European Review for Medical and Pharmacological Sciences 2022; 26: 2926-2943

Characterizing the critical role of metabolism in osteosarcoma based on establishing novel molecular subtypes

W.-B. ZHANG, F.-M. HAN, L.-M. LIU, H.-B. JIN, X.-Y. YUAN, H.-S. SHANG

Department of Tendon Injury, Shenzhen Luohu Hospital of Traditional Chinese Medicine, Shenzhen, China

Abstract. – OBJECTIVE: Osteosarcoma is the third most frequently diagnosed cancer among adolescents. Immunotherapy is an effective curative treatment for metastatic osteosarcoma patients. This study aimed to further reveal the significance of metabolism in tumor progression, and to categorize molecular subtypes for guiding personalized therapy.

MATERIALS AND METHODS: Univariate Cox regression analysis was performed to screen metabolism-related genes associated with osteosarcoma prognosis. A molecular subtyping system was developed by unsupervised consensus clustering. Survival analysis and functional analysis were used to evaluate the performance of subtyping and characterize the TME of subtypes. Stepwise Akaike information criterion (stepAIC) was employed to optimize the prognostic model.

RESULTS: C1 and C2 subtypes showed distinct prognosis, with more favorable survival in C2 subtype. C2 subtype presented a higher immune infiltration and active anti-tumor response. Notably, C2 subtype was predicted to have better immune response to immune checkpoint blockade. In addition, a 5-gene prognostic signature with robust ability to classify patients into high-risk and low-risk groups was developed.

CONCLUSIONS: The study revealed the critical role of metabolism in tumorigenesis by comparing the features between the two subtypes. Oncogenic pathways including epithelial mesenchymal transition (EMT), glycolysis and hypoxia may be closely involved in the correlation with metabolism. Importantly, we developed a novel subtyping system and a 5-gene signature with high potential to be applied in clinical practice.

Key Words:

Osteosarcoma, Metabolism-related genes, Molecular subtypes, Tumor microenvironment, Immunotherapy, Prognostic signature.

Introduction

Osteosarcoma is the most frequently diagnosed malignant tumor in primary bone cancers, and it is also the third common cancer among adolescents[1]. Osteosarcoma occurs with bimodal age distribution in 15-19 ages and 75-79 age, respectively, but is more frequent among adolescents[2]. Age-standardized rate (ASR) has been rising in the past decades, with an incidence of about 3.4 per million worldwide[1,2]. The male population suffers from osteosarcoma more easily than the female population in most countries, with a ratio of 1.43:1 in young individuals aged between 0 and 24[2]. Improved treatments for osteosarcoma have greatly elevated 5-year overall survival (OS) from around 20% to 65%[3]. However, for metastatic patients, the survival chance markedly drops to approximately 30% compared with localized osteosarcoma patients who are estimated to have a 5-year survival rate of 70%-75%[3,4].

Evidence[5-10] has revealed that genetic variations in TP53, RB, CDK, c-Myc and TGFB are common factors in osteosarcoma development. Nevertheless, the mechanism and biology of osteosarcoma metastasis still remain unclear. Fritsche-Guenther et al[11] observed that the metabolism is distinct between malignant and benign osteosarcoma cells, as in the latter state malignant cells have a more activated metabolism. However, metastatic cell lines presented faster metabolism than malignant cell lines, which are more sensitive to glycolytic inhibition[11]. Metabolism in osteosarcoma is modulated by many pathways, such as mTORC1 pathway[12] and AMPK-dependent pathway[13]. Overall, metabolism of osteosarcoma is an essential part to understand the tumorigenesis and metastasis, and it is considered as a potential target for treating osteosarcoma[14-16].

Corresponding Author: Hongsheng Shang, MD; e-mail: 13809866626@163.com

There is a special balance between bone microenvironment and immune microenvironment, where a crosstalk between osteoblasts and monocytes-derived osteoclasts forms a special microenvironment[17]. For metastatic osteosarcoma patients, immunotherapy is a potential strategy to improve survival quality. Evidence[18] has shown that immune checkpoint blockade can partially rescue T cell function in a lung-metastatic mouse model[18]. However, to design a more specialized immunotherapy or to determine whether a patient is suitable to receive targeted therapy could be difficult to some extent.

In this study, we focused on metabolism of osteosarcoma and investigated the relation between metabolism and tumor microenvironment (TME). Two molecular subtypes were developed based on metabolism-related genes and we constructed a 5-gene signature for predicting osteosarcoma prognosis and guiding personalized treatment. Although previous studies[19,20] have proposed some signatures based on metabolism for osteosarcoma, they also show some limitations such as too many genes in one signature or lack of association with TME. Our study proposed a different strategy from previous research and categorized two molecular subtypes and developed a 5-gene signature with a close crosstalk with TME.

Materials and Methods

Data Source and Preprocessing

RNA-seq data of osteosarcoma samples with clinical information (named as TARGET cohort) were downloaded from Therapeutically Applicable Research To Generate Effective Treatments (TARGET, https://ocg.cancer.gov/programs/target) database in August 15, 2021. GSE39058 and GSE21257 cohorts with expression profiles and clinical information were downloaded from Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) database on August 15, 2021. Osteosarcoma samples with survival time and survival status were retained (Table I). For the two GSE cohorts, ComBat function in sva R package was applied to remove batch effects **(Supplementary Figure 1)**. Metabolism-related genes were obtained from previous research[21].

Unsupervised Consensus Clustering

Firstly, univariate Cox regression analysis was conducted in survival (V3.2-7) R package to screen metabolism-related genes in TARGET and GSE cohorts significantly associated with osteosarcoma prognosis ($p < 0.05$). The intersection of the screened genes between the two cohorts was selected for further analysis. Then, unsupervised consensus clustering in ConsensusClusterPlus (V1.52.0) R package was employed to classify samples into different subtypes[22]. Cluster number k from 2 to 10 was included to select the optimal clusters. The optimal k was confirmed by cumulative distribution function (CDF). Consensus clustering is a convenient and useful technique in cancer study providing stable classifications by using gene expression data.

Table I. Clinical information of osteosarcoma patients.

Feature	TARGET	GSE21257	GSE39058
Event			
Alive	55	30	29
Dead	29	23	12
Gender			
Female	37	19	20
Male	47	34	21
NA			
Age			
> 14	44	38	11
< =14	40	15	30
Unknown			
Relapse			
YES	38		20
NO	14		21
Unknown	32		
Metastatic			
YES	21	34	
NO	63	19	

SubMap Analysis

Gene Pattern SubMap analysis, which is a measurement that assesses the similarity of expression profiles between two independent datasets[21], was applied here to validate the effectiveness of molecular subtyping through comparing the gene expression data between C1 and C2 subtypes. In addition, SubMap analysis was also used to predict the response to immunotherapy by comparing the expression profiles of untreated patients with those of treated patients.

Functional Analysis

Differentially expressed genes (DEGs) were identified by limma R package under the conditions of |FC(fold change)| > 1.5 and FDR (false discovery rate) < 0.05[23]. For the screened DEGs, clusterProfiler R package was applied to annotate up-regulated genes (C1/C2) to Gene Ontology (GO, http://geneontology.org/) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.genome.jp/kegg/) pathways[24]. Under FDR < 0.05, GO terms and KEGG pathways were screened. ClusterProfiler enables the automation of biological-term classification and the enrichment of gene sets. Single sample gene set enrichment analysis (ssGSEA) in clusterProfiler was used to enrich significant pathways in "h.all.v7.4.symbols.gmt" and "c2.cp.kegg.v7.4.symbols.gmt".

Gene Set Variation Analysis (GSVA)

GSVA is a gene set enrichment method for calculating the enrichment score of gene sets in one sample or a group of samples based on RNA-seq data[25]. GSVA R package and GSEABase (V1.50.1) were implemented to evaluate the enrichment of human signatures and inflammatory signatures in the two subtypes. Kruskal-Wallis test was conducted to compare the difference of human signatures between the two subtypes.

Assessment of Tumor Microenvironment

The fraction of immune cells in TME was estimated by two measurements. Microenvironment Cell Populations-counter (MCP-counter) allows the quantification of 8 immune cells (T cells, CD8 T cells, cytotoxic lymphocytes, B lineage, NK cells, monocytic lineage, myeloid dendritic cells, neutrophils) and 2 stromal cells (endothelial cells and fibroblasts)[26]. MCP-counter can count the abundance and enrichment scores of 10 cell types in one sample across a cohort. Kruskal-Wallis test was conducted to compare the difference of immune infiltration between the two subtypes.

Estimation of STromal and Immune cells in MAlignant Tumours using Expression data' (ESTIMATE) is a method for quantifying the proportion of immune cells and stromal cells based on the expression of gene signatures[27]. Stromal score, immune score and ESTIMATE score were calculated for each sample, and ESTIMATE score is the combined score of stromal score and immune score. Kruskal-Wallis test was conducted to compare the difference of immune infiltration between the two subtypes.

The fraction of 28 biomarkers of immune-related cells obtained from a previous study was assessed by ssGSEA in GSVA R package[28]. The expression of immune checkpoints screened from previous studies[29,30] was evaluated in the two subtypes. Kruskal-Wallis test was conducted to compare the difference between the two subtypes.

Prediction of Sensitivity to Immunotherapy and Chemotherapy

Tumor Immune Dysfunction and Exclusion (TIDE) analysis was employed to evaluate T cell function, according to gene signatures for each sample[31]. TIDE score was calculated as a surrogate biomarker to predict response to immune checkpoint inhibitors, such as anti-PD-1 and anti-CTLA-4. Specifically, a higher TIDE score represents lower sensitivity to immune checkpoint blockade.

SubMap analysis was performed for comparing the gene expression patterns between subtypes and the estimated outcome of a patient receiving anti-PD-1 inhibitors of nivolumab and pembrolizumab (GSE93157)[32,33]. pRRophetic R package was used to predict the estimated biochemical half-maximal inhibitory concentration (IC50) of chemotherapeutic drugs including cisplatin, AKT inhibitor VIII, doxorubicin, embelin, etoposide, obatoclax mesylate, thapsigargin and vinorelbine[34].

Construction and Validation of a Prognostic Model

The intersection of metabolism-related genes significantly associated with prognosis between TARGET and GSE cohorts was filtered and subjected to Kaplan-Meier survival analysis. Log-rank test was performed in the survival analysis, and genes of $p < 0.05$ in both cohorts remained.

Then, stepwise Akaike information criterion (stepAIC) was conducted to reduce the number of genes[35]. Multivariate Cox regression analysis was performed to calculate the coefficients of the remained genes. The prognostic model was defined as: "Risk Score"=$\sum_{i=1}^{n}$coef(i)*gene(i), where i represents genes and coef represents coefficients of genes. Risk score of each sample in TARGET cohort was calculated, and surv_cutpoint function in survminer (V0.4.9) R package was employed to confirm the optimal cut-off in classifying samples into high-risk and low-risk groups (http://www.sthda.com/english/rpkgs/survminer/). GSE cohort was used to validate the prognostic model. Receiver operating characteristic of the model for predicting 1-year, 3-year and 5-year survival was shown by TimeROC R package[36]. Univariate and multivariate Cox regression analysis were performed to evaluate the independence of risk score as an indicator. Decision curve analysis (DCA) was used to assess the potential benefit that patients could receive based on different indicators.

