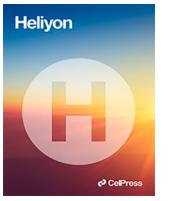
Heliyon 10 (2024) e32231

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Identification of immune-related tumor antigens and immune subtypes in osteosarcoma

Mingshu Zhang, Gongping Xu, Chunyang Xi, Enming Yu *

Department of Orthopedics, The Second Affiliated Hospital of Harbin Medical University, Harbin, China

ARTICLE INFO

Keywords:
Osteosarcoma
Immunotherapy
Tumor antigen
Immune-related genes
Gene vaccine

ABSTRACT

Purpose: The development of tumor vaccines has become a hot topic in immunotherapy for osteosarcoma (OS); however, more tumor antigens with stronger immunogenicity need to be identified.

Methods: We downloaded six sets of gene expression profile data from online databases. The overexpressed genes were analyzed, intersected, and used to calculate the immune infiltration abundance in the TARGET OS dataset based on their expression matrix. Potential tumor antigen genes were identified based on whether they exhibited a high correlation with the antigen-presenting cells (APCs). A total of 1330 immune-related genes (IRGs) from the ImmPort website were retrieved based on their expression, and the Consensus Cluster method was used to obtain immune subtypes of the OS samples. Prognosis, immune microenvironment, and sensitivity to drugs were compared among the immune subtypes.

Results: In total, 680 genes were overexpressed in at least two datasets, of which *TREM2, TNFRSF12A,* and *THY1* were positively correlated with different APCs. Based on the expression matrix of 1330 IRGs in TARGET-OS, two immune subtypes, IS1 and IS2, were identified. The prognosis of the IS1 subtype was better than that of IS2, the expression of immune checkpoint (ICP)-related genes was higher in patients with the IS1 subtype, and immune cell infiltration and sensitivity to 16 drugs were generally higher in IS1 subtype patients.

Conclusion: We identified three APC-correlated genes that can be considered to code for potential novel tumor antigens for OS vaccines. Two immune subtypes in patients with OS were identified to implement personalized treatments using mRNA vaccines.

Key points

TREM2, TNFRSF12A, and *THY1* were positively correlated with different antigen-presenting cells.
These three antigen-presenting cell-correlated genes can be considered to code for potential novel tumor antigens for osteosarcoma vaccines.
Two immune subtypes in patients with osteosarcoma were identified to implement personalized treatments using mRNA vaccines.

* Corresponding author.
E-mail addresses: yu1348498234@163.com, 821028@hrbmu.edu.cn (E. Yu).

https://doi.org/10.1016/j.heliyon.2024.e32231
Received 9 January 2024; Received in revised form 29 May 2024; Accepted 30 May 2024
Available online 31 May 2024

1. Introduction

Osteosarcoma (OS) is a common malignant bone tumor that occurs in children, accounting for approximately 5 % of pediatric tumors [1,2]. The combination of neoadjuvant chemotherapy and surgery remains the preferred treatment for most patients, proving to be effective in raising the survival rate of patients. However, the chemotherapy toxicity, high incidence of tumor metastasis and recurrence are still buffling doctors [3]. According to a meta-analysis, adjuvant chemotherapy (with or without radiotherapy) did not decrease recurrence rate in some types of sarcoma [4]. Many studies also suggested that target therapy and immunotherapy should play important roles in the treatment of sarcom [5–7]. How to kill tumor cells precisely has become an urgent issue that needs to be addressed in the treatment of this disease. Immunotherapy is a highly specific method for killing tumor cells by activation of the host immune system and has proven to be a highly potent antitumor tool [8]. Research has shown that nucleic acid vaccines, genetically modified vaccines, DC and B cell vaccines, and HSP peptide complex vaccines, supplemented with appropriate adjuvants, cytokines, or co-stimulatory molecules to stimulate the body's immune system, produce tumor-specific cytotoxic T lymphocytes (CTLs) and antibodies, and they have good application prospects for preventing tumor formation and clearing tumor lesions [9].

Nucleic acid vaccines (or gene vaccines), including DNA and RNA vaccines, involve cloning of the gene fragments that carry antigen information into a eukaryotic expression plasmid, and the plasmid is then injected into the host so that the gene-encoding antigen can be expressed in the host and presented to T cells, thereby stimulating the host body to produce the corresponding antibodies and CTLs and mediating fluid immunity and cellular immune responses [10]. In recent years, personalized mRNA vaccines containing a variety of new tumor antigens have been developed for melanoma; however, no such vaccines have been developed for OS [11]. During the COVID-19 pandemic, mRNA vaccines were validated for clinical applications. However, compared to traditional protein vaccines, the antigen expression level of mRNA is still low and the immunogenicity is limited. Exploring new tumor antigens with stronger immunogenicity is crucial for the development of mRNA vaccines.

Therefore, we conducted data mining of the gene expression profiles of OS in public databases to identify specific genes that could be used as tumor antigens in the preparation of OS gene vaccines. We also attempted to identify different immune subtypes and predict their effects during mRNA vaccine treatment.

2. Materials and methods

2.1. Data collection and preprocessing

We downloaded the expression profile data of the osteosarcoma (OS) GDC TARGET-OS dataset from the UCSC Xena database (https://xenabrowser.net/datapages), including HTSeq Counts (n = 88) and HTSeq FPKM (n = 88), and we converted FPKM values into TPM values. We also downloaded the phenotype (n = 524) and survival data (n = 288) for OS. A total of 86 patients with both expression and survival data were selected for subsequent analyses. From the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/), we downloaded the OS expression profile datasets GSE19276 [12], GSE99671 [13], GSE155646 [14], and GSE126209 [15], and we standardized the data for subsequent differential gene expression analysis between normal and OS samples. We also downloaded the OS dataset GSE39055 [16] from the GEO dataset, which contains information on patient survival status and survival time for prognostic analysis. Detailed information regarding the five GEO datasets is provided in Table 1.

Additionally, we obtained data from the ImmPort database (https://immport.org). A total of 1330 immune-related genes (IRGs) were identified. We also obtained additional immunogenic cell death (ICD)-related genes (ICDRGs) and immune checkpoint (ICP) genes (ICPRGs) from the published literature [17,18].

2.2. Overexpressed genes in OS tissue

We performed differential analysis between OS and normal tissues using the GSE19276, GSE66971, GSE155646, and GSE126209 datasets. The process was conducted using the "limma" package [19], and it was specified that genes with $|\log2(\text{fold change})| > 1$, and P-value <0.05, were over-expressed, where $\log2(\text{fold change}) > 1$ means that the genes are highly expressed in OS, whereas $|\log2(\text{fold change}) < 0|$ means that genes are expressed at low levels in OS. The "ggplot" package was used to generate the volcano plot to display the differentially expressed genes (DEGs) in each dataset, and the intersection of DEGs was displayed through the Upset plot.

Correlation analysis of DEGs with antigen-presenting cells and survival analysis of the potential genes.

Single-sample gene set enrichment analysis (ssGSEA) was used to perform immune cell infiltration analysis on the transcriptome data of TARGET-OS OS tissues to evaluate the infiltration abundance of the 28 types of immune cells. We used the R package "GSVA" to

Table 1
Osteosarcoma dataset sourced from the GEO database.

Dataset	Platform	Data type	Samples
GSE19276	GPL6848	Expression profiling by array	49
GSE99671	GPL20148	Expression profiling by high-throughput sequencing	51
GSE155646	GPL20988	Expression profiling by high-throughput sequencing	7
GSE126209	GPL20301	Expression profiling by high-throughput sequencing	11
GSE39055	GPL16951	Expression profiling by array	37

perform this [20]. We calculated the correlation between potential tumor antigens and major antigen-presenting cells (APCs) such as dendritic cells, macrophages, and B cells. The Kaplan–Meier (KM) method was used to evaluate the overall survival of patients with TARGET-OS for potential genes of interest, with a median cutoff value and comparison using the logarithmic rank test.

2.3. Identification of immune subtypes

We clustered the samples of TARGET-OS dataset with the 1330 IRGs by using ConsensusClusterPlus R package [21] based on their expressions. The median algorithm was used for partitioning, with 1000 replicates, and 80 % of the patients in the queue were sampled each time. The value range of the clustering set was 2–10. We validated the immune subtypes in the GSE39055 dataset using the same settings.

2.4. Prognostic evaluation of the immune subtypes and differences in the tumor microenvironment

We used the log-rank test to evaluate the prognostic value of different immune subtypes. We also used analysis of variance to determine the correlation between immune subtypes, different immune-related molecules, and cell characteristics.

