Research Paper

A three-gene signature based on tumour microenvironment predicts overall survival of osteosarcoma in adolescents and young adults

Chunkai Wen^{1,2,*}, Hongxue Wang^{1,*}, Han Wang^{1,*}, Hao Mo³, Wuning Zhong¹, Jing Tang¹, Yongkui Lu¹, Wenxian Zhou¹, Aihua Tan¹, Yan Liu¹, Weimin Xie¹

¹Department of Breast, Bone and Soft Tissue Oncology, the Affiliated Tumor Hospital of Guangxi Medical University, Nanning 530021, China
²Graduate School of Guangxi Medical University, Nanning 530021, China
³Department of Bone and Soft Tissue Tumor Surgery, the Affiliated Tumor Hospital of Guangxi Medical University, Nanning 530021, China
*Equal contribution

 Correspondence to: Aihua Tan, Yan Liu, Weimin Xie; email: tanaihua@gxmu.edu.cn, Liuyan@gxmu.edu.cn, xieweimin@gxmu.edu.cn, Liuyan@gxmu.edu.cn, xieweimin@gxmu.edu.cn, Liuyan@gxmu.edu.cn, xieweimin@gxmu.edu.cn, tanaihua@gxmu.edu.cn, Liuyan@gxmu.edu.cn, tanaihua@gxmu.edu.cn, tanaihua@gxmu.edu, tanaihua@gxmu

Copyright: © 2020 Wen et al. This is an open access article distributed under the terms of the <u>Creative Commons Attribution</u> <u>License</u> (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Evidences shows that immune and stroma related genes in the tumour microenvironment (TME) play a key regulator in the prognosis of Osteosarcomas (OSs). The purpose of this study was to develop a TME-related risk model for assessing the prognosis of OSs. 82 OSs cases aged ≤25 years from TARGET were divided into two groups according to the immune/stromal scores that were analyzed by the Estimate algorithm. The differentially expressed genes (DEGs) between the two groups were analyzed and 122 DEGs were revealed. Finally, three genes (COCH, MYOM2 and PDE1B) with the minimum AIC value were derived from 122 DEGs by multivariate cox analysis. The three-gene risk model (3-GRM) could distinguish patients with high risk from the training (TARGET) and validation (GSE21257) cohort. Furthermore, a nomogram model included 3-GRM score and clinical features were developed, with the AUC values in predicting 1, 3 and 5-year survival were 0.971, 0.853 and 0.818, respectively. In addition, in the high 3-GRM score group, the enrichment degrees of infiltrating immune cells were significantly lower and immune-related pathways were markedly suppressed. In summary, this model may be used as a marker to predict survival for OSs patients in adolescent and young adults.

INTRODUCTION

Osteosarcomas (OCs) is the most common primary bone tumor in adolescent and young adults and the incidence is higher at age 15 to 19 years old [1]. OSs has a high potential to metastasize to lungs. Neoadjuvant chemotherapy-Surgical resection-Adjuvant chemotherapy (the so-called sandwich treatment mode) is the standard treatment for early and locally advanced OSs, which has significantly improved the prognosis of patients, with the 5-year survival has exceeded 60% [1, 2].

However, for advanced patients, the application of emerging therapies such as targeted therapy and immunotherapy (immune checkpoint inhibitors) is not optimistic, and the chemotherapeutic regimens included adriamycin, cisplatin, ifosfamide and high-dose methotrexate are still the main options [3, 4]. Unfortunately, about one-third of patients eventually failed due to drug resistance. In the past few decades, the 5-year survival of patients with metastasis at diagnosis or in relapse has not significantly improved, but still remains at only 20% [5, 6]. Therefore, it is remains a major goal for developing new therapy and valid signature to improve or predict the prognosis of OSs. Increasing evidence indicated that the biological behavior of tumor, such as invasion and metastasis, drug resistance and so on, depended not only on the inherent characteristics of tumor cells, but also on the composition and function of tumor microenvironment (TME) [7–9].

OSs has extremely significant heterogeneity, the landscape of driver mutations had proved few mutations recur with high frequency at the intra-tumor level of OSs, resulting in the loss of definite targets for therapy [10, 11]. On the other hand, and more importantly, TME was not only considered essential for the growth of osteosarcoma, but also regulated the proliferation, migration and metastasis of osteosarcoma cells, as well as drug resistance, immunosuppression and immune escape [5, 12]. The TME of OSs is a complex environment composed of active cells (such as tumor cells, stromal cells and immune cells, fibroblasts, etc.), vasculature, extracellular matrix (ECM), and a variety of factors [5, 13, 14]. Among them, the tumor-infiltrating immune cells (TIICs) and stromal cells are closely related to OSs cells by regulating various signal pathways via release of cytokines and other soluble factors, conversely stimulating and facilitating tumor cell metabolism and proliferation in all the stages of carcinogenesis [15-17]. Therefore, TME-related genomic analysis can help to find markers for assessing disease evolution and prognosis of OSs patients, and even for translation in both molecularly targeted therapy and personalized therapy [18, 19].

In the past decade, several methods have been invented to dissect the TME based on gene expression profiles. One of these. Estimate algorithm [20], designed by Yoshihara et al, was used to assess the score of immunity and stromal in the TME based on gene expression profiles, and it has been successfully applied to a variety of solid tumors, such as ovarian cancer [21], bladder cancer [22] and gastric cancer [23]. A similar approach, Cibersort [24], a new biological tool based on the deconvolution technique, can quantify the abundance of specific TIICs. Unfortunately, there are no definitive and clinically applicable prognostic markers for OSs. Thus, this study expects to establish a new predictive model based on TME-related genes and evaluate its effect in estimating the outcome of OSs, so as to provide a reference for further research in the future.

RESULTS

Correlation between immune/stromal score in TME and prognosis of OSs patients in adolescents and young adults

In this study, 82 OSs patients (≤ 25 years of age) from TARGET were enrolled, those with complete gene expression data of samples and clinical information of follow-up. The detailed clinical characteristics of patients were summarized in Supplementary Table 1. The flowchart of the analysis procedure was shown in Figure 1. Based on the normalized matrix data, the TME score was graded using the Estimate algorithm, and the results showed that the immune scores and stromal scores of all patients were ranged from -1508.04 to 2638.38 and from -695.66 to 1962.14, respectively.

According to the median value of immune/stromal score, 82 OSs patients were divided into high and low score group. Kaplan-Meier (K-M) curves showed that the 5-year overall survival (OS) rates of patients with high and low immune scores were 82.6% and 48.7%, respectively (P=0.003) (Figure 2A). As well as, the 5year recurrence-free survival (RFS) rates were 65% and 47.3%. respectively (P=0.006) (Figure 2C). Consistently, the 5-year OS rates of patients with high and low stromal scores were 65% and 47.3%, respectively, the P value approximately reached statistical difference (P=0.053) (Figure 2B), and the 5year RFS rates were 66% and 46.1%, respectively (P=0.012) (Figure 2D), indicating that the TME-related immune and stromal score in the TME were significantly correlated with the survival.

Identification of differentially expressed genes (DEGs)

The DEGs between the high and low immune/stromal scores groups were obtained based on gene expression profiles, and a fold-change > 1 or < -1 and adjusted P<0.05 were used as criterions for screening DEGs. Heatmaps showed the distinct gene expression profiles from the two groups in Figure 3A, 3B. In the high immune score group, 320 genes were down-regulated and 262 genes were up-regulated (Supplementary Tables 2A, 2B). Meanwhile, the high stromal score group had 252 genes down-regulated and 160 genes upregulated (Supplementary Tables 3A, 3B). Furthermore, Venn diagram was used to find the common-DEGs in the two groups (Figure 3C, 3D), showing that 31 genes were generally up-regulated and 12 genes were markedly down-regulated (Supplementary Table 4). The top 50 genes with the most significant difference between the high and low immune score groups, and the high and low stromal score groups, respectively, and other differentially expressed genes crossing with Venn diagram were all included. Finally, a total of 122 genes, defined as the TME-related differentially expressed genes (tmDEGs), were included in the subsequent functional classification analysis.

GO(Gene Ontology) Term analysis confirmed that the tmDEGs mostly involved in regulation of leukocyte activation, leukocyte cell-cell adhesion, positive regulation of cell activation in biological processes (BP), external side of plasma membrane, endocytic vesicel, secretory granule lumen in cellular components (CC), antigen binding, phosphatidylinositol phosphate binding, phosphatidylinositol binding in molecular functions (MF) (Figure 3E). KEGG enrichment analysis showed that tmDEGs mainly enriched in the osteoclast differentiation, B cell receptor signaling pathway cell receptor interaction, viral protein interaction with cytokine and cytokine receptor, cytokine-cytokine receptor interaction and JAK-STAT signaling pathway (Figure 3F).