Statistical Analysis

All statistical analysis was conducted in R (V4.1.1) software. Specific statistical methods were indicated in the corresponding figure legends. Parameters without specific description were default. $p < 0.05$ was considered as a significance. ns, no significance. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Results

Molecular Subtyping for Osteosarcoma Based on Metabolism-Related Genes

We obtained 200 metabolism-related genes from the previous research. Univariate Cox regression analysis on GSE and TARGET cohorts screened 156 genes and 274 genes, respectively ($p < 0.05$, Figure 1A). The intersection of the two cohorts were selected for consensus clustering analysis. From cluster number k = 2 to 10, suitable clusters were confirmed by CDF curve **(Supplementary Figure 2)**. Finally, k = 2 was selected, and samples were neatly classified into C1 and C2 subtypes (Figure 1B). Similarly, samples in TARGET cohort were also clustered into two groups (Figure 1C). Kaplan-Meier survival analysis showed that C1 and C2 subtypes had significantly differential overall survival (OS) in the GSE and TARGET cohorts, with $p = 0.016$ and 0.0068, respectively (Figure 1D and E). SubMap analysis revealed that C1 and C2 subtypes in the two cohorts manifested similar expression pattern ($p < 0.05$, Figure 1F), indicating that the molecular subtyping was effective.

Comparison of Clinical Features Between Two Molecular Subtypes

We also analyzed the relation between the two molecular subtypes and clinical features including survival status, ages, genders, relapse and metastasis. C1 subtype had a significantly higher proportion of samples in dead status ($p = 0.0332$, Figure 2A and B). Two subtypes did not show significant difference on the distribution of ages and genders ($p > 0.05$, Figure 2C and D). As for relapse possibility, C2 subtype was more likely to develop a relapse after treatment ($p = 0.0373$, Figure 2E). However, two subtypes manifested similar proportion in metastasis (Figure 2F). The results further illustrated a worse prognosis of C2 subtype than C1 subtype.

Enriched Pathways of DEGs Identified from C1 and C2 Subtypes

To evaluate whether there were differentially enriched pathways between the two subtypes, we firstly identified differentially expressed genes (DEGs) between them in GSE cohorts using limma R package. $p < 0.05$ and $|FC| > 1.5$ were the threshold to screen DEGs. Here, we filtered 955 DEGs incorporating 762 up-regulated and 193 down-regulated genes. Next, ClusterProfiler R package was applied to conduct functional analysis (KEGG pathways and GO terms) for the up-regulated genes in C1 subtype ($p < 0.05$, **(Supplementary Figure 3)**. Biological process terms related to osteosarcoma, such as ossification and osteoblast differentiation, were annotated **(Supplementary Figure 3)**. Tumor-related pathways including cell cycle and TGF-beta signaling pathway were significantly enriched **(Supplementary Figure 3)**. We conducted the same analysis on TARGET cohort and identified 495 DEGs between C1 and C2 subtypes. Functional analysis revealed that the up-regulated genes were closely associated with focal adhesion, proteoglycans in cancer, ECM-receptor interaction ($p < 0.05$, **(Supplementary Figure 4)**. The above results suggested that tumor-related pathways were greatly activated in C1 subtype. Further-

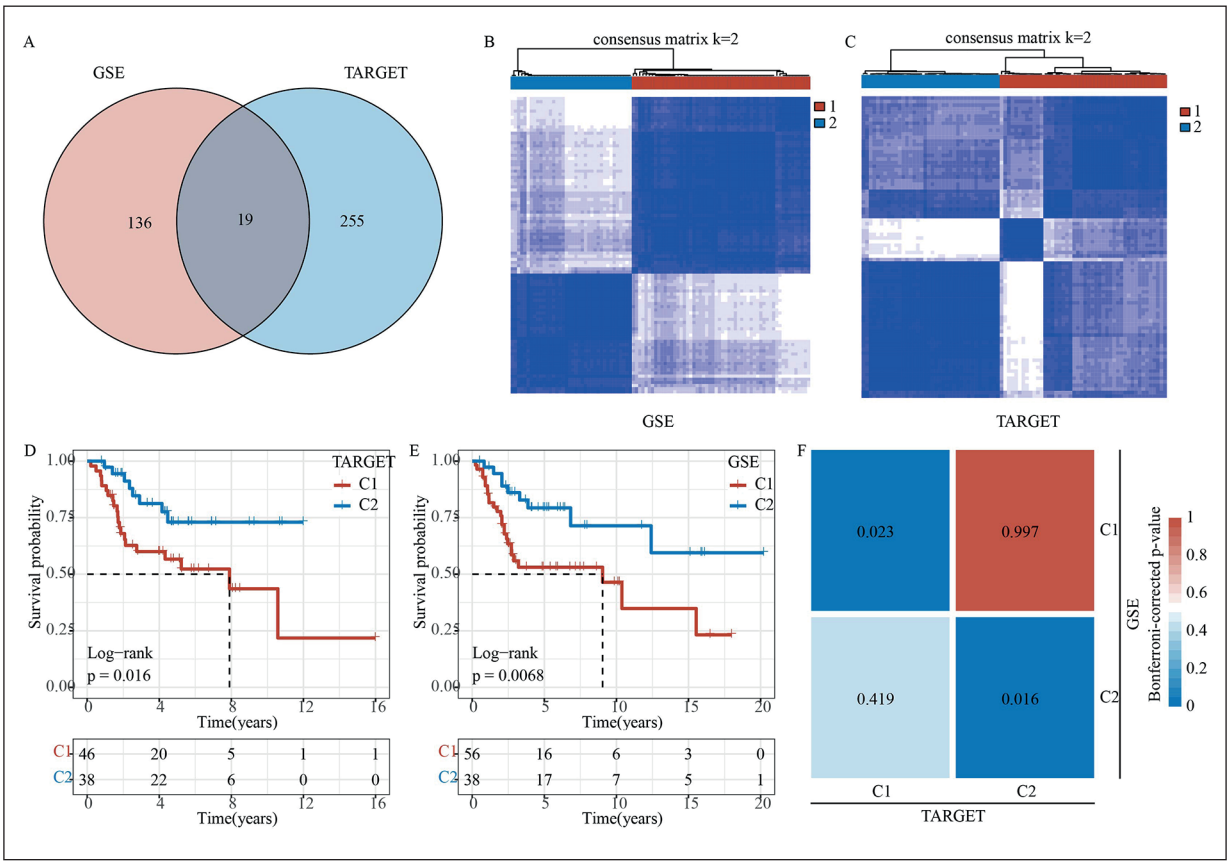

Figure 1. Identification of molecular subtypes based on metabolism-related genes. **A,** Venn plot of prognostic genes in TARGET and GSE cohorts. Blue indicates TARGET cohort and red indicates GSE cohorts. **B-C,** Consensus clustering when cluster number k = 2 in GSE and TARGET cohorts, respectively. Horizontal axis represents clusters where red indicates cluster 1 (C1) and blue indicates cluster 2 (C2). **D-E,** Kaplan-Meier survival plots of C1 and C2 subtypes in GSE and TARGET cohorts. Log-rank test was performed. **F,** SubMap analysis for analyzing the similarity of the same subtypes between two cohorts. *p*-value was corrected by Bonferroni correction. Lower *p*-value indicates higher similarity.

more, we performed ssGSEA to assess specific pathways from "h.all.v7.4.symbols.gmt" and "c2.cp.kegg.v7.4.symbols.gmt" in two subtypes. The results showed that in both GSE and TARGET cohorts C1 subtype was closely related to a high enrichment of tumor-related pathways, including epithelial mesenchymal transition (EMT), glycolysis, hypoxia, cell cycle, focal adhesion and ECM receptor interaction (Figure 3A and B). Meanwhile, immune-related pathways including natural killer cell-mediated cytotoxicity, cytokine-cytokine receptor interaction, toll-like receptor signaling pathway, primary immunodeficiency and autoimmune thyroid disease were noticeably enriched in C2 subtype (Figure 3C and D).

C2 Subtype Is More Immune Infiltrated Than C1 Subtype

We analyzed the immune infiltration of the two subtypes in GSE and TARGET cohorts using three methods (ESTIMATE, MCP-counter and ssGSEA). In GSE cohorts, ESTIMATE score, immune score and stromal score calculated by ESTIMATE were significantly higher in C2 than C1 ($p < 0.05$, Figure 4A). MCP-counter assessed the estimated proportion of 9 immune-related cells, and we observed that 7 out of 9 immune-related cells including B lineage, cytotoxic lymphocytes, monocytic lineage, myeloid dendritic cells, neutrophils, natural killer (NK) cells and T cell showed a higher enrichment score in C2 subtype ($p < 0.05$, Figure 4B). We then used ssGSEA to assess the distribution of 28 immune-related cells obtained from the previous research[28] in GSE cohorts. The data showed that compared with C1 subtype, C2 subtype had higher proportion of most immune-related cells such as activated B cells, activated CD4 T cells, activated CD8 T cells, and activated dendritic cells ($p < 0.001$, Figure 4C). We

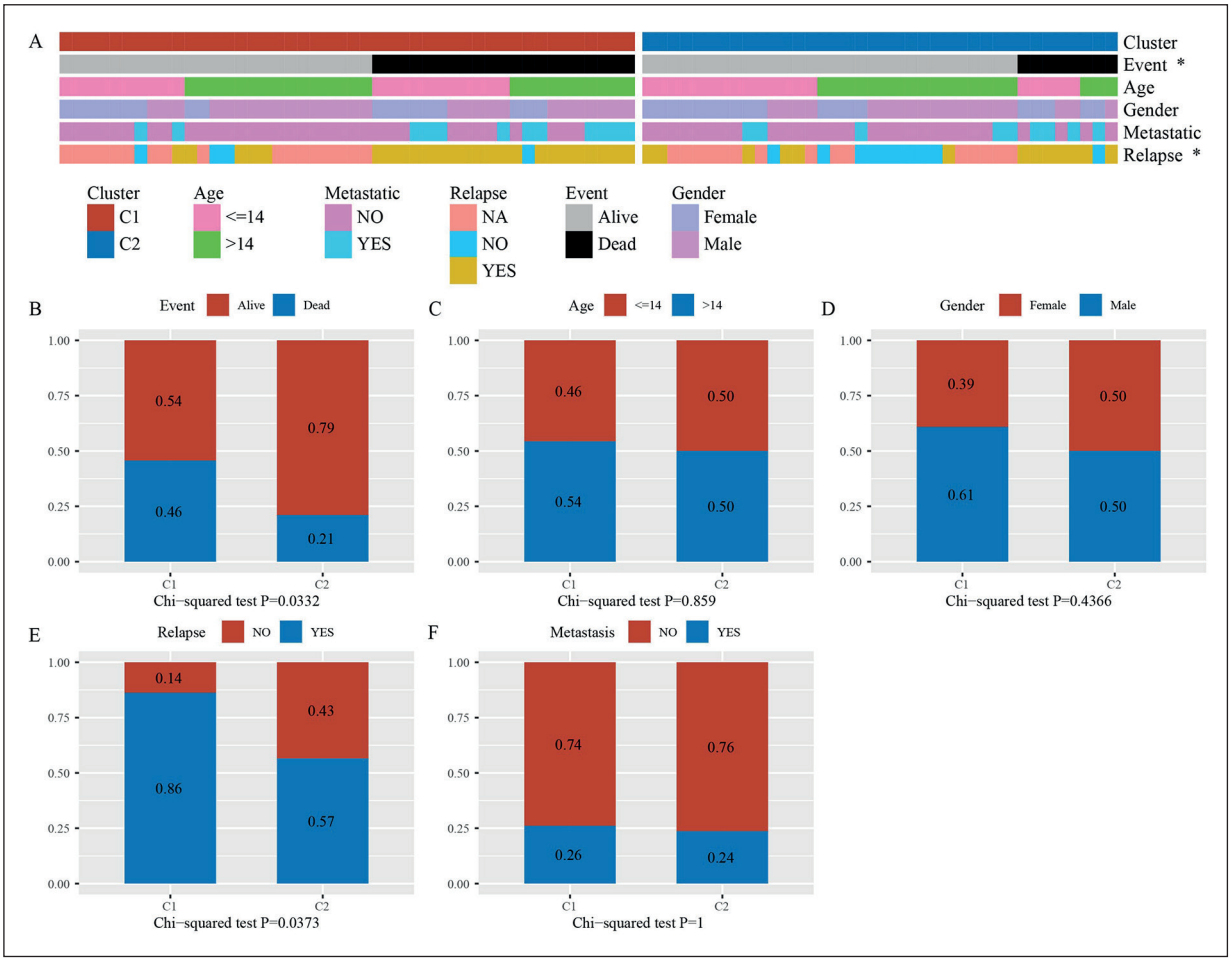

Figure 2. The distribution of two subtypes in clinical features (**A**) including survival status (**B**), ages (**C**), genders (**D**), relapse (**E**), and metastasis (**F**). Chi-square test was performed.

applied the same analysis in TARGET cohorts, and similar results to GSE cohorts was found **(Supplementary Figure 5)**.

