



OPEN

Development of a novel immune-related genes prognostic signature for osteosarcoma

Zuo-long Wu^{2,3,4}, Ya-jun Deng^{3,4}, Guang-zhi Zhang^{3,4}, En-hui Ren^{1,3}, Wen-hua Yuan³ & Qi-qi Xie^{1,3}✉

Immune-related genes (IRGs) are responsible for osteosarcoma (OS) initiation and development. We aimed to develop an optimal IRGs-based signature to assess of OS prognosis. Sample gene expression profiles and clinical information were downloaded from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) and Genotype-Tissue Expression (GTEx) databases. IRGs were obtained from the ImmPort database. R software was used to screen differentially expressed IRGs (DEIRGs) and functional correlation analysis. DEIRGs were analyzed by univariate Cox regression and iterative LASSO Cox regression analysis to develop an optimal prognostic signature, and the signature was further verified by independent cohort (GSE39055) and clinical correlation analysis. The analyses yielded 604 DEIRGs and 10 hub IRGs. A prognostic signature consisting of 13 IRGs was constructed, which strikingly correlated with OS overall survival and distant metastasis ($p < 0.05$, $p < 0.01$), and clinical subgroup showed that the signature's prognostic ability was independent of clinicopathological factors. Univariate and multivariate Cox regression analyses also supported its prognostic value. In conclusion, we developed an IRGs signature that is a prognostic indicator in OS patients, and the signature might serve as potential prognostic indicator to identify outcome of OS and facilitate personalized management of the high-risk patients.

Osteosarcoma (OS) is a primary bone malignant tumor that most commonly affects children, adolescents, and young adults, and it also exhibits a predilection to occur in the metaphysis of long bones, and most commonly occurs in the distal femur (43%), proximal tibia (23%), or humerus (10%)¹. Additionally, osteosarcoma is aggressive and often metastasizes to the lungs². In the past 10 years, the incidence of OS has been annually increasing by 0.3%³, and it has been consistently ranked as the second deadliest cancer in adolescents and children⁴. Despite advances in multimodal therapy, the 5-year survival of osteosarcoma is approximately 60% to 70%, which has remained stagnant over the past three decades, patients with distant metastases still fare poorly, as the 5-year survival rate in these patients does not exceed 20%^{5,6}. In addition, patients with the same clinical or pathological conditions receiving the same treatment regimen may have different clinical outcomes, due to their genetic heterogeneity⁷. Therefore, in-depth exploration of the molecular mechanisms behind the development of OS is crucial to finding effective prognostic biomarkers to guide patient risk stratification, which aligns with the concept advocated by precision medicine.

In recent years, biomolecules and risk models have been used to evaluate the prognosis of OS^{8–11}. However, they have not yet been used in clinical practice because of unavoidable limitations, such as overfitting due to small samples¹². In recent decades, increasing evidence has indicated that the immune response is actively involved in OS occurrence and progression¹³. Immune genes act as pivotal regulator of immune response^{14,15}. They maintain the body's self-tolerance by strictly regulating the immune function and reducing the damage inflicted on the surrounding tissues¹⁶. However, OS cell may use these immune genes to escape the immune system and achieve a favorable environment for their growth^{13,17}. Given the critical role of immune molecules in OS prognosis, these immune-related genes (IRGs) deserve further study.

Here, we identified differentially expressed IRGs (DEIRGs) in OS and normal muscle tissue samples. Subsequently, an IRGs signature that can predict outcome of OS was constructed by using univariate Cox regression and iterative LASSO Cox regression analysis of DEIRGs. In addition, based on an independent cohort, the

¹Breast Disease Diagnosis and Treatment Center, Affiliated Hospital of Qinghai University & Affiliated Cancer Hospital of Qinghai University, No.29 Tongren Road, Xining 810000, Qinghai, China. ²Guanghe Traditional Chinese and Western Medicine Hospital, Lanzhou 730000, Gansu, China. ³Department of Orthopaedics, Second Hospital of Lanzhou University, Lanzhou 730000, Gansu, China. ⁴These authors contributed equally: Zuo-long Wu, Ya-jun Deng and Guang-zhi Zhang. ✉email: jieqq16@lzu.edu.cn

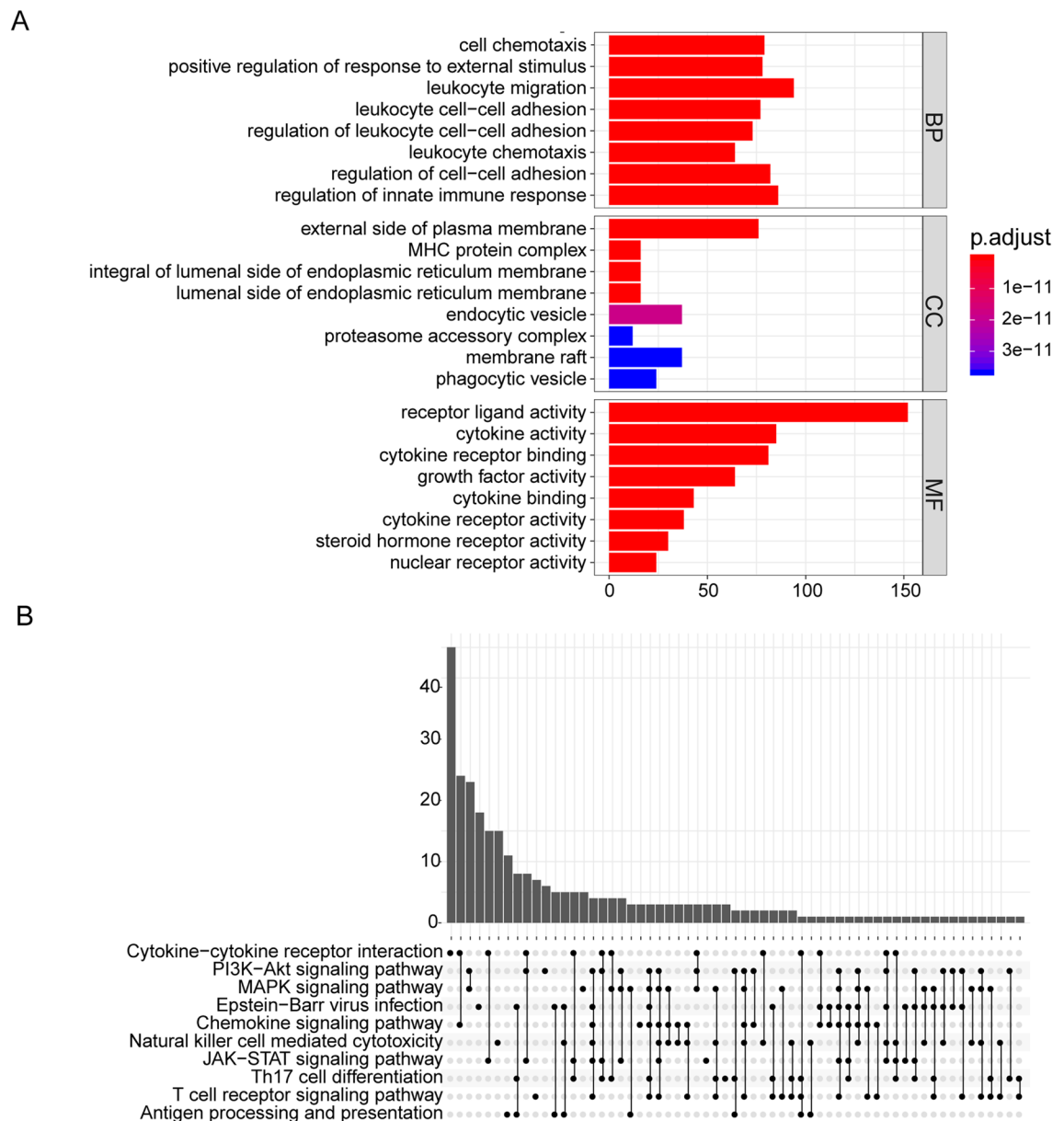


Figure 1. GO and KEGG analysis of DEIRGs. **(A)** GO enrichment analysis of DEIRGs. The color of the bar indicates p.adjust: the redder the color, the smaller the p.adjust value; the bluer the color, the larger the p.adjust value. The horizontal axis represents the number of DEIRGs under the GO term. **(B)** KEGG pathways enrichment analysis of DEIRGs. Significant gene (p.adjust < 0.05) enrichment to the 10 most important paths. p.adjust: adjusted P-value; BP: biological process; CC: cellular component; MF: molecular function.

accuracy of IRGs signature in predicting the prognosis of OS patients was verified. Finally, we also evaluated the independence, repeatability, and clinical value of the IRGs signature in different clinical subgroups. Our results reveal the prognostic value of IRGs signature and provide promising prognostic indicator for OS.

