Construction and validation of an oxidative-stress-related risk model for predicting the prognosis of osteosarcoma

Hanning Wang[1,*], Juntan Li[1,*], Xu Li[1,&]

[1]Department of Orthopedics, The First Hospital of China Medical University, Shenyang, Liaoning Province 110001, P.R. China
[*]Equal contribution

ABSTRACT

Background: Osteosarcoma is the most common bone malignancy in teenagers, and warrants effective measures for diagnosis and prognosis. Oxidative stress (OS) is the key driver of several cancers and other diseases.
Methods: The TARGET-osteosarcoma database was employed as the training cohort and GSE21257 and GSE39055 was applied for external validation. The patients were classified into the high- and low-risk groups based on the median risk score of each sample. ESTIMATE and CIBERSORT were applied for the evaluation of tumor microenvironment immune infiltration. GSE162454 of single-cell sequencing was employed for analyzing OS-related genes.
Results: Based on the gene expression and clinical data of 86 osteosarcoma patients in the TARGET database, we identified eight OS-related genes, including MAP3K5, G6PD, HMOX1, ATF4, ACADVL, MAPK1, MAPK10, and INS. In both the training and validation sets, the overall survival of patients in the high-risk group was significantly worse than that in the low-risk group. The ESTIMATE algorithm revealed that patients in the high-risk group had higher tumor purity but lower immune score and stromal score. In addition, the CIBERSORT algorithm showed that the M0 and M2 macrophages were the predominant infiltrating cells in osteosarcoma. Based on the expression analysis of immune checkpoint, CD274(PDL1), CXCL12, BTN3A1, LAG3, and IL10 were identified as potential immune therapy targets. Analysis of the single cell sequencing data also revealed the expression patterns of OS-related genes in different cell types.
Conclusions: An OS-related prognostic model can accurately provide the prognosis of osteosarcoma patients, and may help identify suitable candidates for immunotherapy.

INTRODUCTION

Osteosarcoma is the most prevalent bone malignancy among adolescents, with a high incidence rate of 8–11/1,000,000/year in the 15–19 years age group [1, 2]. It primarily affects the long bones, including the femur, tibia, and humerus [3]. The second peak of osteosarcoma occurrence is after the age of 50 years [4]. Most patients with osteosarcoma have lung metastases when first diagnosed, and their 5-year survival rate is less than 20%

[5]. However, the 5-year survival rate of osteosarcoma patients without lung metastases is almost 70% [6]. Despite treatment options such as surgery and chemotherapy, the 5-year survival rates are overall dismal because of the genetic complexity and instability of these tumors [7, 8]. High-throughput technologies have been used in recent oxidative stress (OS) studies, which have revealed differentially expressed genes (DEGs) such as COL1A2 and matrix metalloproteases (MMPs) between osteosarcoma and normal samples [9, 10]. These sequence

data provide a method for early diagnosis and treatment, as well as enable accurate prognosis to improve the outcomes of osteosarcoma patients.

OS is characterized by excessive production of reactive oxygen species (ROS) that cannot be effectively quenched by the cellular antioxidative mechanisms [11, 12]. OS is the underlying pathological basis of degenerative diseases, such as Alzheimer's disease [13], diabetes [14], and arthritis [15], and cancers, including bladder cancer [16] and breast cancer [17]. Cancer cells are characterized by increased aerobic glycolysis and high levels of ROS [18]. While the high metabolic rates promote cellular migration, survival, and proliferation, ROS-induced oxidative DNA damage initiates oncogenic transformation and subsequent tumor progression. Given the metabolic and signaling aberrations caused by increased ROS levels, the pathways involved in ROS production are promising therapeutic targets for cancer [19]. An increase in oxidative stress and decrease in antioxidant status is observed in primary bone sarcomas [20]. Several studies have also revealed the anti-cancer potential of differential ROS production pathways for treating osteosarcoma *in vitro* and *in vivo* [21–23].

The human immune system has evolved to eradicate pathogens and tumor cells [24], mainly through effector cells such as basophils, macrophages, neutrophils, eosinophils, and natural killer (NK) cells [25]. In contrast, the humoral immune system is less involved in anti-tumor responses, and may even promote tumor growth under certain conditions. The tumor microenvironment (TME) is a complex network of tumor cells, stromal cells, and immune cells. The non-tumor cell populations also have diverse and complex regulatory effects on tumor progression. For instance, an abundance of cancer-associated fibroblasts (CAFs) in the TME is conducive to tumor growth [26], whereas high levels of tumor-associated macrophages (TAMs) are associated with lower metastasis and improved survival in osteosarcoma [27]. Therefore, the role of TME in osteosarcoma prognosis warrants further investigation.

To this end, we retrieved the clinical and transcriptomic data of osteosarcoma patients from the Therapeutically Applicable Research to Generate Effective Treatments (TARGET) database. The OS-related genes were screened using univariate Cox regression, and a prognostic model was constructed following LASSO and multivariable Cox regression analyses. The accuracy and generalizability of the risk model was validated in an external cohort. The correlation between the risk score and TME was evaluated, and a predictive nomogram was constructed by combining the clinical characteristics and risk score. The flow chart of the study is shown in Figure 1.
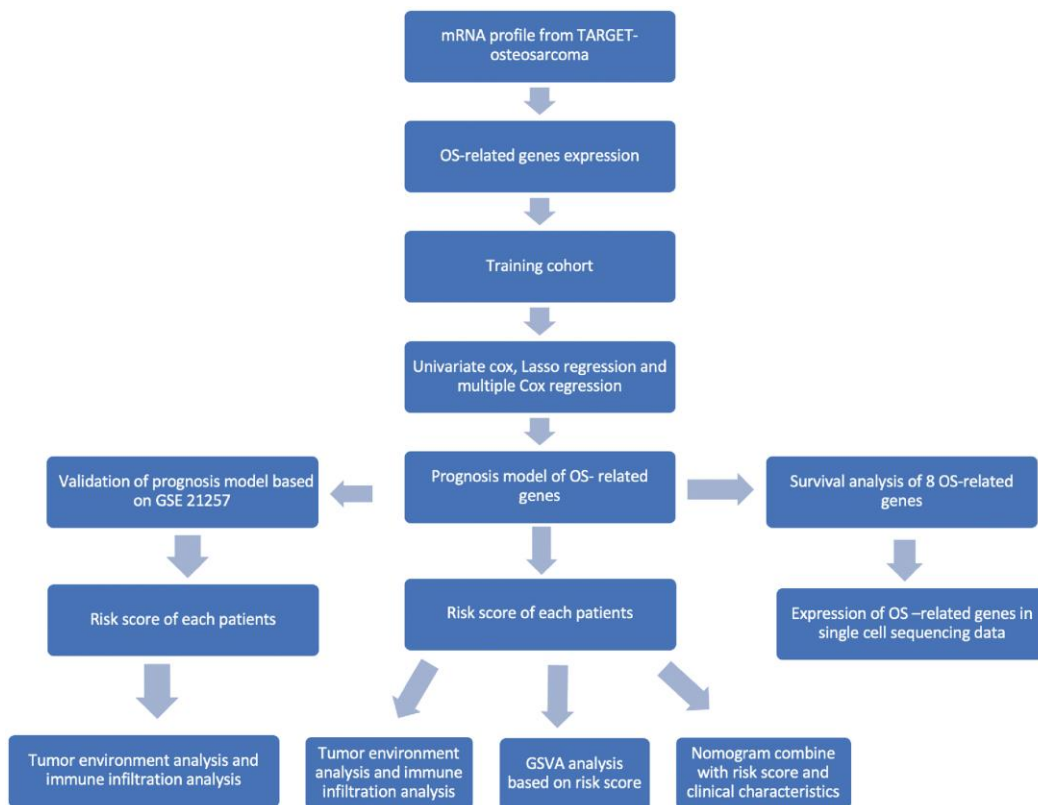

Figure 1. Flowchart of this study.

MATERIALS AND METHODS

Data sources

The mRNA sequencing data and clinical information 86 osteosarcoma patients were retrieved from the TARGET database (https://ocg.cancer.gov/programs/target) to construct an OS-related risk model. A total of 817 OS-related genes were obtained from GeneCards database (https://www.genecards.org/). In addition, the GSE21257 and GSE39055 datasets including 90 osteosarcoma samples was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/) to validate the risk model. Normalization of fragments per kilobase of exon model per million mapped fragments (FPKM) was performed in the TARGET-OS dataset and Limma package was used to normalize gene expression [28]. The R package sva was used to eliminate batch effect of different databases [29]. The clinical data of patients in the TARGET-OS and validation cohorts are summarized in Table 1. The GSE162454 dataset containing the single-cell sequencing data of six osteosarcoma samples was also retrieved from the GEO database [30].

Construction of prognostic model based on OS-related genes

The potential OS-related genes were identified in the TARGET dataset (training cohort) through univariate Cox regression analysis. The GLMNET package [31] from R was then used to perform the least absolute shrinkage and selection operator (LASSO) analysis to remove genes that may lead to overfitting [31, 32]. The seed number for Lasso regression is set to 666, and the penalty coefficient is 0.0868. Finally, the prognostically significant OS-related genes were screened by multivariable Cox regression analysis. The coefficients, hazard ratios (HR), and p values of different genes are listed in Table 2. The risk score of each sample from TARGET-OS was calculated using the Predict algorithm, and the patients were divided into high-risk and low-risk groups on the basis of the median risk score. The Kaplan–Meier method was used to evaluate the difference in the survival status of both risk groups, and receiver operating characteristic (ROC) curves for 1-, 3-, and 5-year survival were plotted using the survival ROC package in R to identify the diagnostic value of the risk model [33].

Validation of OS-related risk model

The performance of the OS-related model was validated in GSE21257 and GSE39055. The risk score for the validation cohort was calculated using the same algorithm as that used for the training cohort, and the samples were divided into the high- and low-risk groups on the basis of the median score for the training cohort. The efficacy of the risk model was evaluated by the Kaplan–Meier method and ROC curve analysis as described.

Gene set enrichment analysis

The DEGs between the high- and low-risk groups in the TARGET cohort were screened using the Limma package. The significantly enriched gene ontology pathways, namely biological processes (BPs), cellular components (CCs), and molecular functions (MFs), associated with DEGs were identified by gene set variation analysis (GSVA) packages in R software with |log2FC|>1 and adj. P value < 0.05 as thresholds [34].

Immune analysis based on OS-related genes

The Estimation of Stromal and Immune cells in Malignant Tumor tissues (ESTIMATE) package in R software was used to calculate tumor purity, stromal score, and immune score in the two risk groups [35]. The CIBERSORT algorithm was used to calculate the proportion of 22 infiltrating immune cell populations in the training and verification cohorts [36] and in the high- and low-risk groups. Some potential immune checkpoints such as PD1 (PDCD1), PDL1 (CD274), IL10, CTL4, BTN3A1, CXCL12, and LAG3, and calculated the CYT score (GZMA and PRF1) were displayed using boxplots to illustrate their expression across high- and low-risk groups, revealing potential immune therapeutic targets. The gene mutation data was obtained from the TCGA database, and the maftools package was used to visualize the mutation data.

Construction of a predictive nomogram based on the risk model

The prognostic model was integrated with the clinical characteristics including gender, race, metastasis, and risk score to build a nomogram using the Survminer package in R. The predictive accuracy of the nomogram was validated by C-index and comparison of the fitting degree between the observed and optimized values was performed.

OS-related gene expression in single-cell sequencing database

The expression of OS-related genes at the single cell level was analyzed in the GSE162454 dataset. Six osteosarcoma samples were analyzed using the Seurat package (Version 4.0) [37]. Poor quality cells (genes >8000, genes <500, or >10% genes mapping to

Table 1. Clinical information of training and validation cohorts.

	TARGET-OS	GSE21257	GSE39055
Samples	86	52	38
Age	14.5	16.7	13.5
Female percent	43.0%	36.5%	44.8%
Metastasis percent	25.6%	64.1%	Na
Survival time	4.11	5.71	4.41

Table 2. Eight genes obtained after multivariable Cox regression analysis.

Gene	Coefficient	HR	HR.95L	HR.95H	*P* value
MAP3K5	−1.817	0.162	0.057	0.465	0.0007
G6PD	−0.662	0.516	0.228	1.167	0.1120
HMOX1	−0.523	0.593	0.408	0.861	0.0061
ATF4	0.8396	2.315	1.130	4.746	0.0218
ACADVL	1.4825	4.404	2.073	9.357	0.0001
MAPK1	−1.135	0.321	0.138	0.746	0.0083
INS	32.215	9.79E+13	1295.499	7.41E+24	0.0117
MAPK10	−1.540	0.214	0.057	0.795	0.0213

mitochondrial genome) were excluded. The expression data were transformed to the log scale and normalized for scaling the sequencing depth. T-Stochastic neighbor Embedding (T-SNE) plots were used to visualize cell clusters.