In addition, we obtained a series of immune checkpoint from previous studies[29,30], and calculated their expression using ssGSEA. In GSE cohorts, immune checkpoints, such as CTLA-4 and PDCD1, were much higher-expressed in C2 subtype ($p < 0.05$, Figure 5). Moreover, the similar results were observed in TARGET cohort **(Supplementary Figure 6)**. The above results indicated that C2 subtype, which had more favorable prognosis, exhibited more activated immune response than C1 subtype.

Different Enrichment of Human Signatures Between C1 and C2 Subtypes

To further characterize the differential features between the two subtypes, we included 15 human signatures from the previous research[37]. 7 out of 15 signatures including CD8 T effector, cell cycle, DNA damage repair (DDR), DNA replication, mismatch repair, nucleotide excision repair, and WNT target were differentially enriched between C1 and C2 subtypes ($p < 0.05$, Figure 6). Apart from CD8 T effector, other signatures related to cell cycle and DNA repair all showed a higher enrichment in C1 subtype. In TARGET cohort, we observed the similar results, although there was no significance in some signatures between the two subtypes **(Supplementary Figure 7)**.

C2 Subtype Is Predicted to Be Sensitive to Immunotherapy by TIDE Analysis

As differential TME between C1 and C2 subtypes was examined, TIDE was applied to predict their sensitivity to immunotherapy. Scores of three aspects (T cell dysfunction, T cell exclusion and TIDE score) were calculated by TIDE based

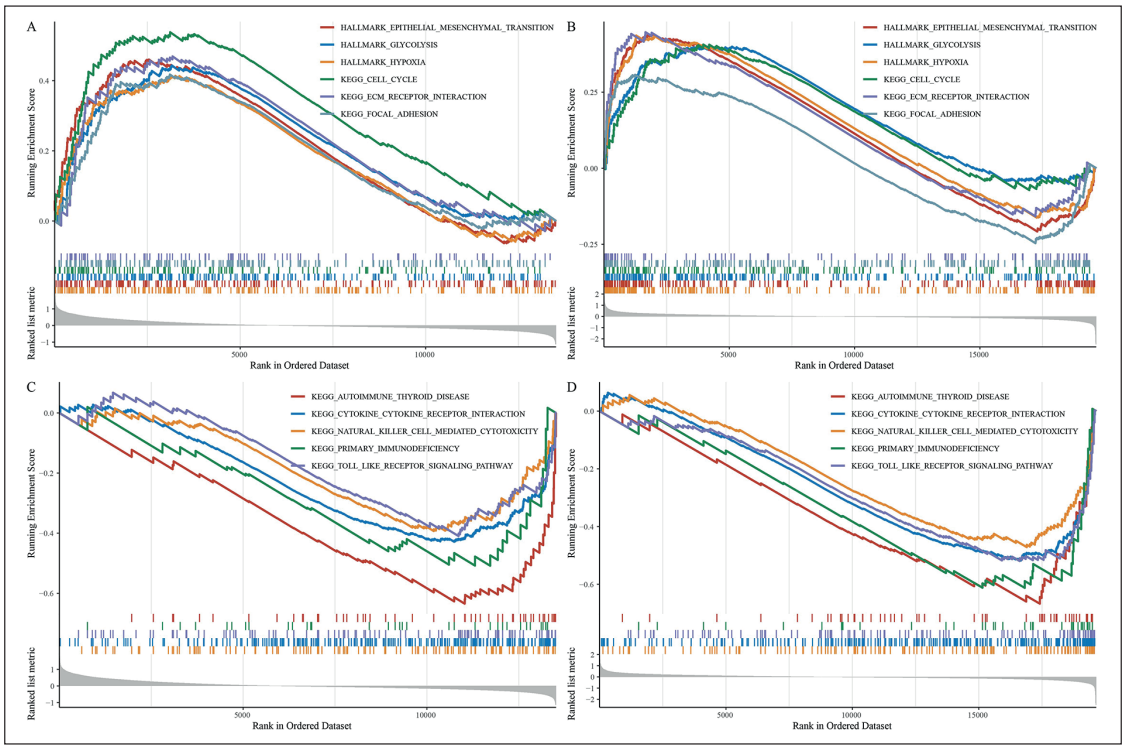

Figure 3. Enriched pathways of C1 subtype in GSE (**A**) and TARGET (**B**) cohorts, and C2 subtype in GSE (**C**) and TARGET (**D**) cohorts. Horizontal axis indicates samples, and vertical axis indicates enrichment score of pathways. Five pathways were labeled in different colors.

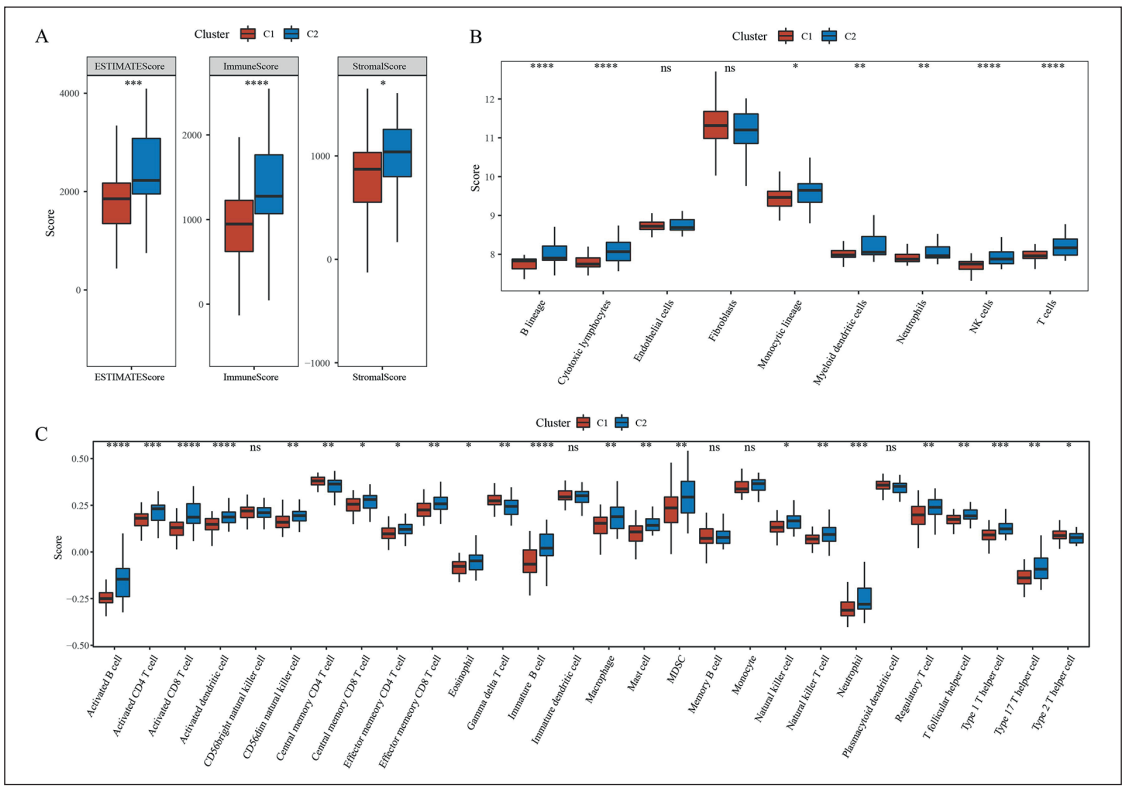

Figure 4. Immune infiltration analysis of two molecular subtypes in GSE cohorts by MCP-counter (**A**), ESTIMATE (**B**) and ssGSEA (**C**). Kruskal-Wallis test was conducted. ns, no significance. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

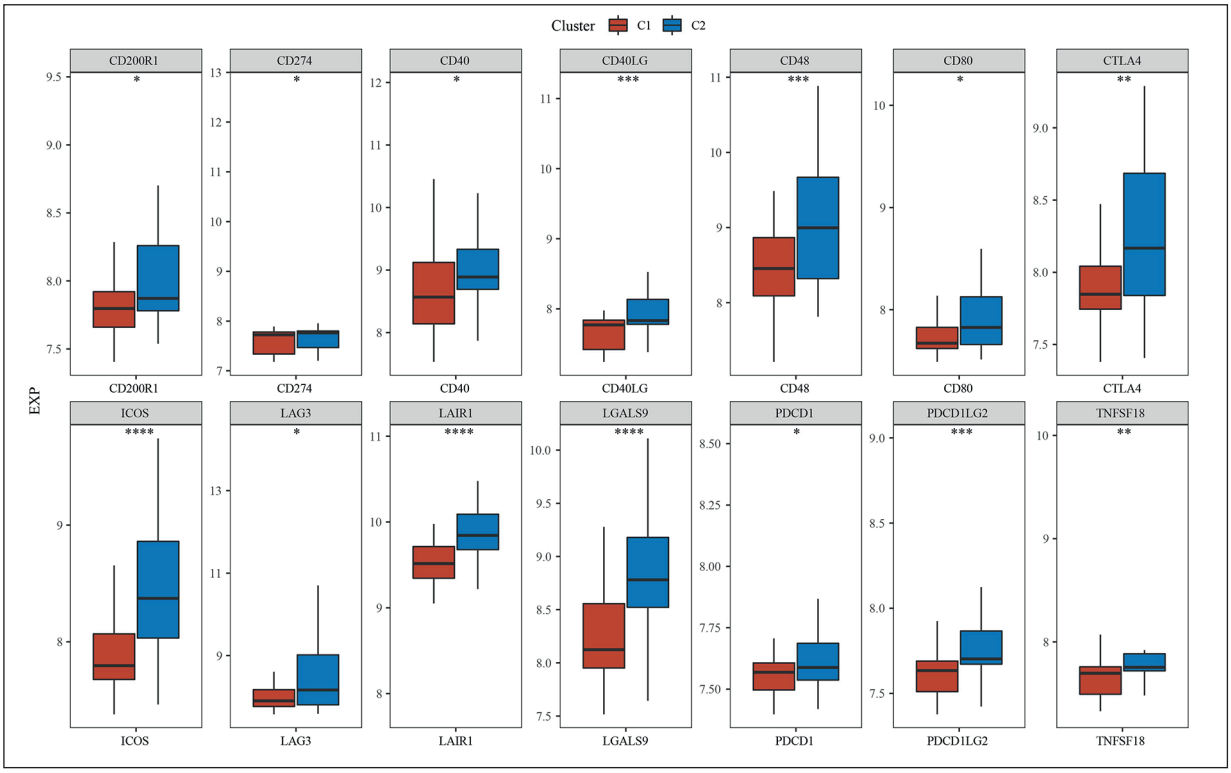

Figure 5. Comparison of immune checkpoints expression between C1 and C2 subtypes in GSE cohorts. Kruskal-Wallis test was conducted. Red indicates C1 subtype and blue indicates C2 subtype. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