Results

GO and KEGG pathways enrichment analysis of DEIRGs. All 604 DEIRGs were screened. GO analysis results showed that DEIRGs are involved in biological functions such as cell chemotaxis, leukocyte adhesion, and innate immune regulation. They were also determined to participate in cellular components such as the external side of the plasma membrane, MHC protein complexes, endoplasmic reticulum membranes, and phagocytic vesicles. Additionally, they are also found to be involved in molecular functions including receptor-ligand, cytokine, steroid receptor, and nuclear receptor activity (Fig. 1A). The KEGG analysis indicated that DEIRGs were mainly enriched in the following signaling pathways: chemokines, PI3K/AKT, MAPK, JAK-STAT, and natural killer (NK) cell-mediated cytotoxicity signaling (Fig. 1B).

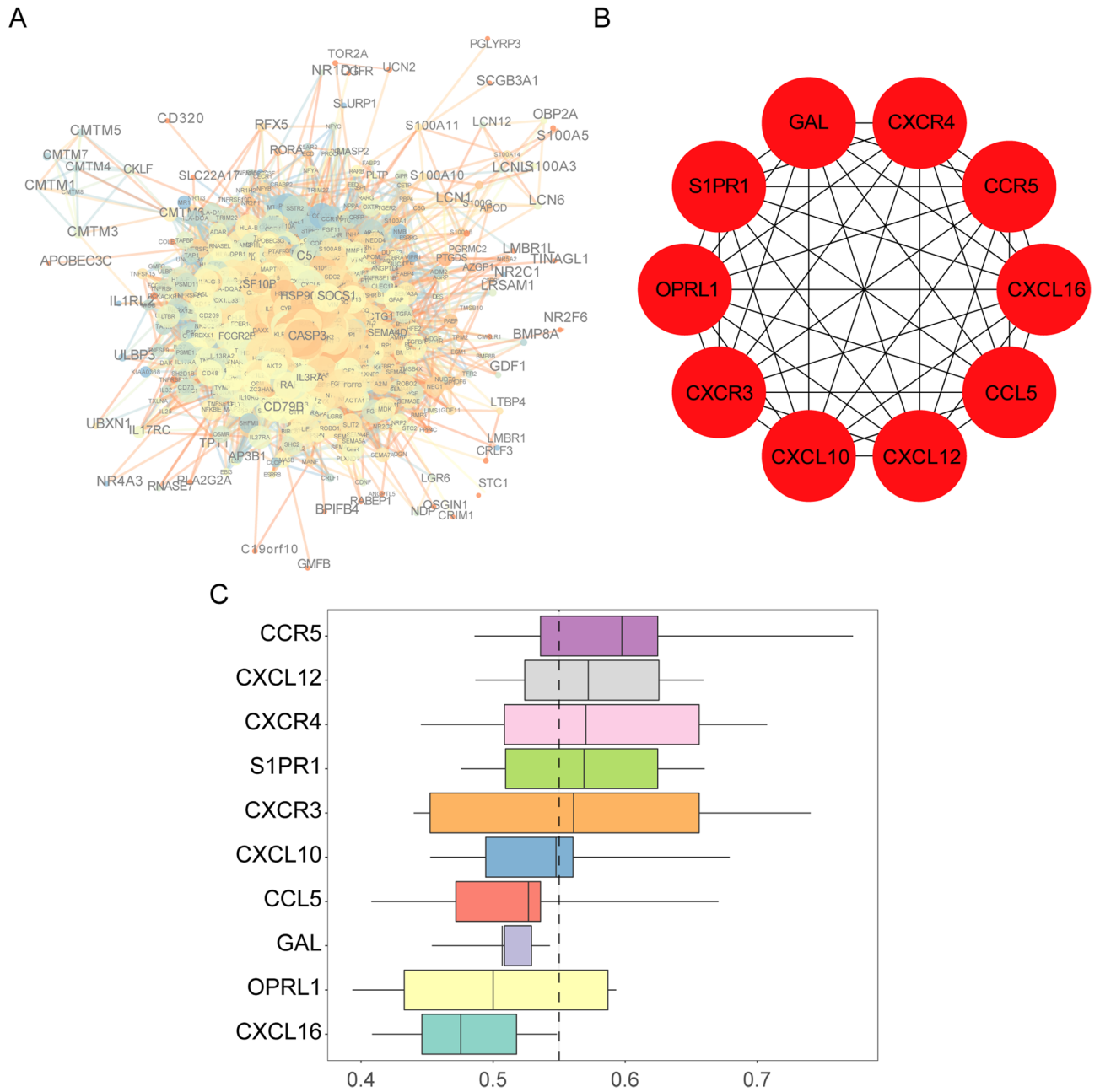


Figure 2. Protein–protein interaction (PPI), hub IRGs, and functional similarity analysis of DEIRGs. **(A)** PPI network. The size of a node represents the clustering coefficient, the color indicates the degree, the width of the line indicates the score; the color of the line represents co-expression. **(B)** Hub IRGs. The hub IRGs were the top 10 DEIRGs scored by the maximum correlation coefficient. **(C)** Functional similarities of 10 hub IRGs. The boxes indicate the middle 50% of the similarities; the upper and lower boundaries represent the 75th and 25th percentiles. The two ends of the line represent the maximum and minimum values. The dashed line represents the cut-off value of similarity.

PPI network construction, hub IRGs screening, and functional similarity analysis of DEIRGs. The result of these analyses were shown in Fig. 2A. CASP3, TNFRSF10B, and HSP90 had a larger weight and a stronger correlation in the PPI network. Ten hub IRGs were obtained, namely CXCR4, CCR5, CXCL16, CCL5, CXCL12, CXCL10, CXCR3, OPRL1, S1PR1, and GAL (Fig. 2B, Table 1). To further recognize the closeness of the interactions between hub IRGs, which were ranked according to average functional similarity, as indicated by the results, CCR5, CXCL12, CXCR4, S1PR1, and CXCR4 were found to be hub genes with cut-off values greater than 0.55, and CCR5, CXCL12, and CXCR4 were the most closely related genes (Fig. 2C).

Identification and assessment of the prognostic signature. To identify the optimal prognostic signature of OS based on IRGs, 82 prognostic-associated IRGs were identified by univariate Cox regression analysis of DEIRGs. Further, we identified the optimal prognostic signature that consisted of 13 prognosis-associated

No	Symbol	Full name	Function
1	CXCL16	C-X-C motif chemokine ligand 16	CXCL16 is highly expressed in osteosarcoma tissues
2	CCL5	C-C motif chemokine ligand 5	High CCL5 expression is associated with osteosarcoma metastasis and poor prognosis of patients with osteosarcoma
3	CCR5	C-C motif chemokine receptor 5	CCR5 controls the proliferation or differentiation of osteosarcoma
4	GAL	Galanin and GMAP prepropeptide	The overexpression of Gal-1 is well established in many types of cancer progression like osteosarcoma, breast, lung, prostate, melanoma, etc
5	S1PR1	Sphingosine-1-phosphate receptor 1	Downregulated S1PR1 suppresses osteosarcoma metastasis and proliferation
6	CXCR4	C-X-C motif chemokine receptor 4	CXCR4-mediated osteosarcoma growth and pulmonary metastasis
7	CXCL12	C-X-C motif chemokine ligand 12	CXCL12 plays a critical role in mediating tumor progression and the immune response in osteosarcoma
8	CXCR3	C-X-C motif chemokine receptor 3	CXCR3 correlates with immune infiltration and predicts poor survival in osteosarcoma
9	CXCL10	C-X-C motif chemokine ligand 10	CXCL10 plays an important role in cancer and autoimmunity
10	OPRL1	Opioid-related nociceptin receptor 1	OPRL1 plays a key role of pain and injury perception

Table 1. Functions of 10 hub IRGs.

IRGs via the iterative LASSO Cox regression analysis (Fig. 3A, Table 2). ROC curve results showed that the accuracy of this signature in diagnosing OS prognosis was high (Fig. 3B, AUC = 0.918). The Kaplan–Meier curve indicated that the overall survival of patients in the high-risk group was markedly worse than that in the low-risk group (Fig. 3C, $p < 0.001$). According to the optimal signature, we obtained the risk score distribution (Fig. 4A), the survival status (Fig. 4B), and the expression characteristics of the immune genes of OS (Fig. 4C). Compared to the low-risk group, the high-risk group had more deaths. In addition, the expression levels of GNRH1, VEGFA, TNFRSF11B, GAL, STC2, BRAF, BMP8A, and CORT were higher in the high-risk group, whereas patients in the low-risk group expressed higher levels of PSMD10, TNFRSF21, GRN, VAV1, and SDC3.

Comparison of the IRGs signature with other known prognostic biomarkers and verification in independent cohort.