These results indicated that the immune score and stromal score could all better reflect the main immune landscapes of TME, showing the activation of immune cells and stromal cells, and the expression of various signal pathways in OSs tissues.

Construction and evaluation of the risk score model

The relationship between tmDEGs and prognosis of 82 OSs patients from TARGET as above was analyzed. The univariate cox regression analysis showed that 32 of 122 tmDEGs were significantly correlated with OS of patients (P<0.05) (Supplementary Table 5). Then Lasso regression analysis was performed on these 32 genes, and 13 genes were screened as candidate genes through relative regression coefficient (Figure 4A, 4B).



Figure 1. The overall design of the present study.

And then the top 30 genes, identified from 122 tmDEGs by XGBoost machine learning according to the order of importance (Figure 4C), were crossed with these 13 genes selected by Lasso regression analysis, and 9 candidate genes were obtained.

Finally, a multivariate cox analysis was performed, and three genes with the minimum Akaike's Information Criterion (AIC) value of 188.9 were identified to construct a risk model (Figure 4D), which was called three-gene risk model (3-GRM). The C index of model was 0.77 (Se=0.042, 95% CI, 0.688~0.852).

According to the median value of 3-GRM score, patients were divided into high score group and low score group. The expression levels of these three genes

were significantly correlated with survival of OSs patients from dataset TARGET, among which, Cochlin (COCH) and Myomesin2 (MYOM2), both overexpressed in the high 3-GRM score group, were associated with poor prognosis. However, PD1EB was significantly down-regulated in high 3-GRM score group, was related to favorable prognosis (Figure 4E–4G). Based on relative coefficients in multivariable cox regression analysis, the scoring formula of 3-GRM was as follows: risk score = $[0.30 \times COCH$ expression level]+ $[0.41 \times MYOM2]$ expression level]- $[2.1 \times PDE1B]$ expression level].

The K-M analysis showed that the OS of high 3-GRM score group was significantly lower than that of low score group (P<0.05) (Figure 5A), and the 5-year



Figure 2. Relationship between the TME-related immune/stromal score and survival of OSs patients. (A) Immune score and overall survival rate. (B) Stromal score and overall survival rate. (C) Immune score and recurrence-free survival rate. (D) Stromal score and recurrence-free survival rate.

survival rates of the high and low score group were 51.3% and 80.1%, respectively (P=0.000). The distribution of genes expression data in the subgroup of difference risk scores were shown in Figure 5B. The prediction ability of 3-GRM was evaluated by

calculating the area under the ROC curve (AUC). The AUC values in predicting 1, 3 and 5-year of survival rate were 0.890, 0.822 and 0.773, respectively, indicating that 3-GRM had a good prediction effect (Figure 5C).



Figure 3. The differences of genes expression and pathways enrichment based on immune scores and stromal scores. (A, B) Heatmap of significantly differentially expressed genes based on immune and stromal scores. (C, D) Venn diagram analysis of aberrantly expressed genes based on immune and stromal scores in osteosarcoma. (E) GO analyses of the tmDEGs in the categories of biological processes (BP), cellular components (CC), and molecular functions (MF). (F) KEGG analysis of tmDEGs genes. tmDEGs, differentially expressed genes of the tumour microenvironment. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

High score of 3-GRM was an independent prognostic factor for OSs patients in adolescent and young adults

Based on the clinical data and survival information of 82 patients, univariate cox regression analysis showed

that the high score of 3-GRM, metastasis at diagnosis and lung metastasis were the risk factors affecting prognosis, but age, gender, race, primary site, specific site of primary tumor, and surgical method were not related to outcome of patients (Figure 6A). As in the TARGET database (as above), 95% of patients (20/21)





with metastasis at diagnosis had lung metastasis, so these two factors were classified as having metastasis at diagnosis for further analysis. Multivariate cox regression analysis confirmed that high 3-GRM score and metastasis at diagnosis were independent risk factors of OSs patients (Figure 6B). K-M analysis showed that the 5-year survival rates of cases with metastases and without metastases at diagnosis were 31.3% and 77.4%, respectively (P<0.001, data not shown).

Validate the new risk model (3-GRM) in GEO dataset

To verify the robustness of 3-GRM, another independent dataset GSE21257 (n=53) (Table 1) from

the GEO database was used. According to the risk scoring formula as above, the patients in GSE21257 dataset were divided into two groups, with 28 cases in high 3-GRM score group and 25 cases in low 3-GRM score group.

Consistent with previous results, this results also confirmed that patients with high 3-GRM score had significantly shorter OS than those with low scores (Figure 7A). The AUC values of the 3-GRM in predicting 1, 3 and 5-year of survival rate were 0.861, 0.710 and 0.694, respectively (Figure 7B). The Figure 7C shows the distribution of the 3-GRM scores and gene expression data in the validation cohorts. At the same time, univariate (Figure 6C) and multivariate



Figure 5. Prognostic analysis of the 3-genes risk model (3-GRM) in TARGET. (A) Prognostic analysis between the high and low 3-GRM score group. (B) Differences of COCH, MYOM2 and PDE1B expression levels between the high and low 3-GRM score group. (C) Timedependent ROC curve analysis of the 3-GRM. (Figure 6D) cox analysis found that huvos grade (histological response grade after chemotherapy) [25], metastasis at diagnosis and the high 3-GRM score were also independent prognostic factors for OSs patients in the GSE21257 dataset, confirming that the 3-GRM was robust and effective.

The relationship between 3-GRM score and TME score, and the difference of TIICs between the high and low 3-GRM score groups

Based on TARGET database, the analysis results found that the 3-GRM score was both significantly correlated with the TME-related immunity score (R=0.398, P<0.01) and the stromal score (R=0.523, P<0.01) (Figure 8A, 8B), respectively, indicating that the higher the 3-GRM score, the lower the immune score and stromal score.

Cibersort algorithm analysis revealed that the most common TIICs in the specimens of 82 OSs patients were macrophages M2(27.8 \pm 10.14)%, macrophages M0 (27.63 \pm 11.35)% and CD4+ memory T cells (17.08 \pm 7.86)%, accounting for more than 70%. In

addition, the proportion of macrophages M0 in the high 3-GRM score group was significantly higher than that in the low score group (P<0.001), while the ratios of monocytes (P=0.05), macrophages M1 (P=0.04) and macrophages M2 (P=0.002) were significantly lower than those in the low score group (Figure 8C).

Furthermore, 64 kinds of TME-related cells were evaluated by Xcell, and the results were consistent with that of Cibersort algorithm. Compared with the low 3-GRM score group, the enrichment degree of TIICs (such as "aDC", "cDC", "iDC", "macrophage M1 and M2", "endothelial cells" and "MSC") were significantly lower, while "muscle cells" and "skeletal muscle cells" were significantly enrichment in the high 3-GRM score group (all P<0.01) (Supplementary Figure 1).

Comparison of immune-related signal pathways between the high and low score group

A total of 1,811 immune-related genes were obtained from the website of IMMPORT (<u>http://www.</u> <u>immport.org/</u>) for differential analysis, with screening criteria were fold-changed >1 or <-1 and



Figure 6. Analysis of prognostic factors of OSs patients. (**A**, **B**) Univariate and multivariate regression analysis of the relation between the 3-GRM and clinicopathological features in TARGET. (**C**, **D**) Univariate and multivariate regression analysis of the relation between the 3-GRM and clinicopathological features in GSE21257.

Clinical features	Cases, n (%)	High 3-GRM score group, n (%)	Low 3-GRM score group, n (%)	P*
Patients(n)	53	28	25	
Median age(range)	17 (14-19)	16.5(14-19)	17(13-19)	0.71
Gender				
Female	19 (35.8)	7 (25.0)	12 (48.0)	0.08
Male	34 (64.2)	21 (75.0)	13 (52.0)	
Histological Subtype				
Osteoblastic	32 (60.4)	19 (67.9)	13 (52.0)	0.24
Others	21 (39.6)	9 (32.1)	12 (48.0)	
Tumor location				
Lower limb	44 (83.0)	23 (82.1)	21 (87.5)	0.88
upper limb	8 (15.1)	5 (17.9)	3 (12.5)	
Unknow	1(1.89)	-	-	
Huvos grade				
I-II	29 (54.7)	14 (58.3)	15 (65.2)	0.85
III-IV	18 (34.0)	10 (41.7)	8 (34.8)	
Unknow	6(11.3)	-	-	
Metastases at diagnosis				
No	39 (73.6)	20 (71.4)	19 (76.0)	0.95
Yes	14 (26.4)	8 (28.6)	6 (24.0)	

Table 1. Clinical baseline data and score grouping of 53 patients with OSs in GES21257 dataset.