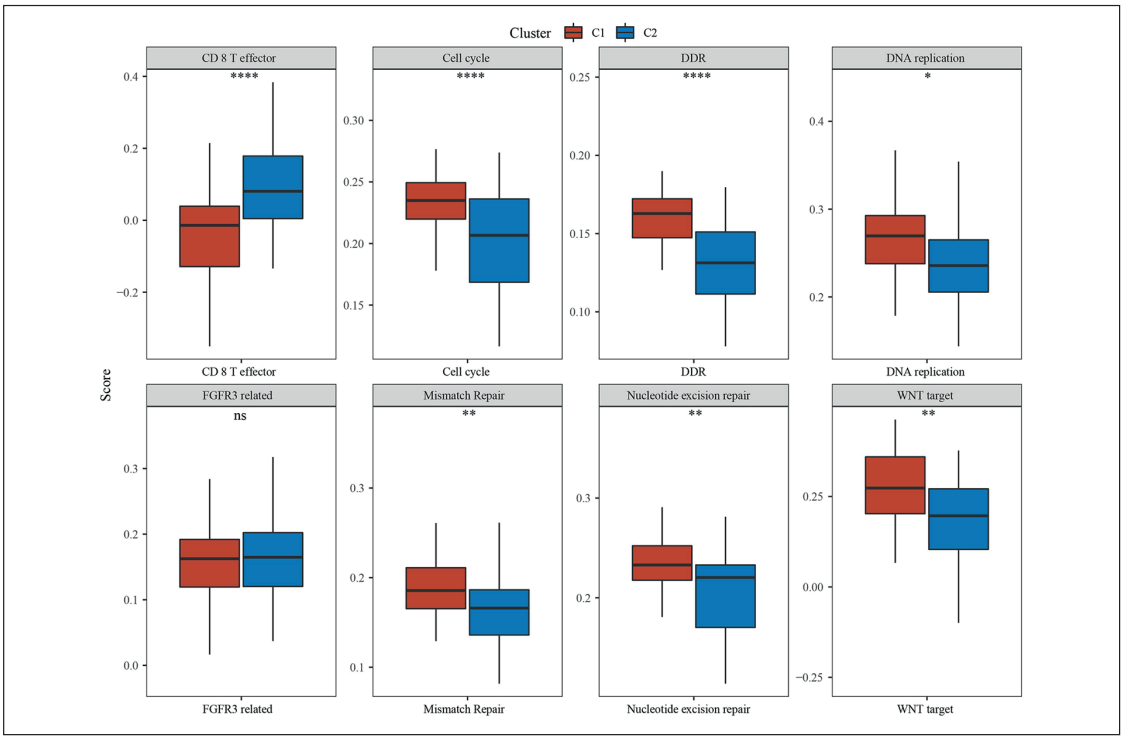

Figure 6. Analysis of human signatures including CD8 T effector, cell cycle, DDR, DNA replication, FGFR3-related pathways, mismatch repair, nucleotide excision repair and WNT target. Student *t*-test was conducted. Red indicates C1 subtype and blue indicates C2 subtype. ns, no significance. *$p < 0.05$, **$p < 0.01$, ****$p < 0.0001$.

on the corresponding gene signatures. In GSE cohorts, C1 subtype had lower score of T cell dysfunction than C2 subtype, indicating a less damage of T cells in C1 subtype ($p < 0.0001$, Figure 7A). However, higher score of T cell exclusion was shown in C1 subtype ($p < 0.0001$, Figure 7A), which could result in a suppressed immune response. Moreover, C2 subtype manifested a lower TIDE score, suggesting a higher sensitivity to immunotherapy ($p < 0.05$). The same conclusion was also found in TARGET cohort ($p < 0.05$, Figure 7B), which indirectly proved the robustness of our molecular subtyping.

Different Responses of Two Subtypes to Immunotherapy and Chemotherapy

To predict the responses to immunotherapy more precisely, we used the expression profiles of another dataset (GSE93157) with treatment data of anti-PD-1 (nivolumab) and anti-PD-L1 (pembrolizumab) and compared the expression pattern between GSE93157 and GSE cohorts by SubMap analysis. Compared with GSE cohorts, C2 subtype showed a similar expression pattern with patients treated by nivolumab in GSE93157 dataset ($p = 0.001$, Figure 8A), while C1 subtype was insensitive to nivolumab and pembrolizumab. In addition, we further examined their sensitivity to chemotherapeutic drugs using estimated IC50. The current data revealed that the two subtypes displayed significantly differential responses to all the 8 chemotherapeutic drugs ($p < 0.0001$, Figure 8B). C2 subtype had a higher estimated IC50 to all the 8 drugs than C1 subtype, indicating that C2 subtype was more sensitive to these chemotherapeutic drugs. This also suggested that C2 subtype could greatly benefit from immunotherapy and chemotherapy compared to C1 subtype, which was also consistent with the result of TIDE prediction.

Constructing a 5-Gene Prognostic Model for Osteosarcoma Based on Metabolism-Related Genes

The intersection of metabolism-related genes associated with prognosis between GSE and TARGET cohorts was selected, and 19 genes remained.

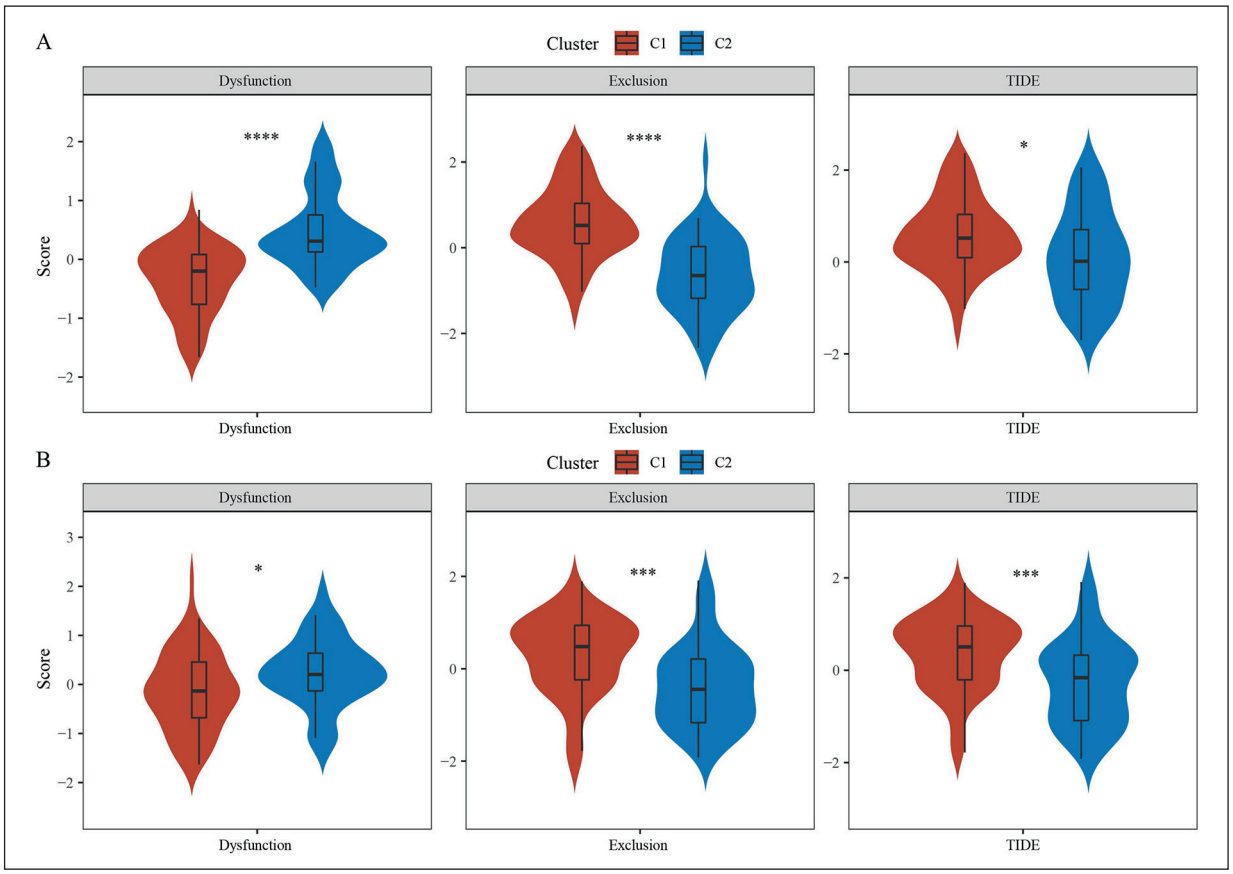

Figure 7. TIDE analysis for predicting the sensitivity of two subtypes to immunotherapy in GSE **(A)** and TARGET **(B)** cohorts. Red indicates C1 subtype and blue indicates C2 subtype. Dysfunction indicates T cell dysfunction and exclusion indicates T cell exclusion. Student *t*-test was conducted. *$p < 0.05$, ***$p < 0.001$, ****$p < 0.0001$.

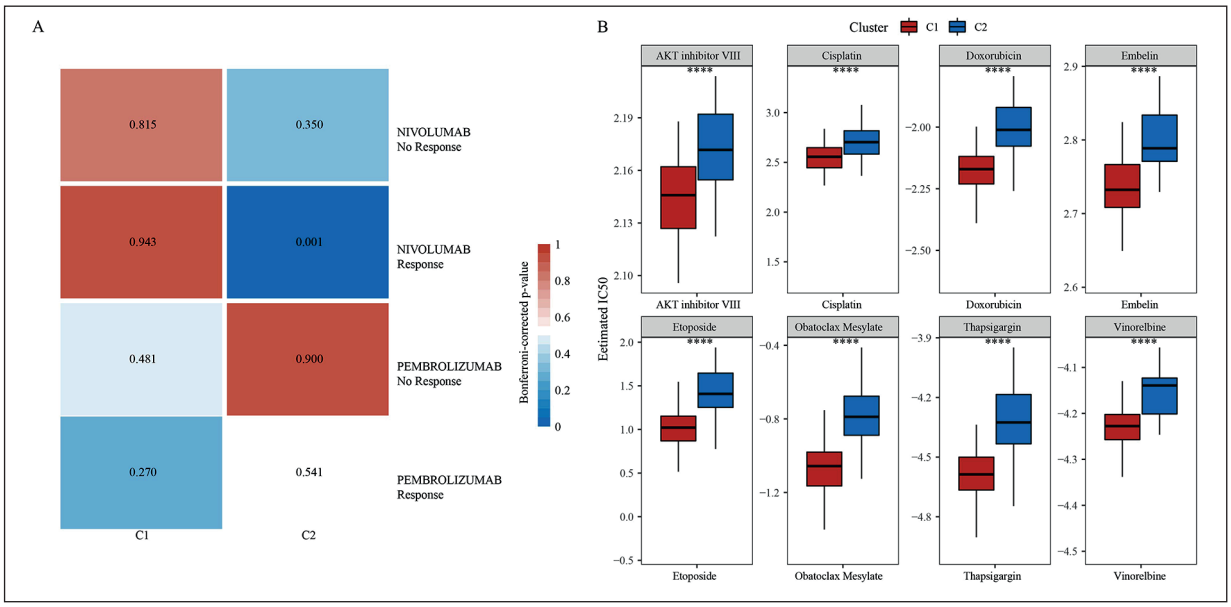

Figure 8. Predicted responses to immunotherapy and chemotherapy of C1 and C2 subtypes. (**A**) SubMap analysis between GSE and GSE93157 cohorts in response to nivolumab and pembrolizumab. (**B**) Estimated IC50 of C1 and C2 subtypes to eight chemotherapeutic drugs in GSE cohorts. Student *t*-test was conducted to test the significance between two groups. ns, no significance. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$.