To determine whether the IRGs signature has a better diagnostic capacity for OS patient survival, we conducted receiver operating characteristic (ROC) analysis of the IRG signature along with other known prognostic biomarkers (SP140, MALAT1, UCA1, and MIR191) in the training cohort. The results showed that the area under the curve (AUC) of the IRGs signature was increased compared to that for other known biomarkers (Fig. 5A), indicating that the IRG signature was a better prognostic biomarker and provided better stability and reliability in predicting the survival of OS patients. To further examine the prognostic value of the IRG signature, we conducted the ROC analysis in another independent cohort (GSE39055). The results showed that the AUCs were 0.92, 0.93, and 0.89 at 1, 3, and 5 years, respectively (Fig. 5B), suggesting that the IRG can also predict the survival of OS patients in other independent cohorts.

Independence of the IRGs signature in survival prediction from clinicopathological factors.

An important feature of a good prognostic biomarker is that it should be independent of clinicopathological prognostic factors. Clinicopathological characteristics, such as the patient's age, sex, and metastasis, are also considered to be the main factors that determine the prognosis of OS patients. To evaluate the independence and applicability of the IRGs signature, we regrouped patients according to different clinicopathological characteristics and performed Kaplan–Meier survival analysis. The Kaplan–Meier curve showed that regardless of sex, age, and metastasis, the survival time of OS patients in the low-risk group was significantly prolonged ($p < 0.05$, Fig. 6A–C). All of results indicated that the IRGs signature showed satisfactory applicability when grouping patients according to different clinicopathological characteristics. Univariate and multivariate COX regression also suggested that the signature is an independent indicator for predicting the prognosis of OS patients (Table 3).

Relationship between the prognostic signature and clinical characteristics. The relationship between clinical characteristics, such as metastasis, age, grade, and the risk score based on the prognosis-associated IRGs signature, was analyzed to validate the accuracy of the prognostic signature further. The results showed that metastasis groups had a significantly higher risk score than non-metastasis groups (Fig. 7C, $p = 0.001$). However, no significant association was observed between age (Fig. 7A, $p = 0.531$), sex (Fig. 7B, $p = 0.485$), and risk score.

Discussion

OS is the most common bone malignancy in children and adolescents, and it is also one of the main causes of cancer-related deaths in this age group¹⁸. Evidence demonstrates that the immune response defines the tumor's microenvironment. In particular, immune cell disorders often cooccur with tumors and are considered an essential driver of OS development^{19,20}. In this study, we analyzed the DEIRGs of the OS and control samples from TARGET and GTEx databases to identify new prognostic biomarkers by constructing a prognostic IRGs signature.

Related studies show that chemotaxis, adhesion of leukocyte, and innate immunity are dysfunctional in the OS microenvironment, thereby reducing the immune response to OS cells^{21–23}. PI3K/AKT signaling pathway²⁴, MAPK signaling pathway²⁵ and JAK-STAT signaling pathway²⁶ have been extensively studied in the OS.

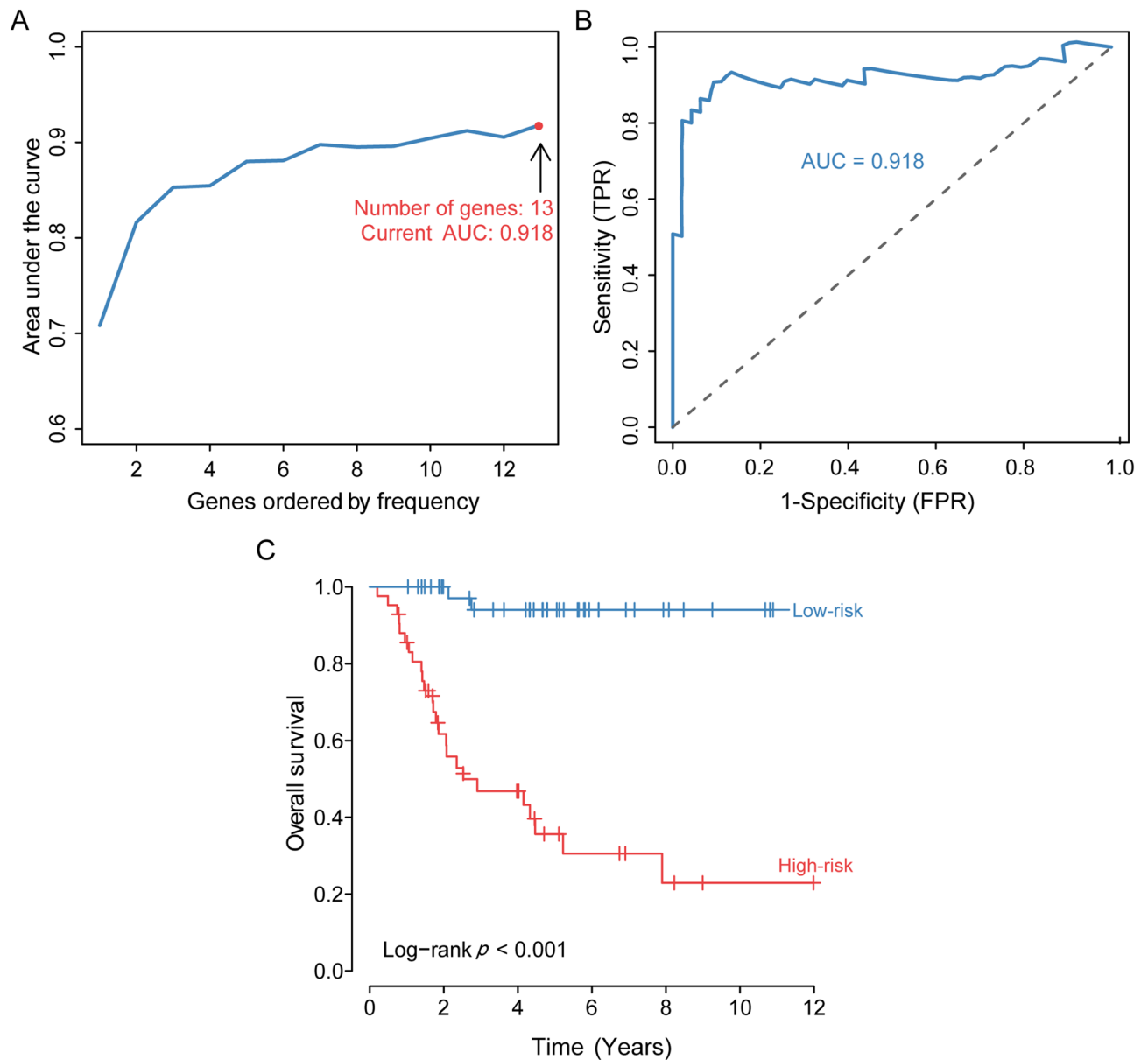


Figure 3. Development and assessment of the prognostic signature. **(A)** Construction of the prognosis-associated IRGs signature. The horizontal axis represents the gene frequency and the vertical axis represents the AUC. **(B)** Time-dependent ROC curve for prognosis-associated DEIRGs signature. The horizontal axis indicates the FDR, and the vertical axis indicates the TPR. **(C)** Kaplan–Meier survival curves of overall survival from the high-risk and low-risk groups. The horizontal axis represents survival time (y), and the vertical axis represents the survival rate (%). ROC: relative operating characteristic curve; AUC: area under the curve. FDR: false positive rate; TPR: true positive rate.

Furthermore, the activation of these signaling pathways is strongly linked to the growth and metastasis of OS cells. Although natural killer cell-mediated cytotoxicity is the host's first-line anti-cancer defense²⁷, the immune response is a seemingly double-edged sword in the OS microenvironment, as a dysregulated immune response is conducive to the occurrence and development of tumors.