Immunohistochemistry and hematoxylin-eosin staining

OS-related genes were experimentally verified by immunohistochemistry (IHC). Three osteosarcoma tissues and peri-tumor tissues were collected from patients at First Affiliated Hospital of China Medical University. A consent form was signed by each participant. The present study was approved by the ethics committee of the hospital with the registration number 2019-285-2.

Hematoxylin-eosin (HE) staining was performed to identify histology type of tissues. Tissues were fixed and paraffinized before IHC staining. Antigens were retrieved using citrate antigen retrieval buffer. Endogenous peroxidase was blocked by 3% hydrogen peroxide solution, incubated at room temperature for 25 min in the dark. In the histochemical area, 3% BSA (Bovine Serum Albumin, Servicebio) was added dropwise to cover the tissue evenly, and it was sealed at room temperature for 30 min. Sections were incubated overnight at 4°C in a humidified chamber with corresponding antibodies. Antibody applications were as follows: ACADVL (Proteintech, 1:200), ATF4 (Servicebio, 1:2000), HMOX1 (Servicebio, 1:500), MAPK10 (Servicebio, 1:200), INS (Servicebio, 1:200). Secondary antibody (HRP-labeled) corresponding to the primary antibody

was added to cover the tissue area and slides were incubated at room temperature for 50 minutes. Freshly prepared DAB (3,3′-Diaminobenzidine) chromogenic were used for color development in the IHC area. Nuclei were counterstained with hematoxylin for about 3 min. Finally, microscopic examination and image acquisition analysis were performed.

Statistical analysis

Statistical analyses were performed using the R software (Version 4.1.3). Two-tailed Student's *t*-test and Wilcox test were performed to compare gene expression or scores between two independent groups. Kaplan–Meier survival analysis, LASSO regression analysis, and multivariable Cox regression analysis were performed to identify prognostic genes. Forest plots of nomogram were visualized using forest algorithm. *P* value < 0.05 was considered statistically significant.

Availability of data and materials

The source data and R code of the manuscript was uploaded in OSF (https://doi.org/10.17605/OSF.IO/238HP).

RESULTS

**Identification of OS-related genes in osteosarcoma**

A total of 817 OS-related genes were identified from the GeneCards database, and their expression matrix was extracted from the TARGET database. Thereafter, 89 OS-related genes were identified by univariate cox

regression analysis, using $p < 0.05$ as the cut off (Supplementary Table 1); of these, 12 genes were screened on the basis of their predictive value using LASSO regression (Supplementary Figure 1 and Supplementary Table 2).

Construction of an OS-related prognostic model

A multivariable Cox regression model was constructed, and eight prognostic OS-related genes were obtained using the stepwise regression function (Table 2). The risk score of each patient was calculated, and the patients were stratified into high-risk and low-risk groups depending on the median score. The patients in the low-risk group showed better overall survival compared with those in the high-risk group (Figure 2A, $p < 0001$). The time-dependent ROC curve was then plotted to evaluate the predictive efficiency of the prognostic model. The areas under curve (AUCs) for 1-, 3-, and 5-year survival were 0.89, 0.91, and 0.91,
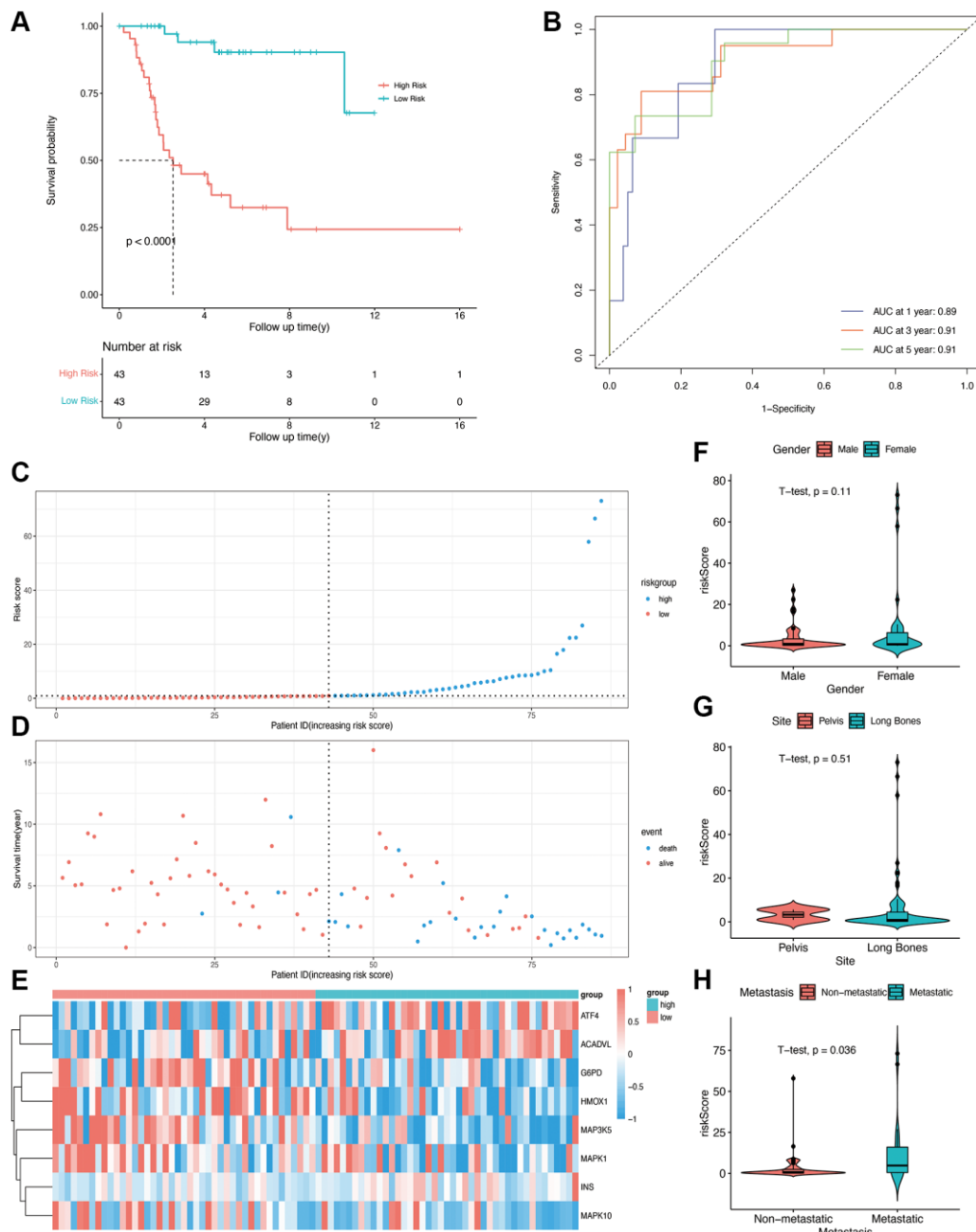
Figure 2. Construction of OS-related prognostic model for osteosarcoma in the TARGET cohort. (**A**) Kaplan-Meier curves showing the overall survival in the high-risk and low-risk groups ($p < 0.0001$). (**B**) AUCs for 1-, 3-, and 5-year survival according to the ROC curves. (**C**) The risk score curve of each patient. (**D**) The distribution of survival status. (**E**) The heatmap of eight OS-related genes between the high- and low-risk groups. (**F**, **G**, **H**) Risk scores in the gender, site, and metastasis subgroups.

respectively (Figure 2B), suggesting that the model can accurately predict the overall survival of osteosarcoma patients. The risk score distribution of patients in the high- and low-risk groups is shown in Figure 2C and 2D, and the expression levels of OS-related genes in each patient are shown in the form of a heatmap in Figure 2E. We calculated the risk score in different subgroups, and did not observe any significant effect of gender or the site of osteosarcoma on risk scores (Figure 2F and 2G). However, patients with metastasis had a higher risk score than those without metastasis ($p = 0.036$, Figure 2H).

Validation of prognostic model

To validate the predictive ability of the OS-related prognostic model, the independent dataset GSE21257 and GSE39055 were used for the validation of our risk

model. The patients in the validation cohort were also grouped into the high- and low-risk groups based on the median score in the training cohort. As shown in Figure 3A, and 3B, compared with low-risk group, the high-risk group had lower survival time. The K-M plot of validation group was shown in Figure 3, there was significant difference in high and low risk group ($p = 0.007$). The areas under the curve for prediction of 1-, 3-, and 5-years survival of patients were 0.63, 0.66 and 0.65, respectively (Figure 3D).

Survival analysis and functional enrichment of the prognostic OS genes

Based on the expression level of each gene in the prognostic model, the patients were divided into high- and low-expression groups. The survival curves corresponding to the genes are shown in Figure 4A-4H .
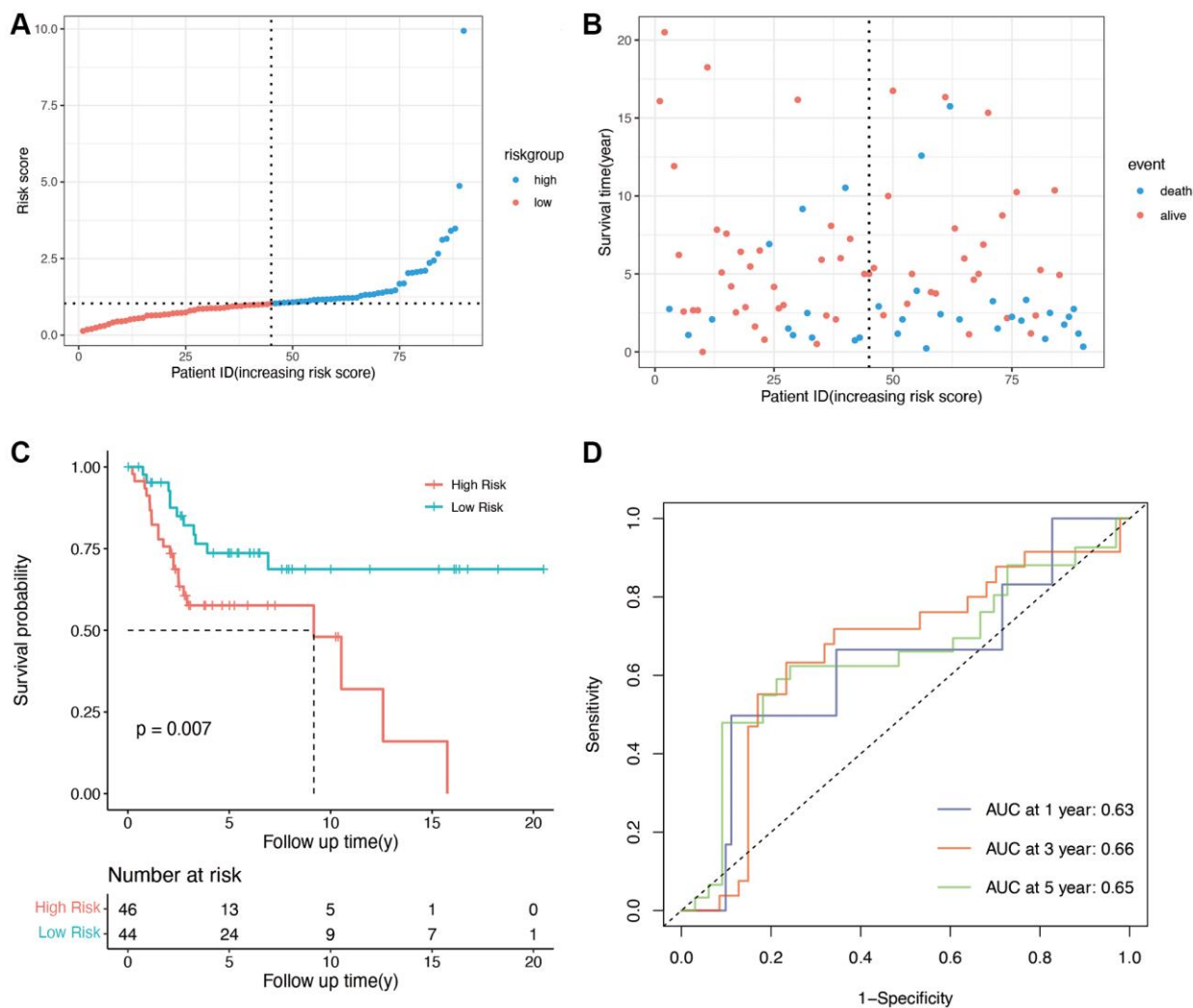
Figure 3. Validation of OS-related genes in the GSE21257 and GSE39055 cohort. (**A**) Validation group was divided into high- and low-risk groups using the median risk score as the cutoff value. (**B**) the relationship between risk score and survival time and status of patients. (**C**) K-M plot of high and low risk groups in validation group ($p = 0.007$). (**D**) AUCs for 1-, 3-, and 5-year survival according to ROC curves.

