



OPEN ACCESS

EDITED BY

Suyan Tian,
Jilin University, China

REVIEWED BY

Jifeng Liu,
Dalian Medical University, China
Tian Tian,
Children's Hospital of Philadelphia,
United States

*CORRESPONDENCE

Daliu Min,
✉ mindaliu1969@163.com
Wei Kong,
✉ weikong@shmtu.edu.cn

[†]These authors have contributed equally to this work and share first authorship

RECEIVED 03 September 2023

ACCEPTED 09 October 2023

PUBLISHED 23 October 2023

CITATION

Jiang Z, Han K, Min D, Kong W, Wang S and Gao M (2023), Identification of the methotrexate resistance-related diagnostic markers in osteosarcoma via adaptive total variation netNMF and multi-omics datasets. *Front. Genet.* 14:1288073. doi: 10.3389/fgene.2023.1288073

COPYRIGHT

© 2023 Jiang, Han, Min, Kong, Wang and Gao. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identification of the methotrexate resistance-related diagnostic markers in osteosarcoma via adaptive total variation netNMF and multi-omics datasets

Zhihan Jiang^{1†}, Kun Han^{2†}, Daliu Min^{2*}, Wei Kong^{1*}, Shuaiqun Wang¹ and Min Gao¹

¹College of Information Engineering, Shanghai Maritime University, Shanghai, China, ²Department of Medical Oncology, The Sixth People's Hospital Affiliated to Shanghai Jiaotong University School of Medicine, Shanghai, China

Osteosarcoma is one of the most common malignant bone tumors with high chemoresistance and poor prognosis, exhibiting abnormal gene regulation and epigenetic events. Methotrexate (MTX) is often used as a primary agent in neoadjuvant chemotherapy for osteosarcoma; However, the high dosage of methotrexate and strong drug resistance limit its therapeutic efficacy and application prospects. Studies have shown that abnormal expression and dysfunction of some coding or non-coding RNAs (e.g., DNA methylation and microRNA) affect key features of osteosarcoma progression, such as proliferation, migration, invasion, and drug resistance. Comprehensive multi-omics analysis is critical to understand its chemoresistant and pathogenic mechanisms. Currently, the network analysis-based non-negative matrix factorization (netNMF) method is widely used for multi-omics data fusion analysis. However, the effects of data noise and inflexible settings of regularization parameters affect its performance, while integrating and processing different types of genetic data is also a challenge. In this study, we introduced a novel adaptive total variation netNMF (ATV-netNMF) method to identify feature modules and characteristic genes by integrating methylation and gene expression data, which can adaptively choose an anisotropic smoothing scheme to denoise or preserve feature details based on the gradient information of the data by introducing an adaptive total variation constraint in netNMF. By comparing with other similar methods, the results showed that the proposed method could extract multi-omics fusion features more effectively. Furthermore, by combining the mRNA and miRNA data of methotrexate (MTX) resistance with the extracted feature genes, four genes, Carboxypeptidase E (*CPE*), LIM, SH3 protein 1 (*LASP1*), Pyruvate Dehydrogenase Kinase 1 (*PDK1*) and Serine beta-lactamase-like protein (*LACTB*) were finally identified. The results showed that the gene signature could reliably predict the prognostic status and immune status of osteosarcoma patients.

KEYWORDS

osteosarcoma, non-negative matrix factorization, adaptive total variation, methotrexate resistance, diagnostic markers

1 Introduction

Osteosarcoma is one of the most common malignant bone cancers, accounting for approximately 30% of all osteosarcomas and mainly affecting children and adolescents, with a peak incidence at age 18 (Sadykova et al., 2020). Neoadjuvant chemotherapy (NAC) consisting of methotrexate, doxorubicin (also known as adriamycin), and cisplatin is referred to as MAP (Benjamin, 2020). The combination of NAC and surgical resection has significantly increased the 5-year survival rate for patients with osteosarcoma from 20% to 70% (Chen et al., 2021). However, up to 20% of patients develop resistance to this treatment regimen (Bacci et al., 2000), and their 5-year survival rate is extremely poor, at around 20% (Prudowsky and Yustein, 2020). Therefore, comprehensive analysis of multi-omics genetic data of osteosarcoma, screening for differentially expressed genes (DEGs) associated with drug resistance and analysis of the impact of DEGs on prognosis are essential for finding new targets to improve overall survival and reverse drug resistance.

Malignant osteosarcoma cells are strongly associated with chemoresistance, recurrence, and metastatic processes (Schivone et al., 2019; Mutsaers and Walkley, 2014), and osteosarcomas have significant heterogeneity at the genomic, transcriptomic, and epigenetic levels resulting from abnormal epigenetic modifications (Sun et al., 2023). For example, methylation levels and miRNA dysfunction have been identified as characteristic events in human osteosarcoma cell lines, with higher methylation events associated with more severe phenotypes (de Azevedo et al., 2019). Abnormal DNA methylation can affect gene expression, cell cycle, and apoptosis and regulate the development and progression of osteosarcoma by inhibiting transcription (Wang et al., 2020). miRNAs (microRNAs) are endogenous small non-coding RNAs that play critical regulatory roles in various biological processes, including differentiation, cell proliferation, cell cycle control, apoptosis, drug resistance, and innate immunity (Mens and Ghanbari, 2018; Patil et al., 2013). Although many studies have identified DNA methylation in osteosarcoma as an important therapeutic target, the reasons why DNA methylation, miRNAs, and target genes combine to lead to chemoresistance and poorer prognosis remain to be determined.

Currently, non-negative matrix factorization (NMF) and its various improvements are widely used in a single type of genetic data analysis. For example, Lei et al. applied NMF to osteosarcoma gene data analysis and identified molecular subgroups with different Ferroptosis-related gene expression patterns (Lei et al., 2021). Jiao et al. proposed a hypergraph regularization constraint-based NMF method (HC-NMF) to select differentially expressed genes and classify tumor samples (Jiao et al., 2020). Leng et al. proposed an adaptive total-variance constraint-based NMF method (ATV-NMF), which can adaptively denoise or maintain feature details based on gradient information (Leng et al., 2017). Zhu et al. applied ATV-NMF to single-cell sequencing data clustering and achieved accurate results in cell subpopulation clustering and the identification of marker genes (Zhu et al., 2021). However, these improved NMF-based methods do not consider the relationship between different types of genetic data and cannot integrate and decompose different types of genetic data simultaneously. To address this issue, Zhang et al. used a joint NMF (jNMF)

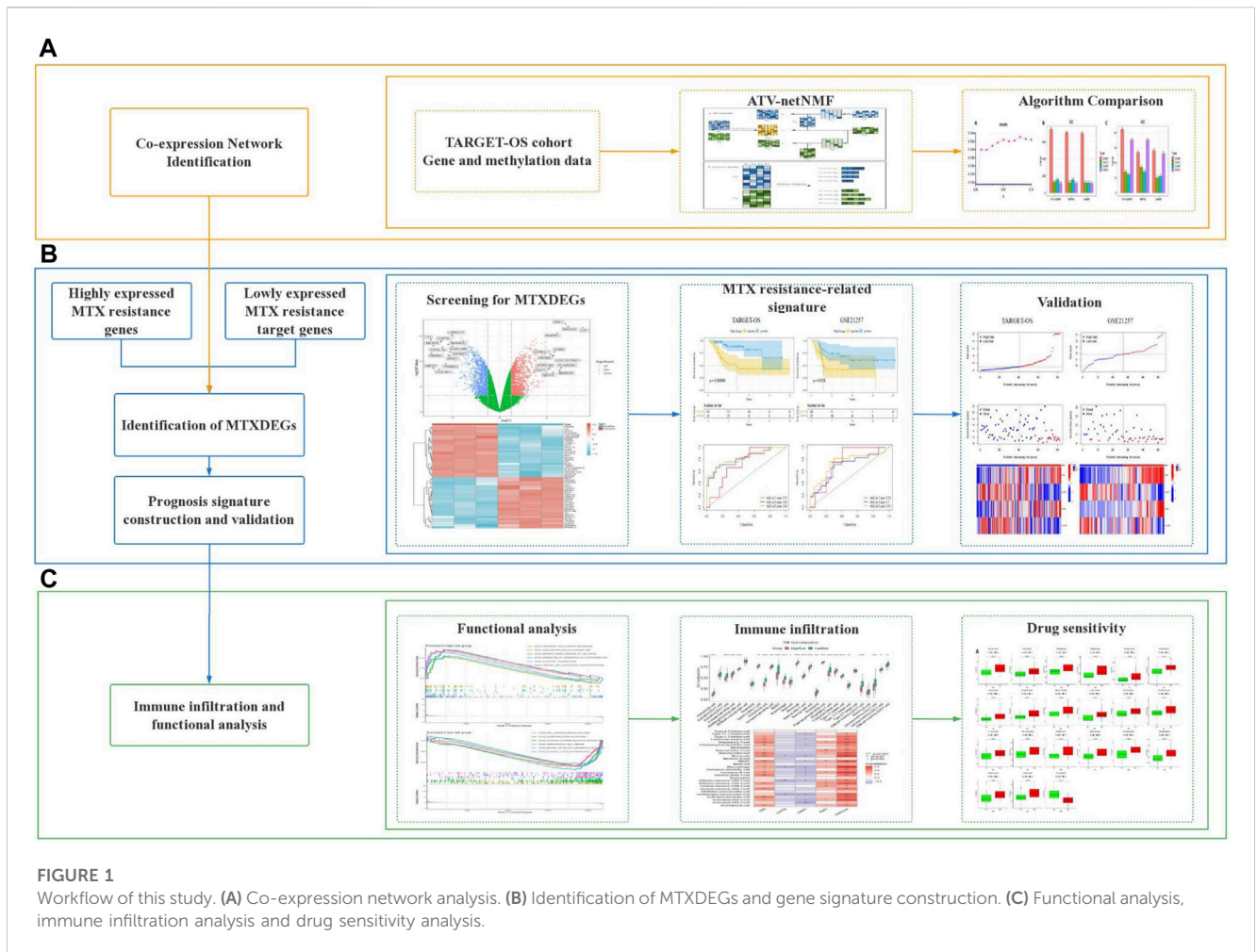
approach to integrate DNA methylation (ME), GE, and miRNA expression data from ovarian cancer to identify ovarian cancer-related multi-dimensional modules (Zhang et al., 2012). Liu et al. proposed a TriNMF-based network-assisted co-clustering method for cancer subtype identification (NCIS) that incorporates molecular interaction networks into the clustering process to improve the identification of cancer subtypes (Liu et al., 2014; Ding et al., 2006). Chen proposed the netNMF method based on NMF using a network framework to identify co-expression modules of two different types of genetic data (Chen and Zhang, 2018). NetNMF uses the decomposed submatrices to construct co-expression networks, which may weaken the connectivity of the nodes in the network. Therefore, Zhuang et al. proposed a hypergraph regularization constraint-based netNMF method (HG-netNMF) (Zhuang et al., 2023), and Ding et al. proposed a graph regularization-based netNMF method (NMFNA), both of which can better mine higher-order features between two genetic data compared to netNMF (Ding et al., 2021). The above NMF-based network analysis method provides an effective way to understand the interactions of different genetic data to understand the pathogenic mechanisms of cancer.

In this study, we proposed an improved NMF network analysis method (ATV-netNMF) to integrate DNA methylation and gene expression data. On this basis, combined with the miRNA and mRNA data of MTX resistance, we built a signature of MTXDEGs that predicted the prognosis of osteosarcoma, and the results revealed that the high-risk group had fewer immune cells and a lower degree of immune infiltration, which could lead to a poor prognosis.

2 Materials and methods

2.1 Workflow of this study

This study is mainly divided into three stages, in the first stage, to efficiently fuse methylation and gene expression data, the proposed ATV-netNMF was applied to two types of genetic data to identify co-expression networks and core gene modules that are strongly associated with variation in both data. Furthermore, the core module was analyzed by KEGG and GO enrichment to compare with other methods based on the number of pathways enriched and pathway significance (Figure 1A). In the second stage, considering the degradation and inhibitory effects of miRNAs, target genes were predicted using upregulated miRNAs intersecting with downregulated genes and core module genes taken as target genes regulated by MTX-resistant miRNAs. Genes highly expressed in MTX-resistant osteosarcoma cells were obtained using the intersection of upregulated genes and core module genes. Finally, the two parts of genes were considered together as MTXDEGs. Then, the extracted MTXDEGs were used to construct the gene signature and calculate the risk scores, which were validated for their predictive performance using an independent dataset (Figure 1B). In the third stage, the risk scores were used to classify the samples into high-risk and low-risk groups for functional analysis, immune infiltration analysis, and drug sensitivity analysis (Figure 1C).