Survival analysis was used to further screen these 19 metabolism-related genes, and 7 genes (CBS, ACSL5, ATP6V0D1, DDAH2, PDE4C, SLC7A1 and PNPO) that were capable of neatly classifying patients into C1 and C2 subtypes were selected (**Supplementary Figure 8** and **Supplementary Figure 9**). Furthermore, multivariate Cox regression analysis and stepAIC were performed to confirm the prognostic genes in TARGET cohort. Finally, 5 genes (CBS, ACSL5, DDAH2, PDE4C, and PNPO) were determined (Figure 9). The prognostic model was defined

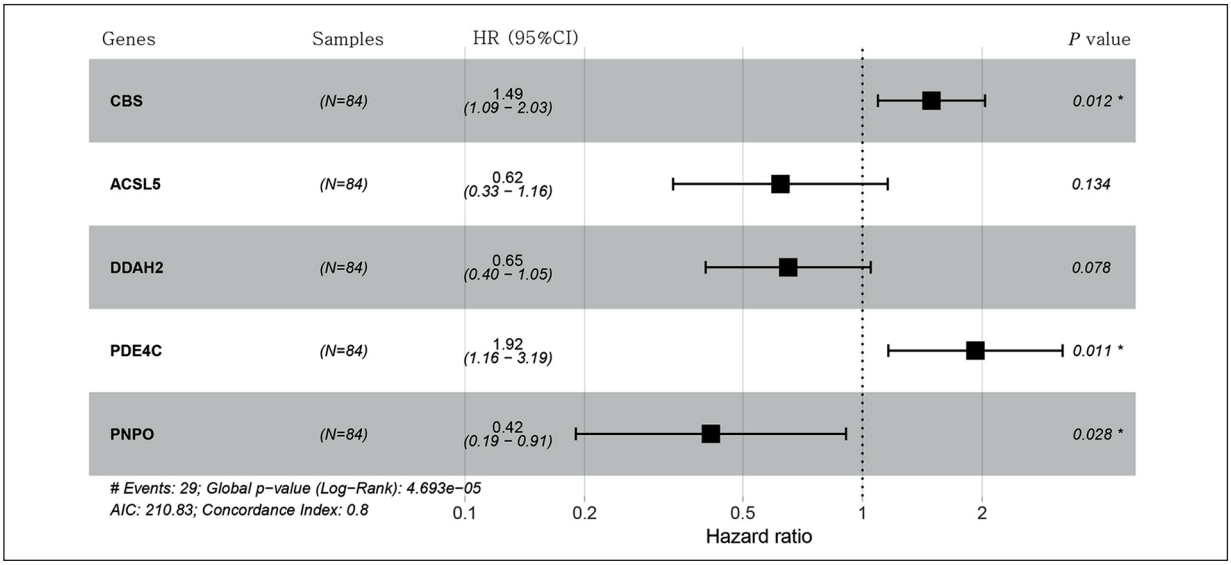

Figure 9. Multivariate Cox regression analysis of 5 prognostic genes. The bottom numbers indicate hazard ratio (HR). The columns from left to right represent genes, total samples, HR (95% confidence interval), and *p*-values.

as: risk score = 0.400*CBS-0.475*ACSL5-0.430*DDAH2+0.654*PDE4C-0.878*PNPO. Risk score of each sample in TARGET cohort was calculated, and median value was selected as a cut-off. Samples were significantly stratified into high-risk and low-risk groups ($p < 0.0001$, Figure 10A), and ROC analysis showed a favorable AUC of 1-year, 3-year and 5-year survival of 0.79, 0.87 and 0.81, respectively (Figure 10B). In GSE cohorts, the samples were also neatly stratified into high-risk and low-risk groups ($p = 0.0048$, Figure 10C). High AUC score for 1-year, 3-year and 5-year survival was found, demonstrating the stability of the 5-gene prognostic model (Figure 10D).

Risk Score is Highly Associated with Clinical Features

The relation between risk score and survival was demonstrated, and then we analyzed the distribution of risk score in other clinical features such as in metastasis, relapse, genders, ages, and molecular subtypes. The result presented that metastatic patients had a higher risk score than non-metastatic patients ($p < 0.05$, Figure 11A). C1 subtype faced a higher risk than C2 subtype, which was consistent with the previous result ($p < 0.01$, Figure 11B). However, risk score was similarly distributed between relapse and non-relapse, females and males, ≤ 14 and > 14

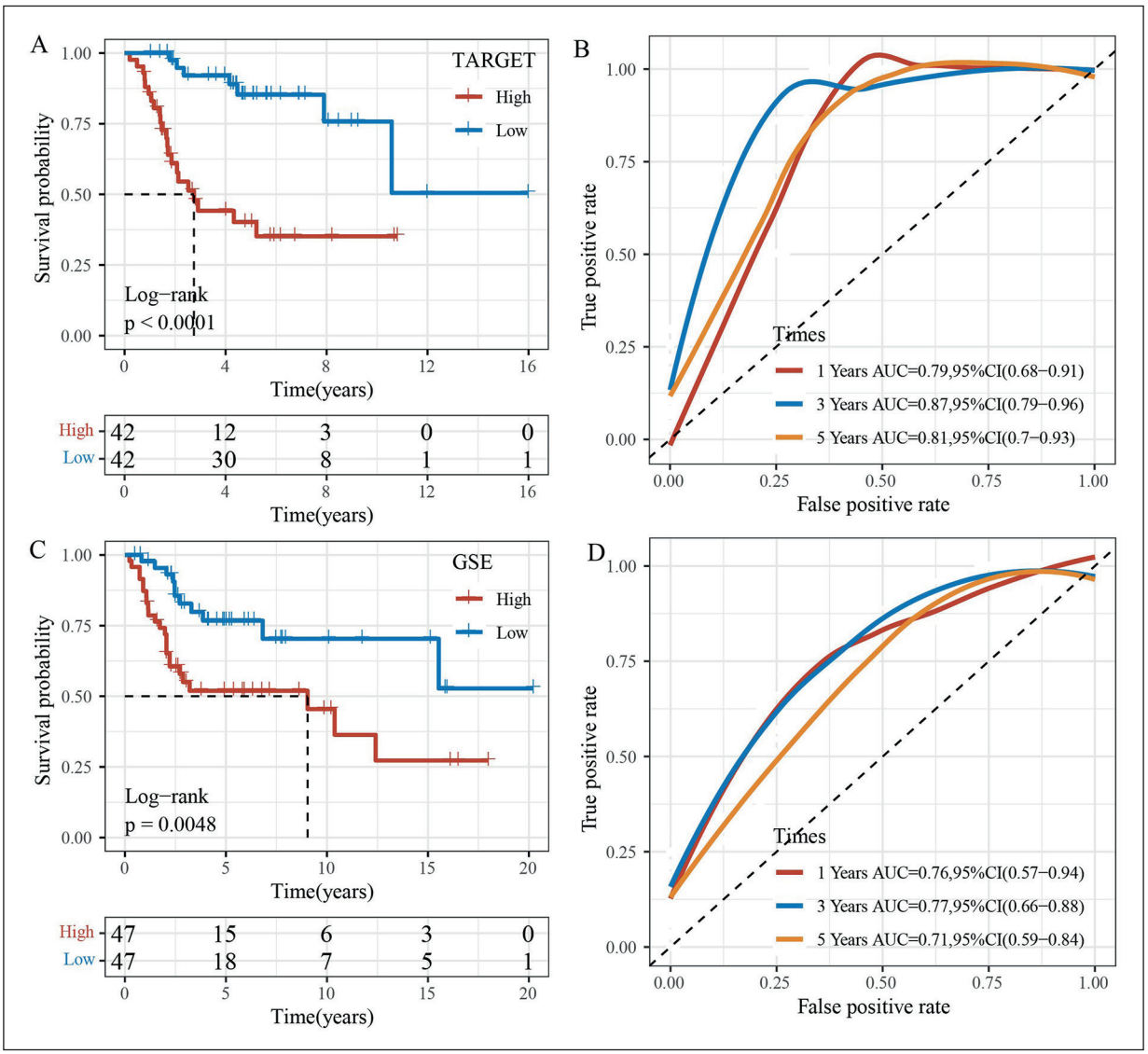

Figure 10. Kaplan-Meier survival plots and ROC analysis of the 5-gene prognostic model in TARGET (**A-B**) and GSE (**C-D**) cohorts. Log-rank test was conducted in survival analysis.

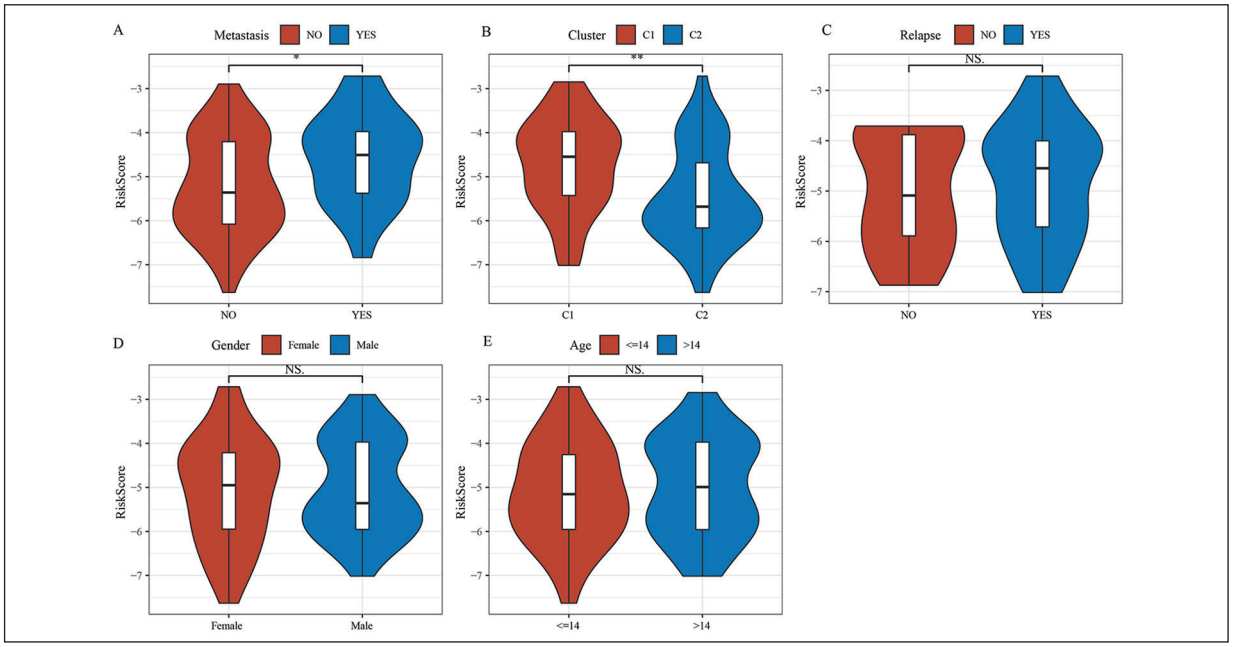

Figure 11. The distribution of risk score in different features including metastasis (**A**), molecular subtypes (**B**), relapse (**C**), genders (**D**) and ages (**E**). Student *t*-test was conducted. NS, no significance. *$p < 0.05$, **$p < 0.01$.

years old (Figure 11C-E). In different clinical features, risk score was stable and effective in stratifying samples into high-risk and low-risk groups, except for the samples with non-relapse ($p < 0.05$, Figure 12). Furthermore, univariate and multivariate Cox regression analysis showed that risk score was the most significant risk factor among other clinical features, with HR = 2.7 (95%CI = 1.8-4.1, p = 1.7e-06) and HR = 2.7 (95%CI = 1.7-4.3, $p <$ 1e-5), respectively (Figure 13).