2.5. Weighted gene Co-expression network analysis

We used the "WGCNA" R package [22] to screen modules for IRGs. The "pickSoftTreshold" function was used to calculate the soft threshold. By choosing a soft threshold, we constructed a scale-free network, prepared the topology matrix, calculated eigengenes

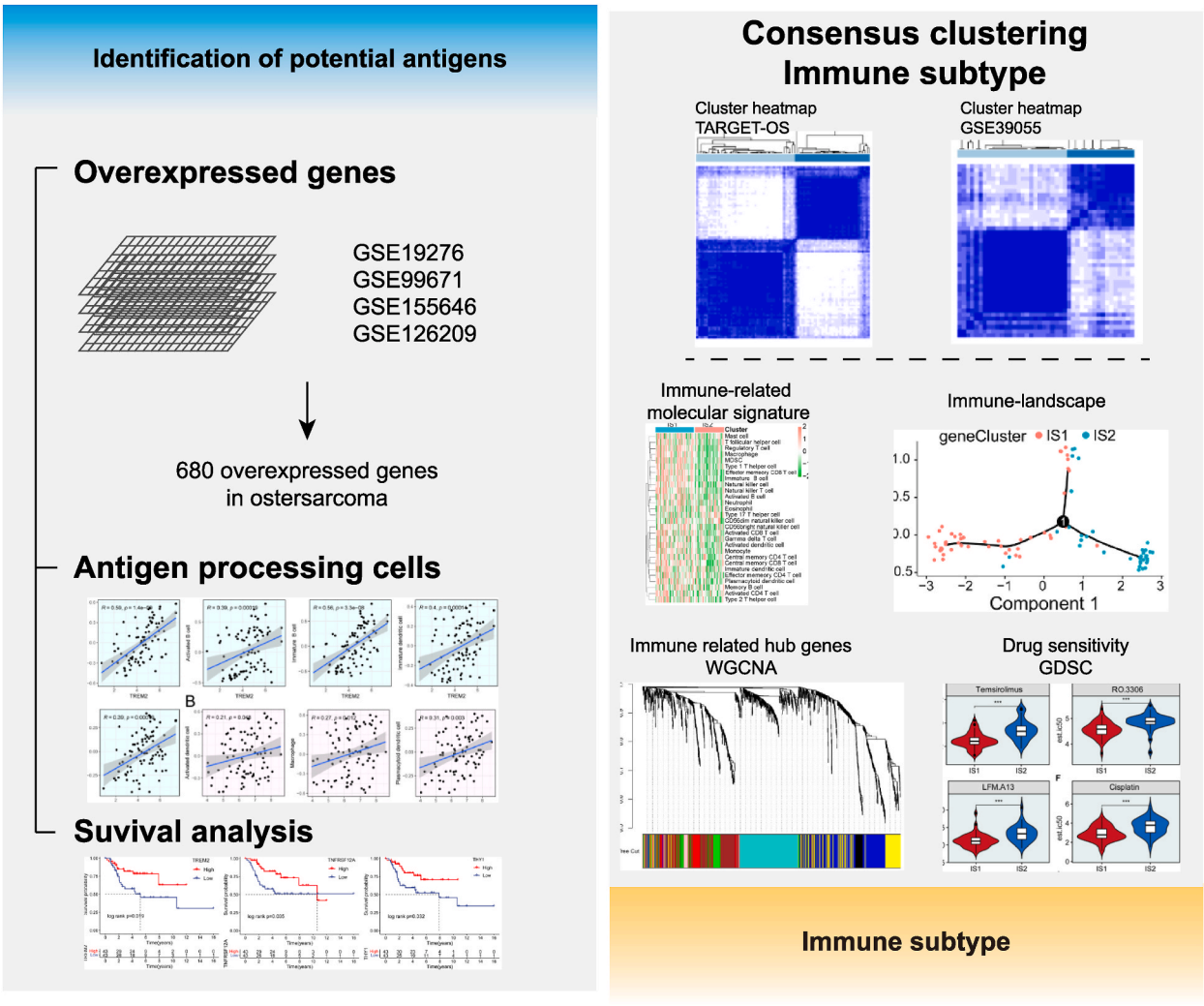


Fig. 1. Flowchart of the process.

through hierarchical clustering, and determined the correlation between them. Functional enrichment analysis was conducted using the Metascape database (www.metascapeorg/) [23] to cluster genes into modules, including both the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Genealogy (GO) [24,25], with Adj. $P < 0.05$ to be considered significantly enriched.

2.6. Construction of the tumor microenvironment immune landscape

R package "Monocle" with Gaussian distribution was used for dimensionality reduction analysis based on graph learning [26]. We calculated the distribution of immune subtypes with the parameters set as: the maximum number of components = 2 and dimensionality reduction method to be "DDRtree." The landscape was visualized using a color-coded functional map.

2.7. Drug sensitivity analysis in cancer

The Genomics of Drug Sensitivity in Cancer (GDSC) database [27] (https://www.cancerrxgene.org/) was searched for tumor drug response data related to specific genes. We used the "pRRophic" algorithm [28] to predict the sensitivity of the immune subtypes to anticancer drugs or small molecule compounds through IC_{50} values.

2.8. Statistical methods

R software (version 4.1.1) was used for all computational analyses. We compared the differences between two groups using the Wilcoxon rank-sum test and the Kruskal–Wallis test for multiple groups. We used the Spearman correlation method for correlation analysis. Statistical significance was set at $P < 0.05$.

3. Results

3.1. Identification of potential immune-related antigens in OS

The workflow chart is depicted in Fig. 1. To identify potential OS immune antigens for mRNA vaccines, we screened for overexpressed genes. Differential analysis of the four datasets from GEO showed that there were 1573 overexpressed genes in GSE19276 (Fig. 2A), 1235 overexpressed genes in GSE99671 (Fig. 2B), 442 overexpressed genes in GSE155646 (Fig. 2C), and 2063 overexpressed

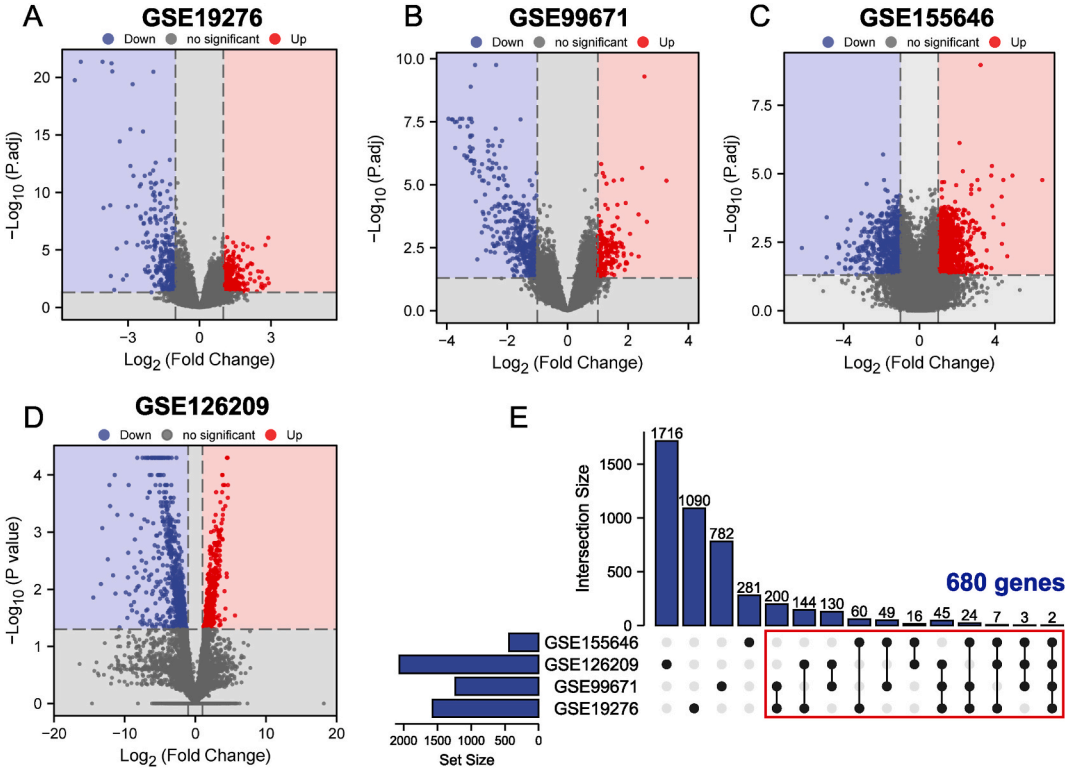


Fig. 2. Screening of overexpression genes in OS. A-D. Differential gene volcano plot of GSE19276, GSE66971, GSE155646, and GSE126209. E. UpSet plot of the intersection of overexpressed genes from four datasets.

genes in GSE126209 (Fig. 2D). The Upset plot showed that 680 genes were overexpressed in at least two datasets (Fig. 2E). We then analyzed the possibility of using these genes as potential OS antigens. Considering the interaction between tumor antigens and antigen-presenting cells, we evaluated 28 immune cells in the TARGET-OS dataset using ssGSEA and analyzed the correlation between APCs and potential tumor antigen genes. The results revealed that the *TREM2, TNFRSF12A,* and *THY1* genes were significantly positively correlated with APCs. Of these, *TREM2* was positively correlated with activated dendritic cells, activated B cells, and immature B cells (Fig. 3A), *TNFRSF12A* was positively correlated with activated dendritic cells, macrophage, and plasmoid dendritic cells (Fig. 3B); and *THY1* was significantly positively correlated with macrophages and plasmoid dendritic cells (Fig. 3C). Prognostic analysis of *TREM2, TNFRSF12A,* and *THY1* revealed that overexpression of these three genes was associated with a better prognosis (Fig. 3D–F). The proteins encoded by these genes may be directly recognized and presented to T cells by APCs, thereby triggering an immune response that kills tumor cells.

3.2. Identification of potential immune subtypes of OS

Immune subtype differentiation can help determine which patients are suitable for vaccination. In the consensus clustering process based on the 1330 IRGs in TARGET-OS, according to the cumulative distribution function and functional trigonometric area, we chose k = 2, where the IRGs were stably aggregated (Fig. 4A–C, E), resulting in two immune subtypes, identified as IS1 and IS2. The principal component analysis (PCA) results showed significant differentiation between subtypes IS1 and IS2 (Fig. 4D). GSE39055 can be divided into the immune subtypes IS1 and IS2, which is consistent with the TARGET-OS dataset (Fig. 4F). The results of the KM survival analysis showed that the prognosis of type IS2 was worse than that of type IS1. The two immune subtypes in the GSE39055 dataset were similar to those in TARGET-OS.