Low expression levels of ACADVL ($p = 0.028$), ATF4 ($p = 0.00047$), and INS ($p = 0.027$) were associated with improved survival, whereas high expression of HMOX1 ($p = 0.0033$) and MAPK10 ($p = 0.036$) showed better prognosis. To further explore the underlying molecular mechanisms of the prognostic model, we identified 106 differentially enriched GO terms between the high- and low-risk groups in the TARGET database with |log2FC|>1 and adj. p value < 0.05 as the cut-offs. The top 50 enriched GO terms obtained from GSVA are shown in Figure 4I, and these included many immune-related terms such as immune response, positive and negative regulation of T cells, and leukocyte-mediated cytotoxicity. This suggested that the OS-related prognostic genes have an impact on TME and immune infiltration in osteosarcoma.
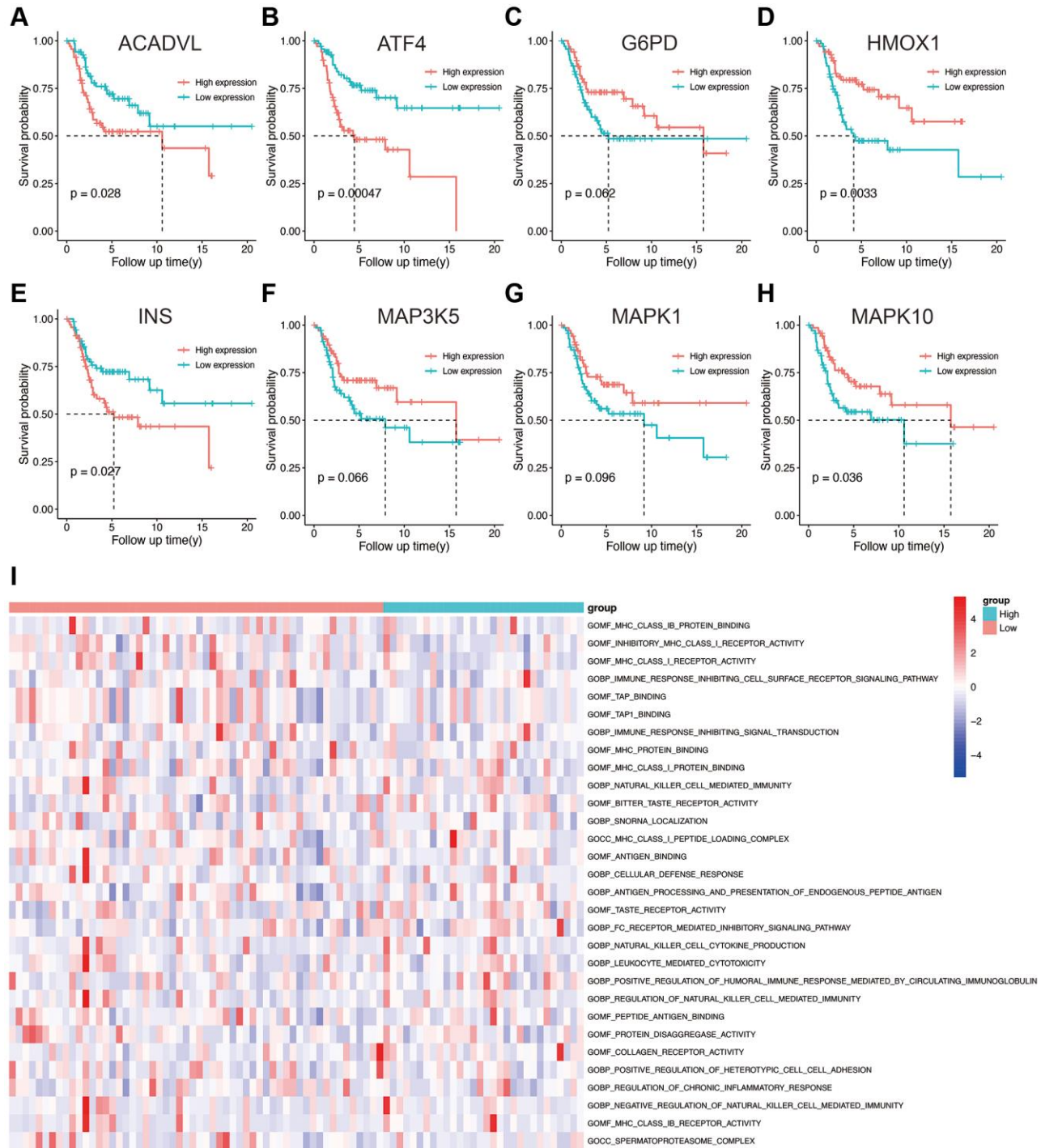
Figure 4. (**A–H**) Kaplan-Meier survival analysis of eight OS-related genes. (**I**) Results of GSVA in the high- and low-risk groups. Abbreviations: GO: gene ontology; BP: biological process; CC: cellular components; MF: molecular functions.

TARGET database with |log2FC|>1 and adj.*p* value < 0.05 as the cut-offs. The top 50 enriched GO terms obtained from GSVA are shown in Figure 4I, and these included many immune-related terms such as immune response, positive and negative regulation of T cells, and leukocyte-mediated cytotoxicity. This suggested that the OS-related prognostic genes have an impact on TME and immune infiltration in osteosarcoma.
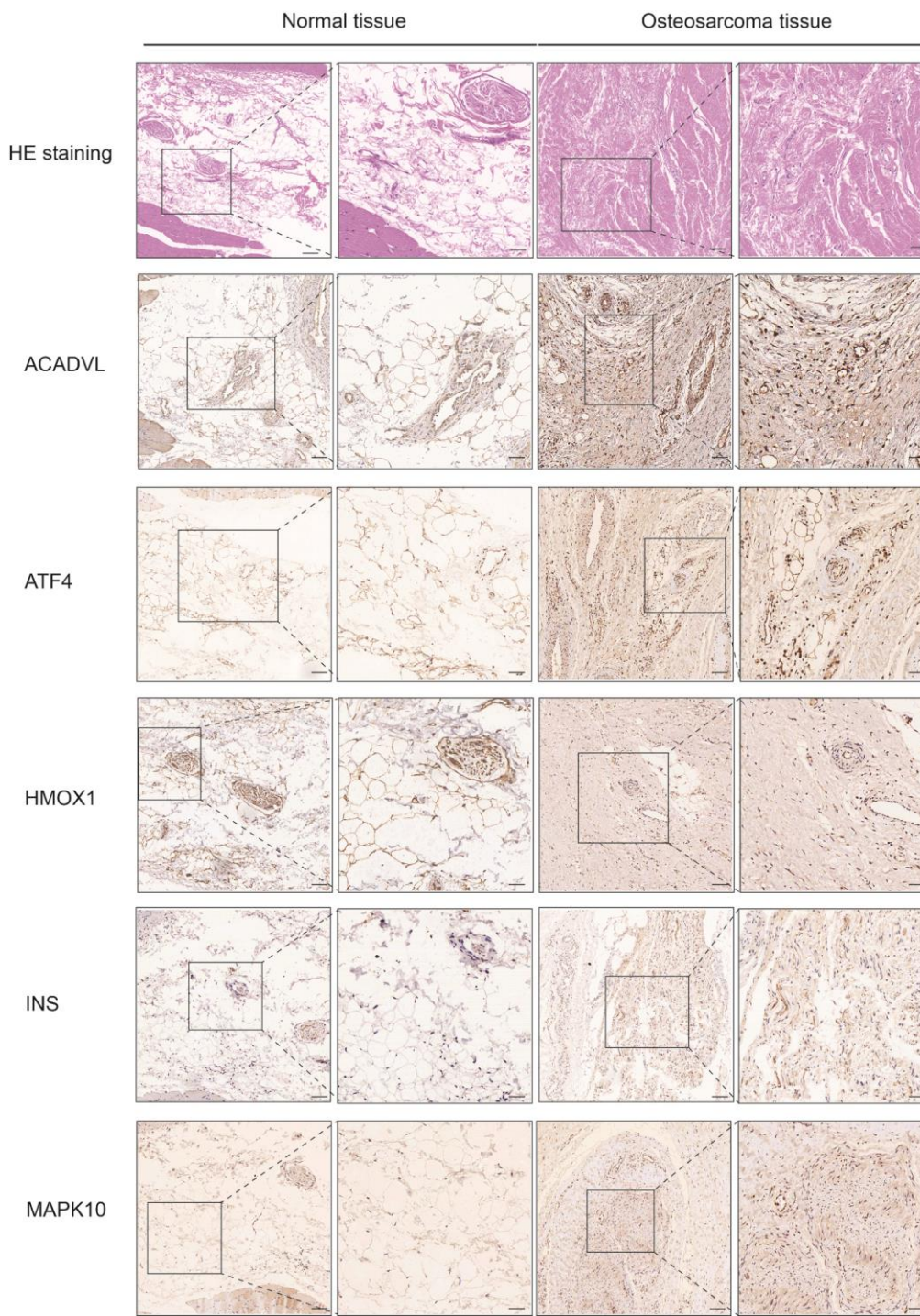
Validation of survival signatures of OS-related genes

We performed IHC among OS-related genes with significantly difference using corresponding antibodies. The IHC results illustrated that the expression of AVADVL, ATF4, INS and MAPK10 was higher in osteosarcoma tissues compared to normal tissues (Figure 5). However, the expression of HMOX1 was

Figure 5. Hematoxylin eosin (HE) staining and immunohistochemistry (IHC) staining of normal and osteosarcoma tissue. Scale bar, 100 μm (left panel) and 50 μm (right panel).

lower in osteosarcoma tissues, which was in line with the prediction of Kaplan-Meier survival curve.

Estimation of tumor immune microenvironment

We calculated the stromal score, immune score, and tumor purity in each sample in the TARGET and GSE21257 datasets using the ESTIMATE algorithm. We also compared the abovementioned parameters between the high- and low-risk groups, and found that patients in the high-risk group with lower stromal score had better overall survival, compared with those with higher stromal scores ($p < 0.0001$, Figure 6A, 6D). However, a higher immune score was associated with
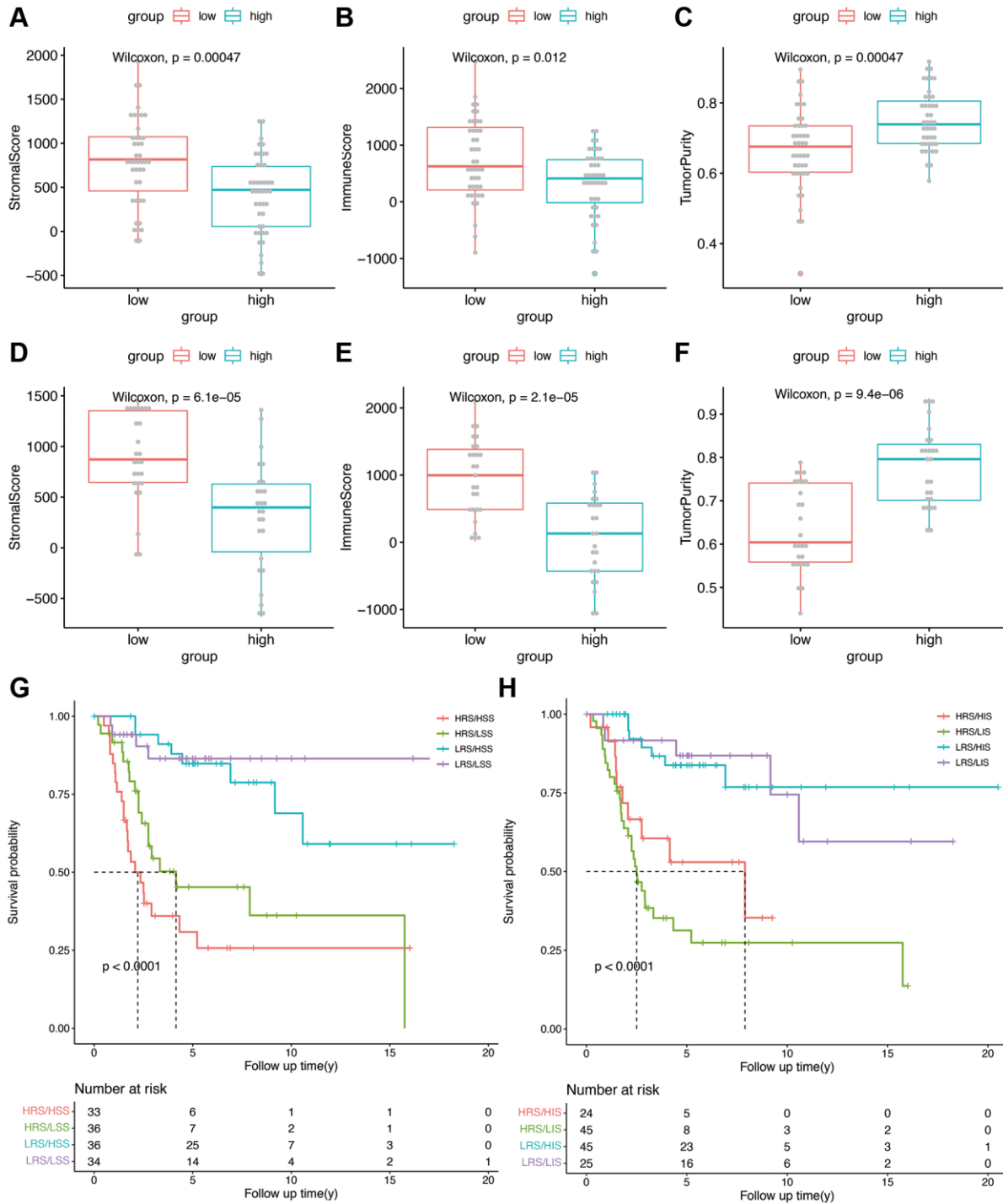
Figure 6. (**A–C**) Stromal score, immune score, and tumor purity in the high- and low-risk groups in the TARGET cohort. (**D–F**) Stromal score, immune score, and tumor purity in the high- and low-risk groups in the validation cohort. (**G**) Kaplan-Meier analysis combining risk score and stromal score. (**H**) Kaplan-Meier analysis combining risk score and immune score.

better overall survival as opposed to a lower immune score in the high-risk group ($p < 0.0001$, Figure 6B, 6E). High risk group of training and validation group had higher tumor purity than low risk group (Figure 6C, 6F). Thus, OS-related prognostic genes are associated with the immune microenvironment of osteosarcoma, and may be able to predict the efficacy of immunotherapy. Then risk score was combined with stromal score and immune score as shown in Figure 6G and 6H,

which illustrated the survival relationship of tumor immune microenvironment score.