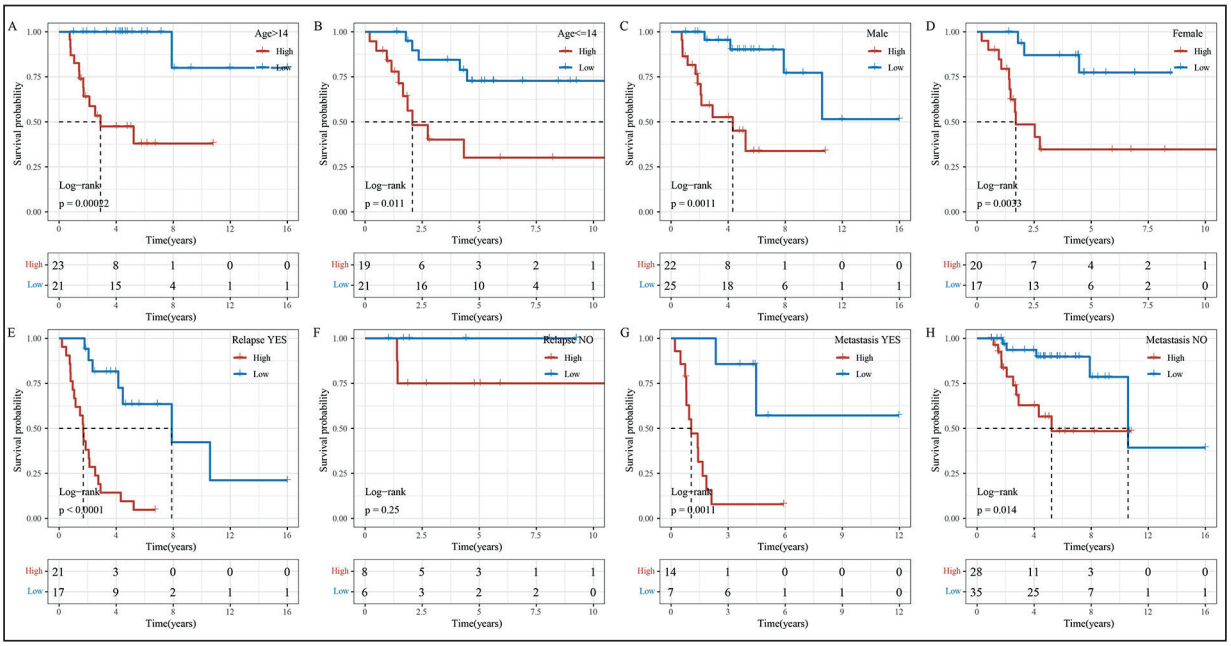

Figure 12. Survival analysis of high-risk and low-risk groups in different clinical features including ages (**A-B**), genders (**C-D**), relapse (**E-F**), metastasis (**G-H**). Log-rank test was conducted.

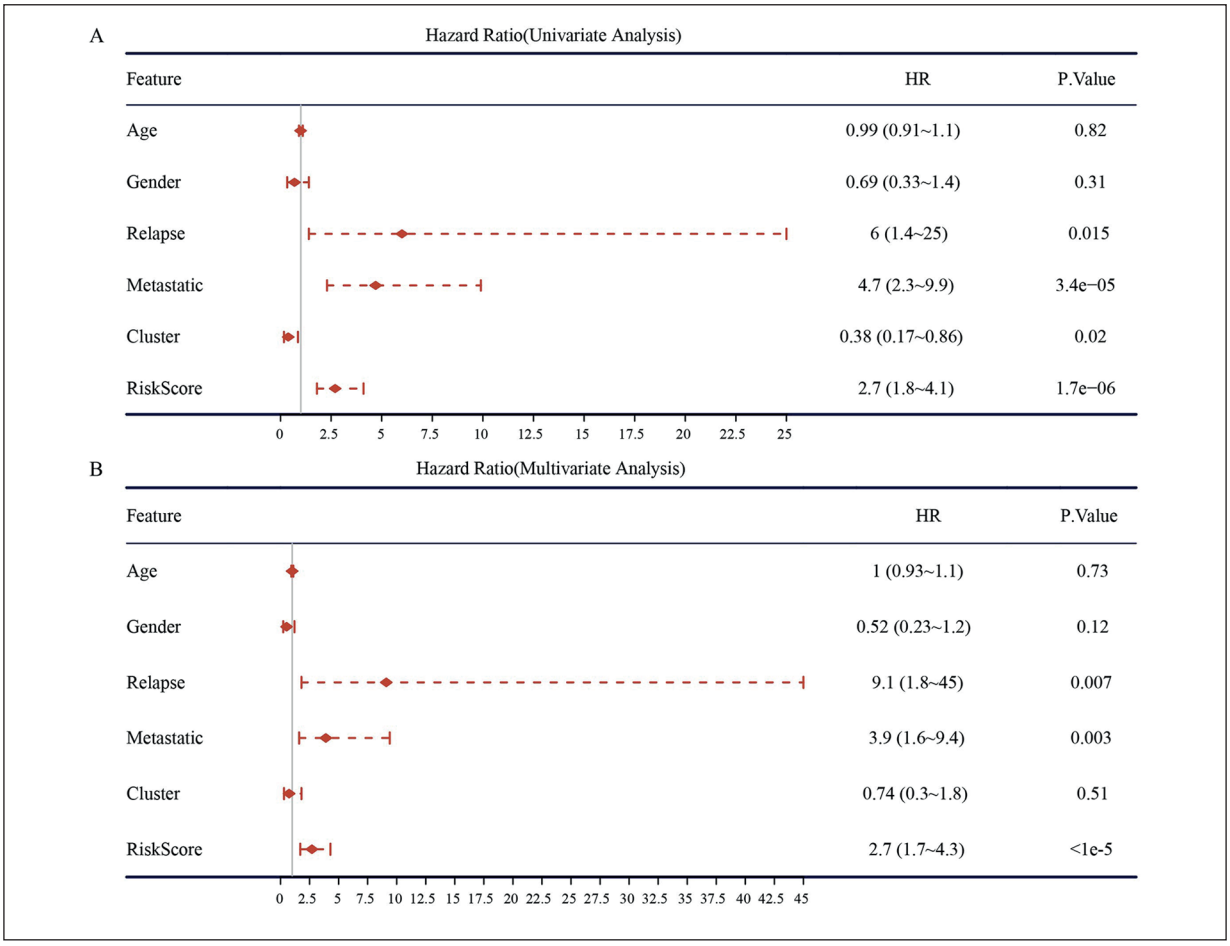

Figure 13. Univariate (**A**) and multivariate (**B**) Cox regression analysis of clinical features, subtypes and risk score. Log-rank test was performed. HR, hazard ratio.

A Nomogram Based on Risk Score for Clinical Application

To promote the applicability of the prognostic model in clinical practice, we developed a nomogram based on risk score, metastasis, and relapse in TARGET cohorts (Figure 14A). Each sample could gain a total score by combining the scores from the three features, and 1-year, 3-year and 5-year survival rate could be predicted. The predicted 1-year, 3-year and 5-year survival rates were corrected by the actual survival (Figure 14B). Decision curve analysis demonstrated that the nomogram exhibited the highest standardized net benefit among these indicators, indicating its optimal performance in assisting clinical prediction (Figure 14C).

Risk Score Is Negatively Correlated with Inflammatory Signatures

We explored the association between risk score and inflammatory signatures including HCK, IgG, interferon, LCK, MHC-I, MHC-II and STAT1. A heatmap presented that these signatures were highly enriched in low-risk group, and strong negative correlations were observed between risk score and inflammatory signatures ($p < 0.001$, Figure 15). The result further demonstrated that risk score was closely associated with inflammatory response, and indirectly indicated a strong correlation between metabolism and inflammatory response.

Discussion

Previous studies[19,38,39] have discovered the close relationship between metabolism and osteosarcoma development and metastasis. Based on metabolism-related genes, Li et al[19] identified a metabolism-related signature composed of 39 prognostic genes for osteosarcoma. The score calculated by the signature was positively associated with im-

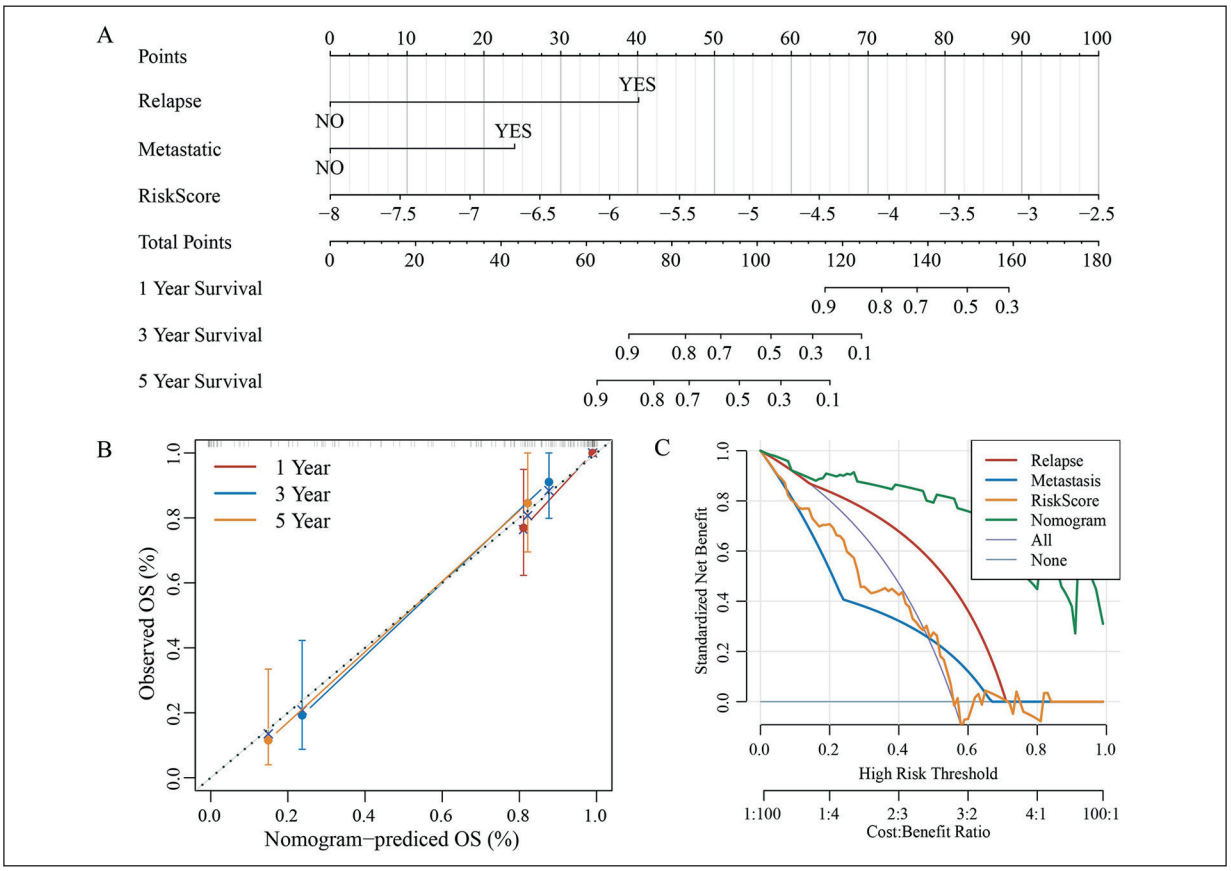

Figure 14. Construction and validation of a nomogram. **A,** A nomogram based on relapse, metastasis and risk score for predicting 1-year, 3-year and 5-year survival rate. **B,** Correction of predicted overall survival (OS) by observed OS. **C,** DCA curve for validating the effectiveness of four indicators including relapse, metastasis, risk score and the nomogram.

mune infiltration, supporting the important effect of metabolism in TME. Although the signature was effective to predict prognosis and provided a guidance to immunotherapy, the number of genes in the signature was too many for clinical use. In another study, Lussier et al[18] developed a 7-gene signature based on metabolism-related genes through "NFM" algorithm and weighted gene co-expression network analysis (WGCNA)[20]. However, the study focused less on the relation between metabolism and TME, but much more on the advantage of the signature.