In total, in our study, we obtained 604 DEIRGs. Of note, we identified 10 hub IRGs, namely CXCR3, CXCR4, CCR5, CCL5, CXCL10, CXCL12, CXCL16, OPRL1, S1PR1, and GAL. Among them, CXCR3²⁸, CXCR4²⁹, CCR5³⁰, CCL5³¹, CXCL16³², CXCL10³³, CXCL12³⁴ and GAL³⁵ have been widely studied in OS, and are involved in the occurrence, metastasis, and angiogenesis of OS. OPRL1 encodes proteins that are endogenous opioid-related neuropeptides and nociceptin/orphanin receptors, which plays a key role in pain perception and nociception^{36,37}. The high expression of OPRL1 in OS may be related to cancerous pain. The coding product of the S1PR1 gene is a receptor protein that is similar to the G-protein-coupled receptor. When S1PR1 was combined with ligand S1P, the growth, invasion, and metastasis of lung cancer, ovarian cancer, and colon cancer are enhanced^{38–40}. Hence, we can speculate that S1PR1 is pivotal in OS. Considering the similarity between molecular functions and cell components of hub IRGs, and through the ranking of semantic similarity, we discovered that CCR5, CXCL12

No	Symbol	Full name	Function
1	BRAF	B-Raf proto-oncogene, serine/threonine kinase	Associated with progression and poor prognosis of several cancers
2	CORT	Cortistatin	Biological activities of anti-inflammation, antioxidation, antitumor activity
3	GAL	Galanin and GMAP prepropeptide	High expression is linked with the initiation and progression of prostate cancer and colorectal cancer
4	GRN	Granulin precursor	Inhibit the growth of hepatocellular carcinoma
5	STC2	Stanniocalcin 2	High expression is associated with progression and poor outcome of colorectal cancer
6	TNFRSF11B	TNF receptor superfamily member 11b	High expression is linked with a worse outcome of colorectal cancer
7	BMP8A	Bone morphogenetic protein 8a	High expression is associated with progression and poor survival of thyroid cancer
8	PSMD10	Proteasome 26S subunit, non-ATPase 10	Abnormal expression is linked with metastasis and worse survival of osteosarcoma
9	VEGFA	Vascular endothelial growth factor A	Promotes cell proliferation and migration, inhibits apoptosis of osteosarcoma
10	VAV1	Vav guanine nucleotide exchange factor 1	High expression is associated with the prognosis of invasive breast cancer
11	GNRH1	Gonadotropin-releasing hormone 1	Associated with vascular invasion and metastasis of osteosarcoma
12	SDC3	Syndecan 3	Overexpression inhibits the proliferation of mesenchymal tumor cells
13	TNFRSF21	TNF receptor superfamily member 21	Abnormal expression is coupled with growth, migration and invasion of osteosarcoma

Table 2. IRGs function in the prognostic signature.

and CXCR4 are the most closely related genes. CCR5, CXCL12 and CXCR4 genes encode chemokine receptors or ligands, which plays a vital part role in the initiation and growth of OS^{29,41,42}. These findings further support the reliability of our study.

Previous research has shown that IRGs are closely related to OS metastasis and prognosis⁴³. For example, Koirala et al.⁴⁴ found that immune cell infiltration and PD-L1 expression in the tumor microenvironment were independent risk factors for OS. Li Bo et al.⁴⁵ reported that CXCL12 acts as a driver in OS metastasis and immune response, and knocking down CXCL12 could effectively inhibit OS progression. Moreover, IRGs signatures have attracted widespread attention and have been used to predict metastasis and prognosis of different tumors^{46–48}. Therefore, in order to further explore the value of IRGs in OS prognosis, we constructed a prognostic signature consisting of 13 prognostic-associated DEIRGs, which has a high diagnostic prognostic efficacy. The high expression level of GNRH1⁴⁹, BRAF⁵⁰, PSMD10⁵¹ and VEGFA⁵² closely correlated with the growth, metastasis, and angiogenesis of OS. The high expression of GAL^{53,54}, TNFRSF11B⁵⁵ and STC2⁵⁶ are linked to prostate cancer and colorectal cancer development and a worse prognosis. The abnormally high expression of BMP8A is an independent factor for the progression and poor prognosis of thyroid carcinoma⁵⁷. CORT is an endogenous cyclic neuropeptide that can regulate the growth and metastasis of lung cancer and thyroid cancer^{58,59}, and it also regulates the inflammatory response by inhibiting the immune infiltration⁶⁰. Granulin a (GRNA) is a 6 kDa peptide hydrolyzed from PGRN, which can effectively inhibit the growth and invasion of human hepatoma cells⁶¹. The high expression of VAV1 is a positive prognostic factor for early invasive breast cancer⁶². Zong et al.⁶³ found that the overexpression of SDC3 can significantly inhibit the proliferation and metastasis of mesenchymal tumor cells. Wu et al.⁶⁴ found that miR20a-5p promotes the proliferation, migration, and invasion of head and neck squamous cell carcinoma by down-regulating TNFRSF21. Another study found that TNFRSF21 also plays an important role in regulating leukocyte infiltration⁶⁵. Obviously, the results of our analysis are consistent with the results of previous studies, which further confirms that this signature has a high value for the prognosis of OS.

To date, a lot of OS prognostic molecules have been found, including MALAT1⁹, UCA1¹⁰ and miR191¹¹. Most of these were based on single-gene prognosis studies. Existing studies have found that the occurrence and development of tumors are not caused by changes in single genes, but are the result of a series of gene changes⁶⁶. In addition, the use of single genes cannot avoid the differences caused by individual heterogeneity. Most importantly, these studies did not use large samples to fully explore the relationship between genes and the prognosis of OS. In this study, 13 prognostic IRGs were identified by univariate cox regression and iterative LASSO cox regression analysis for the risk stratification of OS patients. Extensive analyses proved that this prognostic signature has a higher diagnostic value than pre-existing models. Recently, Shi et al.⁶⁷ also constructed a prognostic signature that consisted of three DEGs (MYC, CPE, and LY86) in OS. However, the DEGs in their study came from metastatic and non-metastatic patients and lacked a normal control sample. Therefore, the gene included in the signature did not reflect the pathological characteristics of the occurrence and development of OS. Our signature was verified by an independent verification set, which has a high diagnostic efficiency compared to that with other biomarkers. However, our research also has some unavoidable limitations and deficiencies. First, in the study, we used normal muscle tissue as a control group. Therefore, compared with normal bone tissue, there may be a certain difference in the expression of IRGs. In addition, due to the lack of protein expression profile data for OS, we used gene expression profile data, which may not fully reflect the biological characteristics of

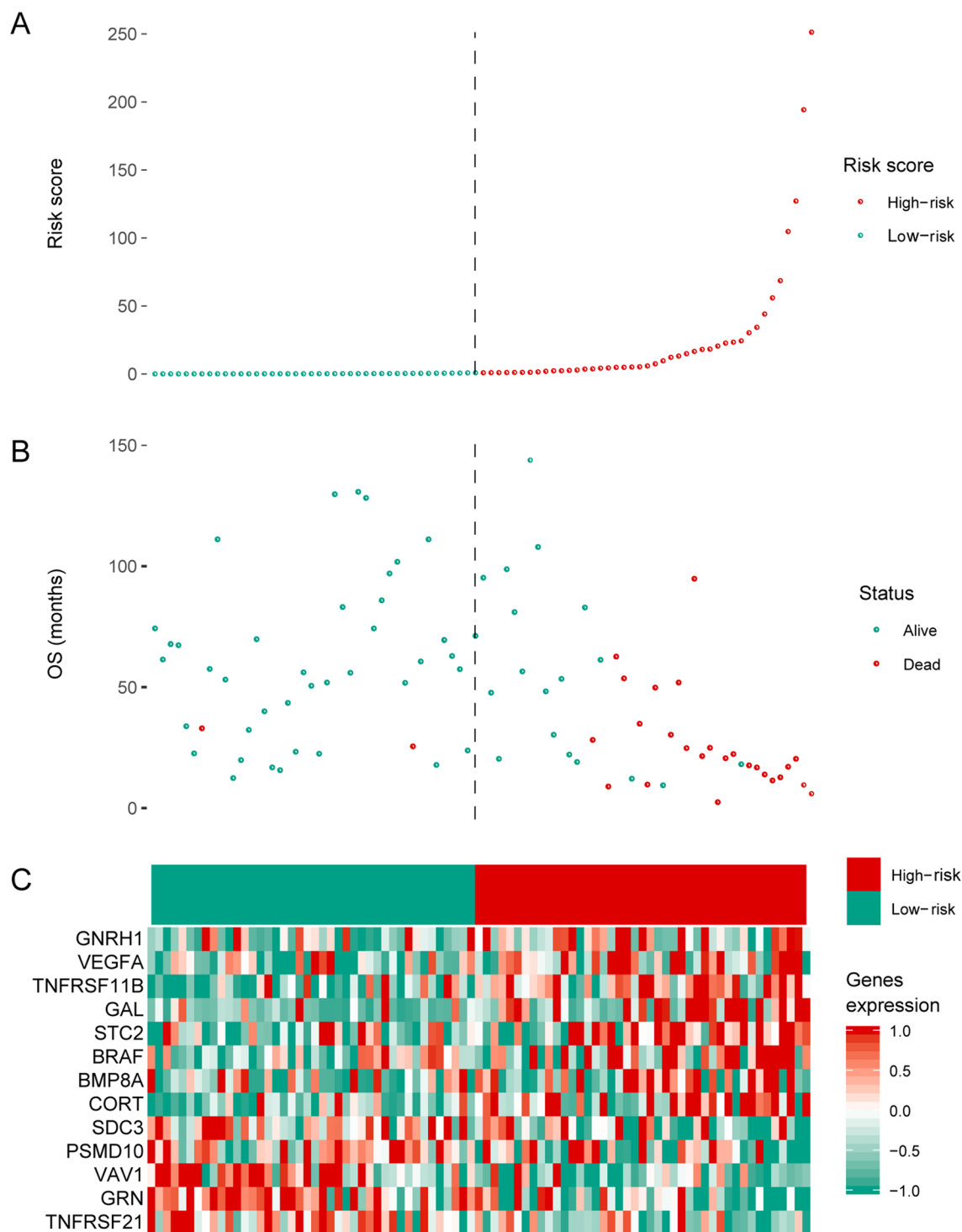


Figure 4. Prognostic analyses of high-risk and low-risk patients. **(A)** Risk score distribution of patients in the prognosis-associated IRGs signature. **(B)** Survival status scatter plots for patients in the prognosis-associated IRG signature. **(C)** Expression patterns of risk genes in the prognosis-associated IRG signature. Red means high expression, green means low expression. OS: overall survival.