3.3. The relationship between OS immune subtypes and immune modulators

ICPRGs and ICDRGs play important roles in regulating the antitumor immunity of the host, which affects the efficacy of mRNA vaccines. Therefore, we evaluated the differential expression of ICPRGs and ICDRGs in the two immune subtypes. Twenty-five ICDRGs were detected in the TARGET-OS dataset, of which nine exhibited differential expression between the two subtypes, and 25 ICDRGs were detected in the GSE39055 dataset, of which 12 exhibited differential expression. We found differential expression of *P2RX7, PANX1*, and *TLR4* in both datasets (Fig. 5A and B). In addition, 46 ICPRGs were found in the TARGET-OS dataset, of which 27 exhibited

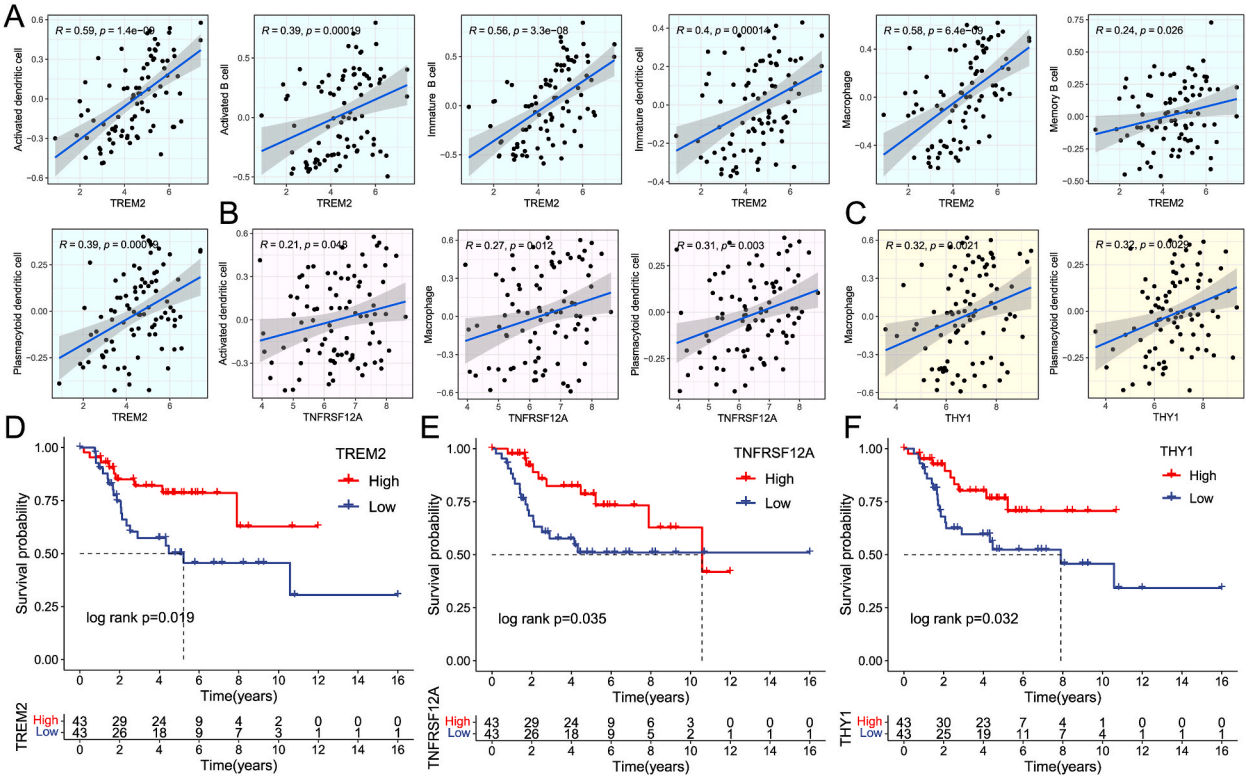


Fig. 3. Identification of tumor antigens related to antigen presentation and prognosis. A. Correlation between *TREM2* and APCs; B. Correlation between *TNFRSF12A* and APCs; C. Correlation between *THY1* and APCs. D. High expression of *TREM2* predicts better prognosis; E. High expression of *TNFRSF12A* predicts better prognosis; F. High expression of *THY1* predicts better prognosis. APCs, antigen-presenting cells.

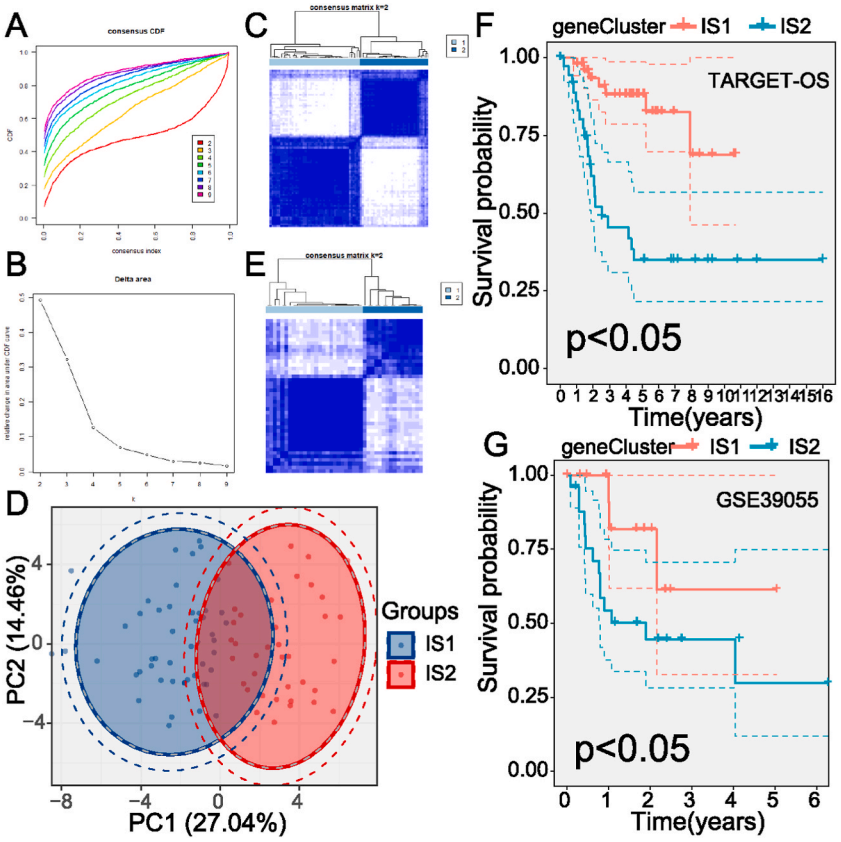


Fig. 4. Identification of potential immune subtypes of OS. A. The cumulative distribution function curve of IRGs in the TARGET-OS dataset. B. The delta area curve. C. TARGET-OS sample K = 2 clustering heat map. D. Principal component analysis of immune subtypes in TARGET-OS. E. GSE39055 sample K = 2 clustering heat map. F. The KM curve on the TARGET-OS dataset. G. The KM curve on the GSE39055 dataset. KM, Kaplan–Meier.

differential expression, and 46 ICPRGs were found in the GSE39055 dataset, of which 13 were differentially expressed. The expression trends of multiple ICPRGs in the two datasets were similar, indicating that the expression of ICPRGs in IS1 patients was higher than that in IS2 (Fig. 5C–D).

3.4. Cellular and molecular characteristics of the immune subtypes

The immune status of the tumor determines the response to mRNA vaccines. After calculating the immune score based on the ESTIAMTE algorithm, we found that, compared to the IS2 subtype, the overall infiltration of immune cells and stromal cells in IS1-type tumor patients was higher, while the tumor purity was lower (Fig. 6A). Moreover, in the GSE39055 dataset, the trends of the two subtypes were consistent with those of TARGET-OS (Fig. 6B). We then compared the distribution differences of 28 types of immune cell infiltration between the IS1 and IS2 subtypes in both the TARGET-OS and GSE39055 datasets, and we found that the abundance of immune cell infiltration in IS1 patients was generally higher than that in patients with IS2 (Fig. 6C–D). The KM curve showed that three out of 22 immune cell types exhibited survival differences, including $CD56^{bright}$ natural killer cells, central memory $CD8^+$ T cells, and gamma-delta T cells, all of which exhibited high scores and better prognosis. (Fig. 6E–G). Subsequently, it was found that $CD56^{bright}$ natural killer cells and central memory $CD8^+$ T cells exhibited higher levels in the IS1 subtype than in the IS2 subtype, whereas gamma-delta T cells did not exhibit a statistically significant difference between the two subtypes (Fig. 6H–J). We speculate that IS1 may be an immune “hot” phenotype with active immune cells, whereas IS2 may be the “cold” phenotype with fewer active cells.