*Comparison of high 3-GRM score group and low 3-GRM score group.

adjusted P<0.05. The results showed that there were 31 DEGs between the two groups (Supplementary Table 6). Identified by GO terms, these DEGs were involved in external side of plasma membrane (CC), humoral immune response(BP), immune response-activating/regulating cell receptor signaling pathway(BP), regulation of lymphocyte activation (BP) and cytokine receptor activity(MF) (all adjusted P<0.01) (Figure 9A).

The enrichment analysis of KEGG pathway explored that of immune-related signaling pathways were markedly down-regulated in the high 3-GRM score group, including cytokine–cytokine receptor interaction, viral protein interaction with cytokine and cytokine receptor, chemokine signaling pathway, Th1/Th2 and Th17 cell differentiation, and T cell receptor signaling pathway (all adjusted P<0.01) (Figure 9B).

Establishment and evaluation of nomogram model for survival based on 3-GRM scoring and clinicopathological features

As the result of multivariate cox regression analysis showing there were two prognostic indicators of OSs (Figure 6B), a nomogram model included 3-GRM score and clinicopathological features (such as metastasis status at diagnosis) were developed (Figure 10A), with the index C was 0.825. The calibration plot for the possibility of 1, 3and 5-year survival showed good agreement between the prediction by risk score and actual observations (Supplementary Figure 2A–2C). The AUC values of nomogram model in predicting 1, 3 and 5-year survival reached to 0.971, 0.853 and 0.818, respectively. In addition, the nomogram model was also verified in GSE21257 dataset, with the AUC values in predicting 1, 3 and 5-year survival were 0.781, 0.840 and 0.795, respectively. The predicted results of the calibration plot were also consistent with the actual observation (Supplementary Figure 3A, 3B).

Recently, prognosis models had been established and used to predict the prognosis of OSs in several studies [26–29]. In our study, the AUC value was used for evaluating the prediction effect of varies models. As shown in Table 2, the calculate effects of nomogram model were always better than those of the 3-GRM and other models [26–29].

DISCUSSION

OSs has a relatively high incidence in the second decade of life, especially at ages 15-19 years [1].

Indeed, genetic investigations have demonstrated the paucity of mutations more specifically involving signal transduction pathways in adolescent and young adults sarcomas compared with adults. The result highlights a major difference between adolescent and young adults and adult. In other words, there may be different landscapes in adult immune osteosarcoma and childhood osteosarcoma [30]. addition, The In clinicopathologic features that may influence the prognosis of OSs patients include age, metastasis at diagnosis, lesion location, and degree of tumor necrosis (Huvos grade) after neoadjuvant chemotherapy, etc. [1, 2]. However, age has not been taken into account in many prognostic model-related studies [26–29]. Therefore, patients under 25 years old were enrolled in this study, and it is hoped that the model could more accurately predict the prognosis of such population.

The TME is an extremely complex system that has not been clearly elucidated [5, 31].

Currently, exploratory researches involved in TME include the molecular events of regulation between OSs and non-tumor cells, the genomic drivers of disease progression, prediction model that best reflects the clinical outcome and its transformational applications [19, 32].

Evidences shows that immune and stromal related genes in TME may play key regulatory roles in the prognosis of OSs. The interaction between Programmed Cell Death Protein 1 (PD-1) and its Ligand 1(PD-L1) is critical for tumor cells survival. PD-1 is expressed in T lymphocytes, while PD-L1 is overexpressed in tumor cells. PD-1 binding with PD-L1 can interfere with and inhibit the ability of T cells to kill cancer cells [33]. The pooled results of a meta-analysis showed that PD-L1 overexpression could predict poor OS (HR 1.45, 95% CI: 1.11-1.90, P<0.01), metastasis-free survival (HR1.58, 95% CI: 1.14-2.19, P<0.01) and event-free survival (HR 2.82, 95% CI: 1.69-4.71, P<0.01) in OSs,



Figure 7. Prognostic analysis of the 3-GRM in GSE21257. (A) Prognostic analysis between the high and low 3-GRM score group. (B) Differences of COCH, MYOM2 and PDE1B expression levels between the high and low 3-GRM score group. (C) Time-dependent ROC curve analysis of the 3-GRM.

and was also significantly correlated with a higher rate of tumor metastasis (OR 2.95, 95% CI: 1.32-6.60, P< 0.01) [34]. CXCL12 (also known as stromal cell-derived factor 1, SDF-1), one of the chemokine protein family, is the main regulator of cell trafficking, affecting both tumor cells and white blood cells [35]. Overexpression of CXCL12 was positively correlated

with the number of tumor infiltrating lymphocytes and the better survival rate of OSs patients [36].

Among various methods, the TME-related gene expression profiling, based on immune score and stromal score obtained by Estimate algorithm, is a common and effective method to evaluate tumor





microenvironment [20–23]. In this study, we analyzed the TME characteristics of OSs patients under 25 years old from TARGET dataset and its relationship with patient prognosis. According to the immune score and stromal score of TME, DEGs were obtained to construct a three-gene risk model (3-GRM). The results of validation were consistent in two independent databases, which confirmed that this 3-GRM was efficient and robust and the high score of 3-GRM was an independent prognostic factor for OSs patients. Further analysis found that the 3-GRM score was all strongly negatively correlated with the immune score and stromal score, respectively, indicating the model could effectively identify the TME status of OSs. Specifically, in this model, both COCH and MYOM2 were overexpressed in the high 3-GRM score group and



Figure 9. Analysis of the expression of immune-related differential genes, enrichment of immune-related pathways in high and low 3-GRD score group. (A) GO analyses of the immune-related differentially expressed genes in the categories of biological processes (BP), cellular components (CC), and molecular functions (MF). (B) KEGG analysis of the immune-related signaling pathways. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.



Figure 10. Nomogram model for predicting the outcome of OSs patients. (A) Nomogram model for predicting the probability of 1, 3, and 5-year overall survival (OS) for adolescent and young adults with osteosarcoma.

Table 2. Comparison of OSs-related prognostic models in TARGET database.	
rubic zi companison of oos related prognostic models in tratezi database	

OSs related prognostic model	Constitution	1-year AUC value	3-year AUC value	5-year AUC value
Nomogram model	3-GRM + metastasis at diagnosis	0.971	0.853	0.818
3-GRM	3-gene signature	0.89	0.822	0.773
Zhang,et al.(2020) [26]	3-gene signature	0.643	0.781	0.809
Li,et al.(2020) [27]	4-gene signature	0.644	0.714	0.738
Liu,et al.(2020) [28]	2-gene signature	-	0.71	0.72
Shi,et al.(2020) [29]	3-gene signature	0.914	0.849	0.822

Note: AUC, area under the ROC curve; OSs, osteosarcomas; 3-GRM, three-gene risk model.

led to poor prognosis. However, PD1EB was significantly down-regulated in the high-scoring group, and analysis showed (Figure 4E–4G) that its overexpression was beneficial to survival. As far as we know, the prognostic risk model based on these three genes has not reported before.

COCH gene is highly expressed in the sensory organs (inner ear and eye), lymph nodes and spleen [37]. It plays an important role in maintaining the shape of cells [38]. And Cochlin, which encoded by COCH, and its domains are the main non-collagenous components of extracellular matrix (ECM) and have high affinity for multi-type collagens [38, 39]. A study shown that COCH was a transition zone-specific genes and was also stroma-specific of the prostate, and involved in the occurrence of prostatic hyperplasia and prostate cancer [40]. Up-expression of COCH was directly related to the stage progression of clear cell renal cell carcinoma (ccRCC) [41]. Interestingly, research indicated that in the circulation, the LCCL domains, one N-terminal factor C homology of COCH, could signal the innate immune cells and amplify the cytokine response in the form of glycosylated polypeptides through unknown pathways [42]. Similar, Cochlin, secreted from follicular dendritic cells in the spleen, was crucial for systemic immune response against bacterial infection by induces secretion of cytokines (IL-1βand IL-6) and enhances the recruitment of immune cells (neutrophils and macrophages) [37, 43, 44].