OS. After all, protein is the executor of the function. Above all, there is still a lack of large sample data sets and clinical samples to verify the accuracy of the results of this analysis.

Conclusion

In summary, we developed an IRGs signature that is a prognostic indicator in OS patients, and further verified it in an independent cohort. Hence, the signature might serve as potential prognostic indicator to identify outcome of OS and facilitate personalized management of the high-risk patients.

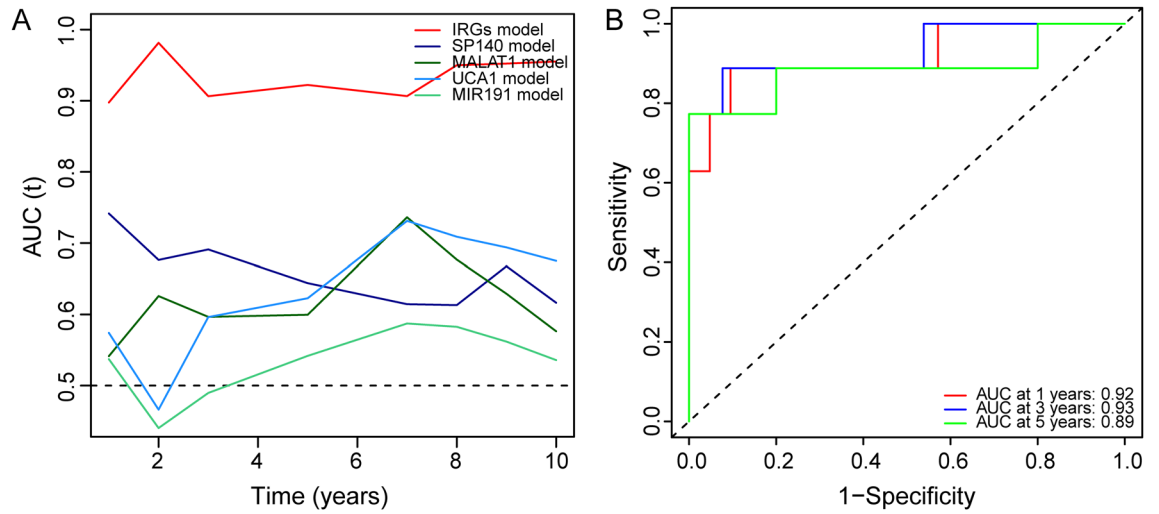


Figure 5. Comparison of IRGs signature with other prognostic biomarkers and verification in an independent cohort. (A) Time-dependent ROC curve of IRGs signature compared to that other prognostic biomarkers. (B) The ROC curve of the IRGs signature predicting survival in an independent cohort.

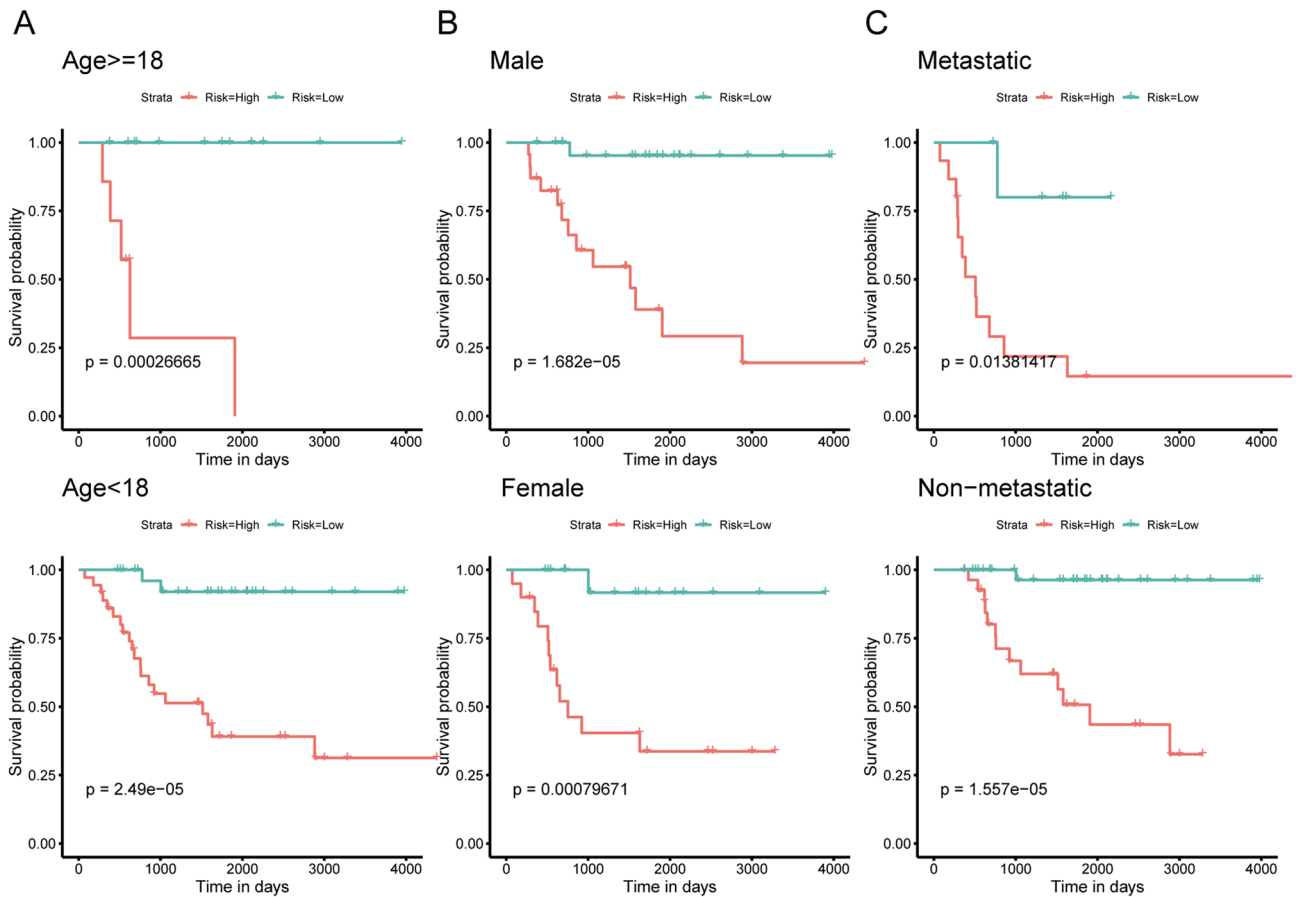


Figure 6. Kaplan–Meier curves of patients with OS in different clinical subgroups. (A) Kaplan–Meier curve for OS patients aged < 18 years and those aged ≥ 18 years. (B) Kaplan–Meier curve of male and female patients with OS. (C) Kaplan–Meier curve of metastatic and non-metastatic OS patients.

Variables	Univariate Cox		Multivariate Cox	
	HR (95% CI)	p value	HR (95% CI)	p value
Gender	0.712 (0.334–1.516)	0.377869	0.608 (0.273–1.352)	0.222178
Age	0.918 (0.347–2.423)	0.863671	2.231 (0.768–6.478)	0.140079
Metastatic	0.212 (0.099–0.454)	6.61E-05	0.182 (0.077–0.434)	0.000119
Stage	0.398 (0.120–1.326)	0.133626	0.270 (0.0727–1.004)	0.050698
Site	2.241 (1.070–4.494)	0.032372	1.669 (0.737–3.779)	0.218905
Risk score	18.088 (4.273–76.561)	8.39E-05	13.196 (3.010–57.860)	0.000624

Table 3. Univariate and multivariate Cox regression models of the IRGs signature in predicting survival.