Tumor-infiltrating immune cells using OS-related risk score

Immune-related pathway GSVA revealed that macrophages were well enriched between high- and low-risk groups (Figure 7A). The CIBERSORT
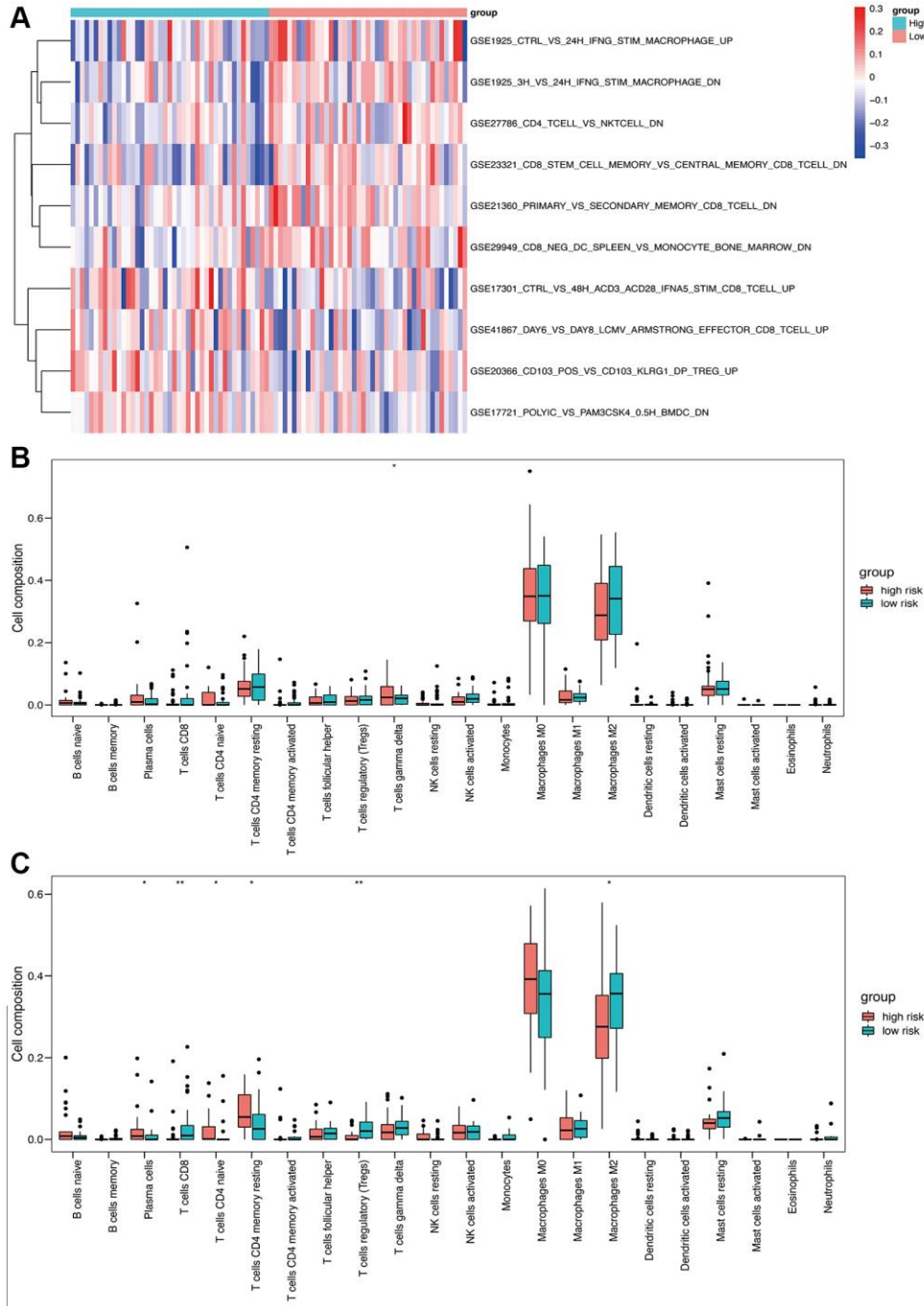
Figure 7. Immune analysis based on high- and low-risk groups. (**A**) Immune-related pathway GSVA between high- and low-risk groups. (**B**) Box plot of 22 immune cells in the high- and low-risk groups of the training cohort. (**C**) Box plot of 22 immune cells in the high- and low-risk groups of the validation cohort.

algorithm was used to evaluate the infiltration of 22 immune cell populations in the TARGET and GSE21257 osteosarcoma samples (Figure 7B and 7C). M0 and M2 macrophages were the top 2 infiltrating cell types in both cohorts. In the TARGET cohort, there was a significant difference in the abundance of infiltrating gamma delta T cells between the high- and low-risk groups (Figure 7B). In contrast, significant differences were observed in the infiltration of plasma cells, CD8 T cells, CD4 naïve T cells, CD4 memory T cells, regulatory T cells, and M2 macrophages between the high- and low-risk groups in the GSE21257 cohort (Figure 7C). These results suggested that the OS-related model is associated with the immunological status of the osteosarcoma microenvironment.

Sensitivity to immunotherapy and tumor mutation burden analysis

To explore sensitivity to immunotherapy, we performed further expression analysis for the possible immunotherapy target. The expression of several representative immune checkpoint genes is shown in Figure 8A–8H. We found that the expression of CD274 (PDL1), CXCL12, BTN3A1, LAG3, and IL10 was significantly decreased in the high-risk group of our model. No significant difference was found in the term of PDCD1(PD1), CTLA4 and CYT score. Figure 8I showed the gene mutation profiles of the high- and low-risk groups. With the application of maftools package, it can be observed that the proportion of mutated genes is not high, with p53 being the most frequently mutated gene. There is no significant difference between the high- and low-risk groups in terms of gene mutations.

Construction and validation of the nomogram

We constructed a nomogram by integrating the OS-related prognostic model and clinical characteristics including gender, race, and metastasis status. As shown in Figure 9A, each item was assigned a score and the total score of each patient predicted the probability of survival in 1, 3, and 5-years. The Forest plot of the nomogram revealed that metastasis and risk score showed significant prognostic value in the cohort (Figure 9B). The accuracy of the nomogram was validated in the TARGET cohort (Figure 9C and 9D), and the results indicated a good fit between the predicted and actual 3- and 5-year survival rates.

Single cell expression of OS-related mRNAs

We identified 23 immune cell subpopulations in the GSE162454 dataset by TSNE analysis at first. Based on the specific markers, we identified NK/T cells, CAFs

(cancer associated fibroblasts), endothelial cells, plasma cells, osteoblastic cells, osteoclasts, B cells, M2 TAMs (tumor associated macrophages), M1 TAMs, monocytes and mast cells (Figure 10A). The expression of marker genes for different cell clusters was displayed in dot plot format in Figure 10B. The expression of OS-related genes was compared in different cell types. Because of the dropout sign of single-cell sequencing, INS was not detected in the GSE162454 dataset. As shown in Figure 10C, HMOX1 were expressed at higher levels in TAMs and osteoclasts. ATF4 and ACADVL are highly expressed in myeloid cells as well as other cell types, which was suggestive of their roles in osteosarcoma progression.

DISCUSSION

Nearly 10 million adolescents are diagnosed with osteosarcoma each year [38]. Because of the anatomical features of bone tissues, chemotherapeutics have not achieved satisfactory response in bone malignancies [39]. OS, which is reflected in the excessive accumulation of ROS, frequently initiates tumorigenesis by promoting lipid oxidation and DNA damage [40]. Therefore, it is crucial to identify the OS-related genes in osteosarcoma and other cancers to increase the repertoire of therapeutic targets. In the present study, we analyzed the high-throughput RNA-sequencing data and clinical information of 86 osteosarcoma patients, and identified eight prognosis-related genes, including MAP3K5, G6PD, HMOX1, ATF4, ACADVL, MAPK1, MAPK10, and INS. An OS-related risk model was developed for predicting patient prognosis. With the combination of expressions of these eight genes, the risk scores had better prognostic performance than with individual gene expressions. Because the OS-related prognostic genes were enriched in immune-related functions, we analyzed the relationship between the risk model and the immunological characteristics of the TME, and observed significant differences in the infiltrating immune cell populations between the high- and low-risk groups. Furthermore, analysis of single cell sequencing results indicated that the OS-related genes were expressed in different cell types, which offers insights into potential treatment strategies for osteosarcoma.

In the field of osteosarcoma, multiple gene sets related to patient prognosis, such as ferroptosis-related genes, immune-related genes, and lipid metabolism genes, have been found [41–43]. However, whether OS-related genes can predict patient outcomes has not been investigated. In this study, the risk scores of patients were calculated based on the expressions of eight OS-related genes, and patients were divided into high- and low-risk groups according to the median number. The

difference in survival time between the two groups was statistically significant.

MAP3K5 is a serine/threonine kinase that mediates the MAP kinase signal transduction pathway [44] in different cancer types, such as pancreatic cancer [45], prostate cancer [46], and thyroid cancer [47]. It is also a ferroptosis-related gene associated with increased ROS accumulation [48]. G6PD is the first and the rate-limiting enzyme in the pentose phosphate pathway, as well as an antioxidant enzyme involved in the ribose 5-phosphate pathway for nucleotide synthesis. Wang et al. found that G6PD inhibition upregulated ROS levels in osteosarcoma cells and induced endoplasmic reticulum (ER) stress, which is consistent with the decreased survival rates observed in the G6PD^{low} group in our cohort. HMOX1 is a cytoprotective enzyme and its overexpression in U-2OS cells has been reported to significantly decrease cell proliferation rates by inducing ferroptosis [49]. Further, ATF4 is a master
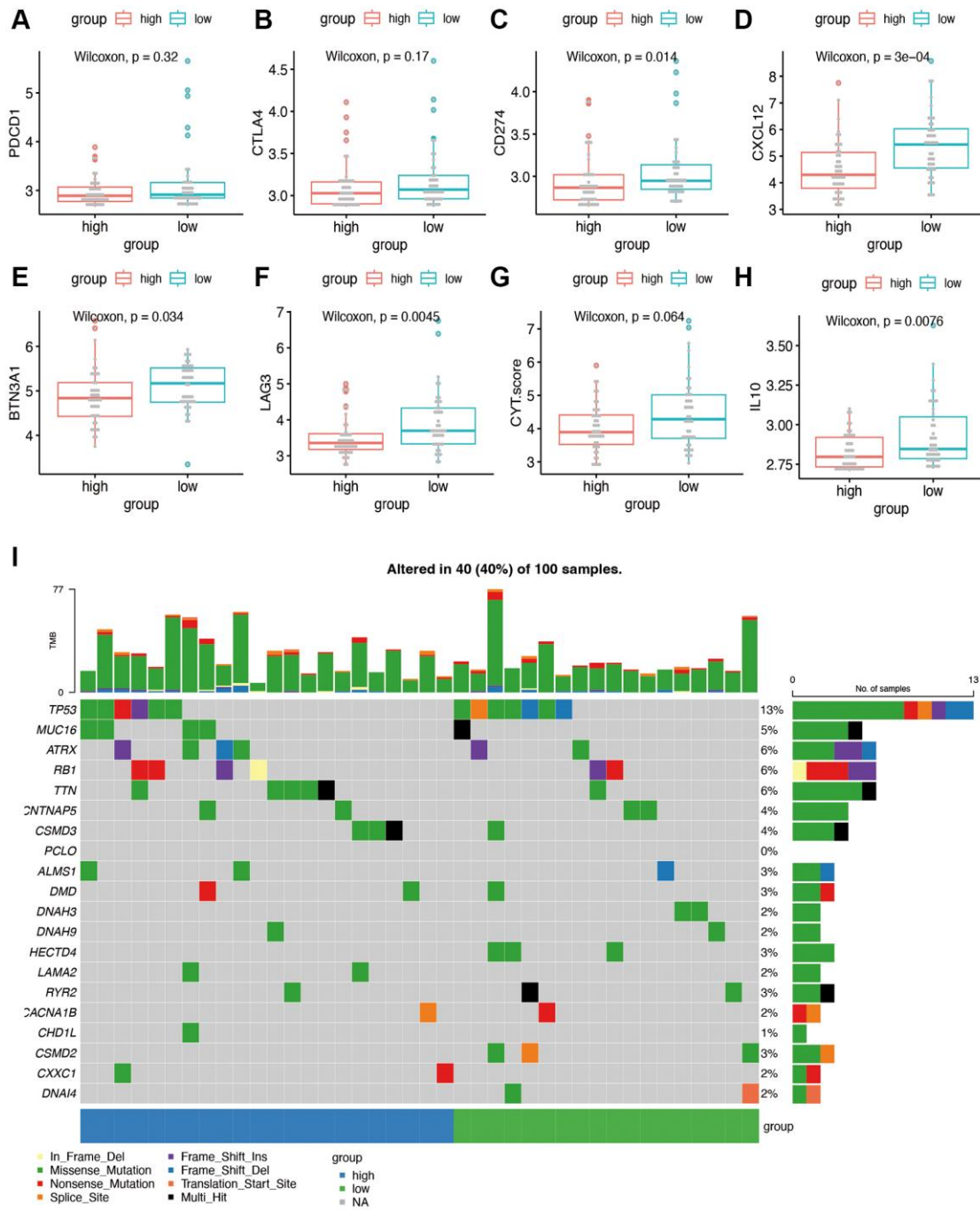
Figure 8. Immunotherapy sensitivity and tumor mutation burden analysis. (**A–H**) Expression levels of representative immune checkpoints between high and low risk group. (**I**) Tumor mutation analysis of high and low risk group.