2.2 Data sources

82 osteosarcoma patients with complete clinical characteristics, methylation data, and gene expression data were obtained from the TARGET database (<https://portal.gdc.cancer.gov/>) as the training cohort, and 53 osteosarcoma patients with RNA-seq and clinical characteristics from GSE21257 in the GEO database (<https://www.ncbi.nlm.nih.gov/geo>) as the validation cohort. GSE16089 (Selga et al., 2009) and GSE223857 (Zhang et al., 2023a) included three methotrexate-resistant and three methotrexate-sensitive osteosarcomas mRNA and miRNA data in the GEO database.

2.3 Adaptive total-variation constrained based netNMF

The typical NMF (Lee and Seung, 1999) decomposes the nonnegative matrix $V^{m \times n}$ into two nonnegative matrices $W^{m \times k}$ and $H^{k \times n}$, where $W^{m \times k}$ is the basis matrix and $H^{k \times n}$ is the loading matrix, such that $V \approx WH$, where $k < \min(m, n)$. To minimize the factorization error between V and WH , which can be written as,

$$\min_{W, H} \|V - WH\|_F^2 \quad (1)$$

$$s.t. W \geq 0, H \geq 0$$

On the basis of two and three-factor NMF ($V \approx FSG$) (Ding et al., 2006), if V is a symmetric similarity matrix, it could be decomposed into GSG^T . For biological networks with the same samples but with two different types of features, combining the above ideas, netNMF (Chen and Zhang, 2018) is defined as,

$$\min_{G_1, G_2, S_{11}, S_{22}} \|R_{11} - G_1 S_{11} G_1^T\|_F^2 + \lambda_1 \|R_{12} - G_1 G_2^T\|_F^2 + \lambda_2 \|R_{22} - G_2 S_{22} G_2^T\|_F^2$$

$$s.t. G_1 \geq 0, G_2 \geq 0, S_{11} \geq 0, S_{22} \geq 0 \quad (2)$$

NMFNA (Ding et al., 2021) applies graph regularization constraints in netNMF that can discover and enhance the inherent geometric data structure and improve the ability to identify modules. Based on netNMF and graph regularization constraints, the objective function of NMFNA is defined as,

$$\min_{G_1, G_2, S_{11}, S_{22}} \|R_{11} - G_1 S_{11} G_1^T\|_F^2 + \alpha \|R_{12} - G_1 G_2^T\|_F^2 + \beta \|R_{22} - G_2 S_{22} G_2^T\|_F^2$$

$$+ \sum_{i=1}^2 \lambda_i Tr(G_i^T L_i G_i)$$

$$s.t. G_1 \geq 0, G_2 \geq 0, S_{11} \geq 0, S_{22} \geq 0 \quad (3)$$

where $R_{11}^{n_1 \times n_1}$ and $R_{22}^{n_2 \times n_2}$ are the autocorrelation matrices of X_1 and X_2 which are symmetric similarity matrices corresponding to the two

features, and $R_{12}^{n_1 \times n_2}$ is the intercorrelation matrix between them, which are all non-negative; $G_1^{n_1 \times k}$ and $G_2^{n_2 \times k}$ are non-negative matrices identifying the feature modules in their respective networks. $S_{11}^{k \times k}$ and $S_{22}^{k \times k}$ are symmetric non-negative decomposition matrices; k is a prespecified dimensionality reduction parameter; L is the graph Laplacian matrix; λ is used to adjust the strength of the graph regularization constraint; α and β are used to balance the first three terms of the objective function, which are set to n_1/n_2 and $(n_1/n_2)^2$ by default.

In order to better remove data noise and retain key feature details, we propose the ATV-netNMF method, which can improve the tolerance of the algorithm to noise and improve the performance of the algorithm by introducing adaptive total-variation constraint on NMFNA. Adaptive total variation (Leng et al., 2017; Levine, 2005) can be adapted based on gradient information for denoising or preserving feature details, which can be illustrated as,

$$E(G) = \|G\|_{ATV} = \int_{\Omega} \frac{1}{p(x,y)} |\nabla G|^p(x,y) dx dy \quad (4)$$

$$p(x,y) = 1 + (1 + |\nabla G|^2)^{-1}, 1 < p(x,y) < 2 \quad (5)$$

where E is the energy function of G , $\|G\|_{ATV} = \int_{\Omega} \frac{1}{p(x,y)} |\nabla G|^p(x,y) dx dy$ represents the adaptive total-variation regularization term, and $(\nabla G)(i,j) = ((\partial_x G)(i,j), (\partial_y G)(i,j))$ is the discrete gradient form with $((\partial_x G)(i,j), (\partial_y G)(i,j))$, is given by

$$(\partial_x G)(i,j) = \begin{cases} G(i+1,j) - G(i,j) & \text{if } i < r \\ G(1,j) - G(r,j) & \text{if } i = r \end{cases} \quad (6)$$

$$(\partial_y G)(i,j) = \begin{cases} G(i,j=1) - G(i,j) & \text{if } j < n \\ G(i,1) - G(i,n) & \text{if } j = n \end{cases} \quad (7)$$

Based on the NMFNA and the adaptive total variation constraint, the objective function of the ATV-netNMF is defined as,

$$\begin{aligned} \min_{G_1, G_2, S_{11}, S_{22}} & \|R_{11} - G_1 S_{11} G_1^T\|_F^2 + \alpha \|R_{12} - G_1 G_2^T\|_F^2 + \beta \|R_{22} - G_2 S_{22} G_2^T\|_F^2 \\ & + \sum_{i=1}^2 [\lambda_i Tr(G_i^T L_i G_i) + 2\|G_i\|_{ATV}] \\ \text{s.t. } & G_1 \geq 0, G_2 \geq 0, S_{11} \geq 0, S_{22} \geq 0 \end{aligned} \quad (8)$$

Where R_{11} and R_{22} are ME and GE co-expression networks, R_{12} is ME-GE co-expression network, other symbolic meanings and parameter settings are the same as NMFNA.

This study uses the multiplicative iterative update algorithm to minimize the objective function of ATV-netNMF. Suppose B_1, B_2, B_3 and B_4 are matrices of Lagrange multipliers which constrain $S_{11} \geq 0, S_{22} \geq 0, G_1 \geq 0$ and $G_2 \geq 0$ respectively, and the Lagrangian function f of ATV-netNMF is

$$\begin{aligned} f = & tr((R_{11} - G_1 S_{11} G_1^T)^T (R_{11} - G_1 S_{11} G_1^T)) \\ & + \alpha tr((R_{12} - G_1 G_2^T)^T (R_{12} - G_1 G_2^T)) \\ & + \beta tr((R_{22} - G_2 S_{22} G_2^T)^T (R_{22} - G_2 S_{22} G_2^T)) + \lambda_1 Tr(G_1^T L_1 G_1) \\ & + \lambda_2 Tr(G_2^T L_2 G_2) + 2\|G_1\|_{ATV} + 2\|G_2\|_{ATV} + tr(B_1^T S_{11}) \\ & + tr(B_2^T S_{22}) + tr(B_3^T G_1) + tr(B_4^T G_2) \end{aligned} \quad (9)$$

Thus, the partial derivatives of f with respect to S_{11}, S_{22}, G_1 , and G_2 are,

$$\frac{\partial f}{\partial S_{11}} = -2G_1^T R_{11} G_1 + 2G_1^T G_1 S_{11} G_1^T G_1 + B_1 \quad (10)$$

$$\frac{\partial f}{\partial S_{22}} = -2G_2^T R_{22} G_2 + 2G_2^T G_2 S_{22} G_2^T G_2 + B_2 \quad (11)$$

$$\begin{aligned} \frac{\partial f}{\partial G_1} = & 4(G_1 S_{11} G_1^T G_1 S_{11} - R_{11} G_1 S_{11}) + 2\alpha(G_1 G_2^T G_2 - R_{12} G_2) \\ & + 2\lambda_1 L_1 G_1 - 2div\left(\frac{\nabla G_1}{|\nabla G_1|^{2-p}}\right) + B_3 \end{aligned} \quad (12)$$

$$\begin{aligned} \frac{\partial f}{\partial G_2} = & 4\beta(G_2 S_{22} G_2^T G_2 S_{22} - R_{22} G_2 S_{22}) + 2\alpha(G_2 G_1^T G_1 - R_{12} G_1) \\ & + 2\lambda_2 L_2 G_2 - 2div\left(\frac{\nabla G_2}{|\nabla G_2|^{2-p}}\right) + B_4 \end{aligned} \quad (13)$$

Let $B_1 S_{11} = 0, B_2 S_{22} = 0, B_3 G_1 = 0, B_4 G_2 = 0$, the iterative formula can be written as,

$$(S_{11})_{kk} \leftarrow (S_{11})_{kk} \frac{(G_1^T R_{11} G_1)_{kk}}{(G_1^T G_1 S_{11} G_1^T G_1)_{kk}} \quad (14)$$

$$(S_{22})_{kk} \leftarrow (S_{22})_{kk} \frac{(G_2^T R_{22} G_2)_{kk}}{(G_2^T G_2 S_{22} G_2^T G_2)_{kk}} \quad (15)$$

$$(g_1)_{ik} \leftarrow (g_1)_{ik} \frac{(\alpha R_{12} G_2 + 2R_{11} G_1 S_{11} + div\left(\frac{\nabla G_1}{|\nabla G_1|^{2-p}}\right))_{ik}}{(2G_1 S_{11} G_1^T G_1 S_{11} + \alpha G_1 G_2^T G_2 + 2\lambda_1 L_1 G_1)_{ik}} \quad (16)$$