3.5. The immune landscape in OS tumors

Using the TARGET-OS dataset, we constructed an immune landscape map for OS (Fig. 7A). Based on the landscape, the patients could be roughly divided into three trajectories. As shown in (Fig. 7B), the horizontal axis (PCA1) and vertical axis (PCA2) axes were significantly negatively correlated with various immune cells. In addition, we observed different intracluster heterogeneities within the same subtype. Based on the trajectory position of the sample group, we further divided all samples into three types of states (Fig. 7C) and analyzed the proportion of states in different IS1 and IS2 immune types. The proportions of these subgroups in different

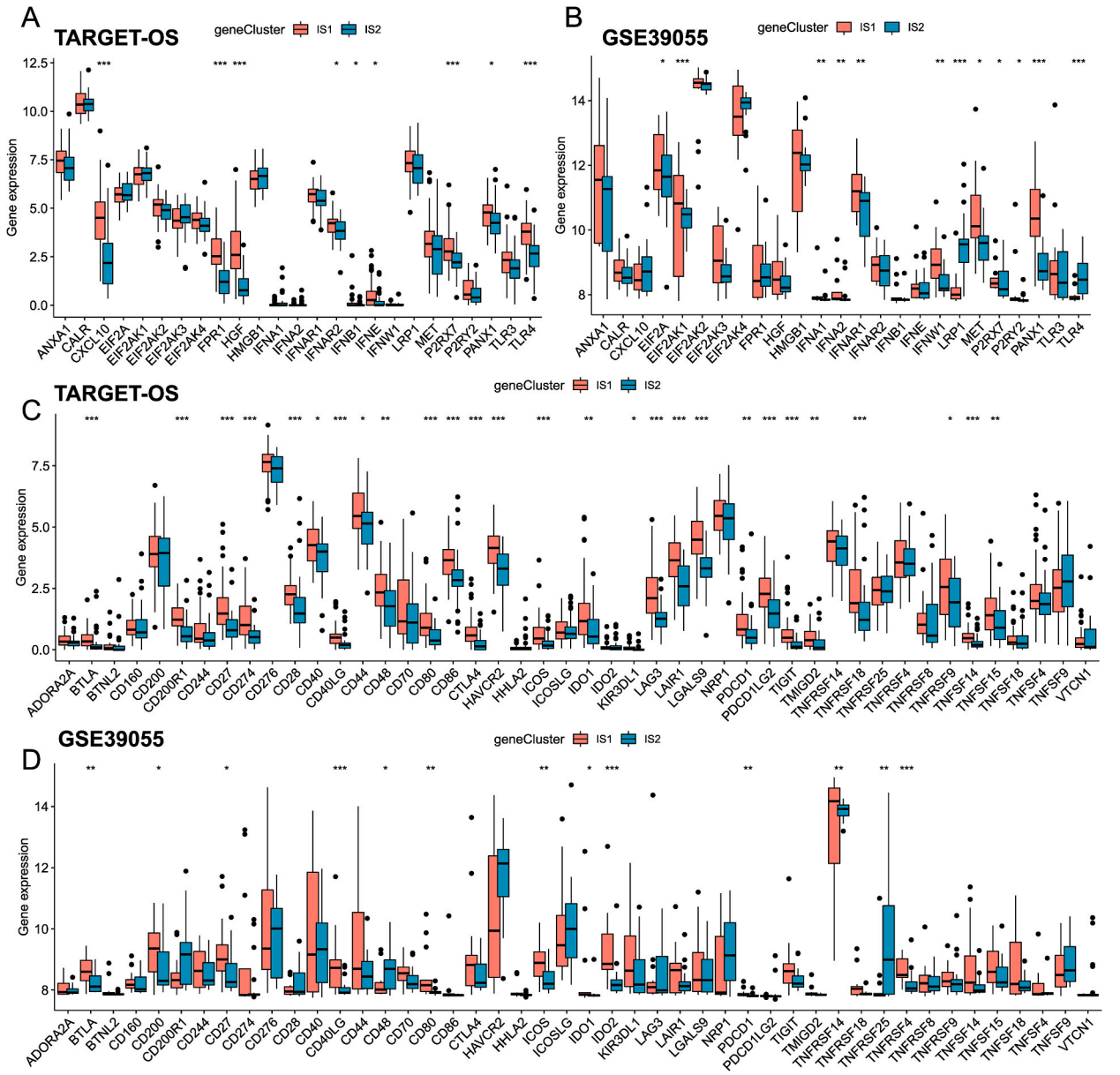


Fig. 5. The relationship between immune subtypes and ICD and ICP regulator genes. (A-B)The differences in the expression of ICD modulators between the immune subtypes in the datasets. A. TARGET-OS dataset. B. GSE39055. (C-D)The differences in the expression of ICP modulators between the immune subtypes in the datasets. C. TARGET-OS. D. GSE39055. *P < 0.05, **P < 0.01,***P < 0.001,****P < 0.0001.

immune subtypes are shown in (Fig. 7D). Prognostic analysis showed significant overall survival differences in the survival curves of these three states, with state 3 predicting the worst prognosis and state 1 the best (Fig. 7E).

3.6. Identification of the module of immune gene co-expression in OS

We performed a WGCNA of the IRGs with a soft threshold set at 4 (Fig. 8A). Each gene module contained at least 20 genes. By clustering and merging the modules with characteristic genes, we obtained nine modules, of which the gray module was unassigned (Fig. 8B–C). By comparing the eigengene scores of the modules, we found that, in most modules, the IS1 type score was significantly higher than the IS2 type score (Fig. 8D). We then conducted KM survival analysis on the other eight modules and found that the turquoise and brown modules exhibited high prognostic scores (Fig. 8E–F). After multivariate Cox regression analysis, only the turquoise module was found to be an independent prognostic factor (Fig. 8G). Functional enrichment analysis of genes using the Metascape database showed that they clustered in pathways such as cytokine signaling in the immune system, enzyme-linked receptor

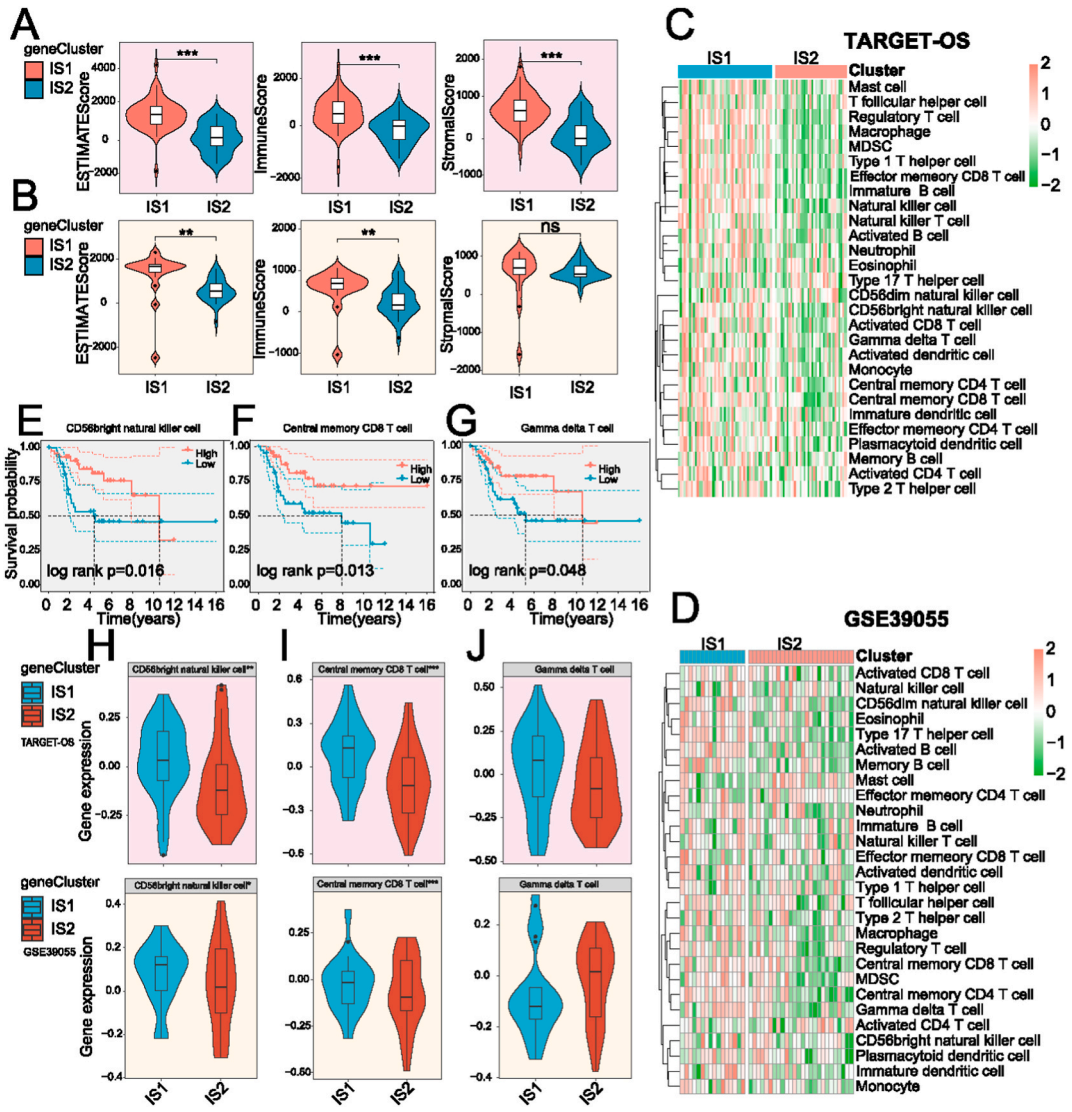


Fig. 6. Immune characteristics of the subtypes. A. Comparison of estimate scores, immune scores, and matrix scores between subtypes in the TARGER-OS dataset; B. Comparison of ESTIMATE scores, immune scores, and matrix scores between subtypes in the GSE39055 dataset; C. Enrichment score heatmap of 28 immune cell markers for OS immune subtypes in the TARGET-OS dataset. D. Differential enrichment score heatmap of 28 immune-related immune cell markers in the GSE39055 dataset. E-G. The high and low scores of $CD56^{bright}$ natural killer cells, central memory $CD8^{+}$ T cells, and gamma-delta T cells in TARGET-OS revealed significant prognostic differences. H-J. Differences in the expression of $CD56^{bright}$ natural killer cells, central memory $CD8^{+}$ T cells, and gamma-delta T cells in TARGET-OS and GSE39055 datasets. *P $<$ 0.05,**P $<$ 0.01,***P $<$ 0.001.

protein signaling pathway, T-cell receptor signaling pathway, and axon guidance (Fig. 9A–B). The genes in the turquoise module included *ABCC4*, *ACKR4*, *ACO1*, *ACVR1B*, *ACVR1C*, *ACVR2A*, *ACVR2B*, *ADAR*, *ADIPOR2*, *ADM*, *AEN*, *AGTR1*, and *AHNAK* (Table S1). The genes in these important modules may serve as biomarkers for predicting the prognosis of patients with OS and could be used to identify suitable patients for mRNA vaccines.