transcriptional regulator of amino acid metabolism and stress responses [50], and induces apoptosis in response to persistent stress conditions through post-transcriptional modifications. ACADVL is an inner mitochondrial membrane protein and catalyzes the first step of the mitochondrial fatty acid beta-oxidation pathway [51]. However, there are currently no reports on the possible role of ACADVL in osteosarcoma.

The mitogen-activated protein kinase (MAPK) family of proteins integrate various upstream signals to regulate multiple cellular functions, including proliferation, differentiation, and survival. MAPK1 is overexpressed in breast cancer [52], lung cancer [53], ovarian cancer [54], and other malignancies. Several studies have shown that MAPK1 is targeted by non-coding RNAs to modulate the invasion and proliferation
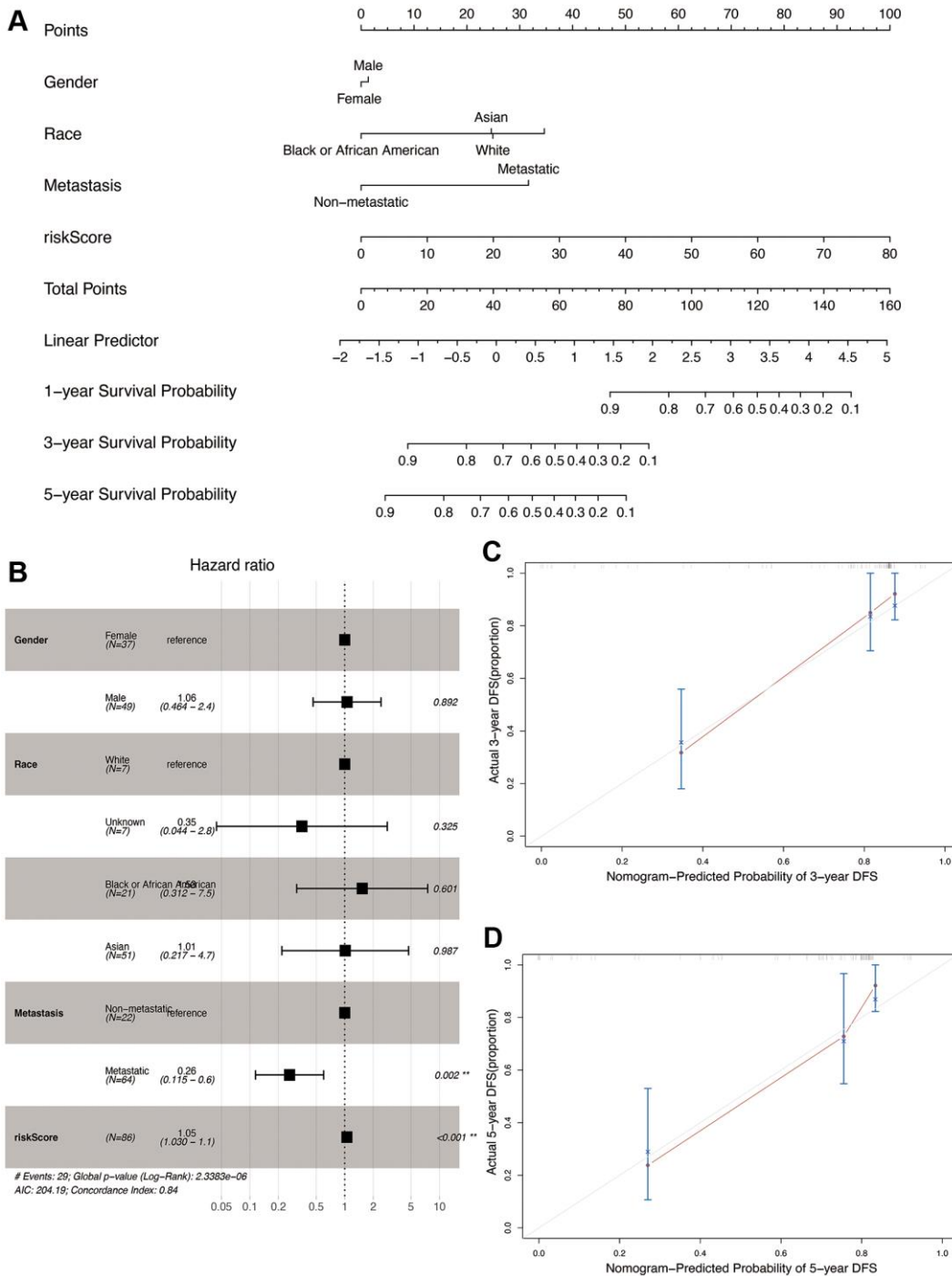
Figure 9. Construction and validation of the nomogram in the TARGET cohort. (**A**) Nomogram combining OS-related risk score and clinical characteristics for predicting 1-, 3-, and 5-year overall survival of osteosarcoma patients. (**B**) Forest plot showing results of multivariable cox regression of the nomogram. (**C**, **D**) The calibration curves for 3- and 5-year overall survival probability.

of osteosarcoma cells [55, 56]. MAPK10 is also associated with the prognosis of renal cell carcinoma [57] and hepatocellular carcinoma [58]. INS encodes insulin, a peptide hormone that plays a vital role in the regulation of carbohydrate and lipid metabolism, and is aberrantly expressed in various diseases. Gene expression microarray analysis in a previous study identified INS as one of the core genes in osteosarcoma with pulmonary metastasis [59]. In this study, we found significant differences between the prognosis of patients with low and high expression levels of ACADVL, ATF4, HMOX1, INS, and MAPK10, indicating that these genes are potential therapeutic targets of osteosarcoma.
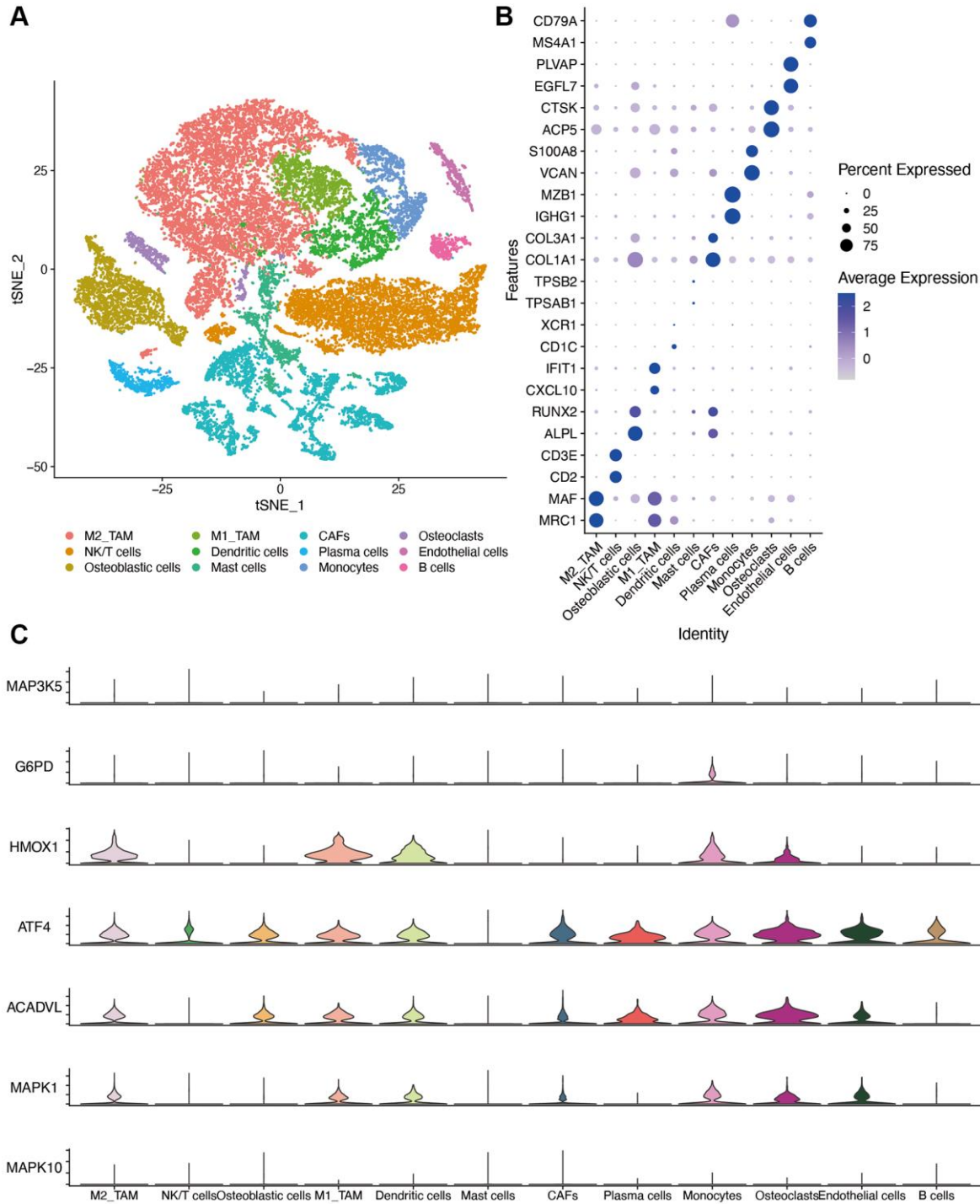
Figure 10. Single-cell sequencing dataset analysis. (**A**) The TSNE plot of 12 cell clusters of GSE162454. (**B**) Dot plot showing the expression levels of marker gene in single cell sequencing data. (**C**) violin plot of OS-related genes in different cell types. Abbreviations: CAF: cancer associated fibroblasts; TAM: tumor associated macrophages.

We further validated the OS-related risk model in an external cohort, and the ROC curve analysis showed high predictive accuracy of the model in both the training and validation cohorts. In addition, the nomogram integrating the risk score and clinical characteristics also reliably predicted the prognosis of osteosarcoma.

The immune microenvironment has been illustrated to have a significant effect on the prognosis of several diseases. Previous studies have shown that lower immune scores are associated with worse prognosis in osteosarcoma [60, 61]. GO analysis based on the risk score in our study illustrated several immune related terms. We found that the high-risk group had significantly higher tumor purity but lower stromal score and immune score in both the training and validation sets, which correlated with a worse prognosis. Myeloid cells, including macrophages, microglia, myeloid-derived suppressor cells, dendritic cells and neutrophils [62], are the predominant infiltrating immune cell populations in the TME, and regulate immune and therapeutic responses [63]. Xiao et al. developed a macrophage-associated risk model for osteosarcoma, which included MAP3K5 as one of the prognosis-related genes [61]. Our results also found that the potential sensitive immune checkpoint such as PDL1, CXCL12, BTN3A1, LAG3, and IL10, were differently expressed between high- and low- risk groups, which provide new therapy target.

Analysis of a single cell sequencing dataset also showed that most of the OS-related genes were highly expressed in myeloid cells in addition to their high expression in osteoblasts and osteoclasts. This warrants further investigation of the function of these genes in osteosarcoma. To summarize, our OS-related risk model is a potential indicator of the prognosis of osteosarcoma patients.

