|PLXNB1	0.27935222	0.12343848	0.63219887	0.00221056
HCK|GAL	0.28460063	0.13388452	0.60498048	0.00109025
HCK|STC1	0.20276464	0.07029161	0.58489914	0.00315527
RAC3|FGFRL1	7.15301218	2.12340384	24.0960209	0.00149745
SEMA3A|SEMA4D	0.33542642	0.16001111	0.70314419	0.00382089
SEMA3A|GAL	0.22876178	0.10756573	0.48651139	0.00012742
SEMA3A|TNFRSF11B	0.31504209	0.1457478	0.68098126	0.003315
SEMA3B|LTBP4	3.63638195	1.74042023	7.59774764	0.00059502
SEMA3B|VEGFC	2.95306523	1.3886214	6.28003733	0.00491149
SEMA3B|C3AR1	3.7431678	1.51781114	9.2312573	0.00415712
SEMA3B|FGFRL1	4.25029395	1.96251825	9.20500926	0.00024253
SEMA5A|C5AR1	3.93004313	1.59685975	9.67225772	0.00289627
C5AR1|PLXNB1	0.26887173	0.10918511	0.66210498	0.00428033
EDNRA|STC2	0.22184761	0.10620662	0.46340204	6.16E-05
PLXNB1|VEGFC	2.88263163	1.38280123	6.009226	0.00473213
PLXNB1|FGFRL1	3.23045085	1.51805887	6.87444531	0.00233969
IGF2|PDGFD	0.33598457	0.1612484	0.7000729	0.00359188
ANGPTL2|SORT1	0.2760354	0.13008073	0.5857558	0.00079861

Supplementary Table 5. Clinical information of patients in the GSE78220 data set.

ID	Gender	Age	OS	Status	Response
GSM2069823	F	66	607	DEAD	progressive
GSM2069824	M	55	927	ALIVE	Partial
GSM2069825	M	62	948	ALIVE	Partial
GSM2069826	M	61	439	ALIVE	Partial
GSM2069827	M	51	882	ALIVE	Partial
GSM2069828	F	55	662	DEAD	progressive
GSM2069829	M	69	NA	NA	Complete
GSM2069830	M	68	1054	ALIVE	Complete
GSM2069831	M	60	387	ALIVE	progressive
GSM2069832	M	59	327	DEAD	progressive
GSM2069833	F	53	917	ALIVE	Complete
GSM2069834	F	27	54	ALIVE	progressive
GSM2069835	M	70	980	DEAD	Partial
GSM2069836	M	19	186	DEAD	progressive
GSM2069837	M	45	1060	ALIVE	Partial
GSM2069838	F	63	337	DEAD	progressive
GSM2069839	M	55	182	DEAD	progressive
GSM2069840	M	63	103	DEAD	progressive
GSM2069841	M	74	262	DEAD	progressive
GSM2069842	M	83	548	ALIVE	Complete
GSM2069843	M	83	548	ALIVE	Complete
GSM2069844	M	82	439	DEAD	Partial
GSM2069845	M	84	269	DEAD	progressive
GSM2069846	M	47	704	ALIVE	progressive
GSM2069847	M	47	171	DEAD	progressive
GSM2069848	F	65	427	ALIVE	Partial
GSM2069849	F	70	364	ALIVE	Partial
GSM2069850	F	57	448	ALIVE	Partial