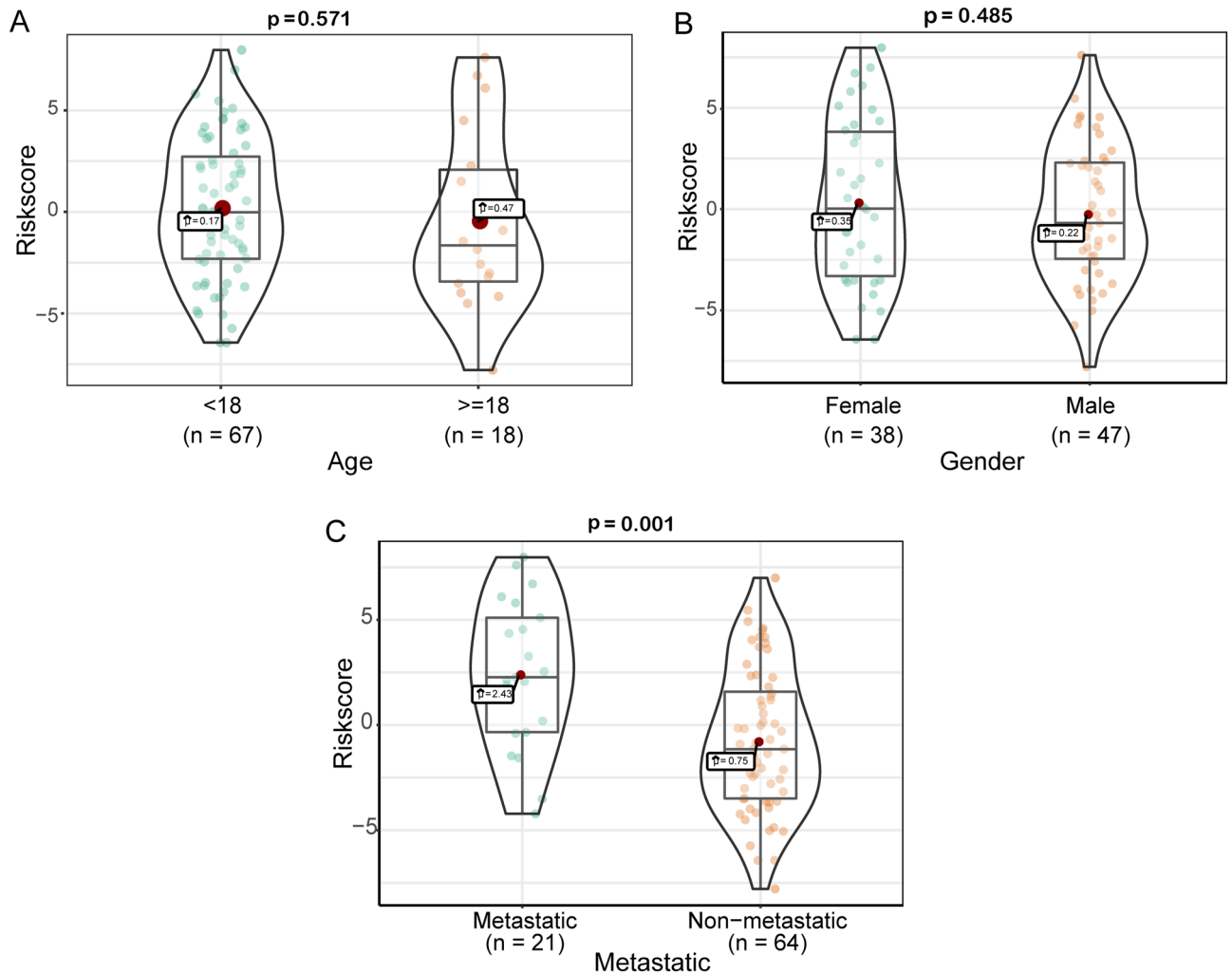


Figure 7. Correlation analysis between prognostic signature and clinical characteristics. (A) Correlation between prognosis-associated IRGs signature and age. (B) Correlation between prognosis-associated IRGs signature and sex. (C) Correlation between prognostic-associated IRGs signature and metastasis. The boxes indicate the middle 50% of the similarities; the upper and lower boundaries indicate 75% and 25%. The two ends of the violins represent the maximum and minimum values. n: number of cases of OS.

Materials and methods

Data processing and screening. Therapeutically Applicable Research to Generate Effective Treatments database (TARGET; <https://ocg.cancer.gov/programs/target>) is an open database for childhood tumors that seeks to identify molecular changes in the occurrence and progression of pediatric cancer using an integrated genomic approach to assist researchers in developing effective treatments. The Genotype-Tissue Expression (GTEx, <https://www.gtexportal.org/home/>) database⁶⁸ provides transcriptome data of various normal human tissues. Gene Expression Omnibus database (GEO, <https://www.ncbi.nlm.nih.gov/geo/>) is a gene expression

database created and maintained by NCBI, which contains high-throughput gene expression data and gene chip expression data submitted by research institutions around the world. We downloaded the gene expression profiles and the corresponding clinical data of OS from the TARGET database, including 88 OS samples, and obtained the normal muscle tissue gene expression profile data set from the GTEx database as a control group, including 396 muscle tissue samples. Then we applied the R software (Version 3.3.3, <https://www.r-project.org/>) `sva` package⁶⁹ to merge the raw data (CEL files) of the two sets. Subsequently, we used the `Limma` package⁷⁰ to screen DEGs between the OS tissue and normal muscle tissue. The cut-off value was $|\log_2$ fold change (\log_2FC) > 1 and $\text{adj. } p < 0.05$. We downloaded and organized the IRGs list from the ImmPort (<https://immport.niaid.nih.gov>) database, selected DEIRGs from DEGs and used them for our analysis.

Functional correlation analysis of DEIRGs. GO is a tool for annotating genes and their products, which aid the integration and utilization of biological data⁷¹. KEGG is a database integrating genomics, chemistry, and system function information, which provides currently known biological metabolic signaling pathways⁷². The `clusterProfiler` package⁷³ was used to perform GO and KEGG enrichment analysis on DEIRGs; $p < 0.05$ was used as a cut-off value for significant gene enrichment. The Search Tool for the Retrieval of Interacting Genes online tool (STRING, <https://www.string-db.org/>, Version: 11.0)⁷⁴ and Cytoscape software⁷⁵ were used to construct the PPI network for DEIRGs, and the hub IRGs were screened using the `cytoHubba` plug-in⁷⁶. The hub IRGs selection criteria shortlisted the top 10 DEGs through the maximum correlation standard algorithm. Based on the semantic similarity of GO terms, `GOSemSim` package⁷⁷ was used to compute closeness of the relationship between the molecular function and cell localization among 10 hub IRGs, and used the average functional similarity to rank the 10 hub IRGs⁷⁸. The results were visualized by the `ggplot2` package⁷⁹.

Identification and assessment of the prognostic signature. To develop the optimal signature for predicting OS prognosis based on IRGs, we performed univariate Cox regression analysis on the obtained DEIRGs, and selected IRGs related to prognosis with a screening criterion of $p < 0.05$. Next, we used the `glmnet` (<https://CRAN.R-project.org/package=glmnet>) package⁸⁰ to perform a machine learning algorithm-iterative LASSO Cox regression analysis on prognostic-associated IRGs to construct the optimal prognosis signature. LASSO is highly dependent on seeds and requires cross-validation to select samples randomly. Once the seeds are replaced, the optimal lambda and resulting features change. Iterative LASSO regression was used to select high-frequency features, such as consensus genes, according to the frequency sequence of features after several runs of LASSO. Then, the consensus genes were sequentially included in the Cox model. After the AUC of ROC reached a peak, the genes were not included. At this point, the model is optimal and contains the least features⁸¹. We counted the consensus genes for which the frequency exceeded 50 after 500 LASSO regressions. Then we fit the expression levels of the consensus genes into a variable through the iterative LASSO Cox regression to construct the optimal prognosis signature of OS. Next, we scored each sample with the optimal signature and divided the patients into a high- or low-risk group, according to the median of the score. Finally, we used R software to draw a risk factor association chart to display the survival status.

Comparison of signature with other known prognostic biomarkers and verification in an independent cohort. Many prognostic markers for patients with OS have been previously determined. SP140 has been identified as a promising prognostic marker for OS patients⁸, and the expression of MALAT1 has been shown to be associated with a worse prognosis for OS patients⁹. UCA1 expression may be an independent prognostic indicator for predicting a poor prognosis in patients with OS¹⁰. In addition, miR191 is highly expressed in the serum of patients with osteosarcoma and is positively correlated with clinical stage¹¹. In order to determine whether our signature has a better ability to predict patient survival than known biomarkers, we conducted a ROC comparative analysis of the signature and other biomarkers. Good prognostic markers should also have a high predictive prognostic performance in other independent cohorts. To test the utility of the signature in this study, we verified it with another independent cohort (GSE39055). Details of the GSE39055 dataset are shown in Supplementary Table 1.

Subgroup survival analysis, signature clinical value evaluation. An important feature of a good prognostic marker is that it should be independent of the currently used clinicopathological prognostic factors. To evaluate the independence and applicability of this signature, we regrouped OS patients according to different clinicopathological characteristics, and then performed Kaplan–Meier survival analysis for their subgroups. We performed univariate and multivariate Cox regressions on clinicopathological characteristics and the signature to evaluate whether the signature is an important prognostic factor.