3.7. Sensitivity of the two immune subtypes to drugs

We found that among the 16 drugs with significant differences, the drug sensitivity of IS1 patients was generally higher than that of IS2 patients (Fig. 10A-P). Based on the above results, it is speculated that IS1-type patients may have a higher sensitivity to these drugs, which further emphasizes the importance of individualized treatment for patients with tumors.

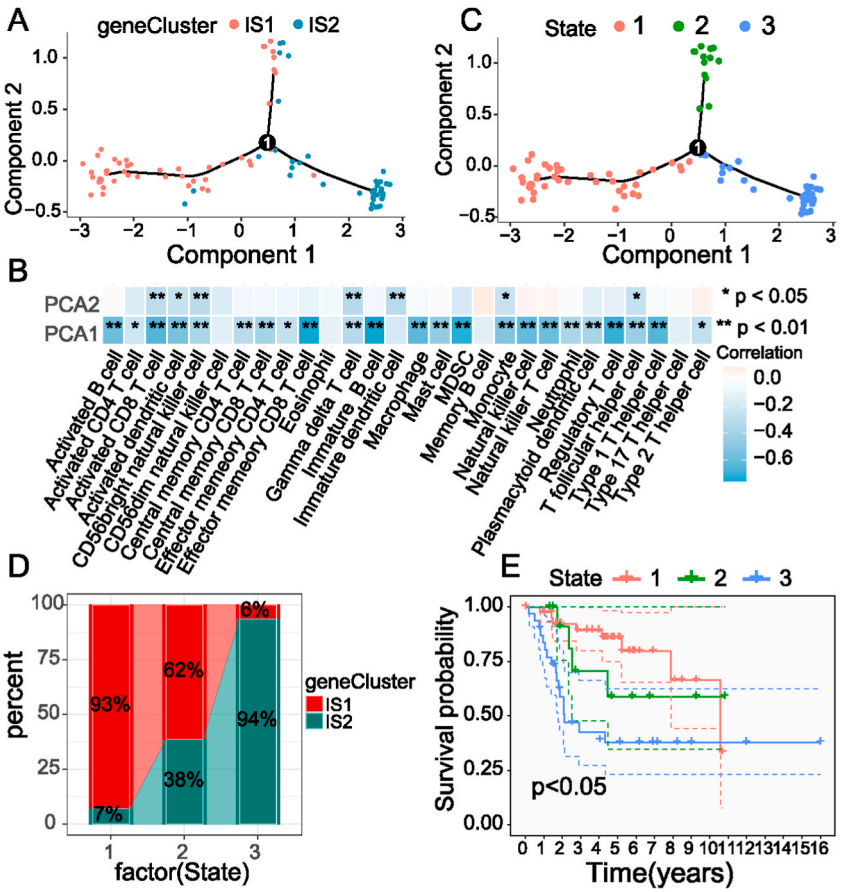


Fig. 7. The immune landscape of OS. A. The immune landscape of OS. The position and color of each patient in the immune spectrum correspond to the identified immune subtypes, representing the overall characteristics of TME. B. The PCA result. C. OS patients reclassified based on their location are divided into states 1, 2, and 3 D. The proportion of immune subtypes in patients with different states. E. Prognostic curves of patients with different states. There are differences in prognosis among patients with different states. *P < 0.05, **P < 0.01, ***P < 0.001. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

4. Discussion

Chemotherapy for OS has always been dominated by cytotoxic chemicals; however, the therapeutic effects of these chemicals are limited. A considerable proportion of patients develop multi-drug resistance during the induction of remission or consolidation treatment, resulting in a poor response to chemotherapy [29]. With improvements in supportive treatment and the development of surgical treatments, the dosage of chemotherapy drugs has significantly decreased, and the treatment effect on OS has also improved significantly. However, the long-term disease-free survival rate of patients with OS remains relatively low [30]. Therefore, research and development of treatments other than cytotoxic drugs have become extremely important.

Surgical care has importance in the treatment of osteosarcoma. Surgery is one of the main methods of treating osteosarcoma. Nurses need to assist the patient during the preoperative preparation by performing the necessary examinations, preparing surgical instruments and equipment. During surgery, nurses need to carefully monitor the patient's vital signs and participate in assisting during surgery. After surgery, nurses are responsible for wound care, dietary management, and pain control.

With the development of tumor immunology and molecular biology research, tumor biological therapy has become the fourth largest treatment method after traditional surgery, radiotherapy, and chemotherapy, and has achieved good results in the treatment of some solid tumors, such as melanoma, renal cell carcinoma, and colorectal cancer [31]. Among the various treatment methods, immunotherapy is the most promising approach to cure malignant tumors, including OS. The development of tumor vaccines, which are part of the immunotherapy arsenal, has become a hot topic in current cancer treatment research [32]. However, currently, the effectiveness of tumor vaccines is not ideal because the immunogenicity of tumor antigens is often too weak to trigger a robust immune response. Thus, the identification of more immunogenic tumor antigens is becoming increasingly urgent [33]. In addition, different immune microenvironments have different sensitivities to tumor vaccines, and personalized treatment is important in immunotherapy.

In this study, we downloaded four osteosarcoma gene expression profile datasets from the Gene Expression Omnibus (GEO)

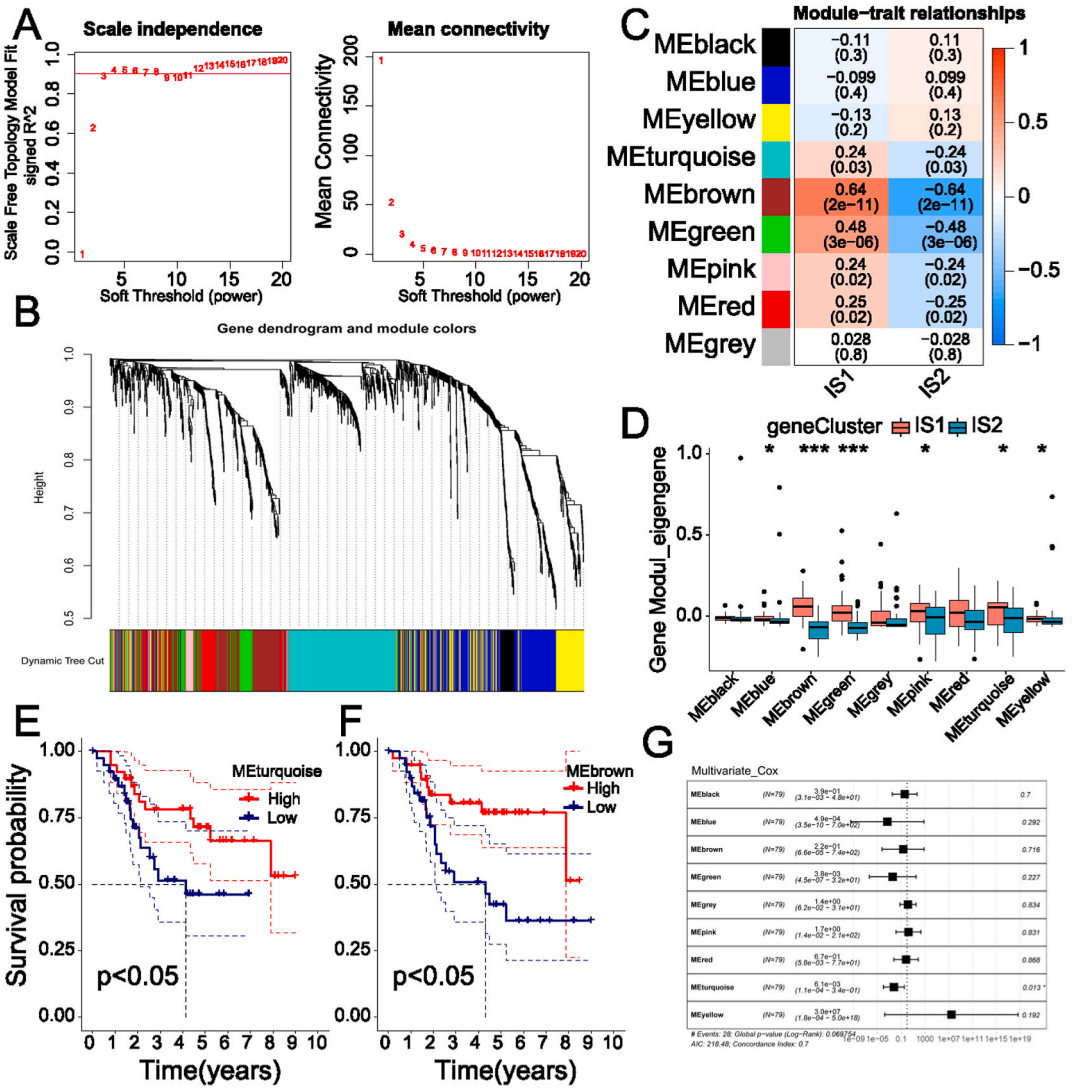


Fig. 8. Identification of OS core immune genes. A. The optimal soft threshold value for WGCNA is determined, with soft power = 4. B. WGCNA clustering tree. C. Nine modules obtained by WGCNA. D. Comparison of eigengene scores between IS1 and IS2 subtypes of different modules; E. The prognosis curve of the terquoise module reveals significant differences. F. The prognosis curve of the brown module reveals significant differences. G. Multivariate Cox analysis of different module ratings. *P < 0.05, **P < 0.01, ***P < 0.001. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