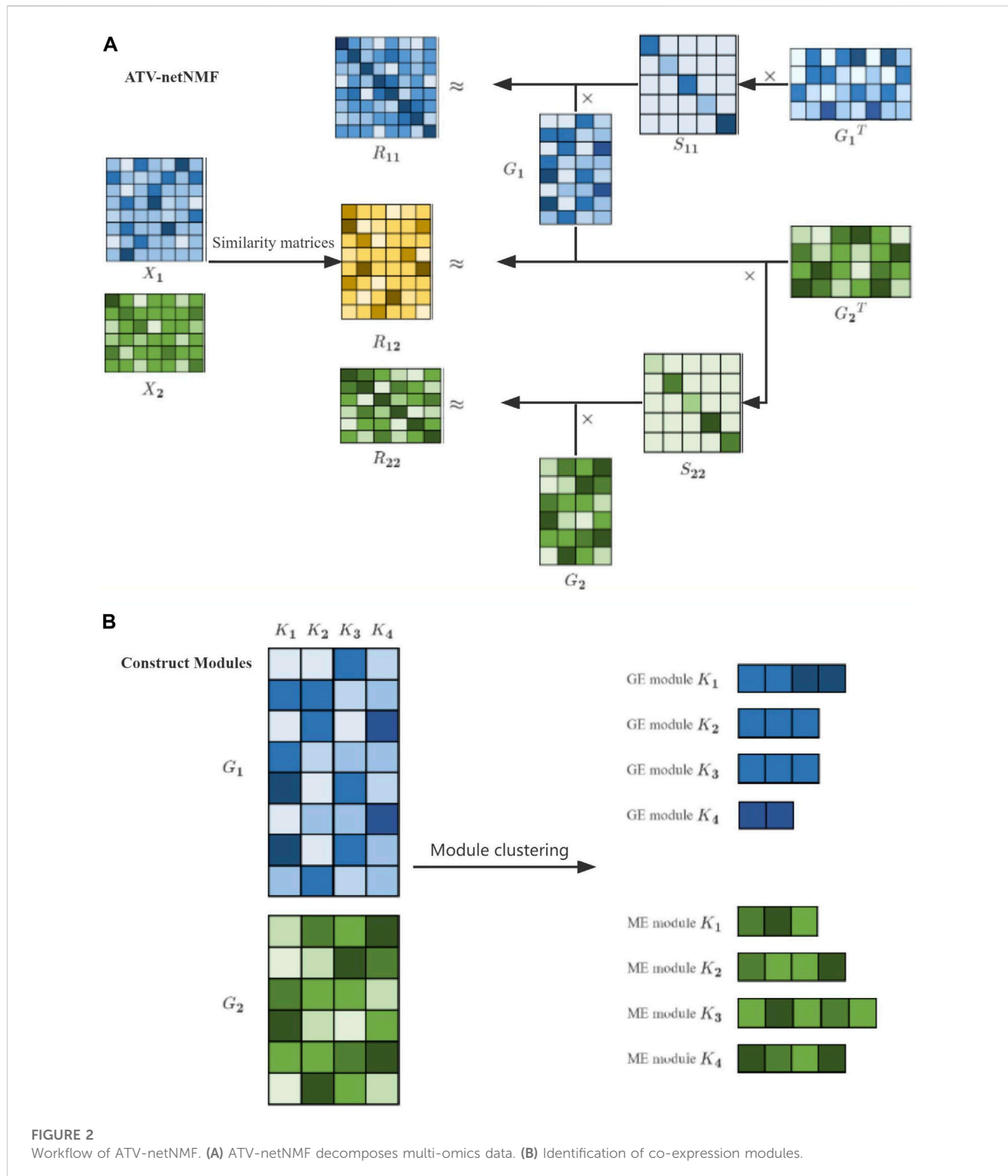
$$(g_2)_{kj} \leftarrow (g_2)_{kj} \frac{(\alpha R_{12} G_1 + 2\beta R_{22} G_2 S_{22} + div\left(\frac{\nabla G_2}{|\nabla G_2|^{2-p}}\right))_{kj}}{(2\beta G_2 S_{22} G_2^T G_2 S_{22} + \alpha G_2 G_1^T G_1 + 2\lambda_2 L_2 G_2)_{kj}} \quad (17)$$

Where div denotes the divergence of the matrix, ∇G denotes the gradient of the matrix, and $|\nabla G|$ denotes the norm of the gradient of the matrix. The adaptive total variation regularization includes a diffusion coefficient $\frac{1}{|\nabla G|^{2-p}}$ in Eqs 18, 19, which controls the data diffusion rate based on gradient information. For data edges, larger values of $|\nabla G|^{2-p}$ and smaller values of $\frac{1}{|\nabla G|^{2-p}}$ help to maintain the edges. In data smoothing regions, smaller values of $|\nabla G|^{2-p}$ and larger values of $\frac{1}{|\nabla G|^{2-p}}$ help to remove noise (Leng et al., 2017). ATV can preserve or enhance data features while removing noise. The overall workflow of ATV-netNMF is shown in (Figure 2A). Firstly, three co-expression networks are constructed using ME and GE data. Then the network is decomposed using the objective function to identify the co-expression modules under the guidance of G_1 and G_2 (Figure 2B), z-scores of each column vector of g_i are calculated as follows.

$$g^* = \frac{g - \bar{g}}{\frac{1}{n-1} \sum_i (g_i - \bar{g})^2} \quad (18)$$

$$\bar{g} = \frac{1}{n} \sum_i g_i \quad (19)$$

The genes with each column weight greater than or equal to the threshold are used as module members, the threshold is set to 2, and the module with the most genes called the core module according to the previous study (Ding et al., 2021).



2.4 Analysis of differentially expressed genes (DEGs) and miRNAs (DE-miRNAs)

The R package “limma” (Chen et al., 2022) was used to find differentially expressed genes and miRNAs between methotrexate-resistant and sensitive cells in the GSE16089 and GSE223857 datasets, with the threshold set to $|\text{Log}_2\text{FC}| > 1$ and $p < 0.05$.

2.5 Prediction of miRNA targeted genes

The target genes of upregulated miRNAs were predicted by TargetScan (<http://www.targetscan.org/>) and miRDB (<http://mirdb.org/miRDB/>) online databases, and the screening threshold was set as the cut-off parameter for TargetScan was the total context++ score < -0.05 , while miRdb score > 50 .

2.6 Construction and validation of a MTX resistance-related signature in osteosarcoma

Statistically, significant ($p < 0.05$) MTXDEGs associated with OS prognosis were obtained by univariate COX regression analysis using the survival package in R. The least absolute shrinkage and selection operator (LASSO) regression analysis was performed on prognosis-related MTXDEGs using the glmnet package in R to reduce the dimensionality of genes in the model. Subsequently, independent prognostic genes were screened using multivariate Cox regression analysis, and regression coefficients for the corresponding genes were generated. A linear combination of gene expression levels and regression coefficients created a signature with the following formula for the risk score.

$$\text{Riskscore} = \sum_{i=1}^n \text{Exp}(i) * \text{coef}(i) \quad (20)$$

The median of Riskscore was used to determine the best critical value to classify patients into high-risk and low-risk groups. Kaplan-Meier survival curves and time-dependent receptor operating characteristic (ROC) curves were used to assess the predictive performance of prognostic signatures on overall survival. GSE21257 was used as a validation set to verify the predictive performance of the signature.

2.7 Construction of the nomogram based on prognostic models

Compared with other clinical characteristics (including metastasis, race, age, and gender), univariate and multivariate COX analyses were performed to determine the independence of our established gene signature in predicting overall survival, and $p < 0.05$ was considered statistically significant. To predict the prognosis of patients with osteosarcoma, a nomogram integrating riskscore and clinical characteristics was constructed, and calibration curves were used to evaluate the predictive accuracy of the nomogram, which was constructed from the “rms” R package (Liu et al., 2023a). To estimate the clinical robustness of the MTX resistance gene signature, decision curve analysis (DCA) was used to calculate the net benefit of the signature for different threshold probabilities in the training and validation datasets.

2.8 Functional analysis

The R package limma was used to obtain differential genes between high-risk and low-risk groups, which were analyzed for GO and KEGG pathway enrichment using David (the Database for Annotation, Visualization and Integrated Discovery, <https://david.ncifcrf.gov>). The threshold for significantly enriched pathways was set to p -value < 0.05 , and the top 20 most significant pathways were selected.

2.9 Gene set enrichment analysis

Gene set enrichment analysis (GSEA) was performed to identify pathways enriched in the high-risk and low-risk groups to explore

the relationship between riskscore and biological function, with the threshold of $p < 0.05$.

2.10 Evaluation of immune cell infiltration

A single sample gene set enrichment analysis (ssGSEA) (Liu et al., 2023b) method was used to analyze the differences in 28 immune cell infiltrates between the high-risk and low-risk groups. Tumor microenvironment analysis was performed on the gene expression data of osteosarcoma using an R package estimate (Zhang et al., 2023b) to obtain the immune score, stromal score, and estimate score for each patient, and the difference in scores between the high-risk and low-risk groups was analyzed. Correlation between immune cells and immune scores was performed using the ggstatsplot R package.

2.11 Drug sensitivity analysis

The OncoPredict (Maeser et al., 2021) R package was used to predict *in vivo* drug response in cancer patients, including half-maximal inhibitory concentration (IC50) values for 189 drugs corresponding to cell lines and a normalized gene expression matrix for 809 tumor cell lines from the Genomics of Drug Sensitivity in Cancer (GDSC) database. IC50 values for the TARGET-OS cohort were predicted using the oncoPredict method with a significance threshold set at $p < 0.001$.

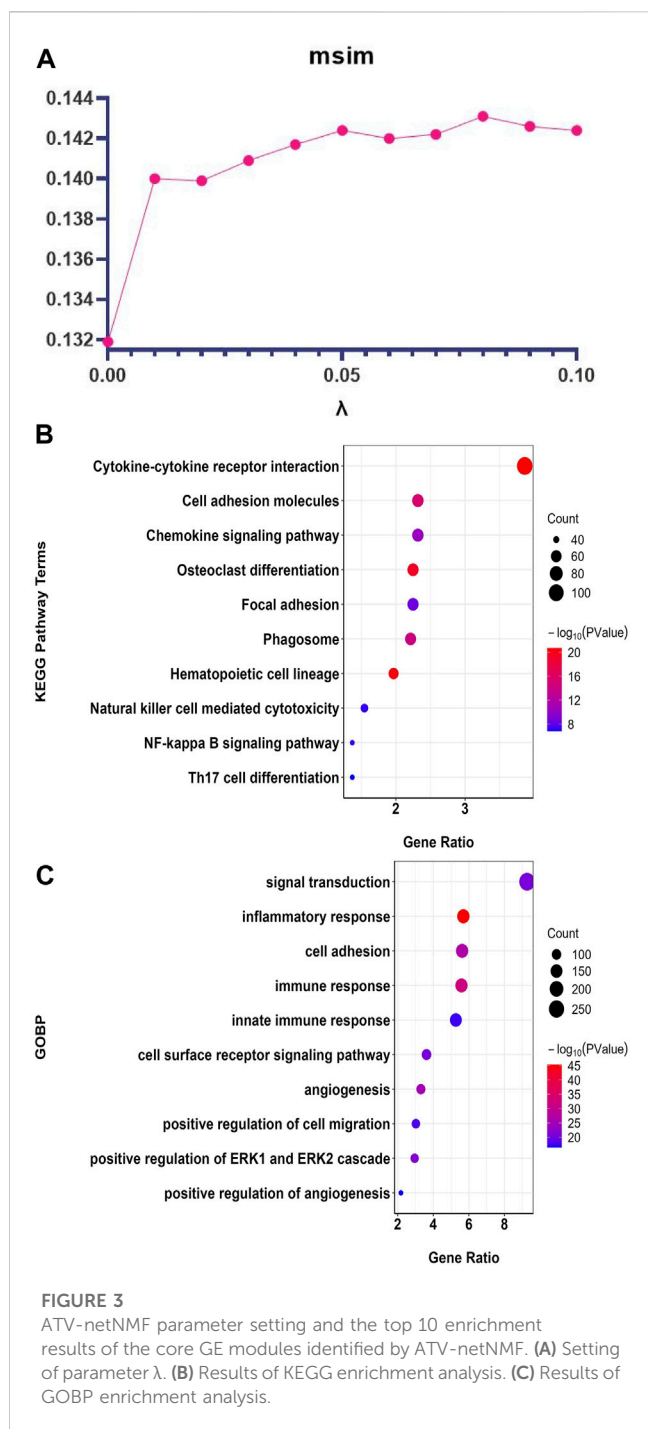
3 Results

3.1 Screening of core modules by ATV-netNMF

The methylation and gene data of the TARGET-OS cohort had the same 82 osteosarcoma samples but with different features. Methylated genes with expression mean values less than 0.25 were filtered out from ME data, 15,819 methylated genes were retained, and 23,683 mRNAs were filtered out in GE data. Three co-expression networks were created from these data: ME network $R_{11}^{15819 \times 15819}$, ME-GE network $R_{12}^{15819 \times 23683}$, and GE network $R_{22}^{23683 \times 23683}$. Based on the previous study (Zhuang et al., 2023), the value of dimensionality reduction k generally is at most one-tenth of the minimum number of samples or features of the network modules. Therefore, in this study, k is set to 8, and the number of iterations is set to 200. By running the λ from 0 to 0.1, the highest module similarity is selected and set to λ_1 and λ_2 , and the module similarity (Wang et al., 2018) is defined as follows.

$$\text{msim} = \sum_{x,y} \frac{|M_x \cap M_y|}{\min(|M_x, M_y|)} \quad (21)$$

M_x denotes the members belonging to module x . According to Figure 3A, λ_1 , λ_2 was set to 0.08. Finally, the core GE module containing 2810 mRNAs closely related to methylation data and the ME core module containing 1013 methylation genes were obtained



by ATV-netNMF integrated analysis. The core gene module was selected for further analysis.