In the current study, we included 200 metabolism-related genes and performed Cox regression analysis to filter genes associated with prognosis. Unlike the above-mentioned studies, the intersected genes between TARGET and GSE cohorts were determined as a basis for consensus clustering. Two molecular subtypes were categorized to have a close association with patients' prognosis in the two cohorts, where C1 subtype had a worse overall survival. GSEA revealed the enriched KEGG pathways of two subtypes respectively, and six oncogenic pathways, including EMT, glycolysis, hypoxia, cell cycle, focal adhesion, and ECM receptor interaction pathways, were highly enriched in C1 subtype. This result indicated that metabolism-related genes were strongly involved in tumorigenesis of osteosarcoma and may serve as modulators on these oncogenic pathways.

EMT, which can promote cancer progression and metastasis, is a critical step in the pathogenesis of many cancer types, including in osteosarcoma. EMT is modulated by complicated signaling pathways such as Wnt/β-catenin signaling pathway[40] and PI3K/AKT/mTOR pathway[41]. The EMT process could be significantly suppressed[42] *via* inhibiting Wnt/β-catenin signaling pathway, suggesting that these EMT-associated pathways may be considered as potential targets to treat osteosarcoma. In the relation of metabolism to EMT, Zhang et al[43] observed that leukocyte-associated immunoglobulin-like receptor-1 (LAIR-1), a collagen receptor, could inhibit EMT

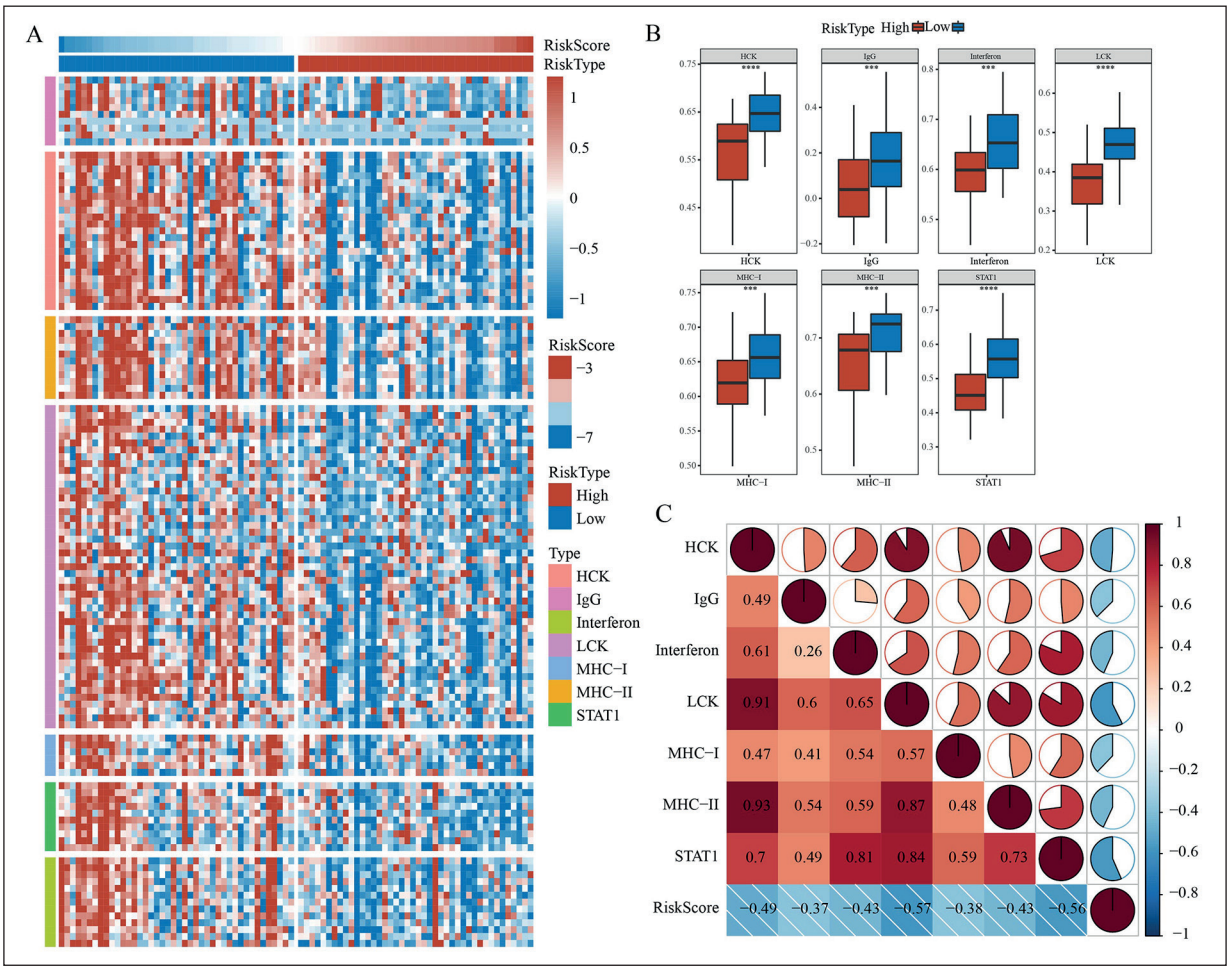

Figure 15. The relation between risk score and inflammatory signatures. **A,** A heatmap presenting the expression of inflammatory signatures in high-risk and low-risk groups. Horizontal axis represents risk score and vertical axis represents each gene in inflammatory signatures. Red indicates relatively high expression and blue represents relatively low expression. **B,** The differential enrichment of inflammatory signatures in high-risk and low-risk groups. Kruskal-Wallis test was conducted. **C,** Pearson correlation analysis among risk score and inflammatory signatures. Red represents negative correlation and blue represents positive correlation. Correlation coefficients were presented in box.

by down-regulating the expression of metabolism-related molecule of glucose transporter 1. Glycolysis, one of metabolism processes, is greatly activated in tumor cells, especially when aerobic glycolysis allows tumor cells to adapt to hypoxia and resists chemotherapeutic drugs[44]. In C1 subtypes, glycolysis and hypoxia pathways were both enriched, which provided an enhanced microenvironment for tumor progression. Shen et al[45] have found that CircECE1 regulates glycolysis through c-Myc/TXNIP axis, and CircECE1 knockdown interferes tumor cell proliferation in osteosarcoma cells.

Immunotherapy is a hopeful treatment for metastatic osteosarcoma patients[46]. The efficiency of immune checkpoint inhibitors such as anti-PD-1 largely depends on the expression of immune checkpoints and T cell status. The component and activity of TME can decide patient's response to immunotherapy. Several clinical trials on immune checkpoint inhibitors are undergoing. A phase-2 trial of anti-PD-1 mAb pembrolizumab revealed that only 1 of 22 osteosarcoma patients shows a partial response[47]. Although the outcome is dismal, it does prove that immune checkpoint blockade can be beneficial to osteosarcoma patients. Therefore, to reach an optimized therapy, characterizing TME for osteosarcoma patient can facilitate the design of a suitable treatment.

In our study, TME of the two subtypes exhibited a great difference that C2 subtype was more active in immune response to tumor cells than

C1 subtype. Anti-tumor immune cells including activated B cells, activated CD4 T cells, activated CD8 T cells and NK cells were significantly enriched in C2 subtype, which contributed to a more favorable prognosis. The predicted immune response to immune checkpoint blockade was also more activated in C2 subtype, which was possibly resulted from a higher expression of immune checkpoints in C2 subtype. Although TIDE prediction manifested that T cell function was better in C1 subtype, T cell exclusion was more severe in C1, and this could lead to a lower efficiency of immunotherapy. In addition, C2 subtype could benefit much from four chemotherapeutic drugs, therefore, this molecular subtyping could be much more helpful in guiding the personalized management to osteosarcoma patients with metastasis.

Based on metabolism-related genes, we constructed a 5-gene prognostic model with robust performance to predict prognosis for osteosarcoma patients. Through correlation analysis with inflammatory signatures, we found that all the seven signatures were higher-expressed in low-risk group, suggesting that inflammatory response was positively correlative with prognosis. For example, interferons are a group of cytokines with anti-tumor property against osteosarcoma in xenograft models[44]. High inflammation of C2 subtype with favorable prognosis indicated that metabolism-related pathways could serve as an important role in regulating inflammatory response in TME. In addition, we used risk score together with clinical features to build a nomogram that could be more conveniently applied in clinical practice. Importantly, the nomogram exhibited a better performance than the 5-gene prognostic model in predicting overall survival.

Some limitations of this study should be noted. Firstly, the samples lacked some clinical follow-up information, particularly some diagnostic detail information, therefore we did not consider factors such as the presence of other health states of the patient to distinguish diagnostic biomarkers. Secondly, the results were obtained only by bioinformatics analysis, which was insufficient; moreover, an experimental validation is needed to confirm these results.

Conclusions

This study proposed two molecular subtypes based on metabolism-related genes. Two novel subtypes showed distinct differences in prognosis, enriched pathways and TME, which strongly supported that metabolism played a critical role in tumorigenesis. Notably, the subtyping had a great significance in providing assistance in clinical decision-making, especially for metastatic osteosarcoma patients. Furthermore, we established a 5-gene prognostic signature to predict survival for osteosarcoma patients. Overall, the study emphasized the importance of metabolism in tumor development, and further revealed the crosstalk between metabolism and TME responsible for anti-tumor activity.

Conflict of Interest

The Authors declare that they have no conflict of interests.

References

1) Mirabello L, Troisi RJ, Savage SA. Osteosarcoma incidence and survival rates from 1973 to 2004: data from the Surveillance, Epidemiology, and End Results Program. Cancer 2009; 115: 1531-1543.
2) Savage SA, Mirabello L. Using epidemiology and genomics to understand osteosarcoma etiology. Sarcoma 2011; 2011: 548151.
3) Casali PG, Bielack S, Abecassis N, Aro HT, Bauer S, Biagini R, Bonvalot S, Boukovinas I, Bovee J, Brennan B, Brodowicz T, Broto JM, Brugières L, Buonadonna A, De Álava E, Dei Tos AP, Del Muro XG, Dileo P, Dhooge C, Eriksson M, Fagioli F, Fedenko A, Ferraresi V, Ferrari A, Ferrari S, Frezza AM, Gaspar N, Gasperoni S, Gelderblom H, Gil T, Grignani G, Gronchi A, Haas RL, Hassan B, Hecker-Nolting S, Hohenberger P, Issels R, Joensuu H, Jones RL, Judson I, Jutte P, Kaal S, Kager L, Kasper B, Kopeckova K, Krákorová DA, Ladenstein R, Le Cesne A, Lugowska I, Merimsky O, Montemurro M, Morland B, Pantaleo MA, Piana R, Picci P, Piperno-Neumann S, Pousa AL, Reichardt P, Robinson MH, Rutkowski P, Safwat AA, Schöffski P, Sleijfer S, Stacchiotti S, Strauss SJ, Sundby Hall K, Unk M, Van Coevorden F, van der Graaf WTA, Whelan J, Wardelmann E, Zaikova O, Blay JY. Bone sarcomas: ESMO-PaedCan-EURACAN Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2018; 29: iv79-iv95.
4) Zhang Y, Yang J, Zhao N, Wang C, Kamar S, Zhou Y, He Z, Yang J, Sun B, Shi X, Han L, Yang Z. Progress in the chemotherapeutic treatment of osteosarcoma. Oncol Lett 2018; 16: 6228-6237.
5) 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.
6) McCoy A, Besch-Williford CL, Franklin CL, Weinstein EJ, Cui X. Creation and preliminary charac-

terization of a Tp53 knockout rat. Dis Model Mech 2013; 6: 269-278.