Correlation analysis of prognostic signature and clinical characteristics. To further evaluate the correlation between the risk score based on the prognosis-associated IRGs signature and clinical characteristics, we classified patients according to age, sex, and distant metastatic status. Then we used the `ggstatsplot` (<https://github.com/IndrajeetPatil/ggstatsplot>) package to analyze the correlation between the risk score and the aforementioned. The results are shown in the `ggplot2` package.

Received: 13 May 2020; Accepted: 13 October 2020
Published online: 27 October 2020

References

- Bielack, S. S. *et al.* 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.* **20**, 776–790 (2002).
- Liu, K. *et al.* The Sp1/FOXC1/HOTTIP/LATS2/YAP/β-catenin cascade promotes malignant and metastatic progression of osteosarcoma. *Mol. Oncol.* Advance online publication; <https://doi.org/10.1002/1878-0261.12760> (2020).
- Wang, Y., Huang, Y., Xiang, P. & Tian, W. LncRNA expression and implication in osteosarcoma: A systematic review and meta-analysis. *Onco Targets Ther.* **10**, 5355–5361 (2017).
- Chen, L. *et al.* Mild microwave ablation combined with HSP90 and TGF-β1 inhibitors enhances the therapeutic effect on osteosarcoma. *Mol. Med. Rep.* **22**, 906–914 (2020).
- Shankar, G. M. *et al.* The role of revision surgery and adjuvant therapy following subtotal resection of osteosarcoma of the spine: A systematic review with meta-analysis. *J. Neurosurg. Spine.* **27**, 97–104 (2017).
- Zhang, M. & Zhang, X. Association of MMP-2 expression and prognosis in osteosarcoma patients. *Int. J. Clin. Exp. Pathol.* **8**, 14965–14970 (2015).
- Rosemann, M. *et al.* A Rb1 promoter variant with reduced activity contributes to osteosarcoma susceptibility in irradiated mice. *Mol. Cancer.* **13**, 182 (2014).
- Hong, W. *et al.* Immune-related prognosis biomarkers associated with osteosarcoma microenvironment. *Cancer Cell Int.* **20**, 83 (2020).
- Wang, J. & Sun, G. FOXO1-MALAT1-miR-26a-5p feedback loop mediates proliferation and migration in osteosarcoma cells. *Oncol. Res.* **25**, 1517–1527 (2017).
- Wen, J. J., Ma, Y. D., Yang, G. S. & Wang, G. M. Analysis of circulating long non-coding RNA UCA1 as potential biomarkers for diagnosis and prognosis of osteosarcoma. *Eur. Rev. Med. Pharmacol. Sci.* **21**, 498–503 (2017).
- Wang, T., Ji, F., Dai, Z., Xie, Y. & Yuan, D. Increased expression of microRNA-191 as a potential serum biomarker for diagnosis and prognosis in human osteosarcoma. *Cancer Biomark.* **15**, 543–550 (2015).
- Sanchez-Diaz, P. C. *et al.* In silico functional analyses and discovery of survival-associated microRNA signatures in pediatric osteosarcoma. *Oncoscience.* **1**, 599–608 (2014).
- Tuohy, J. L. *et al.* Immune dysregulation and osteosarcoma: *Staphylococcus aureus* downregulates TGF-β and heightens the inflammatory signature in human and canine macrophages suppressed by osteosarcoma. *Vet. Comp. Oncol.* **18**, 64–75 (2020).
- Gentles, A. J. *et al.* The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat. Med.* **21**, 938–945 (2015).
- Shen, H. *et al.* Predictive biomarkers for immune checkpoint blockade and opportunities for combination therapies. *Genes Dis.* **6**, 232–246 (2019).
- Topalian, S. L., Drake, C. G. & Pardoll, D. M. Immune checkpoint blockade: A common denominator approach to cancer therapy. *Cancer Cell* **27**, 450–461 (2015).
- Wu, W. *et al.* FGD1 promotes tumor progression and regulates tumor immune response in osteosarcoma via inhibiting PTEN activity. *Theranostics.* **10**, 2859–2871 (2020).
- Anderson, M. E. Update on survival in osteosarcoma. *Orthop. Clin. N. Am.* **47**, 283–292 (2016).
- Kim, R., Emi, M. & Tanabe, K. Cancer immunoediting from immune surveillance to immune escape. *Immunology* **121**, 1–14 (2007).
- Heymann, M. F., Lézet, F. & Heymann, D. The contribution of immune infiltrates and the local microenvironment in the pathogenesis of osteosarcoma. *Cell Immunol.* **343**, 103711 (2019).
- Lauri, D., Needham, L., Martin-Padura, I. & Dejana, E. Tumor cell adhesion to endothelial cells: Endothelial leukocyte adhesion molecule-1 as an inducible adhesive receptor specific for colon carcinoma cells. *J. Natl. Cancer Inst.* **83**, 1321–1324 (1991).
- Byrum, M. L., Pondenis, H. C., Fredrickson, R. L., Wycislo, K. L. & Fan, T. M. Downregulation of CXCR4 expression and functionality after zoledronate exposure in canine osteosarcoma. *J. Vet. Intern. Med.* **30**, 1187–1196 (2016).
- Piazzi, M. *et al.* Expression of the double-stranded RNA-dependent kinase PKR influences osteosarcoma attachment independent growth, migration, and invasion. *J. Cell Physiol.* **235**, 1103–1119 (2020).
- Zhao, G. S. *et al.* TSSC3 promotes autophagy via inactivating the Src-mediated PI3K/Akt/mTOR pathway to suppress tumorigenesis and metastasis in osteosarcoma, and predicts a favorable prognosis. *J. Exp. Clin. Cancer Res.* **37**, 188 (2018).
- Zhu, J. *et al.* Escin induces caspase-dependent apoptosis and autophagy through the ROS/p38 MAPK signalling pathway in human osteosarcoma cells in vitro and in vivo. *Cell Death Dis.* **8**, e3113. <https://doi.org/10.1038/cddis.2017.488> (2017).
- Liu, Y., Guan, J. & Chen, X. Identification of differentially expressed genes under the regulation of transcription factors in osteosarcoma. *Pathol. Oncol. Res.* **25**, 1091–1102 (2019).
- Kubista, B., Trieb, K., Blahovec, H., Kotz, R. & Micksche, M. Hyperthermia increases the susceptibility of chondro- and osteosarcoma cells to natural killer cell-mediated lysis. *Anticancer Res.* **22**, 789–792 (2002).
- Tang, Y., Gu, Z., Fu, Y. & Wang, J. CXCR3 from chemokine receptor family correlates with immune infiltration and predicts poor survival in osteosarcoma. *Biosci. Rep.* **39**, 11. <https://doi.org/10.1042/BSR20192134> (2019).
- Jiang, K. *et al.* SDF-1/CXCR4 axis facilitates myeloid-derived suppressor cells accumulation in osteosarcoma microenvironment and blunts the response to anti-PD-1 therapy. *Int. Immunopharmacol.* **75**, 105818. <https://doi.org/10.1016/j.intimp.2019.105818> (2019).
- Wang, S. W. *et al.* CCL5/CCR5 axis induces vascular endothelial growth factor-mediated tumor angiogenesis in human osteosarcoma microenvironment. *Carcinogenesis* **36**, 104–114 (2015).
- Sun, K. *et al.* High CCL5 expression is associated with osteosarcoma metastasis and poor prognosis of patients with osteosarcoma. *Mol. Med. Rep.* **16**, 6953–6957 (2017).
- Zhang, F., Huang, W., Sheng, M. & Liu, T. MiR-451 inhibits cell growth and invasion by targeting CXCL16 and is associated with prognosis of osteosarcoma patients. *Tumour Biol.* **36**, 2041–2048 (2015).
- Flores, R. J. *et al.* A novel prognostic model for osteosarcoma using circulating CXCL10 and FLT3LG. *Cancer* **123**, 144–154 (2017).
- Pollino, S. *et al.* CXCR4 in human osteosarcoma malignant progression. The response of osteosarcoma cell lines to the fully human CXCR4 antibody MDX1338. *J. Bone Oncol.* **17**, 100239 (2019).
- Miao, J. H. *et al.* Knockdown of galectin-1 suppresses the growth and invasion of osteosarcoma cells through inhibition of the MAPK/ERK pathway. *Oncol. Rep.* **32**, 1497–1504 (2014).
- Toll, L., Bruchas, M. R., Caló, G., Cox, B. M. & Zaveri, N. T. Nociceptin/orphanin FQ receptor structure, signaling, ligands, functions, and interactions with opioid systems. *Pharmacol. Rev.* **68**, 419–457 (2016).
- Kiguchi, N., Ding, H. & Ko, M. C. Central N/OFQ-NOP Receptor System in Pain Modulation. *Adv Pharmacol.* **75**, 217–243 (2016).
- Rostami, N. *et al.* S1PR1 as a novel promising therapeutic target in cancer therapy. *Mol. Diagn. Ther.* **23**, 467–487 (2019).
- Nagahashi, M. *et al.* Sphingosine-1-phosphate in chronic intestinal inflammation and cancer. *Adv. Biol. Regul.* **54**, 112–120 (2014).
- Dai, L. *et al.* Sphingosine kinase 1/sphingosine-1-phosphate (S1P)/S1P receptor axis is involved in ovarian cancer angiogenesis. *Oncotarget.* **8**, 74947–74961 (2017).
- Wang, S. W. *et al.* CCL5 and CCR5 interaction promotes cell motility in human osteosarcoma. *PLoS ONE* **7**, e35101. <https://doi.org/10.1371/journal.pone.0035101> (2012).