The top 10 enrichment terms of the core GE modules identified by ATV-netNMF are shown in Figure 3. The results showed that the genes were most enriched in terms related to tumor cell function and immune function. For example, KEGG was mainly enriched to cell adhesion molecules, osteoclast differentiation, NF-kappa B signaling pathway, Th17 cell differentiation (Figure 3B), and GOBP was mainly enriched to inflammatory response, immune response, cell adhesion, signal transduction (Figure 3C). The above results indicated that ATV-netNMF could effectively screen the gene

network significantly related to the immune microenvironment of osteosarcoma.

3.2 Identification of MTXDEGs

The flowchart for the identification of MTXDEGs is shown in Figure 4A. We analyzed DEGs between methotrexate-resistant and sensitive osteosarcoma cells using GSE16089 and screened 2397 DEGs, of which 1191 were upregulated and 1206 were downregulated (Figures 4B, C). Then DE-miRNAs between methotrexate-resistant and sensitive osteosarcoma cells were analyzed using GSE223857, and 16 DE-miRNAs were screened, of which 9 were upregulated and 7 were downregulated (Figures 4D, E). The upregulated MTX-resistant genes intersected with the core gene module, finally obtaining 172 highly expressed genes in MTX-resistant osteosarcoma cells (Figure 4A). Considering the degradation and translational repression of target genes by miRNAs, the upregulated miRNAs were used to predict the regulated target genes and intersected with the downregulated MTX-resistant genes and the core gene module, and 49 target genes regulated by MTX-resistant miRNAs were finally identified (Figure 4A). They were combined to obtain 221 MTXDEGs for further analysis.

3.3 Construction and validation of a methotrexate resistance-related signature

We selected 82 samples with complete survival information from the TARGET-OS cohort for further analysis. To identify MTXDEGs significantly associated with prognosis, univariate Cox regression analysis was performed on 221 MTXDEGs, and 30 MTXDEGs were significantly associated with overall survival ($p < 0.05$) (Figure 5A). LASSO regression was then performed to screen genes for model building, and 10 genes were screened based on the best λ (Figures 5B, C). Based on the genes generated by LASSO regression, multivariate Cox regression identified 4 genes, *CPE*, *LASPI*, *PDK1*, and *LACTB*, as hub genes; Based on their expression levels, we developed a Riskscore signature of:

$$\begin{aligned}
 \text{Riskscore} = & 0.482850197357161 * CPE \\
 & - 0.655999299291687 * LASPI \\
 & + 0.44675499776076 * PDK1 \\
 & - 0.591342381951838 * LACTB
 \end{aligned}
 \tag{22}$$

The median of Riskscore was used as the threshold for the high-risk and low-risk groups. The Kaplan-Meier curve showed that the overall survival of the high-risk group was significantly lower than that of the low-risk group ($p = 0.00055$) (Figure 6A). The AUC of the 1-, 3-, and 5-year ROC curves were 0.73, 0.81, and 0.83, respectively (Figure 6B). Figure 7 shows the riskscore, survival status, and expression levels of the four candidate genes in the training cohort. The high-risk group had significantly higher risk scores and worse survival status than the low-risk group (Figures 7A–C). The above results show that the signatures can reasonably predict the overall survival of osteosarcoma

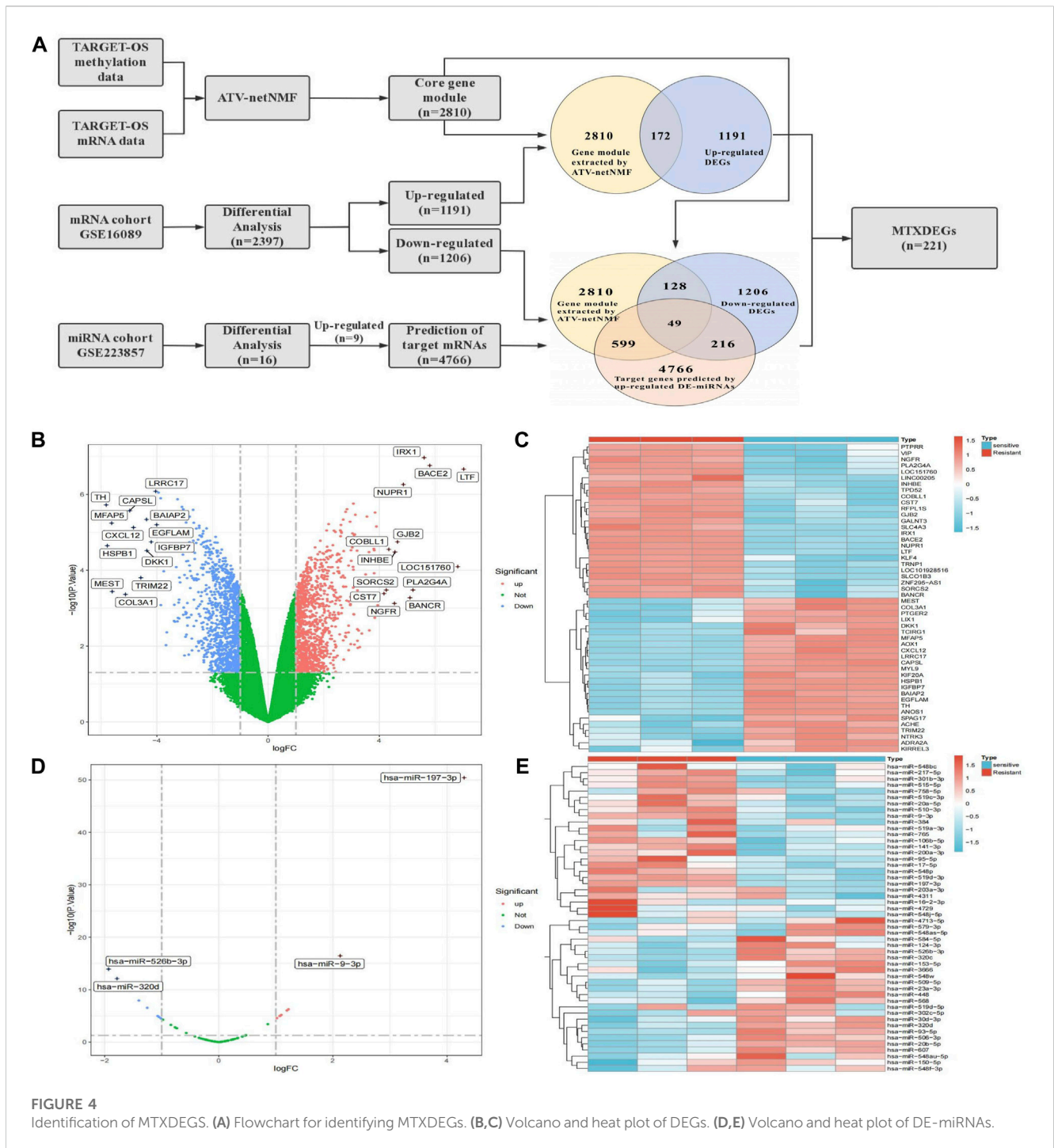
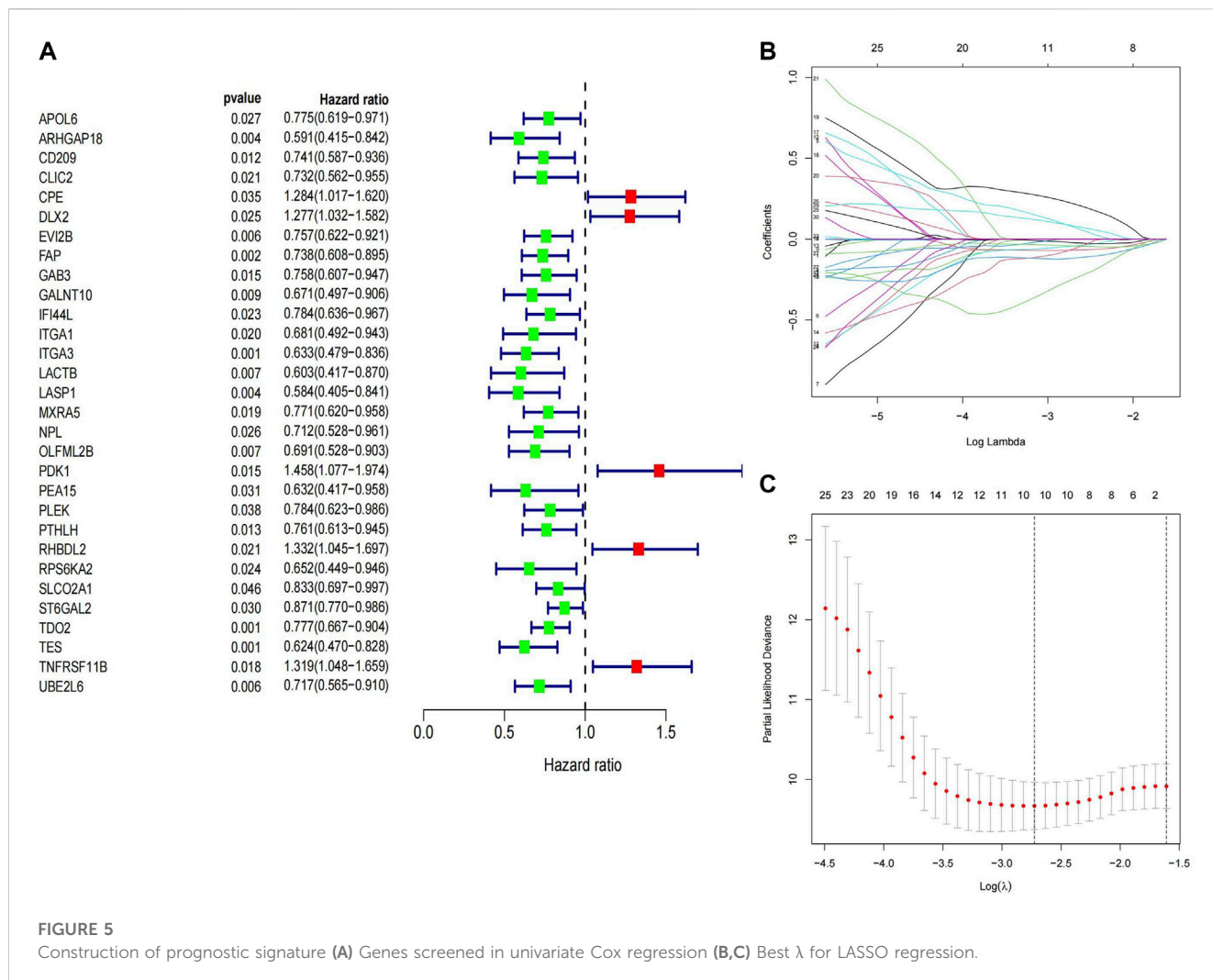


FIGURE 4 Identification of MTXDEGs. (A) Flowchart for identifying MTXDEGs. (B,C) Volcano and heat plot of DEGs. (D,E) Volcano and heat plot of DE-miRNAs.

patients in the training cohort. The validation cohort GSE21257 also classified patients into high-risk and low-risk categories using the median Riskscore. The Kaplan-Meier curves showed that the high-risk group had a shorter survival time compared to the low-risk group ($p = 0.015$) (Figure 6C), with AUC of 0.76, 0.7, and 0.75 for the 1-, 3-, and 5-year ROC curves (Figure 6D), respectively, which was consistent with the results of the training cohort. The riskscore and the expression of the 4 hub genes were also consistent with the training cohort (Figures 7D-F), indicating the prognostic value and reliability of the signature.

3.4 Construction of the nomogram based on MTXDEGs signature in TARGET-OS cohort

We used univariate and multivariate Cox regression analyses to verify whether the Riskscore generated by the 4 genes was an independent prognostic factor. The results showed that metastasis ($p = 0.003$) and Riskscore ($p < 0.001$) were independent prognostic factors for osteosarcoma (Figures 8A, B). We developed a nomogram using Riskscore and clinical data based on these significant factors (Figure 8C). The accuracy of the nomogram was evaluated using calibration curves at 1, 3, and



5 years (Figure 8D). The results showed that the calibration curves were very close to the ideal curve (a straight gray line with a slope of 1 through the origin of the coordinate axis). DCA curve was used to evaluate whether the model contributes to clinical treatment strategies. When the risk threshold probability varied between 0 and 1, the MTX-resistant gene signature achieved a higher net benefit in both the training and validation cohorts than the “treat all” and “treat none” strategies (Figures 8E, F). These results suggest that the MTX resistance gene signature performs well in clinical applications.