Our study has some limitations that ought to be considered. First, the sample size in our study was small compared with that in other similar studies. Second, the clinical data of the patients were not sufficient, and some key information such as the tumor stage and pathological type could not be included in the nomogram. Third, this study was based on sequencing data and bioinformatics analysis, and the results will have to be validated through *in vitro* and *in vivo* studies.

CONCLUSIONS

We constructed a prognostic risk model for osteo-sarcoma based on eight OS-related genes. The model could accurately predict patient prognosis and its performance was validated in an external cohort. In addition, the risk groups classified on the basis of the gene risk score showed significant differences in the immune microenvironment, some immune checkpoints showed difference between high and low risk groups. OS-related genes were highly expressed in myeloid cells. In conclusion, OS-related genes possibly regulate the immune responses in osteosarcoma, and may predict patient response to immunotherapies.

AUTHOR CONTRIBUTIONS

Hanning Wang proposed the research question, designed the experimental plan, analyzed the data, and wrote the manuscript. Juntan Li conducted data analysis and R code operation. Hanning Wang and Juntan Li contributed equally to this article. Xu Li provided guidance on research direction, experimental design, and manuscript revisions.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

ETHICAL STATEMENT AND CONSENT

The present study was approved by the ethics committee of the First Affiliated Hospital of China Medical University with the registration number 2019-285-2. A consent form was signed by each participant.

Editorial note

&This corresponding author has a verified history of publications using a personal email address for correspondence.

REFERENCES

1. Kumar R, Kumar M, Malhotra K, Patel S. Primary Osteosarcoma in the Elderly Revisited: Current Concepts in Diagnosis and Treatment. Curr Oncol Rep. 2018; 20:13.
https://doi.org/10.1007/s11912-018-0658-1
PMID:29492676

2. Nacev BA, Sanchez-Vega F, Smith SA, Antonescu CR, Rosenbaum E, Shi H, Tang C, Socci ND, Rana S, Gularte-Mérida R, Zehir A, Gounder MM, Bowler TG, et al. Clinical sequencing of soft tissue and bone sarcomas delineates diverse genomic landscapes and potential

therapeutic targets. Nat Commun. 2022; 13:3405.
https://doi.org/10.1038/s41467-022-30453-x
PMID:35705560

3. Gianferante DM, Mirabello L, Savage SA. Germline and somatic genetics of osteosarcoma - connecting aetiology, biology and therapy. Nat Rev Endocrinol. 2017; 13:480–91.
https://doi.org/10.1038/nrendo.2017.16
PMID:28338660

4. Tempelaere C, Biau D, Babinet A, Anract P. Osteosarcoma after the age of fifty: A clinicopathological study. Eur J Surg Oncol. 2019; 45:1288–92.
https://doi.org/10.1016/j.ejso.2019.04.010
PMID:31014986

5. Trela A, Szymańska R. Less widespread plant oils as a good source of vitamin E. Food Chem. 2019; 296:160–6.
https://doi.org/10.1016/j.foodchem.2019.05.185
PMID:31202300

6. Gill J, Gorlick R. Advancing therapy for osteosarcoma. Nat Rev Clin Oncol. 2021; 18:609–24.
https://doi.org/10.1038/s41571-021-00519-8
PMID:34131316

7. Czarnecka AM, Synoradzki K, Firlej W, Bartnik E, Sobczuk P, Fiedorowicz M, Grieb P, Rutkowski P. Molecular Biology of Osteosarcoma. Cancers (Basel). 2020; 12:2130.
https://doi.org/10.3390/cancers12082130
PMID:32751922

8. Boulouis HJ, Chomel BB, Guillaume G, Benoît D, Chang CC, Monteil M, Kasten RW, Jack A, Nadia H. Multiple locus variable number tandem repeat analysis for the characterization of wild feline Bartonella species and subspecies. Vet Microbiol. 2020; 247:108788.
https://doi.org/10.1016/j.vetmic.2020.108788
PMID:32768232

9. Ho XD, Nguyen HG, Trinh LH, Reimann E, Prans E, Kõks G, Maasalu K, Le VQ, Nguyen VH, Le NTN, Phung P, Märtson A, Lattekivi F, Kõks S. Analysis of the Expression of Repetitive DNA Elements in Osteosarcoma. Front Genet. 2017; 8:193.
https://doi.org/10.3389/fgene.2017.00193
PMID:29250102

10. Ho XD, Phung P, Le VQ, Nguyen VH, Reimann E, Prans E, Kõks G, Maasalu K, Le NT, Trinh LH, Nguyen HG, Märtson A, Kõks S. Whole transcriptome analysis identifies differentially regulated networks between osteosarcoma and normal bone samples. Exp Biol Med (Maywood). 2017; 242:1802–11.
https://doi.org/10.1177/1535370217736512
PMID:29050494

11. Pisoschi AM, Pop A. The role of antioxidants in the chemistry of oxidative stress: A review. Eur J Med Chem. 2015; 97:55–74.
https://doi.org/10.1016/j.ejmech.2015.04.040
PMID:25942353

12. Ďuračková Z. Some current insights into oxidative stress. Physiol Res. 2010; 59:459–69.
https://doi.org/10.33549/physiolres.931844
PMID:19929132

13. Multhaup G, Ruppert T, Schlicksupp A, Hesse L, Beher D, Masters CL, Beyreuther K. Reactive oxygen species and Alzheimer's disease. Biochem Pharmacol. 1997; 54:533–9.
https://doi.org/10.1016/s0006-2952(97)00062-2
PMID:9337068

14. Muhammad S, Bierhaus A, Schwaninger M. Reactive oxygen species in diabetes-induced vascular damage, stroke, and Alzheimer's disease. J Alzheimers Dis. 2009; 16:775–85.
https://doi.org/10.3233/JAD-2009-0982
PMID:19387112

15. Gelderman KA, Hultqvist M, Olsson LM, Bauer K, Pizzolla A, Olofsson P, Holmdahl R. Rheumatoid arthritis: the role of reactive oxygen species in disease development and therapeutic strategies. Antioxid Redox Signal. 2007; 9:1541–67.
https://doi.org/10.1089/ars.2007.1569
PMID:17678439

16. Miyajima A, Nakashima J, Yoshioka K, Tachibana M, Tazaki H, Murai M. Role of reactive oxygen species in cis-dichlorodiammineplatinum-induced cytotoxicity on bladder cancer cells. Br J Cancer. 1997; 76:206–10.
https://doi.org/10.1038/bjc.1997.363
PMID:9231920

17. Brown NS, Bicknell R. Hypoxia and oxidative stress in breast cancer. Oxidative stress: its effects on the growth, metastatic potential and response to therapy of breast cancer. Breast Cancer Res. 2001; 3:323–7.
https://doi.org/10.1186/bcr315
PMID:11597322

18. Ranjan P, Anathy V, Burch PM, Weirather K, Lambeth JD, Heintz NH. Redox-dependent expression of cyclin D1 and cell proliferation by Nox1 in mouse lung epithelial cells. Antioxid Redox Signal. 2006; 8:1447–59.
https://doi.org/10.1089/ars.2006.8.1447
PMID:16987002

19. Forman HJ, Zhang H. Targeting oxidative stress in disease: promise and limitations of antioxidant therapy. Nat Rev Drug Discov. 2021; 20:689–709.
https://doi.org/10.1038/s41573-021-00233-1
PMID:34194012

20. Nathan FM, Singh VA, Dhanoa A, Palanisamy UD.

Oxidative stress and antioxidant status in primary bone and soft tissue sarcoma. BMC Cancer. 2011; 11:382.
https://doi.org/10.1186/1471-2407-11-382
PMID:21871117

21. Sisubalan N, Ramkumar VS, Pugazhendhi A, Karthikeyan C, Indira K, Gopinath K, Hameed ASH, Basha MHG. ROS-mediated cytotoxic activity of ZnO and CeO$_2$ nanoparticles synthesized using the Rubia cordifolia L. leaf extract on MG-63 human osteosarcoma cell lines. Environ Sci Pollut Res Int. 2018; 25:10482–92.
https://doi.org/10.1007/s11356-017-0003-5
PMID:28963600

22. Pu Y, Tan Y, Zang C, Zhao F, Cai C, Kong L, Deng H, Chao F, Xia R, Xie M, Ge F, Pan Y, Cai S, Huang D. LAMTOR5-AS1 regulates chemotherapy-induced oxidative stress by controlling the expression level and transcriptional activity of NRF2 in osteosarcoma cells. Cell Death Dis. 2021; 12:1125.
https://doi.org/10.1038/s41419-021-04413-0
PMID:34862368

23. Hu H, Yang W, Liang Z, Zhou Z, Song Q, Liu W, Deng X, Zhu J, Xing X, Zhong B, Wang B, Wang S, Shao Z, Zhang Y. Amplification of oxidative stress with lycorine and gold-based nanocomposites for synergistic cascade cancer therapy. J Nanobiotechnology. 2021; 19:221.
https://doi.org/10.1186/s12951-021-00933-1
PMID:34315494

24. Chen C, Xie L, Ren T, Huang Y, Xu J, Guo W. Immunotherapy for osteosarcoma: Fundamental mechanism, rationale, and recent breakthroughs. Cancer Lett. 2021; 500:1–10.
https://doi.org/10.1016/j.canlet.2020.12.024
PMID:33359211

25. Miwa S, Shirai T, Yamamoto N, Hayashi K, Takeuchi A, Igarashi K, Tsuchiya H. Current and Emerging Targets in Immunotherapy for Osteosarcoma. J Oncol. 2019; 2019:7035045.
https://doi.org/10.1155/2019/7035045
PMID:30693030

26. Tao L, Huang G, Song H, Chen Y, Chen L. Cancer associated fibroblasts: An essential role in the tumor microenvironment. Oncol Lett. 2017; 14:2611–20.
https://doi.org/10.3892/ol.2017.6497
PMID:28927027

27. 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; 17:2110–9.
https://doi.org/10.1158/1078-0432.CCR-10-2047
PMID:21372215

28. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43:e47.
https://doi.org/10.1093/nar/gkv007
PMID:25605792

29. Leek JT. svaseq: removing batch effects and other unwanted noise from sequencing data. Nucleic Acids Res. 2014; 42:e161.
https://doi.org/10.1093/nar/gku864
PMID:25294822

30. Liu Y, Feng W, Dai Y, Bao M, Yuan Z, He M, Qin Z, Liao S, He J, Huang Q, Yu Z, Zeng Y, Guo B, et al. Single-Cell Transcriptomics Reveals the Complexity of the Tumor Microenvironment of Treatment-Naive Osteosarcoma. Front Oncol. 2021; 11:709210.
https://doi.org/10.3389/fonc.2021.709210
PMID:34367994

31. Engebretsen S, Bohlin J. Statistical predictions with glmnet. Clin Epigenetics. 2019; 11:123.
https://doi.org/10.1186/s13148-019-0730-1
PMID:31443682

32. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med. 1997; 16:385–95.
https://doi.org/10.1002/(sici)1097-0258(19970228)16:4<385::aid-sim380>3.0.co;2-3
PMID:9044528

33. 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

34. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013; 14:7.
https://doi.org/10.1186/1471-2105-14-7
PMID:23323831

35. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013; 4:2612.
https://doi.org/10.1038/ncomms3612
PMID:24113773

36. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015; 12:453–7.

https://doi.org/10.1038/nmeth.3337
PMID:25822800

37. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, Hoffman P, Stoeckius M, Papalexi E, et al. Integrated analysis of multimodal single-cell data. Cell. 2021; 184:3573–87.e29.
https://doi.org/10.1016/j.cell.2021.04.048
PMID:34062119

38. Stiller CA, Bielack SS, Jundt G, Steliarova-Foucher E. Bone tumours in European children and adolescents, 1978-1997. Report from the Automated Childhood Cancer Information System project. Eur J Cancer. 2006; 42:2124–35.
https://doi.org/10.1016/j.ejca.2006.05.015
PMID:16919776

39. Sauerbrey A, Bielack S, Kempf-Bielack B, Zoubek A, Paulussen M, Zintl F. High-dose chemotherapy (HDC) and autologous hematopoietic stem cell transplantation (ASCT) as salvage therapy for relapsed osteosarcoma. Bone Marrow Transplant. 2001; 27:933–7.
https://doi.org/10.1038/sj.bmt.1703023
PMID:11436103

40. Schieber M, Chandel NS. ROS function in redox signaling and oxidative stress. Curr Biol. 2014; 24:R453–62.
https://doi.org/10.1016/j.cub.2014.03.034
PMID:24845678

41. Qian H, Lei T, Hu Y, Lei P. Expression of Lipid-Metabolism Genes Is Correlated With Immune Microenvironment and Predicts Prognosis in Osteosarcoma. Front Cell Dev Biol. 2021; 9:673827.
https://doi.org/10.3389/fcell.2021.673827
PMID:33937273

42. Liu W, Xie X, Qi Y, Wu J. Exploration of Immune-Related Gene Expression in Osteosarcoma and Association With Outcomes. JAMA Netw Open. 2021; 4:e2119132.
https://doi.org/10.1001/jamanetworkopen.2021.19132
PMID:34342651

43. Lei T, Qian H, Lei P, Hu Y. Ferroptosis-related gene signature associates with immunity and predicts prognosis accurately in patients with osteosarcoma. Cancer Sci. 2021; 112:4785–98.
https://doi.org/10.1111/cas.15131
PMID:34506683

44. Luo Y, Gao S, Hao Z, Yang Y, Xie S, Li D, Liu M, Zhou J. Apoptosis signal-regulating kinase 1 exhibits oncogenic activity in pancreatic cancer. Oncotarget. 2016; 7:75155–64.