42. Neklyudova, O. *et al.* Altered CXCL12 expression reveals a dual role of CXCR4 in osteosarcoma primary tumor growth and metastasis. *J. Cancer Res. Clin. Oncol.* **142**, 1739–1750 (2016).
43. Miwa, S. *et al.* Therapeutic targets for bone and soft-tissue sarcomas. *Int. J. Mol. Sci.* **20**, 170. <https://doi.org/10.3390/ijms20010170> (2019).
44. Koirala, P. *et al.* Immune infiltration and PD-L1 expression in the tumor microenvironment are prognostic in osteosarcoma. *Sci. Rep.* **6**, 30093. <https://doi.org/10.1038/srep30093> (2016).
45. Li, B. *et al.* Epigenetic regulation of CXCL12 plays a critical role in mediating tumor progression and the immune response in osteosarcoma. *Cancer Res.* **78**, 3938–3953 (2018).
46. Wu, J., Zhao, Y., Zhang, J., Wu, Q. & Wang, W. Development and validation of an immune-related gene pairs signature in colorectal cancer. *Oncoimmunology*. **8**, 1596715. <https://doi.org/10.1080/2162402X.2019.1596715> (2019).
47. Lin, P. *et al.* Development of a prognostic index based on an immunogenomic landscape analysis of papillary thyroid cancer. *Aging* **11**, 480–500 (2019).
48. Tan, W. *et al.* Construction of an immune-related genes nomogram for the preoperative prediction of axillary lymph node metastasis in triple-negative breast cancer. *Artif. Cells Nanomed. Biotechnol.* **48**, 288–297 (2020).
49. von Alten, J. *et al.* GnRH analogs reduce invasiveness of human breast cancer cells. *Breast Cancer Res. Treat.* **100**, 13–21 (2006).
50. Kim, B. *et al.* Clinical meaning of BRAF mutation in Korean patients with advanced colorectal cancer. *World J. Gastroenterol.* **20**, 4370–4376 (2014).
51. Cheng, L. *et al.* Gankyrin promotes osteosarcoma tumorigenesis by forming a positive feedback loop with YAP. *Cell Signal.* **65**, 109460. <https://doi.org/10.1016/j.cellsig.2019.109460> (2020).
52. Wu, H., Zhang, J., Dai, R., Xu, J. & Feng, H. Transferrin receptor-1 and VEGF are prognostic factors for osteosarcoma. *J. Orthop. Surg. Res.* **14**, 296. <https://doi.org/10.1186/s13018-019-1301-z> (2019).
53. Shih, T. C. *et al.* Targeting galectin-1 impairs castration-resistant prostate cancer progression and invasion. *Clin. Cancer Res.* **24**, 4319–4331 (2018).
54. Nagayoshi, K. *et al.* Galanin plays an important role in cancer invasiveness and is associated with poor prognosis in stage II colorectal cancer. *Oncol. Rep.* **33**, 539–546 (2015).
55. Marziani, D. *et al.* Expression of basic fibroblast growth factor, its receptors and syndecans in bladder cancer. *Int. J. Immunopathol. Pharmacol.* **22**, 627–638 (2009).
56. Ieta, K. *et al.* Clinicopathological significance of stanniocalcin 2 gene expression in colorectal cancer. *Int. J. Cancer.* **125**, 926–931 (2009).
57. Liu, K., Gao, M., Qin, D., Wang, H. & Lu, Q. Serous BMP8A has clinical significance in ultrasonic diagnosis of thyroid cancer and promotes thyroid cancer cell progression. *Endocr. Metab. Immune Disord. Drug Targets.* **20**, 591–598 (2019).
58. Cassoni, P. *et al.* Ghrelin and cortistatin in lung cancer: Expression of peptides and related receptors in human primary tumors and in vitro effect on the H345 small cell carcinoma cell line. *J. Endocrinol. Investig.* **29**, 781–790 (2006).
59. Cassoni, P. *et al.* Cortistatin-14 inhibits cell proliferation of human thyroid carcinoma cell lines of both follicular and parafollicular origin. *J. Endocrinol. Investig.* **25**, 362–368 (2002).
60. Delgado-Maroto, V. *et al.* The neuropeptide cortistatin attenuates experimental autoimmune myocarditis via inhibition of cardiomyogenic T cell-driven inflammatory responses. *Br. J. Pharmacol.* **174**, 267–280 (2017).
61. Chen, X. *et al.* Interaction between granulin A and enolase 1 attenuates the migration and invasion of human hepatoma cells. *Oncotarget.* **8**, 30305–30316 (2017).
62. Grassilli, S. *et al.* High nuclear level of Vav1 is a positive prognostic factor in early invasive breast tumors: A role in modulating genes related to the efficiency of metastatic process. *Oncotarget.* **5**, 4320–4336 (2014).
63. Zong, F. *et al.* Effect of syndecan-1 overexpression on mesenchymal tumour cell proliferation with focus on different functional domains. *Cell Prolif.* **43**, 29–40 (2010).
64. Wu, H. *et al.* Upregulated miR20a5p expression promotes proliferation and invasion of head and neck squamous cell carcinoma cells by targeting of TNFRSF21. *Oncol. Rep.* **40**, 1138–1146 (2018).
65. Schmidt, C. S. *et al.* Resistance to myelin oligodendrocyte glycoprotein-induced experimental autoimmune encephalomyelitis by death receptor 6-deficient mice. *J. Immunol.* **175**, 2286–2292 (2005).
66. Matin, A. & Nadeau, J. H. Sensitized polygenic trait analysis. *Trends Genet.* **17**, 727–731 (2001).
67. Shi, Y. *et al.* A risk signature-based on metastasis-associated genes to predict survival of patients with osteosarcoma. *J. Cell Biochem.* **121**, 3479–3490 (2020).
68. multitissue gene regulation in humans. Human genomics. the genotype-tissue expression (GTEx) pilot analysis. *Science* **348**, 648–660 (2015).
69. Parker, H. S. *et al.* Preserving biological heterogeneity with a permuted surrogate variable analysis for genomics batch correction. *Bioinformatics* **30**, 2757–2763 (2014).
70. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47. <https://doi.org/10.1093/nar/gkv007> (2015).
71. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: An R package for comparing biological themes among gene clusters. *OMICS.* **16**, 284–287 (2012).
72. Harris, M. A. *et al.* The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* **32**, D258–D261 (2004).
73. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: New perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).
74. Szklarczyk, D. *et al.* STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).
75. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P. L. & Ideker, T. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).
76. Chin, C. H. *et al.* cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **8**(Suppl 4), S11 (2014).
77. Yu, G. *et al.* GOSemSim: An R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics* **26**, 976–978 (2010).
78. Sevilla, J. L. *et al.* Correlation between gene expression and GO semantic similarity. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2**, 330–338 (2005).
79. Ito, K. & Murphy, D. Application of ggplot2 to pharmacometric graphics. *CPT Pharmacometrics Syst. Pharmacol.* **2**, e79. <https://doi.org/10.1038/psp.2013.56> (2013).
80. Engebretsen, S. & Bohlin, J. Statistical predictions with glmnet. *Clin. Epigenetics.* **11**, 123. <https://doi.org/10.1186/s13148-019-0730-1> (2019).
81. Tibshirani, R. The lasso method for variable selection in the Cox model. *Stat. Med.* **16**, 385–395 (1997).

Acknowledgements

This work benefits from TARGET and GTEx database. We thank the contributors who uploaded the data in the TARGET and GTEx. We thank FigureYa (Xiao Ya Hua Tu) for the figure technology support. We also thank Editage (<https://www.editage.cn>) for English language editing.

Author contributions

Conceptualization, Z.-L.W., G.-Z.Z.; methodology, W.-H.Y.; software, E.-H.R.; writing—original draft preparation, Z.-L.W.; writing—review and editing, Y.-J.D. and Q.-Q.X.; visualization, Y.-J.D.; supervision, Q.-Q.X.; All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by The Key Research & Development and Transformation Project of Qinghai Province for 2018 (2018-SF-113).

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-75573-w>.

Correspondence and requests for materials should be addressed to Q.X.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020