MYOM2 (myomesin 2) is a type of muscle fiber related protein and is an essential component of cytoskeleton [45, 46]. It is also the main component of the M-bands, which are located in the center of the sarcomere and are essential for the stability of sarcomere contraction [45, 47]. Study had revealed that in relapsed/refractory diffuse large B-cell lymphoma patients, late oncogenic events was composed of clonally represented recurrent mutations/gene alterations including MYOM2 [48]. One report showed that MYOM2 was the only significantly up-regulated gene in localized invasive periodontitis, suggesting that it was associated with inflammation [49]. Tumor necrosis factor (TNF) blockers have a high efficacy in treating Ankylosing Spondylitis (AS), a study reported that IgG Galactosylation status combined with MYOM2 rs2294066 polymorphism (T allele in rs2294066 leads to MYOM2 overexpression) could precisely predicts anti-TNF response of AS, indicating that MYOM2 might associated with the pharmacology of TNF blockers [50]. In addition, the results of genomic screening represented MYOM2 mutations was probably causative for arthrogryposis [51].

PDE1B (Phosphodiesterase 1B) gene is located on chromosomes 12q13 and is a member of the cyclic nucleotide phosphodiesterase (PDE) family [52]. Interestingly, studies show that PDE1B may play a regulatory role in differentiation of multiple immune cell types via degrading intracellular levels of cAMP and cGMP [53, 54]. Exploration shown that granulocyte macrophage colony stimulating factor (GM-CSF) could shift the differentiation from a macrophage to a cell phenotype and also up-regulated dendritic PDE1B.Yet, in the presence of GM-CSF, IL-4 treatment suppressed the up-regulation of PDE1B2 (one of PDE1B variants with unique N-terminal sequences) [55]. Further research found that inhibiting PDE1B2 upregulation did not prevent HL-60 cells differentiation, but could change some aspects of macrophage like phenotype, such as increased cell proliferation, phagocytosis and leukocyte adhesion molecule CD11b expression, accompanied by a lower basal levels of cAMP in cells [56]. Moreover, PDE1B2 might involve in the occurrence and prognosis of breast cancer [57, 58].

Regrettably, the mechanism by which these three genes affect the prognosis of patients with OSs is still unclear, and it is worth exploring in the future. We speculate that COCH overexpression may induce the secretion of multiple cytokines and enhance the recruitment of

immune cells, leading to the imbalance of the local inflammatory environment of TME and mediating the immune escape of tumor cells [59-61]. Similarly, MYOM2 overexpression may also mediate the malignant phenotype of tumor cells by affecting the inflammatory response of TME [49-51]. However, under the mediation of cytokines such as GM-CSF, the PD1EB up-regulation of may promote the differentiation of macrophages into M1 subtype macrophages, thereby enhancing the anti-tumor immune response and helping to improve the prognosis of patients [53-56, 62, 63].

Moreover, we also discussed the properties of microenvironmental cell levels and immune-related signal pathways between the high 3-GRM group and the low 3-GRM group. In general, the degree of enrichment of immune infiltrating cells in the high score group was significantly reduced, and immune regulation-related signal pathways such as cytokine interaction signaling pathways and T cell receptor signaling pathways were significantly down-regulated, indicating that there were general immunosuppression in the TME, which might affect the prognosis of patients. Several important studies identified that there was a complex associations between TIICs and cancer survival. On the one hand, the characteristics of TIICs, including the density, composition and activation of immune cells, were closely related to the response of immunotherapy and chemotherapy [64, 65]. The presence of either a preexisting or induced immune response might indicate a more favourable prognosis [66]. On the other hand, OSs cells could regulate the recruitment and differentiation of immune infiltrating cells (TIICs) to establish a local immune tolerance environment that was conducive to tumor growth, drug resistance and metastasis. OSs cells also could control the T-lymphocyte responses via the PD-1/PDL-1 system to affect the balance between M1 and M2 macrophage subtypes, then leading to immune tolerance [58, 59]. A study used whole genome, T cell receptor sequencing and other methods to analyze the immune status of OSs, and confirmed that there were likely multiple immune-suppressive features in OSs, which might lead to poor response to immune checkpoint inhibitors and neoadjuvant chemotherapy [67].

In order to further comprehensively explore the impact of potential factors, such as 3-GRM score, clinicopathological characteristics, huvos grade, etc., on the prognosis of patients, this study established a nomogram model, which composed of 3-GRM scores and with metastasis at diagnosis. The verification analysis in two independent datasets identified that the nomogram model had higher prediction efficiency than 3-GRM and other models [26–29]. This nomogram model was consistent with the multivariate analysis results of several large-scale clinical studies that the most adverse factors for prognosis were pulmonary metastases at diagnosis [68].

However, our research also has certain limitations. First of all, due to the limited number of cases and clinical data available in the data set, this study cannot analyze the correlation between the 3-GRM and clinical staging, tumor location, histological grade and other factors. Secondly, the role and mechanism of the three prognostic genes in OSs need to be further explored. Third, the 3-GRM needs to be further validated in multi-center clinical trials and prospective studies.

In summary, this study aimed at the microenvironment of OSs and obtained a risk model (3-GRM) consisting of three immune/stroma-related genes through strict classification and screening criteria. The verification results proved that the model had high sensitivity and specificity, and good robustness. In addition, combining the 3-GRM score with the patient's clinical characteristics might further improve the prediction efficiency. Therefore, this 3-GRM could be used as a marker to predict the outcome for OSs patients in adolescent and young adults.

MATERIALS AND METHODS

Data preparation

The gene expression profiles and clinical data of TARGET osteosarcoma patients were downloaded from the UCSC Xena website (http://xena.ucsc.edu/). Of those, a total of 82 patients under 25 years old and survive for more than 1 month were enrolled. The GSE21257 database was served as a validation, including 53 OSs patients and their clinical information (age, gender, histological subtype and huvos grade, tumor location, metastasis at diagnosis, and survival). All samples were obtained from primary lesion, and gene expression profiling were detected by microarray. The expression matrix data for the validation set GSE21257 was downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), R package "limma" was used for quality control and normalization [69]. Gene expression value with multiple probes for genes was calculated as the average of the probes.

Identification of DEGs related to the immune and stromal score of the TME

The immune and stromal scores were calculated based on mRNA expression profile by "estimate" package (<u>http://r-forge.rproject.org</u>) [20]. According to the median of scores, 82 OSs patients from TARGET dataset were

divided into two groups: high and low immune/stromal score group. The differential expression genes of the TME (tmDEGs) between these two groups were identified with "limma" in R package (version 3.6.1; <u>https://www.rproject.org/</u>) [69]. The filtering criteria were the fold-change> 1 or <-1 and adjusted P<0.05.

Enrichment analysis of DEGs

The potential functional enrichment of tmDEGs and immune-related differential genes were analyzed by the "clusterprofile" in R [70]. Functional enrichment included gene ontology (GO) categories in biological processes (BP), molecular functions (MF), cellular components (CC) and Kyoto Encyclopedia of Gene and Genome enriched by (KEGG) Pathway. The threshold of false detection rate (FDR) < 0.05 was considered significant.

Construction of risk model of Oss

The genes related to survival of OSs patients were screened by univariate cox and lasso regression, and Xgboost was used to further narrow the screening range, then multivariate cox regression analysis was used to determine the genes used to establish the prognosis model. The Lasso regression was analyzed using "glmnet" R software package. XGBoost machine learning [71], used Anaconda software, and its script was from MichaelMW (github.com/MichaelMW/bnfo. course). The formula of risk scoring was as follows: risk score= [0.30×COCH expression level]+[0.41×MYOM2 expression level]-[2.1×PDE1B expression level]. In this paper, the evaluation of the 3-GRM involved ROC curve. "Survival ROC" R software package was used to visualize the specificity and sensitivity of the 3-GRM.

Enrichment analysis of cells related to TME

We used two bioinformatics methods to analyze cells in TME, in which, Cibersort (<u>https://cibersortx.stanford.edu</u>) calculated the proportion of 22 kinds of TIICs in microenvironment based on deconvolution (Figure 9A) [24]; XCELL (<u>https://xcell.ucsf.edu/</u>) calculated the independent enrichment scores of 64 kinds of immune and stromal cells based on gene enrichment(Figure 9B) [72]. We further applied the Wilcox test to compare the differences of cells characteristics between high score and low score groups.

Construction and validation of the nomogram model

Using "RMS" in R to draw the nomograms of 1, 3, 5year OS of OSs patients [73]. 3-GRM score and metastasis at diagnosis, which were two independent prognostic factors identified by multivariate cox regression analysis, were incorporated into the nomogram model. "Bootstrap" obtained 1, 3, 5-year calibration plots through re-sampling 1000 times, and the calibration curve was visualized to evaluate the consistency between actual and predicted survival rates.