database, analyzed the differentially expressed genes in each dataset, and used the intersection to obtain 680 genes that were over-expressed in at least two datasets. We used TARGET-OS as the training dataset and evaluated the estimated score of 28 immune cell types for each gene using ssGSEA. We analyzed the correlation between the expression of presenting immune cells and potential tumor antigen genes and identified three genes closely related to antigen-presenting cells: *TREM2, TNFRSF12A,* and *THY1*. These three genes positively correlated with activated dendritic cells, activated B cells, immature B cells, macrophages, and plasmoid dendritic cells, indicating a high probability that antigen-presenting cells can recognize these gene products and present them to T cells, triggering a specific immune response. Dendritic cells play a crucial role in antigen capture, processing, presentation, and activation of lymphocytes to produce immune responses. Tumor-related antigens can be recognized and captured by the dendritic cells, triggering an immune response and killing cancer cells [34]. However, not all tumor antigens are recognized and presented by immune cells. In this study, we identified three such entities.

The *TREM2* gene encodes a membrane protein that is used to form a receptor signaling complex related to inflammatory cytokine production and immune response [35]. Xiong et al. [36] analyzed the publicly available scRNA-seq dataset of melanoma samples from patients undergoing immune checkpoint therapy (ICT) treatment and determined that *TREM2* could be used as a presentation antigen of macrophages, which is consistent with the results of the present study.

TNFRSF12A belongs to the tumor necrosis factor receptor family. Zhang et al. [37] pointed out a significant correlation between

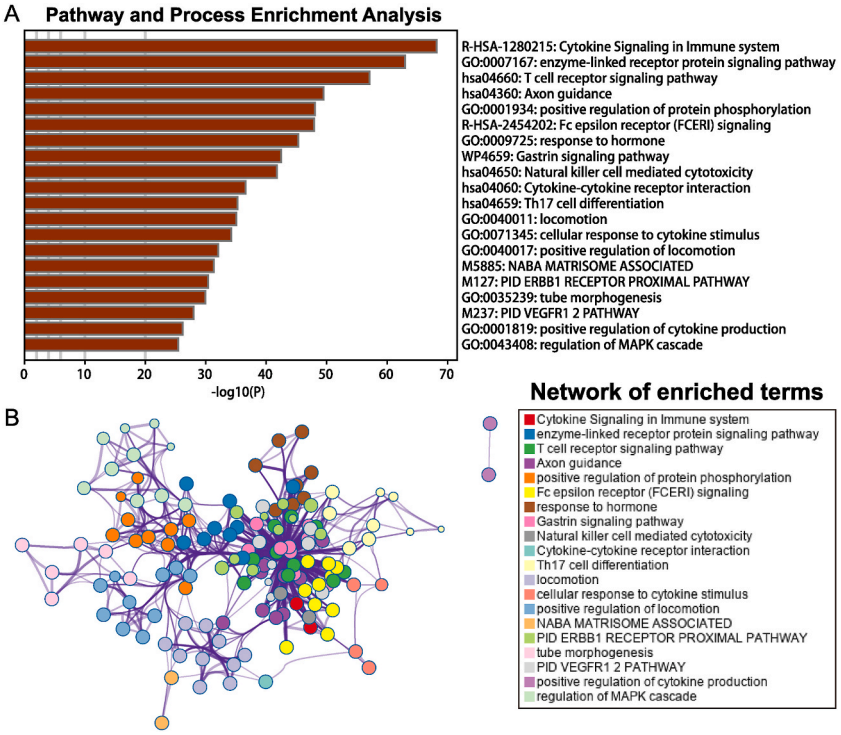


Fig. 9. Functional enrichment analysis of the blue module genes in WGCNA. A. The bar chart shows the TOP20 enrichment function and pathway of genes, including the enrichment analysis results of GO, KEGG, WP, Reactome databases. B. The network diagram shows the relationship between the enriched pathways. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

TNFRSF12A and immune cell infiltration. *TNFRSF12A* is also a prognostic biomarker of glioma and a potential immunotherapeutic target. In this study, it served as a sensitive antigen for OS, activating T cells to produce immune responses, which is consistent with previous research.

THY1 is a tumor cell surface antigen. A study by Lung et al. [38] showed that in nasopharyngeal carcinoma, this gene, as a tumor suppressor, has antitumor invasion activity. The tumor antigen may be captured by immune cells and presented to T cells, triggering a specific immune response, thereby playing a role in killing tumor cells and engaging in antitumor invasion activity.

The ROC analysis in this study showed that patients with high expression of these three genes in OS samples had a better prognosis, indicating that these three genes have strong immunogenicity as tumor antigens and have potential applications in immunotherapy.

Recognition of tumor-specific or tumor-related antigens by the host immune system is a prerequisite for the application of recombinant tumor vaccines and other immunotherapy methods. However, not all patients are sensitive to these vaccines. Different individuals' immune microenvironments and immune states lead to different reactions to tumor antigens, which results in nucleic acid vaccines (or mRNA vaccines) being effective in some individuals but ineffective in others. Therefore, identifying patients who are sensitive to a vaccine is also crucial for the application of mRNA vaccines.

We retrieved 1330 IRGs from an online database and conducted consensus clustering of the samples based on their expression profiles in the TARGET-OS gene dataset. Two immune subtypes, IS1 and IS2, were identified with significant differentiation between the two subtypes, and the prognosis of IS2 was worse than that of IS1. ICP- and ICD-related genes can predict the therapeutic effects of immune-targeted drugs. The efficacy of mRNA vaccines can be poor in patients with upregulated ICP-related genes and better in patients with upregulated ICD-related genes. We sought to identify ICP- and ICD-related genes in the overexpressed gene sets of both the TARGET-OS and GSE39055 datasets and analyzed their differential expression between the two immune subtypes. In terms of ICD-related genes, the expression trends of ICD genes between IS1 and IS2 of the two datasets were not the same, and no conclusion could be drawn. However, we found that the expression trend of multiple ICP genes in patients with the IS1 subtype was the same, with most of them being higher in type IS1 than in IS2 samples, indicating that mRNA vaccines have less of an effect on IS1 than on IS2.

Previous studies have shown that different immune subtypes indicated various biological heterogeneity and treatment response [39,40]. Therefore, it is necessary to investigate the underlying molecular subtypes according to immune patterns to evaluate and improve individualized medical decisions for patients with OS. Based on our analysis and research results, we speculated that IS1 subtype might exhibit an immune "hot" phenotype, associated with active immune cells, while IS2 subtype might represent an immune "cold" phenotype, characterized by fewer active immune cells. In the immune "hot" IS1 phenotype, there might be more anti-tumor immune cells, such as $CD56^{bright}$ natural killer cells, central memory $CD8^{+}$ T cells, and gamma-delta T cells, which typically exert a positive effect against tumor growth [41–43]. Conversely, the immune "cold" IS2 phenotype might be more inclined to contain immune cells that promote tumor development, such as Th17 cells [44]. Therefore, we considered the IS1 subtype as tending towards an

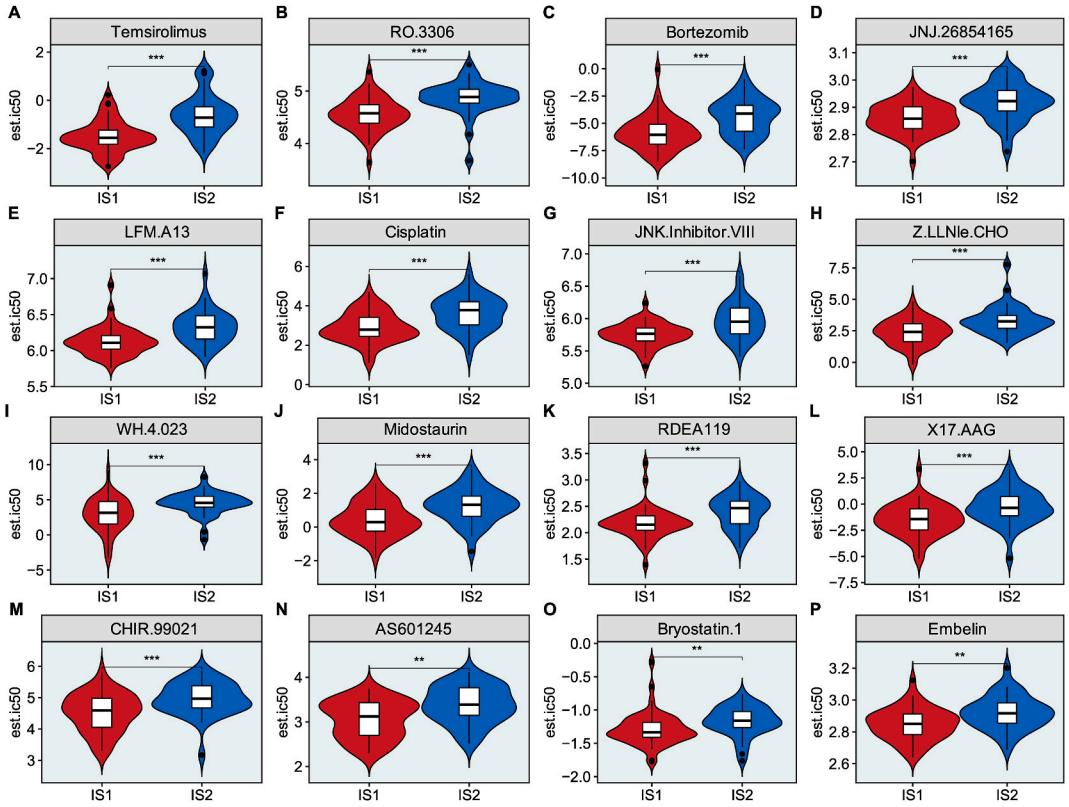


Fig. 10. Drug sensitivity of the two subtypes. A-P. Based on the GDSC database, the *P*-value ranking of IS1 and IS2 patients ranges from small to large for different drugs in Top16.

anti-tumor phenotype, while the IS2 subtype leans towards a pro-tumor phenotype. mRNA vaccines are more suitable for vaccination with the immune "cold" phenotype, which is consistent with the results of the ICP gene analysis mentioned above [45]. These findings may have important implications for guiding the future development and clinical application of mRNA vaccines that target different immune phenotypes.