3.5 Functional analysis

A total of 1033 DEGs were identified between the high-risk and low-risk groups, which were analyzed for GO and KEGG enrichment. Figure 9 shows the top 20 most significantly enriched KEGG pathways and GO biological processes. KEGG analysis showed that DEGs were mainly enriched in Cytokine-cytokine receptor interaction, Cell adhesion molecules (CAMs), Phagosome, and other pathways related to tumor immune cell function and apoptosis (Figure 9A). GOBP analysis showed that DEGs were mainly enriched in signal transduction, immune

response, inflammatory response, and other pathways related to immune function (Figure 9B). We also performed GSEA analysis to identify the underlying biological processes in the high-risk and low-risk groups. The results showed that ascorbate_and_aldarate_metabolism, drug_metabolism_cytochrome_p450, and other pathways related to metabolic function were enriched in the high-risk group (Figure 9C), and cell_adhesion_molecules_cams, chemokine_signaling_pathway and other pathways related to immune function were enriched in the low-risk group (Figure 9D). The above results indicated that Riskscore is significantly associated with osteosarcoma immune status.

3.6 Immune cell infiltration

To study the differences in immune infiltration between the two groups of patients, we calculated the immune infiltration scores in the high- and low-risk groups using the ESTIMATE method. We noticed that the ImmuneScore, StromalScore, and ESTIMATEScore were significantly lower in the high-risk group than in the low-risk group. Riskscore and three immune scores were significantly negatively correlated (Figure 10A). We performed ssGSEA analysis and calculated the infiltration abundance of 28 immune

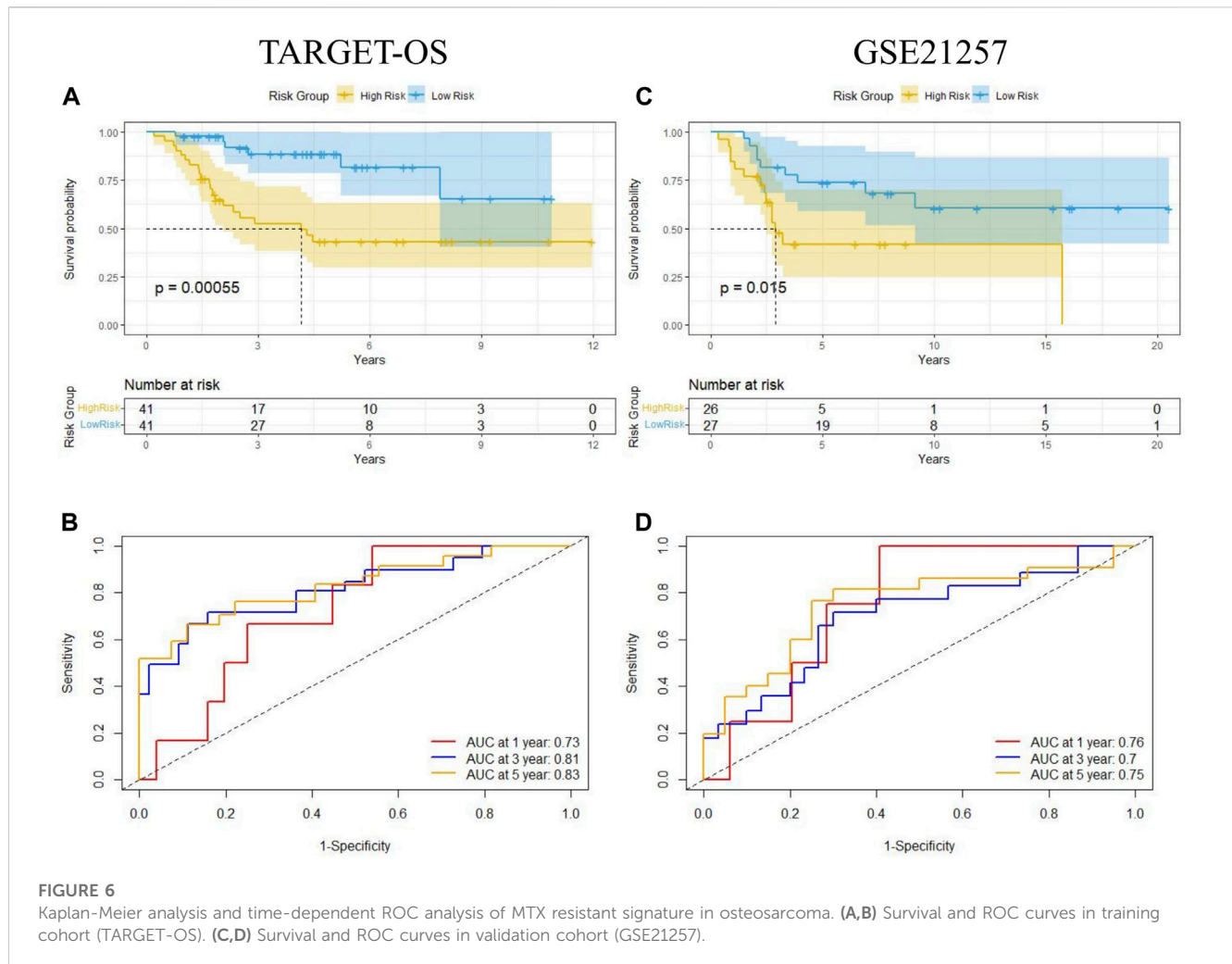


FIGURE 6 Kaplan-Meier analysis and time-dependent ROC analysis of MTX resistant signature in osteosarcoma. (A,B) Survival and ROC curves in training cohort (TARGET-OS). (C,D) Survival and ROC curves in validation cohort (GSE21257).

cells. Surprisingly, the infiltration abundance of all 28 immune cells was lower in the high-risk group, with 23 immune cells being the most significant (Figure 10B). Among them, the hub genes *CPE* and *PDK1* were significantly positively correlated with 18 and 12 immune cells, and *LACTB* and *LASPI* were significantly negatively correlated with 4 and 13 immune cells. Riskscore was significantly positively correlated with 24 immune cells (Figure 10C). The heat map (Figure 10D) and the above results indicated that in the high-risk group, the infiltration of immune cells was significantly lower, and there were fewer immune cells in the tumor immune microenvironment, which may lead to a poor prognosis.

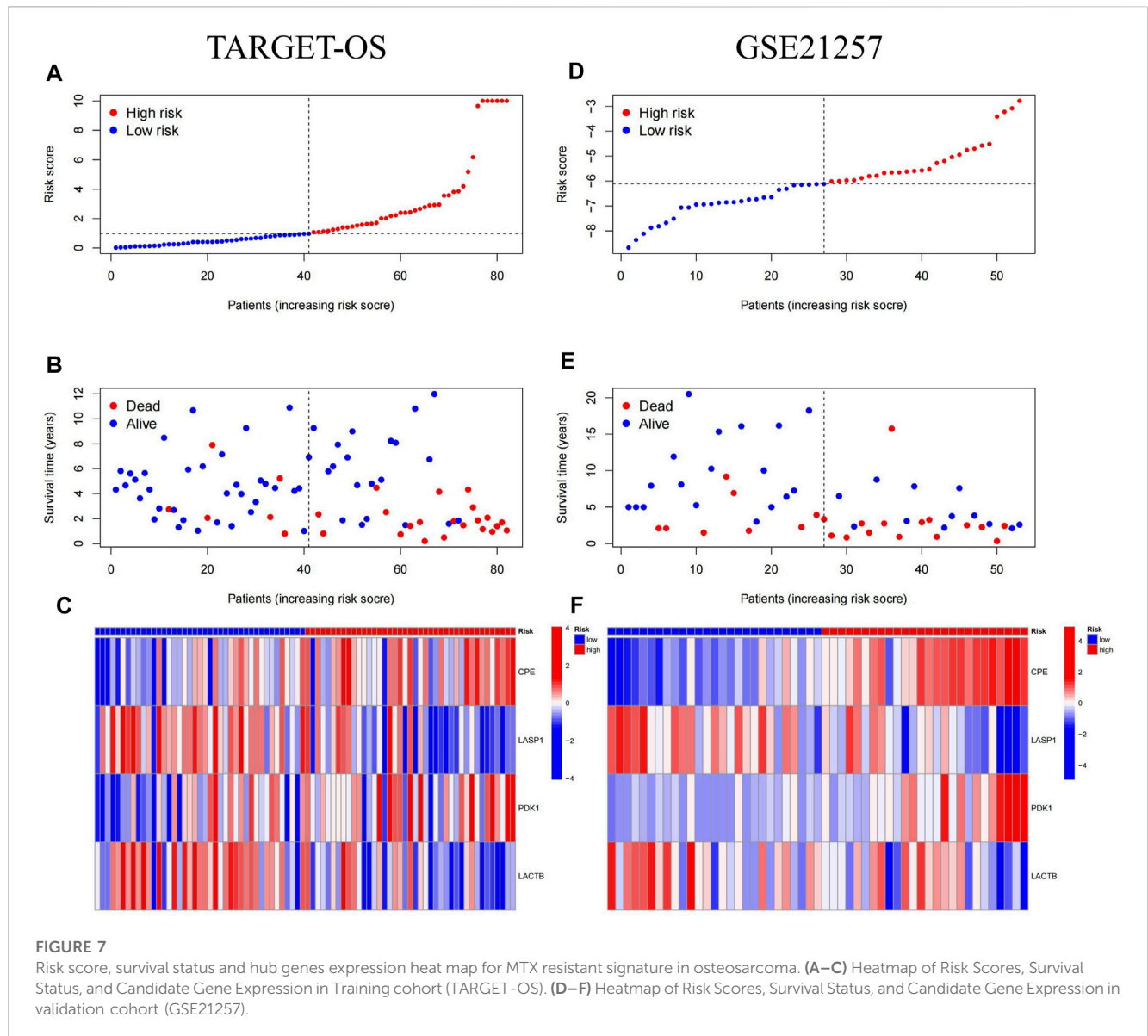
3.7 Drug sensitivity analysis

We further analyzed the response to chemotherapy and targeted therapy in the high-risk and low-risk groups. With a threshold $p < 0.001$, the results showed that the high-risk group showed higher resistance to 20 drugs, and the high-risk group was more sensitive to only one targeted drug, BI-2536 (Figure 11) [a small molecule inhibitor against PLK1 with a dual role in inducing apoptosis and impairing autophagy in neuroblastoma cells (Li et al., 2016)].

4 Discussion

4.1 Comparison of different multi-omics NMF methods

ATV-netNMF successfully constructed co-expression networks between methylation and gene expression data, identified characteristic modules and characterized genes, and identified osteosarcoma biomarkers. To verify the performance of ATV-netNMF, we compared it with NMFNA and netNMF. We performed GO and KEGG enrichment analyses using the core modules identified by each of them, and the number of pathways and p -values of the pathways obtained are shown in Figures 12A, B and Table 1. It can be seen that the ME and GE core modules identified by ATV-netNMF enriched more pathways, and the p -values of the significant pathways were lower compared to the other methods, which suggests that the modules identified by ATV-netNMF may contain more biological information related to osteosarcoma, and enriched to more significant pathways. This is because the basis vectors obtained from ATV-netNMF decomposition are more sparse than netNMF and NMFNA, eliminating some noise in the data and enhancing some features and details.



To further evaluate the applicability of ATV-netNMF and the contribution between individual modules, we downloaded gene expression data and miRNA data from cancer samples in the High-Risk Wilms Tumor (TARGET-WT), Breast Invasive Carcinoma (TCGA-BRCA), and Lung Adenocarcinoma (TCGA-LUAD) datasets from the UCSC Xena database (<https://xenabrowser.net/datapages/>) for analysis. Specifically, we calculated the Pearson correlation coefficients between the reconstructed matrices R_{11} , R_{12} and R_{22} and the original matrices R_{11} , R_{12} and R_{22} in netNMF, NMFNA, netNMF + ATV, and ATV-netNMF, respectively. As can be seen from Table 2, the reconstruction of both R11 and R22 outperforms the netNMF method for all four datasets under the separate constraints of graph regularization and adaptive total variation. The proposed ATV-netNMF method also outperforms the previous methods in reconstructing all four datasets. The above results confirm that graph regularization and adaptive full mutation help improve the reconstruction performance of the methods, and the proposed

ATV-netNMF can be used to detect co-expression modules from multiple diseases and genetic data.