Statistical methods

All analyses were performed using R software (version 3.5.2). Survival analyses used the Kaplan-Meier (K-M) method and tested the relationship between them using a log-rank test. The spearman rank correlation analyzes the correlation between the two variables. The Venn diagram was drawn using the "VennDiagram" package, and the violin curve was drawn using the "violot" package. Immune-related genes were recruited at IMMPORT (http://www.immport.org/) [74]. According to the lowest AIC value, cox multiple factors determined the genes included in the model [75]. Analysis of variance was used to test the relationship between 3-GRM score and clinical characteristics, and Wilcox test was used to analyze the characteristics of cells in microenvironment in different risk groups. All statistical tests are two-sided tests, P value <0.05 is considered statistically significant.

AUTHOR CONTRIBUTIONS

C.W. and H.W. designed the study and performed data analysis; H.M., W.Z., J.T., and Y.L. revised the paper; W.Z. performed literature search and data collection; H.W. constructed figures and improved the language; A.T., Y.L. and W.X. directed the overall project. All authors reviewed the manuscript.

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest to disclose.

FUNDING

The study was funded by the Institute of Cancer Prevention and Treatment, Guangxi Zhuang Autonomous Region (2018GXNSFBA138019). In addition, we are very grateful to TARGET database and GEO databases for providing valuable data resources to enable us to conduct this research.

REFERENCES

 Casali PG, Bielack S, Abecassis N, Aro HT, Bauer S, Biagini R, Bonvalot S, Boukovinas I, Bovee JV, Brennan B, Brodowicz T, Broto JM, Brugières L, et al, and ESMO Guidelines Committee, PaedCan and ERN EURACAN. Bone sarcomas: ESMO-PaedCan-EURACAN clinical practice guidelines for diagnosis, treatment and followup. Ann Oncol. 2018 (Suppl 4); 29:iv79–95. <u>https://doi.org/10.1093/annonc/mdy310</u> PMID:<u>30285218</u>

- George S. Developments in systemic therapy for soft tissue and bone sarcomas. J Natl Compr Canc Netw. 2019; 17:625–28. <u>https://doi.org/10.6004/jnccn.2019.5020</u> PMID:<u>31117040</u>
- Dyson KA, Stover BD, Grippin A, Mendez-Gomez HR, Lagmay J, Mitchell DA, Sayour EJ. Emerging trends in immunotherapy for pediatric sarcomas. J Hematol Oncol. 2019; 12:78. <u>https://doi.org/10.1186/s13045-019-0756-z</u> PMID:<u>31311607</u>
- 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, et al. Pembrolizumab in advanced soft-tissue sarcoma and bone sarcoma (SARC028): a multicentre, two-cohort, single-arm, open-label, phase 2 trial. Lancet Oncol. 2017; 18:1493–501. https://doi.org/10.1016/S1470-2045(17)30624-1
- PMID:<u>28988646</u>
 5. Corre I, Verrecchia F, Crenn V, Redini F, Trichet V. The osteosarcoma microenvironment: a complex but targetable ecosystem. Cells. 2020; 9:976.

https://doi.org/10.3390/cells9040976 PMID:32326444

 Youn P, Milano MT, Constine LS, Travis LB. Long-term cause-specific mortality in survivors of adolescent and young adult bone and soft tissue sarcoma: a population-based study of 28,844 patients. Cancer. 2014; 120:2334–42. https://doi.org/10.1002/cncr.28733 PMID:24752471

7. Ehnman M, Chaabane W, Haglund F, Tsagkozis P. The tumor microenvironment of pediatric sarcoma:

mesenchymal mechanisms regulating cell migration and metastasis. Curr Oncol Rep. 2019; 21:90. <u>https://doi.org/10.1007/s11912-019-0839-6</u> PMID:31418125

- Wu T, Dai Y. Tumor microenvironment and therapeutic response. Cancer Lett. 2017; 387:61–68. <u>https://doi.org/10.1016/j.canlet.2016.01.043</u> PMID:<u>26845449</u>
- 9. Hirata E, Sahai E. Tumor microenvironment and differential responses to therapy. Cold Spring Harb Perspect Med. 2017; 7:a026781. <u>https://doi.org/10.1101/cshperspect.a026781</u> PMID:<u>28213438</u>
- 10. Saraf AJ, Fenger JM, Roberts RD. Osteosarcoma: accelerating progress makes for a hopeful future. Front Oncol. 2018; 8:4.

https://doi.org/10.3389/fonc.2018.00004 PMID:29435436

- Duffaud F. Role of TKI for metastatic osteogenic sarcoma. Curr Treat Options Oncol. 2020; 21:65. <u>https://doi.org/10.1007/s11864-020-00760-w</u> PMID:<u>32601895</u>
- Anwar MA, El-Baba C, Elnaggar MH, Elkholy YO, Mottawea M, Johar D, Al Shehabi TS, Kobeissy F, Moussalem C, Massaad E, Omeis I, Darwiche N, Eid AH. Novel therapeutic strategies for spinal osteosarcomas. Semin Cancer Biol. 2020; 64:83–92. <u>https://doi.org/10.1016/j.semcancer.2019.05.018</u> PMID:<u>31152785</u>
- Hui L, Chen Y. Tumor microenvironment: sanctuary of the devil. Cancer Lett. 2015; 368:7–13. <u>https://doi.org/10.1016/j.canlet.2015.07.039</u> PMID:<u>26276713</u>
- 14. Arneth B. Tumor Microenvironment. Medicina (Kaunas). 2019; 56:15. https://doi.org/10.3390/medicina56010015 PMID:<u>31906017</u>
- Di Martino JS, Mondal C, Bravo-Cordero JJ. Textures of the tumour microenvironment. Essays Biochem. 2019; 63:619–29. <u>https://doi.org/10.1042/EBC20190019</u> PMID:31654075
- Jerez S, Araya H, Thaler R, Charlesworth MC, López-Solís R, Kalergis AM, Céspedes PF, Dudakovic A, Stein GS, van Wijnen AJ, Galindo M. Proteomic analysis of exosomes and exosome-free conditioned media from human osteosarcoma cell lines reveals secretion of proteins related to tumor progression. J Cell Biochem. 2017; 118:351–60.

https://doi.org/10.1002/jcb.25642 PMID:27356893

- Cortini M, Avnet S, Baldini N. Mesenchymal stroma: role in osteosarcoma progression. Cancer Lett. 2017; 405:90–99. <u>https://doi.org/10.1016/j.canlet.2017.07.024</u> PMID:28774797
- Suehara Y, Alex D, Bowman A, Middha S, Zehir A, Chakravarty D, Wang L, Jour G, Nafa K, Hayashi T, Jungbluth AA, Frosina D, Slotkin E, et al. Clinical genomic sequencing of pediatric and adult osteosarcoma reveals distinct molecular subsets with potentially targetable alterations. Clin Cancer Res. 2019; 25:6346–56.

https://doi.org/10.1158/1078-0432.CCR-18-4032 PMID:<u>31175097</u>

 Roberts RD, Lizardo MM, Reed DR, Hingorani P, Glover J, Allen-Rhoades W, Fan T, Khanna C, Sweet-Cordero EA, Cash T, Bishop MW, Hegde M, Sertil AR, et al. Provocative questions in osteosarcoma basic and translational biology: a report from the Children's Oncology Group. Cancer. 2019; 125:3514–25. https://doi.org/10.1002/cncr.32351 PMID:31355930