To further identify the immune status, we constructed an immune landscape from which we observed that all samples could be divided into three states. Among the three states, state 1 had the best prognosis, state 2 had a medium prognosis, and state 3 had the worst prognosis. We analyzed the distribution proportion of these three states in the IS1 and IS2 subtypes and found that IS1 comprised the majority of state 1 and some state 2, while IS2 comprised the majority of state 3 and some state 2, indicating that IS1 patients have a better prognosis, whereas IS2 patients have a worse prognosis, which is consistent with the survival analysis results of the two immune subtypes mentioned above.

We further searched for potential tumor antigen genes based on the two immune subtypes, clustered and modularized the genes using WGCNA, and analyzed the differences between each module in the IS1 and IS2 subtypes. We found that the correlation score in most modules was higher in the IS1 type and that the turquoise module was an independent prognostic factor in survival analysis. Functional enrichment analysis revealed that the genes in this module were enriched in cytokine signaling in the immune system, the T-cell receptor signaling pathway, and other tumor antigen presentation and immune response channels, indicating that the genes in this module might serve as tumor antigens; however, the potential of the individual genes still needs to be explored.

We explored the sensitivity of the two immune subtypes to 16 drugs in the GDSC database. We found that the sensitivity of the IS1 subtype to tumor drug therapy was higher than that of IS2. Since the 16 drugs were primarily protease inhibitors or chemotherapeutic agents and do not include gene or mRNA vaccines, the conclusion that the IS2 subtype is more sensitive to mRNA vaccines has not been ruled out.

This study had some limitations. First, there were inherent limitations of bioinformatic analyses, including potential biases introduced by data preprocessing and algorithmic assumptions. Therefore, more experimental evidence is required to explore the biological mechanisms involved in our analysis. Second, the retrospective publicly available datasets may have inherent variability and limitations in terms of sample size, patient demographics, and selection bias. Some studies used many cancer cohorts to improve the reliability of the results [46–48]. In future studies, more OS cohorts will be used for the analysis. Additionally, although we identified three genes with high tumor antigen potential, their immunogenicity still needs to be validated through subsequent experiments and clinical trials.

There are several knowledge gaps and challenges in the current treatment of sarcoma, such as poor immune efficacy in some

patients and incomplete understanding of drug resistance mechanisms [4,5]. Current research is also identifying new therapeutic targets and developing personalized treatment plans to address these challenges through the integration of multi omics data [6,7]. In addition, with the advancement of technology and the deepening of research and development, there may be significant progress in this field in the next five years. For example, vaccine therapy for OS is expected to be further improved, precision therapy based on immune typing will become a commonly used clinical approach, and in-depth research on drug resistance mechanisms will also bring new solutions [49,50]. In the coming years, we hope experiments will be conducted to clarify the detailed pathways between these three antigens and APCs, and to develop mRNA vaccines that can be used for the treatment of osteosarcoma by utilizing these three genes. We also hope to have specific methods to distinguish immune subtypes in patients, to better implement personalized treatment.

5. Conclusion

Three potential tumor antigen genes, *TREM2*, *TNFRSF12A*, and *THY1*, were positively correlated with the expression of immune-presenting cells and can be considered to be potential novel tumor antigens for OS vaccines. There are two subtypes of immune cells in patients with OS, and the prognosis, immune infiltration, immune microenvironment, and sensitivity to mRNA vaccines of the two subtypes are significantly different. In clinical practice, attention should be paid to the identification of patient immune subtypes, and personalized treatments should be implemented.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Data availability statement

The datasets used during the current study are available from the corresponding author on reasonable request.

CRediT authorship contribution statement

Mingshu Zhang: Writing – original draft, Data curation, Conceptualization. **Gongping Xu:** Visualization, Resources, Methodology, Investigation, Formal analysis, Data curation. **Chunyang Xi:** Visualization, Validation, Software, Resources, Methodology, Formal analysis. **Enming Yu:** Writing – review & editing, Visualization, Supervision, Software, Project administration, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e32231.

References

[1] Z. Shoaib, T.M. Fan, J.M.K. Irudayaraj, Osteosarcoma mechanobiology and therapeutic targets, Br. J. Pharmacol. 179 (2) (2022) 201–217, https://doi.org/10.1111/bph.15713.
[2] L.R. Sadykova, A.I. Ntekim, M. Muyangwa-Semenova, et al., Epidemiology and risk factors of osteosarcoma, Cancer Invest. 38 (5) (2020) 259–269, https://doi.org/10.1080/07357907.2020.1768401.
[3] P.S. Meltzer, L.J. Helman, New horizons in the treatment of osteosarcoma, N. Engl. J. Med. 385 (22) (2021) 2066–2076, https://doi.org/10.1056/NEJMra2103423.
[4] Alessandro Rizzo, Margherita Nannini, Annalisa Astolfi, et al., Impact of chemotherapy in the adjuvant setting of early stage uterine leiomyosarcoma: a systematic review and updated meta-analysis, Cancers 12 (7) (2020) 1899, https://doi.org/10.3390/cancers12071899.
[5] Annalisa Astolfi, Margherita Nannini, Valentina Indio, et al., Genomic database analysis of uterine leiomyosarcoma mutational profile, Cancers 12 (8) (2020) 2126, https://doi.org/10.3390/cancers12082126.

[6] Alessandro Rizzo, Veronica Mollica, Valentina Tateo, et al., Hypertransaminasemia in cancer patients receiving immunotherapy and immune-based combinations: the MOUSEION-05 study, Cancer Immunol. Immunother. 72 (6) (2023) 1381–1394, https://doi.org/10.1007/s00262-023-03366-x.

[7] Alessandro Rizzo, Maria Abbondanza Pantaleo, Maristella Saponara, Margherita Nannini, Current status of the adjuvant therapy in uterine sarcoma: a literature review, World J Clin Cases 7 (14) (2019) 1753–1763, https://doi.org/10.12998/wjcc.v7.i14.1753.

[8] C. Chen, L. Xie, T. Ren, Y. Huang, J. Xu, W. Guo, Immunotherapy for osteosarcoma: fundamental mechanism, rationale, and recent breakthroughs, Cancer Lett. 500 (2021) 1–10, https://doi.org/10.1016/j.canlet.2020.12.024.

[9] F. Cersosimo, S. Lonardi, G. Bernardini, et al., Tumor-associated macrophages in osteosarcoma: from mechanisms to therapy, Int. J. Mol. Sci. 21 (15) (2020) 5207, https://doi.org/10.3390/ijms21155207. Published 2020 Jul 23.

[10] F. Jafari, S. Javdansirat, S. Sanaie, et al., Osteosarcoma: a comprehensive review of management and treatment strategies, Ann. Diagn. Pathol. 49 (2020) 151654, https://doi.org/10.1016/j.anndiagpath.2020.151654.

[11] K. Yahiro, Y. Matsumoto, Immunotherapy for osteosarcoma, Hum. Vaccines Immunother. 17 (5) (2021) 1294–1295, https://doi.org/10.1080/21645515.2020.1824499.

[12] L. Endo-Munoz, A. Cumming, S. Sommerville, I. Dickinson, N.A. Saunders, Osteosarcoma is characterised by reduced expression of markers of osteoclastogenesis and antigen presentation compared with normal bone, Br. J. Cancer 103 (1) (2010) 73–81, https://doi.org/10.1038/sj.bjc.6605723.

[13] X.D. Ho, P. Phung, V. Q Le, et al., Whole transcriptome analysis identifies differentially regulated networks between osteosarcoma and normal bone samples, Exp. Biol. Med. (Maywood, NJ, U. S.) 242 (18) (2017) 1802–1811, https://doi.org/10.1177/1535370217736512.

[14] S. Simpson, M. Dunning, S. de Brot, et al., Molecular characterisation of canine osteosarcoma in high risk breeds, Cancers 12 (9) (2020) 2405, https://doi.org/10.3390/cancers12092405. Published 2020 Aug 25.

[15] R. Liu, Y. Hu, T. Liu, Y. Wang, Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma cancer, BMC Cancer 21 (1) (2021) 1345, https://doi.org/10.1186/s12885-021-09042-6. Published 2021 Dec 18.

[16] A.D. Kelly, B. Haibe-Kains, K.A. Janeway, et al., MicroRNA paraffin-based studies in osteosarcoma reveal reproducible independent prognostic profiles at 14q32, Genome Med. 5 (1) (2013) 2, https://doi.org/10.1186/gm406. Published 2013 Jan 22.