45. Qiu CJ, Wang XB, Zheng ZR, Yang CZ, Lin K, Zhang K, Tu M, Jiang KR, Gao WT. Development and validation of a ferroptosis-related prognostic model in pancreatic cancer. Invest New Drugs. 2021; 39:1507–22.
https://doi.org/10.1007/s10637-021-01114-5
PMID:34195903

46. Pressinotti NC, Klocker H, Schäfer G, Luu VD, Ruschhaupt M, Kuner R, Steiner E, Poustka A, Bartsch G, Sültmann H. Differential expression of apoptotic genes PDIA3 and MAP3K5 distinguishes between low- and high-risk prostate cancer. Mol Cancer. 2009; 8:130.
https://doi.org/10.1186/1476-4598-8-130
PMID:20035634

47. Wang Y, Yang J, Chen S, Wang W, Teng L. Identification and Validation of a Prognostic Signature for Thyroid Cancer Based on Ferroptosis-Related Genes. Genes (Basel). 2022; 13:997.
https://doi.org/10.3390/genes13060997
PMID:35741758

48. Lei G, Zhuang L, Gan B. Targeting ferroptosis as a vulnerability in cancer. Nat Rev Cancer. 2022; 22:381–96.
https://doi.org/10.1038/s41568-022-00459-0
PMID:35338310

49. Lin H, Chen X, Zhang C, Yang T, Deng Z, Song Y, Huang L, Li F, Li Q, Lin S, Jin D. EF24 induces ferroptosis in osteosarcoma cells through HMOX1. Biomed Pharmacother. 2021; 136:111202.
https://doi.org/10.1016/j.biopha.2020.111202
PMID:33453607

50. Wortel IMN, van der Meer LT, Kilberg MS, van Leeuwen FN. Surviving Stress: Modulation of ATF4-Mediated Stress Responses in Normal and Malignant Cells. Trends Endocrinol Metab. 2017; 28:794–806.
https://doi.org/10.1016/j.tem.2017.07.003
PMID:28797581

51. Tcheng M, Roma A, Ahmed N, Smith RW, Jayanth P, Minden MD, Schimmer AD, Hess DA, Hope K, Rea KA, Akhtar TA, Bohrnsen E, D'Alessandro A, et al. Very long chain fatty acid metabolism is required in acute myeloid leukemia. Blood. 2021; 137:3518–32.
https://doi.org/10.1182/blood.2020008551
PMID:33720355

52. Niu L, Zhou Y, Zhang W, Ren Y. Long noncoding RNA LINC00473 functions as a competing endogenous RNA to regulate MAPK1 expression by sponging miR-198 in breast cancer. Pathol Res Pract. 2019; 215:152470.
https://doi.org/10.1016/j.prp.2019.152470

PMID:31201066

53. Jun T, Zheng FS, Ren KM, Zhang HY, Zhao JG, Zhao JZ. Long non-coding RNA UCA1 regulates the proliferation, migration and invasion of human lung cancer cells by modulating the expression of microRNA-143. Eur Rev Med Pharmacol Sci. 2018; 22:8343–52.
https://doi.org/10.26355/eurrev_201812_16532
PMID:30556875

54. Xu M, Zhou K, Wu Y, Wang L, Lu S. Linc00161 regulated the drug resistance of ovarian cancer by sponging microRNA-128 and modulating MAPK1. Mol Carcinog. 2019; 58:577–87.
https://doi.org/10.1002/mc.22952
PMID:30556928

55. Wu J, Zhang C, Chen L. MiR-511 mimic transfection inhibits the proliferation, invasion of osteosarcoma cells and reduces metastatic osteosarcoma tumor burden in nude mice via targeting MAPK1. Cancer Biomark. 2019; 26:343–51.
https://doi.org/10.3233/CBM-190534
PMID:31524148

56. Jerez S, Araya H, Hevia D, Irarrázaval CE, Thaler R, van Wijnen AJ, Galindo M. Extracellular vesicles from osteosarcoma cell lines contain miRNAs associated with cell adhesion and apoptosis. Gene. 2019; 710:246–57.
https://doi.org/10.1016/j.gene.2019.06.005
PMID:31176732

57. Tsai YC, Huang CY, Hsueh YM, Fan YC, Fong YC, Huang SP, Geng JH, Chen LC, Lu TL, Bao BY. Genetic variants in MAPK10 modify renal cell carcinoma susceptibility and clinical outcomes. Life Sci. 2021; 275:119396.
https://doi.org/10.1016/j.lfs.2021.119396
PMID:33774030

58. Li H, Li Y, Zhang Y, Tan B, Huang T, Xiong J, Tan X, Ermolaeva MA, Fu L. MAPK10 Expression as a Prognostic Marker of the Immunosuppressive Tumor Microenvironment in Human Hepatocellular Carcinoma. Front Oncol. 2021; 11:687371.
https://doi.org/10.3389/fonc.2021.687371
PMID:34408980

59. Shi Z, Zhou H, Pan B, Lu L, Wei Z, Shi L, Yao X, Kang Y, Feng S. Exploring the key genes and pathways of osteosarcoma with pulmonary metastasis using a gene expression microarray. Mol Med Rep. 2017; 16:7423–31.
https://doi.org/10.3892/mmr.2017.7577
PMID:28944885

60. Zhang C, Zheng JH, Lin ZH, Lv HY, Ye ZM, Chen YP, Zhang XY. Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma. Aging (Albany NY). 2020; 12:3486–501.
https://doi.org/10.18632/aging.102824
PMID:32039832

61. Xiao KW, Liu ZB, Zeng ZH, Yan FF, Xiao LF, Li JL, Cai L. Construction and Validation of a Macrophage-Associated Risk Model for Predicting the Prognosis of Osteosarcoma. J Oncol. 2021; 2021:9967954.
https://doi.org/10.1155/2021/9967954
PMID:34188683

62. Engblom C, Pfirschke C, Pittet MJ. The role of myeloid cells in cancer therapies. Nat Rev Cancer. 2016; 16:447–62.
https://doi.org/10.1038/nrc.2016.54
PMID:27339708

63. De Leo A, Ugolini A, Veglia F. Myeloid Cells in Glioblastoma Microenvironment. Cells. 2020; 10:18.
https://doi.org/10.3390/cells10010018
PMID:33374253
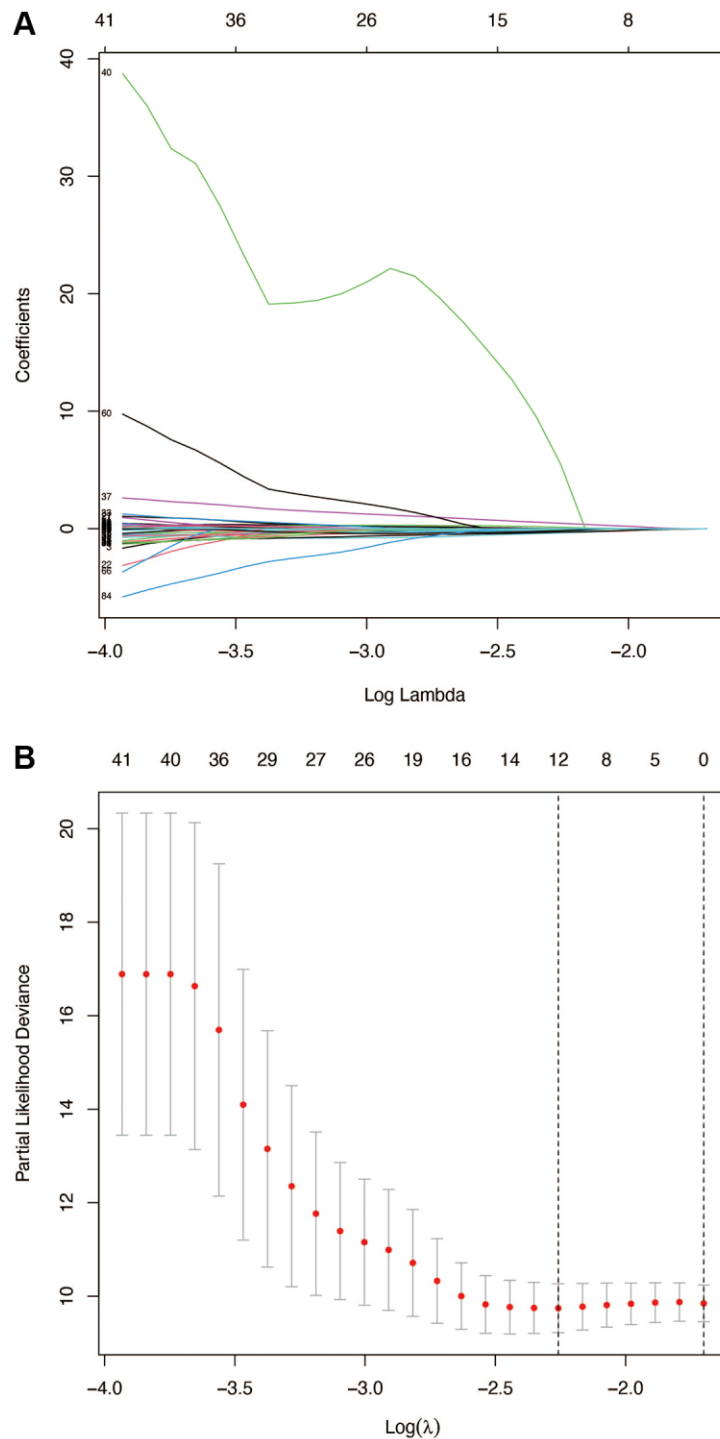
Supplementary Figure 1. LASSO regression analysis with optimal lambda.

Supplementary Tables

Supplementary Table 1. 89 OS-related genes in osteosarcoma identified by univariate cox regression analysis.

Gene	HR	HR.95L	HR.95H	*p*-value
CAT	0.332	0.164	0.671	0.002
HMOX1	0.721	0.530	0.981	0.037
MAPK1	0.388	0.178	0.848	0.018
G6PD	0.376	0.188	0.752	0.006
TXN2	0.422	0.204	0.871	0.020
GPX1	0.636	0.412	0.980	0.040
ACADVL	2.886	1.374	6.065	0.005
MAP3K5	0.286	0.120	0.682	0.005
CCL2	0.553	0.363	0.842	0.006
INS	15608084289545.4	15654.959	15561349814940800000000	0.004
ATF4	1.845	1.087	3.131	0.023
EIF2AK3	2.416	1.197	4.878	0.014
FMO3	0.388	0.157	0.964	0.041
IL10	0.020	0.001	0.629	0.026
ACE	0.305	0.095	0.981	0.046
PPARG	0.426	0.257	0.708	0.001
NUDT1	1.965	1.146	3.369	0.014
TLR4	0.495	0.261	0.938	0.031
MAOB	1.519	1.015	2.274	0.042
SFXN4	3.580	1.593	8.042	0.002
ATM	1.735	1.054	2.857	0.030
VCAM1	0.589	0.405	0.858	0.006
GPX7	2.302	1.262	4.197	0.007
KCNJ5	0.189	0.061	0.587	0.004
CD36	1.601	1.156	2.217	0.005
CYP2C19	5205.861	6.620	4093638.921	0.012
MB	1.169	1.015	1.347	0.030
BLVRB	0.519	0.312	0.862	0.011
MAP2K4	0.523	0.287	0.953	0.034
EME2	4.008	1.028	15.627	0.046
AGTR1	0.516	0.294	0.903	0.020
CYGB	1.624	1.034	2.549	0.035
MRPS23	0.309	0.114	0.837	0.021
PIK3CG	0.289	0.112	0.748	0.010
ADH5	0.344	0.147	0.807	0.014
CBS	2.188	1.269	3.771	0.005
ANXA5	0.461	0.223	0.952	0.036
TERT	4.444	2.120	9.314	0.000
MAP2K7	2.826	1.346	5.931	0.006
EGFR	0.610	0.401	0.928	0.021
TPD52	2.577	1.509	4.399	0.001
NDUFV1	2.627	1.082	6.378	0.033
TLR2	0.357	0.162	0.786	0.011
CYP11B2	402875.664	13.324	12181370764.833	0.014