 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

- 21. Ding Q, Dong S, Wang R, Zhang K, Wang H, Zhou X, Wang J, Wong K, Long Y, Zhu S, Wang W, Ren H, Zeng Y. A nine-gene signature related to tumor microenvironment predicts overall survival with ovarian cancer. Aging (Albany NY). 2020; 12:4879–95. <u>https://doi.org/10.18632/aging.102914</u> PMID:<u>32208363</u>
- 22. Li F, Teng H, Liu M, Liu B, Zhang D, Xu Z, Wang Y, Zhou H. Prognostic Value of Immune-Related Genes in the Tumor Microenvironment of Bladder Cancer. Front Oncol. 2020; 10:1302.
 https://doi.org/10.3389/fonc.2020.01302
 PMID:32850407
- Cai WY, Dong ZN, Fu XT, Lin LY, Wang L, Ye GD, Luo QC, Chen YC. Identification of a tumor microenvironmentrelevant gene set-based prognostic signature and related therapy targets in gastric cancer. Theranostics. 2020; 10:8633–47. <u>https://doi.org/10.7150/thno.47938</u> PMID:<u>32754268</u>
- 24. 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; 12:453–57. <u>https://doi.org/10.1038/nmeth.3337</u> PMID:<u>25822800</u>
- Gorlick R, Liao AC, Antonescu C, Huvos AG, Healey JH, Sowers R, Daras M, Calleja E, Wexler LH, Panicek D, Meyers PA, Yeh SD, Larson SM. Lack of correlation of functional scintigraphy with (99m)technetiummethoxyisobutylisonitrile with histological necrosis following induction chemotherapy or measures of pglycoprotein expression in high-grade osteosarcoma. Clin Cancer Res. 2001; 7:3065–70. PMID:11595696
- 26. 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

https://doi.org/10.18632/aging.10282 PMID:<u>32039832</u>

- Li Y, Ge F, Wang S. Four genes predict the survival of osteosarcoma patients based on TARGET database. J Bioenerg Biomembr. 2020; 52:291–99. <u>https://doi.org/10.1007/s10863-020-09836-6</u> PMID:32514876
- Liu S, Liu J, Yu X, Shen T, Fu Q. Identification of a twogene (PML-EPB41) signature with independent prognostic value in osteosarcoma. Front Oncol. 2020; 9:1578. https://doi.org/10.3389/fonc.2019.01578

PMID:32039036

- 29. Shi Y, He R, Zhuang Z, Ren J, Wang Z, Liu Y, Wu J, Jiang S, Wang K. A risk signature-based on metastasisassociated genes to predict survival of patients with osteosarcoma. J Cell Biochem. 2020; 121:3479–90. https://doi.org/10.1002/jcb.29622 PMID:31898371
- Heymann MF, Schiavone K, Heymann D. Bone sarcomas in the immunotherapy era. Br J Pharmacol. 2020. [Epub ahead of print]. https://doi.org/10.1111/bph.14999 PMID:31975481
- Alfranca A, Martinez-Cruzado L, Tornin J, Abarrategi A, Amaral T, de Alava E, Menendez P, Garcia-Castro J, Rodriguez R. Bone microenvironment signals in osteosarcoma development. Cell Mol Life Sci. 2015; 72:3097–113. <u>https://doi.org/10.1007/s00018-015-1918-γ</u>

PMID:25935149

 Senthebane DA, Rowe A, Thomford NE, Shipanga H, Munro D, Mazeedi MA, Almazyadi HA, Kallmeyer K, Dandara C, Pepper MS, Parker MI, Dzobo K. The role of tumor microenvironment in chemoresistance: to survive, keep your enemies closer. Int J Mol Sci. 2017; 18:1586.

https://doi.org/10.3390/ijms18071586 PMID:28754000

 Chocarro de Erauso L, Zuazo M, Arasanz H, Bocanegra A, Hernandez C, Fernandez G, Garcia-Granda MJ, Blanco E, Vera R, Kochan G, Escors D. Resistance to PD-L1/PD-1 blockade immunotherapy. A tumor-intrinsic or tumor-extrinsic phenomenon? Front Pharmacol. 2020; 11:441.

https://doi.org/10.3389/fphar.2020.00441 PMID:<u>32317979</u>

- 34. Wang F, Yu T, Ma C, Yuan H, Zhang H, Zhang Z. Prognostic value of programmed cell death 1 ligand-1 in patients with bone and soft tissue sarcomas: a systemic and comprehensive meta-analysis based on 3,680 patients. Front Oncol. 2020; 10:749. <u>https://doi.org/10.3389/fonc.2020.00749</u> PMID:32582532
- 35. 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

- 36. Li B, Wang Z, Wu H, Xue M, Lin P, Wang S, Lin N, Huang X, Pan W, Liu M, Yan X, Qu H, Sun L, et al. Epigenetic regulation of CXCL12 plays a critical role in mediating tumor progression and the immune response in osteosarcoma. Cancer Res. 2018; 78:3938–53. <u>https://doi.org/10.1158/0008-5472.CAN-17-3801</u> PMID:<u>29735547</u>
- 37. Komatsu K, Li JD. Cleaved cochlin: one guard with two roles in surveillance. Cell Host Microbe. 2019; 25:475–76. https://doi.org/10.1016/j.chom.2019.03.011 PMID:<u>30974079</u>
- 38. Bhattacharya SK. Focus on molecules: cochlin. Exp Eye Res. 2006; 82:355–56. <u>https://doi.org/10.1016/j.exer.2005.09.023</u> PMID:<u>16297912</u>
- 39. Nagy I, Trexler M, Patthy L. The second von willebrand type a domain of cochlin has high affinity for type I, type II and type IV collagens. FEBS Lett. 2008; 582:4003–07. <u>https://doi.org/10.1016/j.febslet.2008.10.050</u> PMID:19013156
- 40. van der Heul-Nieuwenhuijsen L, Hendriksen PJ, van der Kwast TH, Jenster G. Gene expression profiling of the human prostate zones. BJU Int. 2006; 98:886–97. <u>https://doi.org/10.1111/j.1464-410X.2006.06427.x</u> PMID:<u>16978289</u>
- Santorelli L, Capitoli G, Chinello C, Piga I, Clerici F, Denti V, Smith A, Grasso A, Raimondo F, Grasso M, Magni F. In-depth mapping of the urinary *N*-glycoproteome: distinct signatures of ccRCC-related progression. Cancers (Basel). 2020; 12:239. https://doi.org/10.3390/cancers12010239 PMID:31963743
- 42. Moussion C, Sixt M. A conduit to amplify innate immunity. Immunity. 2013; 38:853–54. <u>https://doi.org/10.1016/j.immuni.2013.05.005</u> PMID:<u>23706666</u>
- Jung J, Yoo JE, Choe YH, Park SC, Lee HJ, Lee HJ, Noh B, Kim SH, Kang GY, Lee KM, Yoon SS, Jang DS, Yoon JH, et al. Cleaved cochlin sequesters pseudomonas aeruginosa and activates innate immunity in the inner ear. Cell Host Microbe. 2019; 25:513–25.e6. <u>https://doi.org/10.1016/j.chom.2019.02.001</u> PMID:<u>30905438</u>
- Py BF, Gonzalez SF, Long K, Kim MS, Kim YA, Zhu H, Yao J, Degauque N, Villet R, Ymele-Leki P, Gadjeva M, Pier GB, Carroll MC, Yuan J. Cochlin produced by follicular

dendritic cells promotes antibacterial innate immunity. Immunity. 2013; 38:1063–72. <u>https://doi.org/10.1016/j.immuni.2013.01.015</u> PMID:23684986

- 45. Agarkova I, Perriard JC. The M-band: an elastic web that crosslinks thick filaments in the center of the sarcomere. Trends Cell Biol. 2005; 15:477–85. <u>https://doi.org/10.1016/j.tcb.2005.07.001</u> PMID:<u>16061384</u>
- 46. Lu Y, Ye Y, Bao W, Yang Q, Wang J, Liu Z, Shi S. Genome-wide identification of genes essential for podocyte cytoskeletons based on single-cell RNA sequencing. Kidney Int. 2017; 92:1119–29. <u>https://doi.org/10.1016/j.kint.2017.04.022</u> PMID:<u>28709640</u>
- Wiesen MH, Bogdanovich S, Agarkova I, Perriard JC, Khurana TS. Identification and characterization of layer-specific differences in extraocular muscle mbands. Invest Ophthalmol Vis Sci. 2007; 48:1119–27. <u>https://doi.org/10.1167/iovs.06-0701</u> PMID:<u>17325154</u>
- 48. Camicia R, Winkler HC, Hassa PO. Novel drug targets for personalized precision medicine in relapsed/refractory diffuse large B-cell lymphoma: a comprehensive review. Mol Cancer. 2015; 14:207. <u>https://doi.org/10.1186/s12943-015-0474-2</u> PMID:<u>26654227</u>
- 49. Sørensen LK, Havemose-Poulsen A, Sønder SU, Bendtzen K, Holmstrup P. Blood cell gene expression profiling in subjects with aggressive periodontitis and chronic arthritis. J Periodontol. 2008; 79:477–85. <u>https://doi.org/10.1902/jop.2008.070309</u> PMID:<u>18315430</u>
- 50. Liu J, Zhu Q, Han J, Zhang H, Li Y, Ma Y, He D, Gu J, Zhou X, Reveille JD, Jin L, Zou H, Ren S, Wang J. IgG galactosylation status combined with MYOM2rs2294066 precisely predicts anti-TNF response in ankylosing spondylitis. Mol Med. 2019; 25:25. <u>https://doi.org/10.1186/s10020-019-0093-2</u> PMID:<u>31195969</u>
- 51. Pehlivan D, Bayram Y, Gunes N, Coban Akdemir Z, Shukla A, Bierhals T, Tabakci B, Sahin Y, Gezdirici A, Fatih JM, Gulec EY, Yesil G, Punetha J, et al, and Baylor-Hopkins Center for Mendelian Genomics. The genomics of arthrogryposis, a complex trait: candidate genes and further evidence for oligogenic inheritance. Am J Hum Genet. 2019; 105:132–50. <u>https://doi.org/10.1016/j.ajhg.2019.05.015</u> PMID:31230720
- 52. Francis SH, Turko IV, Corbin JD. Cyclic nucleotide phosphodiesterases: relating structure and function. Prog Nucleic Acid Res Mol Biol. 2001; 65:1–52.