4.2 Biological functional analysis

Patients with localized osteosarcoma can be cured with neoadjuvant chemotherapy and surgical resection in up to 70% of cases, but survival rates for chemotherapy-resistant and metastatic patients are less than 20%. There is a significant correlation between response to chemotherapy and the prognosis of osteosarcoma, and one of the main challenges is inherent or acquired resistance. Methotrexate is used as a common strategy for chemotherapy in osteosarcoma, and patients with MTX resistance often experience tumor recurrence and metastasis. Therefore, the discovery of reliable biomarkers and the search for new therapeutic targets are essential to improve the clinical prognosis of osteosarcoma. To provide new prognostic predictors and

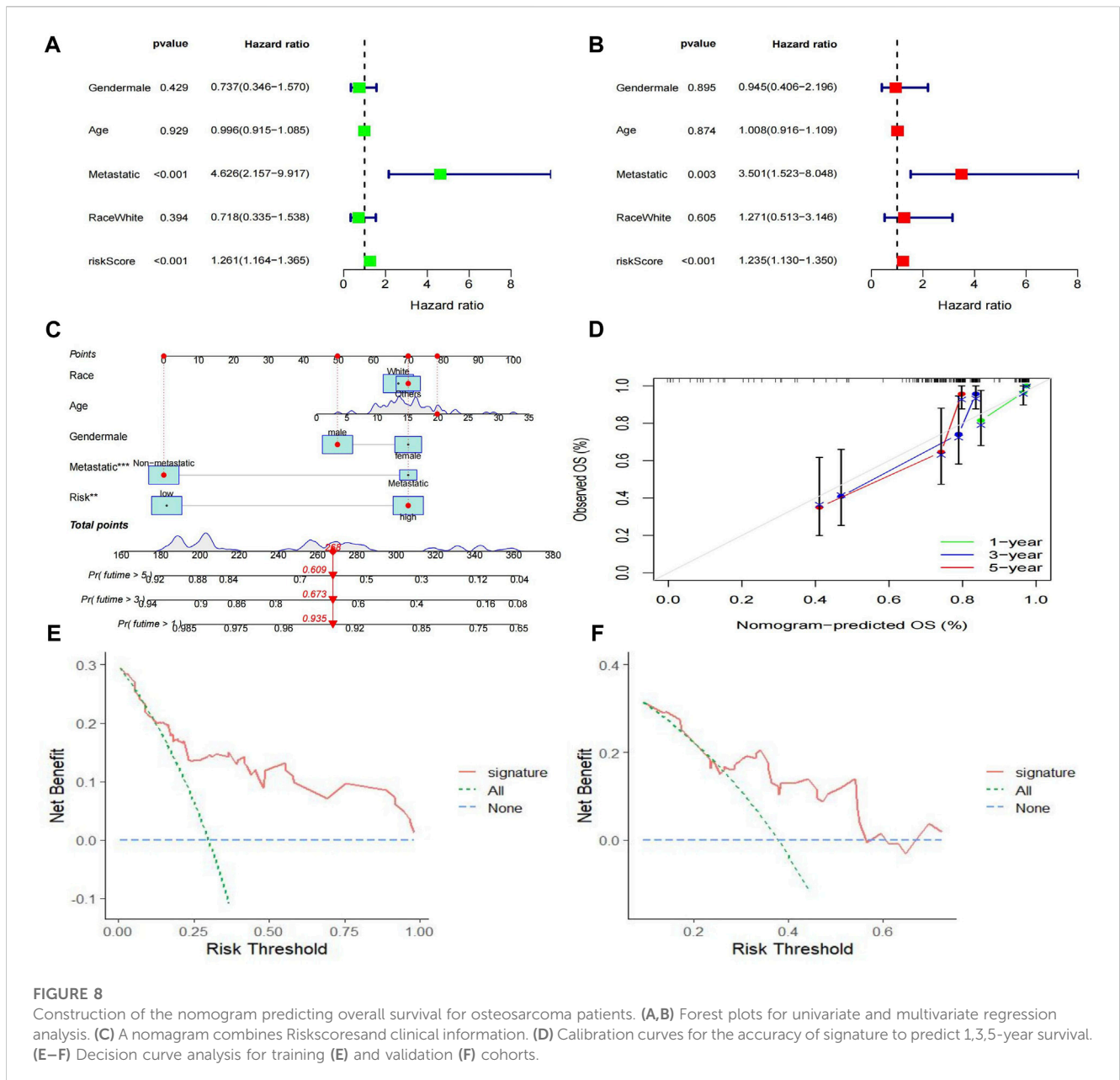


FIGURE 8

Construction of the nomogram predicting overall survival for osteosarcoma patients. (A,B) Forest plots for univariate and multivariate regression analysis. (C) A nomogram combines RiskScores and clinical information. (D) Calibration curves for the accuracy of signature to predict 1,3,5-year survival. (E–F) Decision curve analysis for training (E) and validation (F) cohorts.

immunotherapies for chemotherapy-resistant patients with osteosarcoma, our study identified meaningful predictive biomarkers by comprehensively analyzing multi-omics genetic data of osteosarcoma.

In this study, first, ME, GE, and ME-GE networks were constructed based on two genetic data. Then, these networks were decomposed under the constraints of total adaptive variance and graph regularization, while co-expression modules were efficiently identified, which is the highlight of ATV-netNMF. Finally, the core networks were analyzed for GO and KEGG enrichment. Compared with NMFNA and netNMF, ATV-netNMF could identify more osteosarcoma-related GO terms and pathways, indicating that ATV-netNMF could effectively detect modules and characterize genes. Based on that, combined with MTX-resistant mRNA and miRNA data, we established a new 4-gene prognostic signature for osteosarcoma, including two high-risk MTXDEGs (CPE, PDK1) and

two low-risk MTXDEGs (LASPI, LACTB). We can categorize patients into high-risk and low-risk subgroups based on the risk scores derived from this predictive model. CPE is highly expressed in MTX-resistant cells. PDK1, LASPI, and LACTB were lowly expressed in MTX-resistant cells and were target genes regulated by MTX-resistant miRNAs. CPE is a prohormone processing enzyme that is usually overexpressed in osteosarcoma cell lines (Shi et al., 2020), and the downregulation of CPE inhibits the migration and invasive ability of osteosarcoma cells. Overexpression of a splice variant of CPE, CPE-ΔN, promotes the growth and metastasis of osteosarcoma cells (Li et al., 2016). PDK1 is a key rate-limiting enzyme of the tricarboxylic acid cycle. PDK1 expression was suppressed in DXR-resistant osteosarcoma cells (Zhang et al., 2021), consistent with our experimental findings that PDK1 is downregulated in MTX-resistant osteosarcoma. PDK1 was overexpressed in osteosarcoma, multiple myeloma, acute myelogenous leukemia, and breast cancer (Zhang

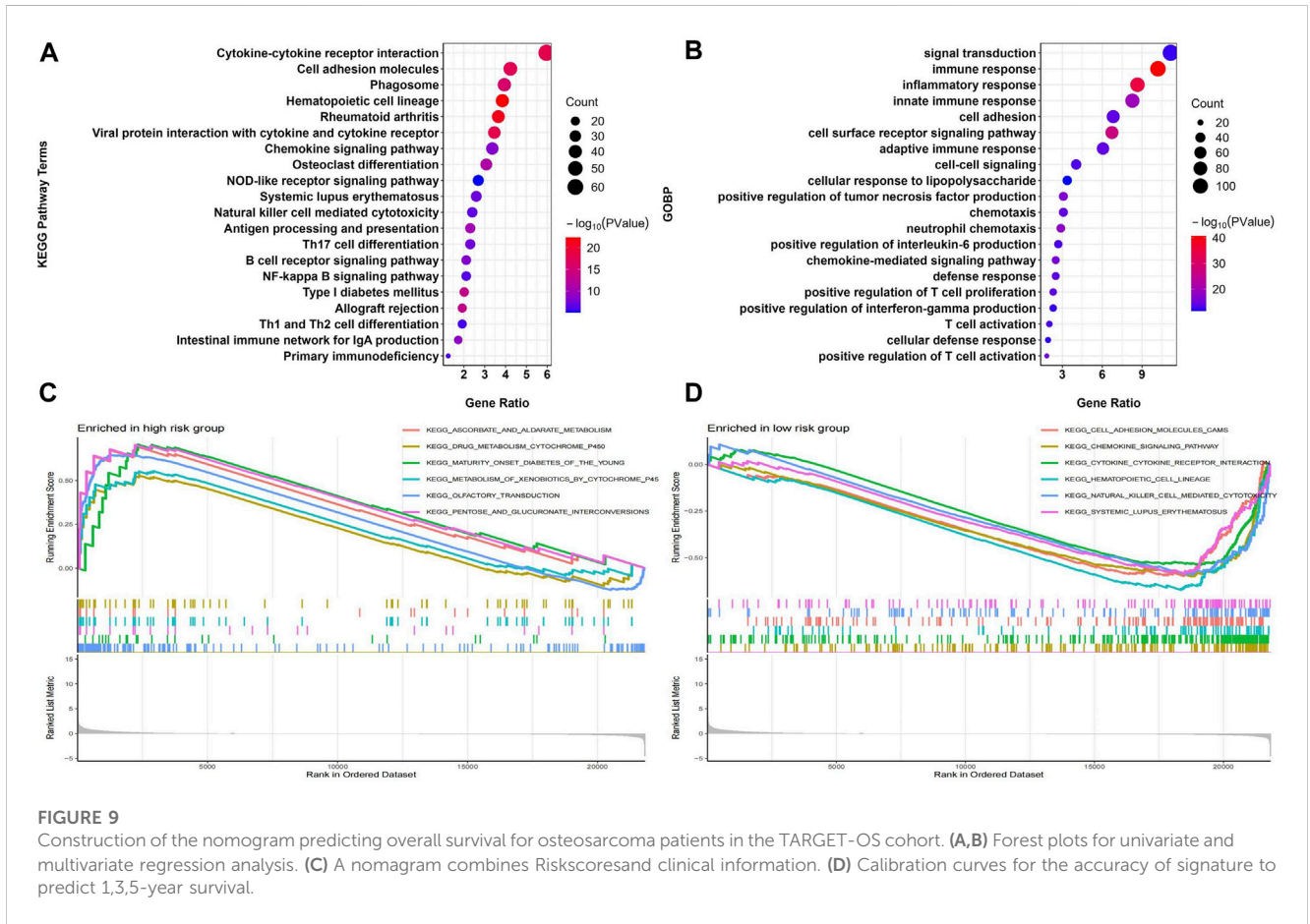


FIGURE 9 Construction of the nomogram predicting overall survival for osteosarcoma patients in the TARGET-OS cohort. (A,B) Forest plots for univariate and multivariate regression analysis. (C) A nomogram combines RiskScore and clinical information. (D) Calibration curves for the accuracy of signature to predict 1,3,5-year survival.