7) Chen X, Bahrami A, Pappo A, Easton J, Dalton J, Hedlund E, Ellison D, Shurtleff S, Wu G, Wei L, Parker M, Rusch M, Nagahawatte P, Wu J, Mao S, Boggs K, Mulder H, Yergeau D, Lu C, Ding L, Edmonson M, Qu C, Wang J, Li Y, Navid F, Daw NC, Mardis ER, Wilson RK, Downing JR, Zhang J, Dyer MA. Recurrent somatic structural variations contribute to tumorigenesis in pediatric osteosarcoma. Cell Rep 2014; 7: 104-112.
8) Kohlmeyer JL, Gordon DJ, Tanas MR, Monga V, Dodd RD, Quelle DE. CDKs in Sarcoma: Mediators of Disease and Emerging Therapeutic Targets. Int J Mol Sci 2020; 21: 3018.
9) Ueda T, Healey JH, Huvos AG, Ladanyi M. Amplification of the MYC Gene in Osteosarcoma Secondary to Paget's Disease of Bone. Sarcoma 1997; 1: 131-134.
10) Lamora A, Talbot J, Mullard M, Brounais-Le Royer B, Redini F, Verrecchia F. TGF-β Signaling in Bone Remodeling and Osteosarcoma Progression. J Clin Med 2016; 5: 96.
11) Fritsche-Guenther R, Gloaguen Y, Kirchner M, Mertins P, Tunn PU, Kirwan JA. Progression-Dependent Altered Metabolism in Osteosarcoma Resulting in Different Nutrient Source Dependencies. Cancers (Basel) 2020; 12: 1371.
12) Wang DW, Wu L, Cao Y, Yang L, Liu W, E XQ, Ji G, Bi ZG. A novel mechanism of mTORC1-mediated serine/glycine metabolism in osteosarcoma development. Cell Signal 2017; 29: 107-114.
13) Morishita M, Kawamoto T, Hara H, Onishi Y, Ueha T, Minoda M, Katayama E, Takemori T, Fukase N, Kurosaka M, Kuroda R, Akisue T. AICAR induces mitochondrial apoptosis in human osteosarcoma cells through an AMPK-dependent pathway. Int J Oncol 2017; 50: 23-30.
14) Martins AS, Batista de Carvalho ALM, Marques MPM, Gil AM. Response of Osteosarcoma Cell Metabolism to Platinum and Palladium Chelates as Potential New Drugs. Molecules 2021; 26: 4805.
15) Gao S, Tu DN, Li H, Jiang JX, Cao X, You JB, Zhou XQ. Pharmacological or genetic inhibition of LDHA reverses tumor progression of pediatric osteosarcoma. Biomed Pharmacother 2016; 81: 388-393.
16) Gorska-Ponikowska M, Perricone U, Kuban-Jankowska A, Lo Bosco G, Barone G. 2-methoxyestradiol impacts on amino acids-mediated metabolic reprogramming in osteosarcoma cells by its interaction with NMDA receptor. J Cell Physiol 2017; 232: 3030-3049.
17) Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. Nat Rev Cancer 2014; 14: 722-735.
18) Lussier DM, Johnson JL, Hingorani P, Blattman JN. Combination immunotherapy with α-CTLA-4 and α-PD-L1 antibody blockade prevents immune escape and leads to complete control of metastatic osteosarcoma. J Immunother Cancer 2015; 3: 21.
19) Li LQ, Zhang LH, Yuan YB, Lu XC, Zhang Y, Liu YK, Wen J, Khader MA, Liu T, Li JZ, Zhang Y. Signature based on metabolic-related gene pairs can predict overall survival of osteosarcoma patients. Cancer Med 2021; 10: 4493-4509.
20) Zhu N, Hou J, Ma G, Guo S, Zhao C, Chen B. Co-expression network analysis identifies a gene signature as a predictive biomarker for energy metabolism in osteosarcoma. Cancer Cell Int 2020; 20: 259.
21) Yang C, Huang X, Liu Z, Qin W, Wang C. Metabolism-associated molecular classification of hepatocellular carcinoma. Mol Oncol 2020; 14: 896-913.
22) Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010; 26: 1572-1573.
23) 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.
24) Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. Omics 2012; 16: 284-287.
25) Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 2013; 14: 7.
26) Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, de Reyniès A. Erratum to: Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol 2016; 17: 249.
27) 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, Mills GB, Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun 2013; 4: 2612.
28) Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H, Trajanoski Z. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. Cell Rep 2017; 18: 248-262.
29) Li W, Wang H, Ma Z, Zhang J, Ou-Yang W, Qi Y, Liu J. Multi-omics Analysis of Microenvironment Characteristics and Immune Escape Mechanisms of Hepatocellular Carcinoma. Front Oncol 2019; 9: 1019.
30) Zheng M, Hu Y, Gou R, Liu O, Nie X, Li X, Liu Q, Hao Y, Liu J, Lin B. Identification of immune-enhanced molecular subtype associated with BRCA1 mutations, immune checkpoints and clinical outcome in ovarian carcinoma. J Cell Mol Med 2020; 24: 2819-2831.
31) Jiang P, Gu S, Pan D, Fu J, Sahu A, Hu X, Li Z, Traugh N, Bu X, Li B, Liu J, Freeman GJ, Brown MA, Wucherpfennig KW, Liu XS. Signatures of T cell dysfunction and exclusion predict cancer im-

munotherapy response. Nat Med 2018; 24: 1550-1558.

32) Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, Hodi FS, Martín-Algarra S, Mandal R, Sharfman WH, Bhatia S, Hwu WJ, Gajewski TF, Slingluff CL, Jr., Chowell D, Kendall SM, Chang H, Shah R, Kuo F, Morris LGT, Sidhom JW, Schneck JP, Horak CE, Weinhold N, Chan TA. Tumor and Microenvironment Evolution during Immunotherapy with Nivolumab. Cell 2017; 171: 934-949.e16.

33) Roh W, Chen PL, Reuben A, Spencer CN, Prieto PA, Miller JP, Gopalakrishnan V, Wang F, Cooper ZA, Reddy SM, Gumbs C, Little L, Chang Q, Chen WS, Wani K, De Macedo MP, Chen E, Austin-Breneman JL, Jiang H, Roszik J, Tetzlaff MT, Davies MA, Gershenwald JE, Tawbi H, Lazar AJ, Hwu P, Hwu WJ, Diab A, Glitza IC, Patel SP, Woodman SE, Amaria RN, Prieto VG, Hu J, Sharma P, Allison JP, Chin L, Zhang J, Wargo JA, Futreal PA. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. Sci Transl Med 2017; 9: eaah3560.

34) Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. PLoS One 2014; 9: e107468.

35) Zhang Z. Variable selection with stepwise and best subset approaches. Ann Transl Med 2016; 4: 136.

36) Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. Stat Med 2013; 32: 5381-5397.

37) Mariathasan S, Turley SJ, Nickles D, Castiglioni A, Yuen K, Wang Y, Kadel EE, III, Koeppen H, Astarita JL, Cubas R, Jhunjhunwala S, Banchereau R, Yang Y, Guan Y, Chalouni C, Ziai J, Şenbabaoğlu Y, Santoro S, Sheinson D, Hung J, Giltnane JM, Pierce AA, Mesh K, Lianoglou S, Riegler J, Carano RAD, Eriksson P, Höglund M, Somarriba L, Halligan DL, van der Heijden MS, Loriot Y, Rosenberg JE, Fong L, Mellman I, Chen DS, Green M, Derleth C, Fine GD, Hegde PS, Bourgon R, Powles T. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. Nature 2018; 554: 544-548.

38) Leite TC, Watters RJ, Weiss KR, Intini G. Avenues of research in dietary interventions to target tumor metabolism in osteosarcoma. J Transl Med 2021; 19: 450.

39) Hu XK, Rao SS, Tan YJ, Yin H, Luo MJ, Wang ZX, Zhou JH, Hong CG, Luo ZW, Du W, Wu B, Yan ZQ, He ZH, Liu ZZ, Cao J, Wang Y, Situ WY, Liu HM, Huang J, Wang YY, Xia K, Qian YX, Zhang Y, Yue T, Liu YW, Zhang HQ, Tang SY, Chen CY, Xie H. Fructose-coated Angstrom silver inhibits osteosarcoma growth and metastasis via promoting ROS-dependent apoptosis through the alteration of glucose metabolism by inhibiting PDK. Theranostics 2020; 10: 7710-7729.

40) Tian H, Zhou T, Chen H, Li C, Jiang Z, Lao L, Kahn SA, Duarte MEL, Zhao J, Daubs MD, Buser Z, Brochmann EJ, Wang JC, Murray SS. Bone morphogenetic protein-2 promotes osteosarcoma growth by promoting epithelial-mesenchymal transition (EMT) through the Wnt/β-catenin signaling pathway. J Orthop Res 2019; 37: 1638-1648.

41) Ma H, Su R, Feng H, Guo Y, Su G. Long noncoding RNA UCA1 promotes osteosarcoma metastasis through CREB1-mediated epithelial-mesenchymal transition and activating PI3K/AKT/mTOR pathway. J Bone Oncol 2019; 16: 100228.

42) Yi Z, Pu Y, Gou R, Chen Y, Ren X, Liu W, Dong P. Silencing of RIPK4 inhibits epithelial-mesenchymal transition by inactivating the Wnt/β-catenin signaling pathway in osteosarcoma. Mol Med Rep 2020; 21: 1154-1162.

43) Zhang J, Zhang Y, Cheng S, Mu Y, Liu Y, Yi X, Jiang D, Ding Y, Zhuang R. LAIR-1 overexpression inhibits epithelial-mesenchymal transition in osteosarcoma via GLUT1-related energy metabolism. World J Surg Oncol 2020; 18: 136.

44) Hsu PP, Sabatini DM. Cancer cell metabolism: Warburg and beyond. Cell 2008; 134: 703-707.

45) Shen S, Yao T, Xu Y, Zhang D, Fan S, Ma J. CircECE1 activates energy metabolism in osteosarcoma by stabilizing c-Myc. Mol Cancer 2020; 19: 151.

46) 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.

47) Tawbi HA, Burgess M, Bolejack V, Van Tine BA, Schuetze SM, Hu J, D'Angelo S, Attia S, Riedel RF, Priebat DA, Movva S, Davis LE, Okuno SH, Reed DR, Crowley J, Butterfield LH, Salazar R, Rodriguez-Canales J, Lazar AJ, Wistuba, II, Baker LH, Maki RG, Reinke D, Patel S. Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. Lancet Oncol 2017; 18: 1493-1501.