https://doi.org/10.1016/s0079-6603(00)65001-8 PMID:<u>11008484</u>

 Azevedo MF, Faucz FR, Bimpaki E, Horvath A, Levy I, de Alexandre RB, Ahmad F, Manganiello V, Stratakis CA. Clinical and molecular genetics of the phosphodiesterases (PDEs). Endocr Rev. 2014; 35:195–233.

https://doi.org/10.1210/er.2013-1053 PMID:24311737

- 54. Hertz AL, Beavo JA. Cyclic nucleotides and phosphodiesterases in monocytic differentiation. Handb Exp Pharmacol. 2011; 204:365–90. <u>https://doi.org/10.1007/978-3-642-17969-3 16</u> PMID:<u>21695649</u>
- 55. Bender AT, Ostenson CL, Wang EH, Beavo JA. Selective up-regulation of PDE1B2 upon monocyte-tomacrophage differentiation. Proc Natl Acad Sci USA. 2005; 102:497–502. <u>https://doi.org/10.1073/pnas.0408535102</u> PMID:<u>15625104</u>
- 56. Bender AT, Beavo JA. PDE1B2 regulates cGMP and a subset of the phenotypic characteristics acquired upon macrophage differentiation from a monocyte. Proc Natl Acad Sci USA. 2006; 103:460–65. <u>https://doi.org/10.1073/pnas.0509972102</u> PMID:<u>16407168</u>
- 57. Wittliff JL, Sereff SB, Daniels MW. Expression of genes for methylxanthine pathway-associated enzymes accompanied by sex steroid receptor status impacts breast carcinoma progression. Horm Cancer. 2017; 8:298–313. https://doi.org/10.1007/s12672-017-0309-2

https://doi.org/10.1007/s12672-017-0309-2 PMID:<u>28971320</u>

- 58. Chatterjee G, Rosner A, Han Y, Zelazny ET, Li B, Cardiff RD, Perkins AS. Acceleration of mouse mammary tumor virus-induced murine mammary tumorigenesis by a p53 172H transgene: influence of FVB background on tumor latency and identification of novel sites of proviral insertion. Am J Pathol. 2002; 161:2241–53. <u>https://doi.org/10.1016/S0002-9440(10)64500-2</u> PMID:<u>12466138</u>
- Rothlin CV, Ghosh S. Lifting the innate immune barriers to antitumor immunity. J Immunother Cancer. 2020; 8:e000695. <u>https://doi.org/10.1136/jitc-2020-000695</u> PMID:32273348
- 60. Mortezaee K. Immune escape: a critical hallmark in solid tumors. Life Sci. 2020; 258:118110. <u>https://doi.org/10.1016/j.lfs.2020.118110</u> PMID:<u>32698074</u>
- 61. Rogovskii VS. The linkage between inflammation and immune tolerance: interfering with inflammation in cancer. Curr Cancer Drug Targets. 2017; 17:325–32.

https://doi.org/10.2174/1568009617666170109110816 PMID:<u>28067176</u>

62. Müller DA, Silvan U. On the biomechanical properties of osteosarcoma cells and their environment. Int J Dev Biol. 2019; 63:1–8.
 https://doi.org/10.1387/ijdb.190019us
 PMID:30919911

 Heymann MF, Lézot F, Heymann D. The contribution of immune infiltrates and the local microenvironment in the pathogenesis of osteosarcoma. Cell Immunol. 2019; 343:103711. <u>https://doi.org/10.1016/j.cellimm.2017.10.011</u> PMID:29117898

- 64. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, Diehn M, West RB, Plevritis SK, Alizadeh AA. The prognostic landscape of genes and infiltrating immune cells across human cancers. Nat Med. 2015; 21:938–945. <u>https://doi.org/10.1038/nm.3909</u> PMID:26193342
- 65. Fridman WH, Zitvogel L, Sautès-Fridman C, Kroemer G. The immune contexture in cancer prognosis and treatment. Nat Rev Clin Oncol. 2017; 14:717–34. <u>https://doi.org/10.1038/nrclinonc.2017.101</u> PMID:<u>28741618</u>
- 66. Turan T, Kannan D, Patel M, Matthew Barnes J, Tanlimco SG, Lu R, Halliwill K, Kongpachith S, Kline DE, Hendrickx W, Cesano A, Butterfield LH, Kaufman HL, et al. Immune oncology, immune responsiveness and the theory of everything. J Immunother Cancer. 2018; 6:50.

https://doi.org/10.1186/s40425-018-0355-5 PMID:29871670

 Wu CC, Beird HC, Andrew Livingston J, Advani S, Mitra A, Cao S, Reuben A, Ingram D, Wang WL, Ju Z, Hong Leung C, Lin H, Zheng Y, et al. Immuno-genomic landscape of osteosarcoma. Nat Commun. 2020; 11:1008. <u>https://doi.org/10.1038/s41467-020-14646-w</u>

PMID:32081846

- 68. Bielack SS, Kempf-Bielack B, Delling G, Exner GU, Flege S, Helmke K, Kotz R, Salzer-Kuntschik M, Werner M, Winkelmann W, Zoubek A, Jürgens H, Winkler K. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. J Clin Oncol. 2002; 20:776–90. https://doi.org/10.1200/JCO.2002.20.3.776 PMID:11821461
- 69. 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

- 70. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16:284–87. <u>https://doi.org/10.1089/omi.2011.0118</u> PMID:<u>22455463</u>
- 71. Li W, Yin Y, Quan X, Zhang H. Gene expression value prediction based on XGBoost algorithm. Front Genet. 2019; 10:1077. https://doi.org/10.3389/fgene.2019.01077
 PMID:31781160
- 72. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017; 18:220. https://doi.org/10.1186/s13059-017-1349-1 PMID:29141660
- 73. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. J Clin Oncol. 2008; 26:1364–70.

https://doi.org/10.1200/JCO.2007.12.9791 PMID:<u>18323559</u>

- 74. Bhattacharya S, Andorf S, Gomes L, Dunn P, Schaefer H, Pontius J, Berger P, Desborough V, Smith T, Campbell J, Thomson E, Monteiro R, Guimaraes P, et al. ImmPort: disseminating data to the public for the future of immunology. Immunol Res. 2014; 58:234–39. <u>https://doi.org/10.1007/s12026-014-8516-1</u> PMID:<u>24791905</u>
- 75. Burns RJ, Deschênes SS, Schmitz N. Associations between depressive symptoms and social support in adults with diabetes: comparing directionality hypotheses with a longitudinal cohort. Ann Behav Med. 2016; 50:348–57.

https://doi.org/10.1007/s12160-015-9760-x PMID:<u>26631086</u>

SUPPLEMENTARY MATERIALS

Supplementary Figures



Supplementary Figure 1. Comparison of the level of TIICs between low risk group and high risk group by XCell. The horizontal axis represents the type of microenvironment cells, and the vertical axis represents the relative enrichment score. Blue and red represent low and high risk score, respectively. The Wilcoxon rank-sum test assessed data. *P<0.05, **P<0.01, ***P<0.001. NS, no significance; TIICs, tumour-infiltrating immune cells.



Supplementary Figure 2. Calibration plot of the nomogram model for predicting the probability of OS at 1, 3, and 5 years. (A) 1-year calibration curve. (B) 3-year calibration curve. (C) 5-year calibration curve.