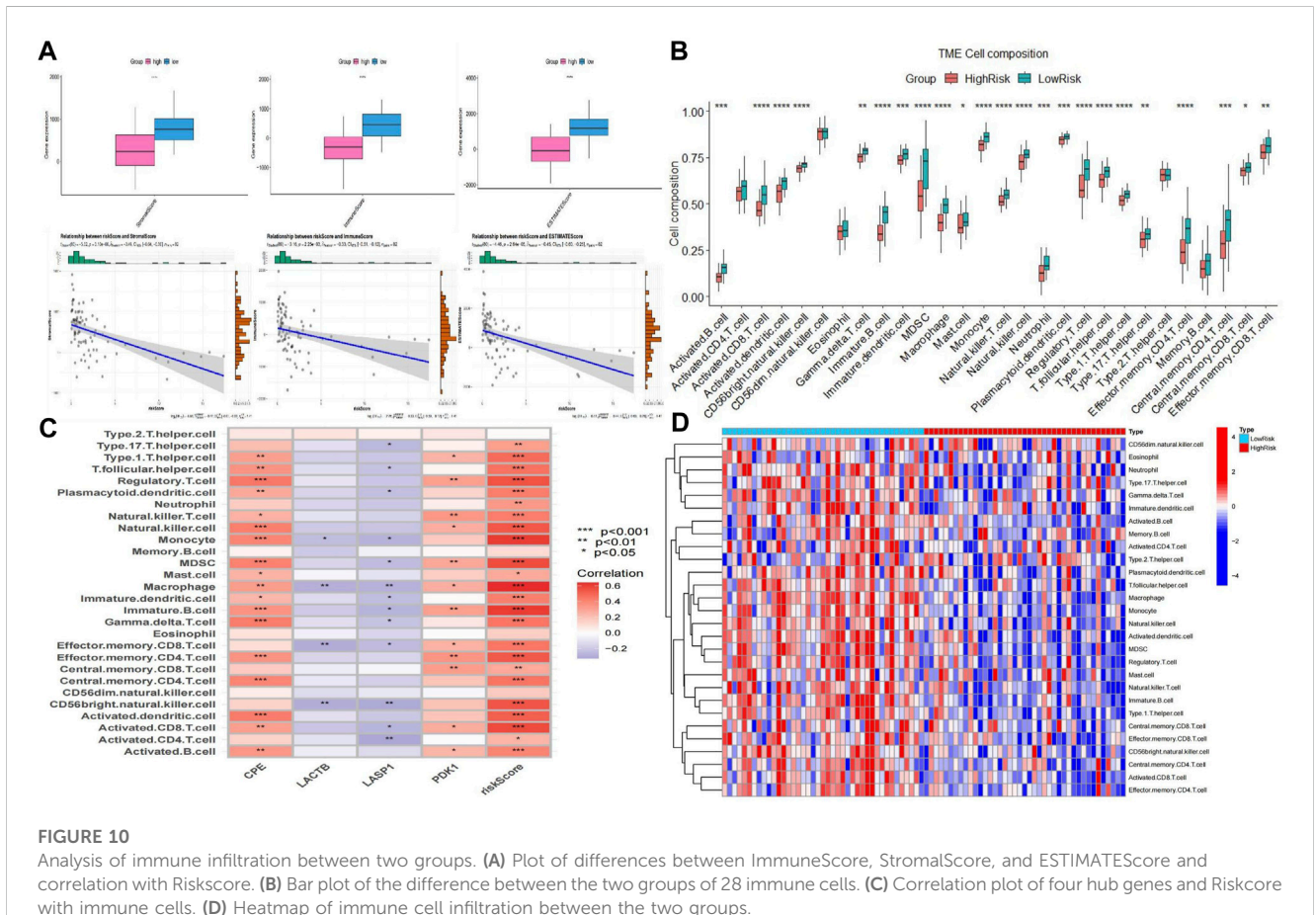
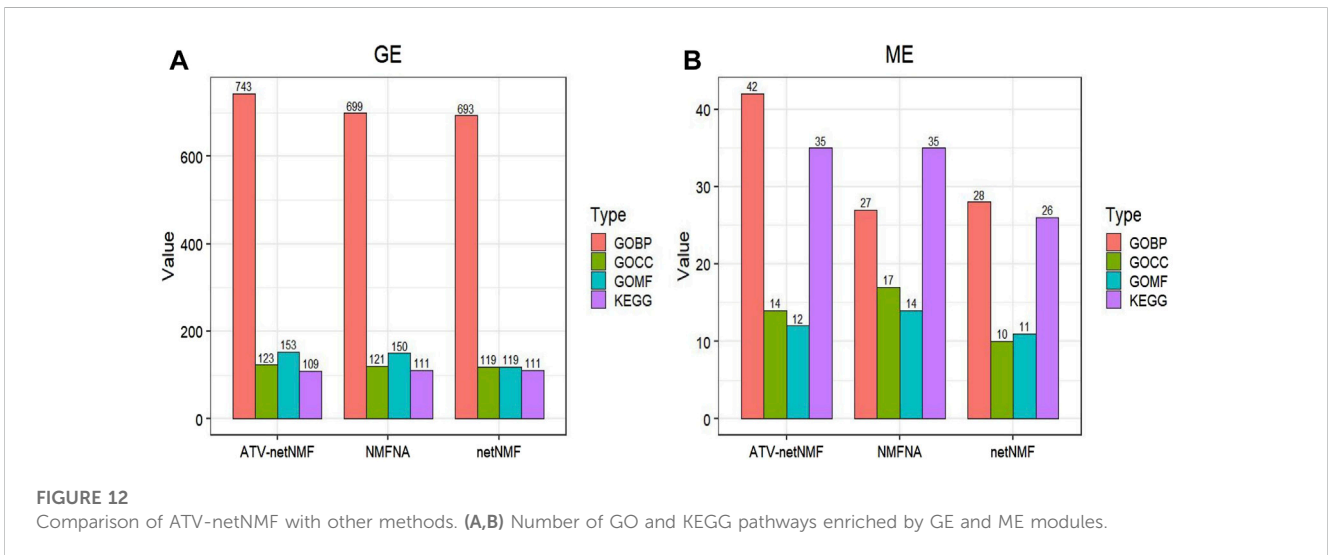
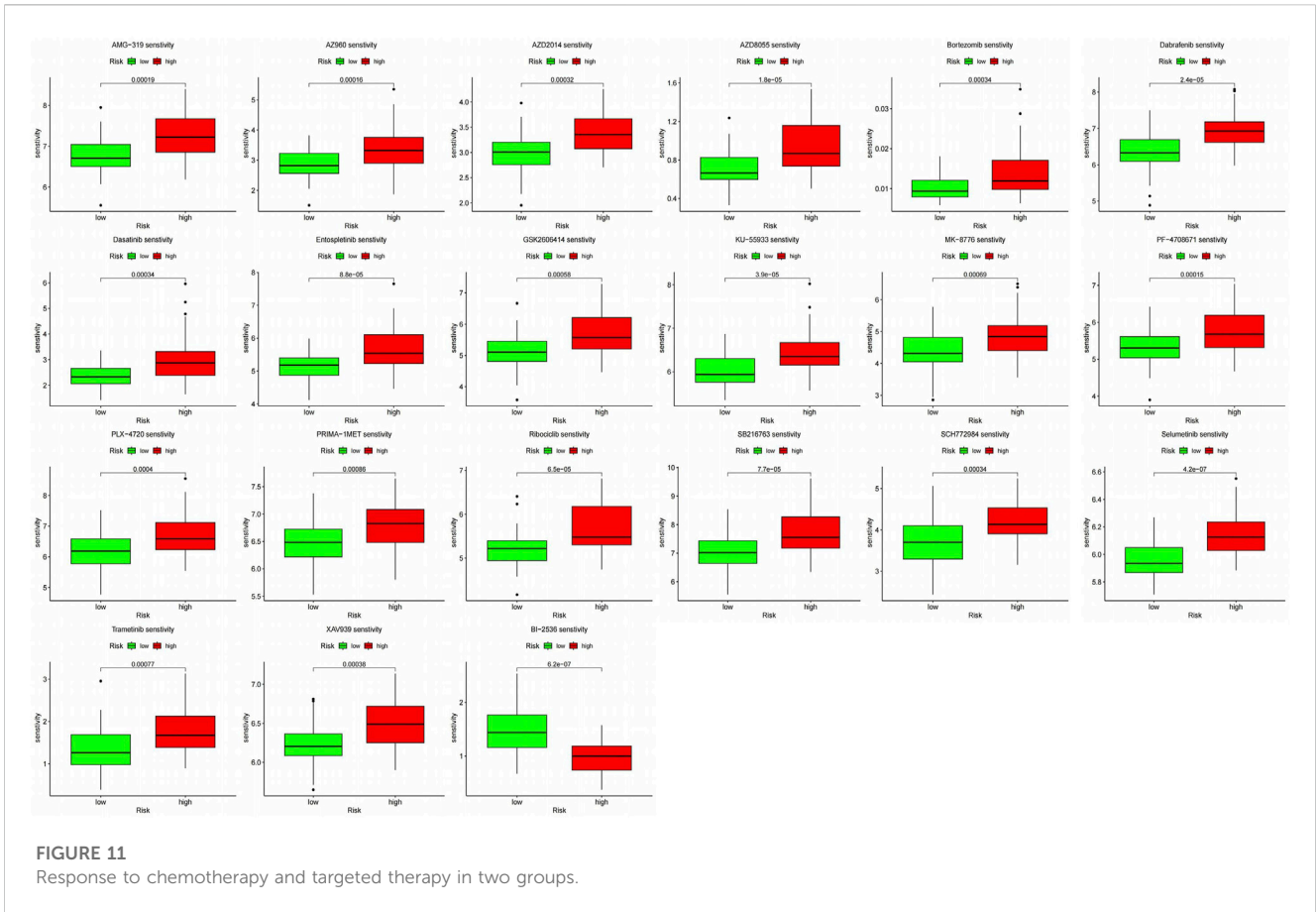


FIGURE 10 Analysis of immune infiltration between two groups. (A) Plot of differences between ImmuneScore, StromalScore, and ESTIMATEScore and correlation with RiskScore. (B) Bar plot of the difference between the two groups of 28 immune cells. (C) Correlation plot of four hub genes and RiskScore with immune cells. (D) Heatmap of immune cell infiltration between the two groups.



et al., 2020). *LASPI* is an actin-binding protein, and overexpression of *LASPI* is associated with poor prognosis in patients with gastric cancer (Keckesova et al., 2017). After the downregulation of *LASPI*, the resistance of osteosarcoma cells to cisplatin was reduced, the IC50 decreased, and the knockdown of *LASPI* could result in the inhibition of the proliferation of osteosarcoma cells (Chang et al., 2022). *LACTB* is a mitochondrial protein that is highly expressed in skeletal muscle, heart, and liver (Keckesova et al., 2017). In breast and

colorectal cancers, low expression of *LACTB* predicts a poorer prognosis for patients (Zhang et al., 2018; Li et al., 2019). However, in glioblastoma, *LACTB* overexpression inhibits cancer cell proliferation (Hu et al., 2022).

Based on these four genetic features in the risk model, we performed a comprehensive analysis and assessment of the two subgroups, which showed a significant difference in survival time between patients in the high-risk and low-risk groups. In

TABLE 1 KEGG pathways that appear in three methods in the GE module.

KEGG pathway	ATV-netNMF	NMFNA	netNMF
	<i>p</i> -value	<i>p</i> -value	<i>p</i> -value
Cytokine-cytokine receptor interaction	2.00E-21	8.06E-22	5.61E-14
Osteoclast differentiation	3.33E-20	1.98E-20	3.71E-18
Cell adhesion molecules	6.05E-16	1.43E-15	1.35E-15
Rap1 signaling pathway	4.33E-11	1.61E-10	4.25E-10
NF-kappa B signaling pathway	4.58E-08	1.64E-07	3.86E-05
Th1 and Th2 cell differentiation	2.07E-07	2.15E-07	1.70E-07
Pathways in cancer	6.19E-07	3.73E-07	3.26E-06
JAK-STAT signaling pathway	6.83E-07	2.14E-06	2.04E-06

The bold value indicates the Minimum values.