PGD	0.364	0.183	0.726	0.004
MAPKAPK3	0.508	0.275	0.940	0.031
INSR	2.307	1.231	4.326	0.009
NDUFB9	1.478	1.053	2.074	0.024
CRAT	0.471	0.278	0.798	0.005
TNFRSF1A	0.360	0.199	0.653	0.001
PTPN1	0.432	0.189	0.984	0.046
LOX	1.745	1.222	2.491	0.002
SUOX	0.294	0.090	0.959	0.042
BMP2	1.363	1.007	1.844	0.045
ALDH1A1	0.619	0.399	0.959	0.032
NOL3	1.999	1.190	3.359	0.009
RPS27A	2.068	1.044	4.096	0.037
AOC3	2.115	1.299	3.443	0.003
CYP2C8	2.816	1.142	6.942	0.025
MYC	2.246	1.439	3.505	0.000
CYP2A6	0.001	0.000	0.706	0.040
PRNP	0.506	0.285	0.898	0.020
PAH	69.244	1.223	3921.661	0.040
CYP27A1	0.584	0.375	0.910	0.018
FAAH	1.696	1.115	2.580	0.014
OGDH	0.288	0.126	0.660	0.003
NDUFB10	0.529	0.290	0.966	0.038
TG	2.133	1.092	4.166	0.027
PTS	2.438	1.203	4.942	0.013
MAPKAP1	0.257	0.085	0.771	0.015
CASP1	0.588	0.358	0.967	0.036
AS3MT	34.423	4.283	276.687	0.001
CHEK1	1.916	1.056	3.477	0.032
ADRB1	2.475	1.362	4.497	0.003
EPHX2	2.183	1.140	4.180	0.019
SCARB1	1.571	1.093	2.258	0.015
PXN	0.390	0.164	0.924	0.032
PRKD2	0.377	0.146	0.975	0.044
KCNMA1	2.430	1.382	4.272	0.002
UGT1A6	4651.907	3.283	6591366.350	0.023
TREM2	0.661	0.453	0.965	0.032
IGF1R	1.584	1.142	2.196	0.006
MUC1	1.457	1.106	1.921	0.008
PTGIS	1.500	1.069	2.105	0.019
ITGAM	0.424	0.233	0.773	0.005
KRAS	0.446	0.204	0.976	0.043
MSR1	0.604	0.370	0.986	0.044
PTPN3	1.973	1.076	3.618	0.028

Supplementary Table 2. OS-related gene expression set obtained after lasso regression.

id	TXN2	MAP3K5	G6PD	HMOX1	ATF4	CAT	ACADVL	MAPK1	INS	EIF2AK3	CCL2	MAPK10
1	6.12	4.11	4.93	7.20	8.54	6.11	6.66	6.40	2.77	5.52	4.87	3.68
2	6.39	3.84	5.51	8.24	9.14	5.30	8.13	5.80	2.77	4.42	7.48	3.08
3	6.13	3.91	5.62	6.75	7.95	6.30	8.47	6.07	2.77	5.04	4.67	2.97
4	7.13	3.79	4.87	4.69	8.43	5.59	7.85	6.38	2.77	5.42	3.31	3.33
5	6.22	3.11	5.22	4.15	8.38	6.77	6.60	6.09	2.77	4.67	3.83	3.08
6	6.17	3.54	5.45	5.70	8.09	5.48	7.28	5.60	2.77	5.72	3.22	2.99
7	6.60	4.56	4.66	5.80	9.94	5.39	7.16	6.20	2.87	5.32	2.99	2.90
8	5.57	3.41	5.60	5.31	7.59	5.62	6.70	6.54	2.77	5.58	3.19	3.42
9	6.47	3.79	5.73	6.75	9.30	5.52	9.03	5.67	2.77	4.94	5.04	2.97
10	6.92	4.82	5.48	9.44	8.81	6.08	7.95	6.24	2.77	4.47	4.14	3.75
11	6.91	3.92	4.88	6.29	9.37	6.12	7.97	7.25	2.77	5.17	4.75	2.97
12	5.15	3.52	5.37	5.95	9.93	4.88	7.78	6.36	2.77	5.22	2.84	2.91
13	6.51	3.54	5.05	7.92	8.49	6.56	7.71	7.11	2.77	5.35	4.34	2.94
14	6.62	3.28	4.94	5.97	9.20	5.39	8.80	6.77	2.77	5.87	3.39	2.98
15	6.48	3.28	4.48	6.40	9.62	4.86	7.51	7.05	2.77	5.63	3.07	4.60
16	6.99	4.34	5.37	9.99	9.27	5.76	8.37	6.14	2.77	4.82	6.01	3.17
17	5.83	3.81	5.73	4.86	7.75	5.75	6.86	6.13	2.77	5.14	3.06	3.03
18	6.48	5.00	5.72	6.00	8.07	6.40	7.47	6.32	2.77	4.76	6.43	2.96
19	6.49	3.75	5.17	6.96	9.58	5.84	8.15	6.93	2.76	5.12	3.55	3.18
20	7.92	6.01	4.68	6.24	9.72	6.36	7.71	7.67	2.77	4.93	3.47	2.99
21	6.47	4.71	5.20	10.35	8.05	7.15	7.81	6.58	2.77	4.55	7.25	2.94
22	7.25	3.87	5.76	7.05	8.31	5.74	7.23	5.81	2.77	4.25	6.28	3.24
23	6.55	3.90	5.05	6.71	8.39	6.10	7.51	6.29	2.76	4.70	4.15	3.99
24	5.92	3.74	5.13	7.53	8.44	5.84	7.31	6.62	2.77	5.48	4.74	2.95
25	6.32	3.96	5.44	6.59	9.22	6.27	7.22	6.79	2.77	5.77	4.74	2.95
26	6.60	3.92	4.86	6.00	8.59	6.34	6.78	6.44	2.77	4.83	4.09	3.15
27	6.43	4.58	5.64	9.58	8.13	6.36	7.07	6.70	2.77	4.33	5.03	3.65
28	7.04	3.69	5.02	5.84	10.56	5.85	6.84	6.66	2.76	3.70	3.42	3.58
29	6.78	4.20	6.28	6.99	9.23	5.58	7.75	6.28	2.76	4.64	5.39	4.12
30	7.23	3.39	5.79	6.31	9.64	6.04	8.40	6.77	2.78	5.54	3.66	3.86
31	5.59	3.65	4.90	8.92	9.08	5.90	7.26	6.35	2.77	5.13	4.33	3.42
32	6.91	4.14	5.81	7.61	8.36	5.98	7.67	7.17	2.77	5.57	3.52	3.08
33	6.75	4.32	5.33	6.56	9.29	5.38	7.73	6.07	2.77	5.14	5.03	3.18
34	6.06	3.34	5.13	6.11	9.32	4.67	7.51	5.84	2.77	5.93	3.21	2.93
35	6.50	3.57	5.16	6.23	8.58	5.80	7.50	5.66	2.76	5.06	4.39	2.93
36	6.72	5.07	6.89	8.91	8.60	6.39	7.21	6.22	2.77	5.06	5.46	3.35
37	7.12	3.41	5.21	6.22	9.34	6.58	7.30	6.13	2.77	3.98	3.86	3.88
38	6.70	4.47	4.51	5.19	8.83	4.71	7.77	5.53	2.76	4.76	3.23	3.00
39	6.38	3.37	5.40	4.78	8.76	5.20	8.02	6.41	2.77	4.76	3.11	3.38
40	6.84	4.08	5.56	9.65	9.51	5.54	7.76	6.45	2.77	4.66	7.13	3.37
41	6.66	3.75	5.81	6.89	8.45	6.41	7.31	6.06	2.77	4.70	4.88	2.97
42	6.60	4.16	5.75	8.96	8.21	5.87	7.92	6.04	2.77	4.79	6.27	3.10
43	7.26	4.04	6.49	8.30	9.29	6.08	7.92	6.06	2.76	4.67	5.12	3.49
44	6.16	3.97	5.68	6.13	7.67	5.84	6.87	6.06	2.77	4.48	4.43	3.34
45	6.29	3.63	5.36	8.31	7.82	5.58	7.25	6.80	2.77	4.65	4.15	3.43
46	5.87	3.97	5.46	6.90	9.09	6.81	8.08	6.05	2.76	5.27	3.75	2.97

47	6.68	3.94	5.11	7.01	9.32	6.25	6.82	6.51	2.77	5.18	4.21	2.93
48	7.04	4.16	5.04	6.34	9.20	6.07	7.70	6.61	2.77	4.79	4.38	3.39
49	7.04	4.43	5.81	6.63	8.56	5.86	7.11	6.93	2.77	5.40	5.07	3.60
50	6.78	3.15	5.54	6.68	8.48	6.05	8.23	6.18	2.77	5.13	4.74	2.97
51	6.72	4.13	5.61	5.31	9.40	5.55	6.19	6.18	2.77	4.81	4.41	3.01
52	6.74	3.83	5.19	7.17	9.58	5.71	7.98	6.03	2.77	5.38	4.36	3.00
53	7.47	4.28	5.43	5.97	8.71	5.49	8.13	5.28	2.77	3.50	3.59	4.08
54	6.95	3.74	6.29	8.93	9.02	5.82	7.66	6.21	2.77	5.09	3.97	3.00
55	7.60	4.31	6.01	6.81	10.15	5.50	7.47	6.37	2.76	5.00	4.75	4.29
56	6.55	3.78	4.95	6.18	8.77	5.48	6.57	7.02	2.77	5.22	4.53	3.63
57	5.48	4.03	5.39	5.10	8.00	5.67	7.72	6.00	2.77	5.09	3.99	3.08
58	6.37	3.67	5.84	5.73	8.85	6.01	7.48	6.81	2.77	4.03	4.90	3.18
59	6.32	3.64	6.14	9.41	8.44	6.10	7.63	6.33	2.77	4.13	5.72	3.03
60	7.77	5.23	7.33	6.85	9.68	5.34	7.65	5.80	2.77	3.69	7.09	3.09
61	6.23	3.11	5.73	5.53	8.62	6.07	6.96	7.81	2.77	4.48	3.68	4.80
62	7.21	3.67	6.63	7.03	8.33	5.43	7.35	5.39	2.77	3.37	5.21	3.81
63	7.33	4.24	5.79	6.87	8.36	5.90	6.92	6.39	2.77	4.21	7.27	2.93
64	6.95	3.56	6.01	7.80	9.42	5.95	7.80	6.27	2.77	4.99	5.73	3.23
65	6.49	3.80	4.71	6.65	8.74	5.80	7.78	6.08	2.77	4.83	3.92	2.94
66	6.80	4.21	5.47	5.21	7.83	6.43	7.17	5.48	2.77	4.68	4.77	2.95
67	6.20	5.08	5.28	6.50	8.69	4.82	6.93	6.47	2.77	5.26	4.74	2.96
68	6.89	3.57	5.67	8.12	9.34	4.93	7.95	5.56	2.77	6.16	4.27	3.11
69	6.15	5.14	6.63	6.07	8.53	5.94	8.59	6.52	2.77	5.66	4.33	3.20
70	7.44	4.88	5.21	6.24	9.27	6.29	8.12	6.59	2.76	5.63	5.68	3.24
71	5.73	3.32	4.42	8.36	8.82	5.89	6.94	5.72	2.77	5.23	5.05	3.23
72	6.76	3.26	5.69	8.65	8.33	5.69	7.71	6.66	2.77	4.68	4.66	3.03
73	6.76	3.80	7.09	8.26	9.01	6.10	8.18	6.30	2.77	5.31	5.48	2.98
74	7.12	4.02	5.79	8.12	9.69	4.22	7.62	6.41	2.77	4.72	6.94	3.16
75	6.70	4.43	6.36	8.24	8.77	5.63	7.50	6.06	2.77	4.56	6.13	3.34
76	7.19	3.80	6.18	8.38	8.53	5.39	7.44	6.13	2.77	4.59	5.36	3.08
77	6.92	3.58	5.96	7.89	10.56	5.28	8.08	6.48	2.76	5.67	3.77	3.29
78	6.88	3.85	6.99	9.26	9.08	6.03	7.17	6.94	2.76	5.02	6.37	3.33
79	6.16	4.46	5.90	6.58	8.27	6.79	7.66	6.97	2.77	4.73	4.42	3.06
80	6.70	4.83	5.34	7.89	8.77	6.54	8.53	6.66	2.76	4.32	6.84	3.04
81	6.93	3.76	4.97	6.30	9.48	6.64	7.78	6.34	2.77	5.12	5.96	2.99
82	6.32	3.41	4.82	6.75	8.94	5.66	7.95	6.12	2.77	4.62	5.12	2.90
83	6.50	4.25	5.37	5.46	9.00	6.13	6.99	6.07	2.77	5.19	7.73	3.09
84	6.68	4.29	6.19	6.19	8.26	6.09	6.51	6.36	2.77	4.32	4.18	3.00
85	6.97	4.17	5.47	9.51	8.98	6.02	8.15	6.29	2.76	4.75	4.77	2.95
86	6.88	5.08	5.57	7.06	8.15	5.19	7.69	6.90	2.76	3.93	5.34	3.06