[17] X. Huang, T. Tang, G. Zhang, T. Liang, Identification of tumor antigens and immune subtypes of cholangiocarcinoma for mRNA vaccine development, Mol. Cancer 20 (1) (2021) 50, https://doi.org/10.1186/s12943-021-01342-6. Published 2021 Mar 8.

[18] X. Huang, G. Zhang, T. Tang, T. Liang, Identification of tumor antigens and immune subtypes of pancreatic adenocarcinoma for mRNA vaccine development, Mol. Cancer 20 (1) (2021) 44, https://doi.org/10.1186/s12943-021-01310-0. Published 2021 Mar 1.

[19] M.E. Ritchie, B. Phipson, D. Wu, et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies, Nucleic Acids Res. 43 (7) (2015) e47, https://doi.org/10.1093/nar/gkv007.

[20] S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data, BMC Bioinf. 14 (2013) 7, https://doi.org/10.1186/1471-2105-14-7. Published 2013 Jan 16.

[21] M.D. Wilkerson, D.N. Hayes, ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking, Bioinformatics 26 (12) (2010) 1572–1573, https://doi.org/10.1093/bioinformatics/btq170.

[22] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008) 559, https://doi.org/10.1186/1471-2105-9-559. Published 2008 Dec 29.

[23] Y. Zhou, B. Zhou, L. Pache, et al., Metascape provides a biologist-oriented resource for the analysis of systems-level datasets, Nat. Commun. 10 (1) (2019) 1523, https://doi.org/10.1038/s41467-019-09234-6. Published 2019 Apr 3.

[24] M. Kanehisa, S. Goto, KEGG: kyoto encyclopedia of genes and genomes, Nucleic Acids Res. 28 (1) (2000) 27–30, https://doi.org/10.1093/nar/28.1.27.

[25] Gene Ontology Consortium, Gene ontology consortium: going forward, Nucleic Acids Res. 43 (Database issue) (2015) D1049–D1056, 10.1093/nar/gku1179.

[26] P. Perešíni, M. Kuźniar, D. Kostić, Monocle: dynamic, fine-grained data plane monitoring[C]. Proceedings of the 11th ACM Conference on Emerging Networking Experiments and Technologies, 2015, pp. 1–13.

[27] W. Yang, J. Soares, P. Greninger, et al., Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells, Nucleic Acids Res. 41 (Database issue) (2013) D955–D961, https://doi.org/10.1093/nar/gks1111.

[28] P. Geeleher, N. Cox, R.S. Huang, pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels, PLoS One 9 (9) (2014) e107468, https://doi.org/10.1371/journal.pone.0107468. Published 2014 Sep. 17.

[29] F. Jafari, S. Javdansirat, S. Sanaie, et al., Osteosarcoma: a comprehensive review of management and treatment strategies, Ann. Diagn. Pathol. 49 (2020) 151654, https://doi.org/10.1016/j.anndiagpath.2020.151654.

[30] S. Tsukamoto, C. Errani, A. Angelini, A.F. Mavrogenis, Current treatment considerations for osteosarcoma metastatic at presentation, Orthopedics 43 (5) (2020) e345–e358, https://doi.org/10.3928/01477447-20200721-05.

[31] J.F. Miller, M. Sadelain, The journey from discoveries in fundamental immunology to cancer immunotherapy, Cancer Cell 27 (4) (2015) 439–449, https://doi.org/10.1016/j.ccell.2015.03.007.

[32] H.G. Pratt, E.M. Justin, B.A. Lindsey, Applying osteosarcoma immunology to understand disease progression and assess immunotherapeutic response, Adv. Exp. Med. Biol. 1258 (2020) 91–109, https://doi.org/10.1007/978-3-030-43085-6_6.

[33] Y. Zhou, D. Yang, Q. Yang, et al., Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma [published correction appears in Nat Commun. 2021 Apr 30;12(1):2567], Nat. Commun. 11 (1) (2020) 6322, https://doi.org/10.1038/s41467-020-20059-6. Published 2020 Dec 10.

[34] A. Harari, M. Graciotti, M. Bassani-Sternberg, L.E. Kandalaft, Antitumour dendritic cell vaccination in a priming and boosting approach, Nat. Rev. Drug Discov. 19 (9) (2020) 635–652, https://doi.org/10.1038/s41573-020-0074-8.

[35] M. Binnewies, J.L. Pollack, J. Rudolph, et al., Targeting TREM2 on tumor-associated macrophages enhances immunotherapy, Cell Rep. 37 (3) (2021) 109844, https://doi.org/10.1016/j.celrep.2021.109844.

[36] D. Xiong, Y. Wang, M. You, A gene expression signature of TREM2^{hi} macrophages and γδ T cells predicts immunotherapy response, Nat. Commun. 11 (1) (2020) 5084, https://doi.org/10.1038/s41467-020-18546-x. Published 2020 Oct 8.

[37] Y. Zhang, X. Yang, X.L. Zhu, et al., A novel immune-related prognostic biomarker and target associated with malignant progression of glioma, Front. Oncol. 11 (2021) 643159, https://doi.org/10.3389/fonc.2021.643159. Published 2021 Apr 16.

[38] H.L. Lung, A.K. Cheung, Y. Cheng, et al., Functional characterization of THY1 as a tumor suppressor gene with antiinvasive activity in nasopharyngeal carcinoma, Int. J. Cancer 127 (2) (2010) 304–312, https://doi.org/10.1002/ijc.25047.

[39] X. Shen, X. Gu, R. Ma, X. Li, J. Wang, Identification of the immune signatures for ovarian cancer based on the tumor immune microenvironment genes, Front. Cell Dev. Biol. 10 (2022) 772701, https://doi.org/10.3389/fcell.2022.772701.

[40] S. Han, Q. Wang, M. Shen, X. Zhang, J. Wang, Immunogenic cell death related mRNAs associated signature to predict immunotherapeutic response in osteosarcoma, Heliyon 10 (6) (2024) e27630, https://doi.org/10.1016/j.heliyon.2024.e27630.

[41] B. Cozar, M. Greppi, S. Carpentier, E. Narni-Mancinelli, L. Chiossone, E. Vivier, Tumor-infiltrating natural killer cells, Cancer Discov. 11 (1) (2021) 34–44, https://doi.org/10.1158/2159-8290.CD-20-0655.

[42] B. Fan, B. Fan, N. Sun, H. Zou, X. Gu, A radiomics model to predict gammadelta T-cell abundance and overall survival in head and neck squamous cell carcinoma, Faseb. J. 38 (5) (2024) e23529, https://doi.org/10.1096/fj.202301353RR.

[43] S. Wang, H. Li, T. Chen, et al., Human gammadelta T cells induce CD8(+) T cell antitumor responses via antigen-presenting effect through HSP90-MyD88-mediated activation of JNK, Cancer Immunol. Immunother. 72 (6) (2023) 1803–1821, https://doi.org/10.1007/s00262-023-03375-w.

[44] P. Li, X. Shi, Y. Xu, B. Zhong, Y. Lu, Y. Sun, Interleukin-22 promotes osteosarcoma cell proliferation and invasion via STAT3 activation, Med. Sci. Mon. Int. Med. J. Exp. Clin. Res. 24 (2018) 7802–7808, https://doi.org/10.12659/MSM.910962.

[45] H. Xu, X. Zheng, S. Zhang, et al., Tumor antigens and immune subtypes guided mRNA vaccine development for kidney renal clear cell carcinoma, Mol. Cancer 20 (1) (2021) 159, https://doi.org/10.1186/s12943-021-01465-w.

[46] L. Yang, S. Wang, Q. Zhang, et al., Clinical significance of the immune microenvironment in ovarian cancer patients, Mol Omics 14 (5) (2018) 341–351, https://doi.org/10.1039/c8mo00128f.
[47] S. Wang, Q. Zhang, C. Yu, Y. Cao, Y. Zuo, L. Yang, Immune cell infiltration-based signature for prognosis and immunogenomic analysis in breast cancer, Briefings Bioinf. 22 (2) (2021) 2020–2031, https://doi.org/10.1093/bib/bbaa026.
[48] S. Wang, Y. Xiong, Q. Zhang, et al., Clinical significance and immunogenomic landscape analyses of the immune cell signature based prognostic model for patients with breast cancer, Briefings Bioinf. 22 (4) (2021), https://doi.org/10.1093/bib/bbaa311.
[49] Barbara Marcelli, Anne de Jong, Thomas Janzen, Mariela Serrano, Jan Kok, Oscar P. Kuipers, Complete genome sequences of 28 lactococcal bacteriophages isolated from failed dairy fermentation processes, Microbiol Resour Announc 9 (12) (2020) e01535, https://doi.org/10.1128/MRA.01535-19, 19.
[50] Vadahanambi Sridhar, Inwon Lee, Hyun Park, Metal organic frameworks derived Fe-N-C nanostructures as high-performance electrodes for sodium ion batteries and electromagnetic interference (EMI) shielding, Molecules 26 (4) (2021) 1018, https://doi.org/10.3390/molecules26041018.

Mingshu Zhang received her Bachelor Degree from the Harbin Medical University in 2015. Her research interests include bioinformatics, machine learning, and detection of mRNA expression.

Gongping Xu received his Ph.D from the Harbin Medical University in 1998. His research interests include biological network modeling, high-throughput biological data analysis and machine learning.

Chunyang Xi received his Ph.D from the Harbin Medical University in 2000. His current research interests include computer vision and graphics and machine learning.

Enming Yu received his Ph.D from the Harbin Medical University in 2021. His current research interests include big data mining and machine learning.