Supplementary Figure 3. (A) 1, 3 and 5-year ROC curve for survival prediction of osteosarcoma patients by nomogram model in GSE21257. (B) 1, 3 and 5-year calibration plot for survival prediction of osteosarcoma patients by nomogram model in GSE21257.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 2, 3.

	level	Total
patient(n)	-	82
Survival time (months)	-	49.88 (±35.01)
age (years)	-	14.50 (±3.92)
Survival status (n[%])	Alive	56 (68.3)
	Dead	26 (31.7)
Gender(n[%])	Male	45 (54.9)
	Female	37 (45.1)
Race(n[%])	While	48 (58.5)
	Asian	7 (8.5)
	Black or African American	7 (8.5)
	Unknown	20 (24.4)
Disease at diagnosis (n[%])	Non-metastatic	61 (74.4)
	Metastatic	21 (25.6)
Metastasis at lung (n[%])	Yes	27 (32.9)
	Unknown	55 (67.1)
Primary tumor site(n[%])	Leg/Foot	74 (90.2)
	Arm/hand	6 (7.3)
	Pelvis	2 (2.4)
Specific tumor site(n[%])	Femur	44 (53.7)
	Tibia /Fibula	28 (34.1)
	Humerus	6 (7.3)
	Pelvis	4 (4.9)
Specific tumor side(n[%])	Left	10 (12.2)
	Right	10 (12.2)
	Unknown	62 (75.6)
Specific tumor region(n[%])	Distal	26 (31.7)
	Proximal	22 (26.8)
	Unknown	34 (41.5)
Definitive Surgery(n[%])	Limb sparing	38 (46.3)
	Amputation	4 (4.9)
	Unknown	40 (48.8)
Primary site progression(n[%])	No	19 (23.2)
	Yes	15 (18.3)
	Unknown	48 (58.5)

Supplementary Table 2A. 262 differential genes up-regulated in the high immune score group.

Supplementary Table 2B. 320 differential genes down-regulated in the high immune score group.

Supplementary Table 3A. 160 differential genes up-regulated in the high stromal score group.

Supplementary Table 3B. 251 differential genes down-regulated in the high stromal score group.

Supplementary Table 4. The intersection of differential genes between high and low score groups.

Gene	Regulation
CCL14	up
CCR2	up
FCGR1A	up
FCGR1B	up
FCGR1C	up
FCGR2B	up
GAPT	up
GIMAP5	up
IGHG2	up
IGHM	up
IGHV1-24	up
IGKV3-11	up
IL10	up
IL2RA	up
LILRA1	up
LILRA2	up
LLNLR-470E3.1	up
MS4A14	up
P2RY10	up
PLD4	up
RP11-24F11.2	up
RP11-455F5.5	up
RP11-494O16.3	up
RP11-733O18.1	up
RP11-84C10.2	up
TLR7	up
TLR8	up
TNFSF8	up

TRBV29-1	up
TRBV7-3	up
VENTX	up
AC003988.1	down
AC013470.6	down
COCH	down
EMILIN3	down
FAM60BP	down
GS1-309P15.4	down
HAUS6P1	down
HK2P1	down
MAP3K15	down
RP1-290I10.2	down
RP11-283C24.1	down
RPRML	down

Id	HR	HR.95 L	HR.95 H	P value
ARHGAP9	0.51087512	0.279370094	0.934220925	0.029190049
ARL11	0.2925651	0.106229156	0.805751841	0.017415939
CARD11	0.366473556	0.135386517	0.991995883	0.048179537
CCL2	0.554934323	0.361183981	0.852618385	0.007196217
СОСН	1.444447813	1.154257554	1.807594396	0.001310445
COLEC10	2.165998823	1.213980939	3.864600135	0.008886882
DENND1C	0.40798314	0.170180199	0.978082315	0.044469242
FCGR2B	0.330313897	0.122811646	0.888411433	0.028209966
GPR65	0.367016467	0.149575271	0.9005572	0.028620701
HSPD1P5	5.996700537	1.928439911	18.64741397	0.00197144
HTR2B	0.173056522	0.037741629	0.793515297	0.023968604
IGHG2	0.695364235	0.487357095	0.992150157	0.045132177
IL2RA	0.247279044	0.077194996	0.792109969	0.018656486
ITGAM	0.472112561	0.266834219	0.835313669	0.009935059
KBTBD11-OT1	3.569424862	1.869198572	6.816179958	0.000115665
LAG3	0.438454408	0.197341701	0.974159375	0.042945491
LILRB2	0.516296389	0.28565598	0.933157295	0.028593342
LILRB4	0.560448573	0.358927426	0.875114524	0.010874652
LRRC25	0.629703223	0.399797335	0.991817892	0.045998229
MIR4666A	2.388143135	1.302508685	4.378648448	0.004886248
MS4A4A	0.609618914	0.409457223	0.907628928	0.014799605
MYOM2	1.673891374	1.229811511	2.278326645	0.001056289
NCF1B	0.256873893	0.086789752	0.760276362	0.014088033
NCF4	0.630684051	0.422394571	0.941684386	0.024212517
PCED1B	0.373265768	0.164832238	0.845267499	0.01812482
PDE1B	0.128492357	0.028553448	0.578223886	0.007499985
PIK3R5	0.345663883	0.141728729	0.843043756	0.019527423
RASGRP4	0.262925485	0.098455778	0.702140718	0.007685921
SNX20	0.267559459	0.087584958	0.817355696	0.020671466
TNFSF8	0.256172732	0.088599698	0.740685017	0.011933582
TRBV7-3	0.061055107	0.004092557	0.910855004	0.042593103
VSIG4	0.663373939	0.489320843	0.899338314	0.008210926

Supplementary Table 5. Screening of genes affecting prognosis of osteosarcoma patients by univariate Cox regression analysis.

Supplementary T	Table 6. Immune-related	differential genes in	high and low risk group.
-----------------	-------------------------	-----------------------	--------------------------

Gene	Low risk	High risk	Log FC	P Value	FDR
C3	2.101113494	0.993065337	-1.081193548	3.42E-06	0.000164068
CCL23	0.163230344	0.078820756	-1.050261781	0.007313622	0.035887774
CD3D	1.257976912	0.581421327	-1.113449547	3.45E-05	0.000757924
CD3E	1.234124531	0.559110455	-1.142282751	2.65E-06	0.000162902
CNTFR	0.763638404	1.740009475	1.188133599	0.008324904	0.039384637
COLEC10	0.116194113	0.342232389	1.558439322	3.88E-05	0.000805077
CSPG5	0.331461945	0.664455344	1.003329	5.66E-06	0.000232374
CX3CR1	1.310084234	0.647589283	-1.016508559	0.000112751	0.001950043
CXCR3	0.864951829	0.425292057	-1.024165878	1.88E-05	0.000505711
CXCR6	0.42387118	0.211257015	-1.004626627	0.000203797	0.003205921
FCGR2B	0.925072469	0.353461012	-1.388015301	6.26E-09	4.34E-06
IGHM	1.600302821	0.779517085	-1.03769238	0.004409497	0.027274265
IL2RA	0.770377786	0.294716765	-1.386236974	3.42E-06	0.000164068
PAK6	0.013952663	0.047304117	1.761425248	0.005420618	0.02981382
PRLHR	0.015912746	0.035923823	1.174758031	0.009123249	0.041138437
PROK2	0.357443064	0.932191954	1.382913603	0.001419738	0.011887491
PTHLH	0.467021326	0.15219946	-1.617525187	0.000558506	0.006400523
RETN	0.403612052	0.074929169	-2.429369894	0.000337595	0.004686353
RNASE2	1.131581553	0.441860267	-1.356678453	3.89E-05	0.000805077
TLR7	1.040862022	0.519912736	-1.001437434	2.79E-06	0.000162902
TRAV1-2	0.179271881	0.061862702	-1.535007475	0.000217345	0.003229565
TRAV14DV4	0.209995614	0.055742782	-1.913502272	0.000922779	0.008931487
TRBC2	1.572979597	0.750764535	-1.067067551	7.74E-05	0.001432748
TRBV12-4	0.278282339	0.138663227	-1.00496411	0.005634189	0.030482407
TRBV20-1	0.625931807	0.136296403	-2.199257999	1.42E-05	0.00039359
TRBV28	0.980494657	0.45568394	-1.105476244	0.000227234	0.003329603
TRBV29-1	0.230535009	0.09849979	-1.226793302	0.005011945	0.029217714
TRBV4-2	0.288890712	0.060070621	-2.265792342	0.000564217	0.006400523
TRBV6-5	0.302632378	0.119943755	-1.335208298	0.007216178	0.035887774
TRBV7-3	0.259605434	0.063285289	-2.0363785	8.82E-06	0.00026592
TRDC	1.285825752	0.538065431	-1.256841623	0.000919254	0.008931487