TABLE 2 Pearson’s correlation coefficients between the three original matrices and the three reconstructed matrices obtained by different algorithms in different data sets.

		netNMF	NMFNA	netNMF + ATV	ATV-netNMF
TARGET-OS	corr (R_{11}, R'_{11})	0.8737	0.8768	0.8792	0.8813
	corr (R_{12}, R'_{12})	0.9038	0.9060	0.8958	0.9001
	corr (R_{22}, R'_{22})	0.9170	0.9224	0.9194	0.9233
TARGET-WT	corr (R_{11}, R'_{11})	0.8678	0.8682	0.8694	0.8762
	corr (R_{12}, R'_{12})	0.9237	0.9225	0.9166	0.9193
	corr (R_{22}, R'_{22})	0.9167	0.9163	0.9200	0.9204
TCGA-BRCA	corr (R_{11}, R'_{11})	0.9285	0.9293	0.9291	0.9296
	corr (R_{12}, R'_{12})	0.9190	0.9287	0.9213	0.9249
	corr (R_{22}, R'_{22})	0.7579	0.7627	0.7764	0.7708
TCGA-LUAD	corr (R_{11}, R'_{11})	0.9094	0.9052	0.9133	0.9141
	corr (R_{12}, R'_{12})	0.8943	0.9145	0.9162	0.9165
	corr (R_{22}, R'_{22})	0.7894	0.7902	0.7892	0.7930

The bold value indicates the Maximum values.

addition, our KEGG and GO enrichment analyses showed that many immune and tumor-related pathways were enriched. We further performed GSEA analysis and found that metabolism-related pathways such as ascorbate and aldate metabolism, cytochrome P450 in drug metabolism to xenobiotics, adolescent diabetes mellitus, glucose metabolism, etc., were enriched in the high-risk group. In the low-risk group, pathways related to immune function, such as cell adhesion molecules, hematopoietic factor signaling pathways, cytokine-receptor interactions, natural killer cell-mediated cytotoxicity, and hematopoietic cell lines, were enriched.

We further analyzed the infiltration status of various immune cells using ESTIMATE and ssGSEA methods to investigate the immune infiltration differences between the two subgroups. The findings showed that the immune, stromal, and estimate scores were significantly lower in the high-risk group than in the low-risk

group. A significant negative correlation existed between our calculated risk scores and the three immune scores. In addition, the infiltration abundance of all 28 immune cells was lower in the high-risk group, with significant differences in the infiltration abundance of 23 immune cells. Riskscore showed a significant positive correlation with 24 immune cells. Osteosarcoma is considered a “cold tumor” in terms of immunogenicity. In the high-risk group, the infiltration of immune cells was significantly lower. The lower number of immune cells in the tumor immune microenvironment may lead to a worse prognosis. The prognosis of high-risk patients may be improved by increasing the immune reactivity.

The need for high iron is an important feature of many cancer cells (Torti and Torti, 2013), and many cancer cells also have higher basal levels of intracellular unstable iron compared to normal cells. In osteosarcoma cell lines, higher levels of iron in the cells enhanced

ascorbate-induced pharmacological toxicity. They made the cells more sensitive to ascorbic acid, thereby increasing the resistance of MNNG/HOS and U2OS cells to ascorbate-induced drug toxicity (Schoenfeld et al., 2017; Zhou et al., 2020). P450 is a major phase I drug-metabolizing enzyme that activates a variety of potent chemical carcinogens. Previous studies have confirmed that resistance to chemotherapy in osteosarcoma is associated with cytochrome P450, and our results may provide evidence for these previous findings (Ferrari et al., 2009). The researchers found that low expression of monocytes in patients with osteosarcoma reduced the expression of cell adhesion molecules and chemokine receptors, and they also exhibited decreased chemotactic function, i.e., the ability of monocytes to enter the tumor site and initiate an anti-tumor immune response (Mason, 1970). Our findings support the notion that patients with higher monocyte expression have monocytes that can migrate to areas of inflammation that respond to chemotactic proteins, thereby improving survival (Tuohy et al., 2016). Lower expression of regulatory T cells predicts shorter overall survival (Biller et al., 2010); However, a higher degree of T cell infiltration predicts increased survival (Scott et al., 2018). In summary, there is a cross-talk between immune-metabolic responses and tumor-related pathways that lead to tumorigenesis and chemoresistance. The above evidence suggests that immunomodulation has beneficial effects on prognosis. Immune dysfunction promotes tumor progression and drug resistance; therapeutic strategies to reverse immune dysfunction can improve patient prognosis, and identifying relevant biomarkers would further improve clinical response.

Survival of patients in the high-risk and low-risk groups was significantly correlated with their sensitivity to chemotherapy, and changes in therapeutic strategies are necessary to improve outcomes in patients who are insensitive to chemotherapeutic agents. Therefore, we analyzed the sensitivity of patients in both groups to commonly used chemotherapeutic and targeted drugs. The results showed that the high-risk group resisted most drugs, and BI-2536 may be considered a therapeutic candidate for the high-risk group.

This study has some limitations. Although this study is based on multiple datasets and multi-omics data, further experimental validation still needs to be improved. In subsequent studies, we need to conduct more experiments to clarify the underlying molecular mechanisms of MTXDEGs.

5 Conclusion

We proposed an adaptive total variant constrained-based netNMF multi-omics analysis method that integrates and efficiently identifies co-expression modules and characteristic genes in osteosarcoma methylation and gene expression data. Combined with the methotrexate-resistant multi-omics data, we identified a four-gene-based prognostic model with predictive solid ability for patient survival, immune microenvironment, and

immunotherapeutic efficacy, which provides direction for new therapeutic strategies. In conclusion, the MTX resistance-associated model based on ATV-netNMF offers new targets for researchers to explore the mechanism of action of chemoresistance in osteosarcoma.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

ZJ: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Software, Validation, Visualization, Writing—original draft, Writing—review and editing. KH: Conceptualization, Formal Analysis, Methodology, Validation, Writing—original draft, Writing—review and editing. DM: Funding-acquisition, Project administration, Resources, Supervision, Validation, Writing—review and editing. WK: Conceptualization, Methodology, Project administration, Resources, Supervision, Validation, Writing—review and editing. SW: Resources, Supervision, Validation, Writing—review and editing. MG: Data curation, Software, Writing—original draft.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Natural Science Foundation of Shanghai (No. 21ZR1449000).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Bacci, G., Ferrari, S., Bertoni, F., Ruggieri, P., Picci, P., Longhi, A., et al. (2000). Long-term outcome for patients with nonmetastatic osteosarcoma of the extremity treated at the Istituto Ortopedico Rizzoli

according to the Istituto Ortopedico Rizzoli/osteosarcoma-2 protocol: an updated report. *J. Clin. Oncol.* 18, 4016–4027. doi:10.1200/jco.2000.18.24.4016

- Benjamin, R. S. (2020). Adjuvant and neoadjuvant chemotherapy for osteosarcoma: a historical perspective. *Adv. Exp. Med. Biol.* 1257, 1–10. doi:10.1007/978-3-030-43032-0_1
- Billar, B. J., Guth, A., Burton, J. H., and Dow, S. W. (2010). Decreased ratio of CD8+ T cells to regulatory T cells associated with decreased survival in dogs with osteosarcoma. *J. Veterinary Intern. Med.* 24, 1118–1123. doi:10.1111/j.1939-1676.2010.0557.x
- Chang, N., Ge, N., Zhao, Y., Yang, L., Qin, W., and Cui, Y. (2022). HSA_CIRC_0007142 contributes to cisplatin resistance in esophageal squamous cell carcinoma via mir-494-3p/lasp1 axis. *J. Clin. Laboratory Analysis* 36, e24304. doi:10.1002/jcla.24304
- Chen, H., Zhang, X., Wang, X., Quan, X., Deng, Y., Lu, M., et al. (2021). MRI-based radiomics signature for pretreatment prediction of pathological response to neoadjuvant chemotherapy in osteosarcoma: a multicenter study. *Eur. Radiol.* 31, 7913–7924. doi:10.1007/s00330-021-07748-6
- Chen, J., and Zhang, S. (2018). Discovery of two-level modular organization from matched genomic data via joint matrix tri-factorization. *Nucleic Acids Res.* 46, 5967–5976. doi:10.1093/nar/gky440
- Chen, X., Yuan, Q., Liu, J., Xia, S., Shi, X., Su, Y., et al. (2022). Comprehensive characterization of extracellular matrix-related genes in PAAD identified a novel prognostic panel related to clinical outcomes and immune microenvironment: a silico analysis with *in vivo* and *in vitro* validation. *Front. Immunol.* 13. doi:10.3389/fimmu.2022.985911
- de Azevedo, J., Fernandes, T., Fernandes, J., de Azevedo, J., Lanza, D., Bezerra, C., et al. (2019). Biology and pathogenesis of human osteosarcoma (review). *Oncol. Lett.* doi:10.3892/ol.2019.11229
- Ding, C., Li, T., Peng, W., and Park, H. (2006). Orthogonal nonnegative matrix T-factorizations for clustering. *Proc. 12th ACM SIGKDD Int. Conf. Knowl. Discov. data Min.* doi:10.1145/1150402.1150420
- Ding, Q., Sun, Y., Shang, J., Li, F., Zhang, Y., and Liu, J.-X. (2021). NMFNA: a non-negative matrix factorization network analysis method for identifying modules and characteristic genes of pancreatic cancer. *Front. Genet.* 12, 678642. doi:10.3389/fgene.2021.678642
- Ferrari, S., Palmerini, E., Staals, E., Abate, M. E., Longhi, A., Cesari, M., et al. (2009). Sex- and age-related chemotherapy toxicity in patients with non-metastatic osteosarcoma. *J. Chemother.* 21, 205–210. doi:10.1179/joc.2009.21.2.205
- Hu, Y., Liu, H., Zhu, Z., Qi, X., Yuan, W., Tian, M., et al. (2022). LACTB suppresses migration and invasion of glioblastoma via downregulating RHOC/Cofilin Signaling Pathway. *Biochem. Biophysical Res. Commun.* 629, 17–25. doi:10.1016/j.bbrc.2022.09.002
- Jiao, C.-N., Gao, Y.-L., Yu, N., Liu, J.-X., and Qi, L.-Y. (2020). Hyper-graph regularized constrained NMF for selecting differentially expressed genes and tumor classification. *IEEE J. Biomed. Health Inf.* 24, 3002–3011. doi:10.1109/jbhi.2020.2975199
- Keckesova, Z., Donaher, J. L., De Cock, J., Freinkman, E., Lingrell, S., Bachovchin, D. A., et al. (2017). LACTB is a tumour suppressor that modulates lipid metabolism and Cell State. *Nature* 543, 681–686. doi:10.1038/nature21408
- Lee, D. D., and Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. *Nature* 401, 788–791. doi:10.1038/44565
- Lei, T., Qian, H., Lei, P., and Hu, Y. (2021). Ferroptosis-related gene signature associates with immunity and predicts prognosis accurately in patients with osteosarcoma. *Cancer Sci.* 112, 4785–4798. doi:10.1111/cas.15131
- Leng, C., Cai, G., Yu, D., and Wang, Z. (2017). Adaptive total-variation for non-negative matrix factorization on manifold. *Pattern Recognit. Lett.* 98, 68–74. doi:10.1016/j.patrec.2017.08.027
- Levine, S. E. (2005). An adaptive variational model for image decomposition. *Lect. Notes Comput. Sci.*, 382–397. doi:10.1007/11585978_25
- Li, H.-T., Dong, D.-Y., Liu, Q., Xu, Y.-Q., and Chen, L. (2019). Overexpression of LACTB, a mitochondrial protein that inhibits proliferation and invasion in glioma cells. *Oncol. Res. Featur. Preclin. Clin. Cancer Ther.* 27, 423–429. doi:10.3727/096504017x15030178624579
- Li, Y., Li, X., Fan, S., Li, L., Wang, L., Du, Z., et al. (2016). Silencing of carboxypeptidase E inhibits cell proliferation, tumorigenicity, and metastasis of osteosarcoma cells. *Oncotargets Ther.* 9, 2795–2803. doi:10.2147/ott.s98991
- Liu, J., Li, Y., Ma, J., Wan, X., Zhao, M., Zhang, Y., et al. (2023b). Identification and immunological characterization of lipid metabolism-related molecular clusters in nonalcoholic fatty liver disease - lipids in Health and Disease. *Biomed. Cent.* Available at: <https://lipidworld.biomedcentral.com/articles/10.1186/s12944-023-01878-0>. doi:10.1186/s12944-023-01878-0
- Liu, J., Zhong, L., Deng, D., Zhang, Y., Yuan, Q., and Shang, D. (2023a). The combined signatures of the tumour microenvironment and nucleotide metabolism-related genes provide a prognostic and therapeutic biomarker for gastric cancer. *Nat. News* 13, 6622. doi:10.1038/s41598-023-33213-z
- Liu, Y., Gu, Q., Hou, J. P., Han, J., and Ma, J. (2014). A network-assisted co-clustering algorithm to discover cancer subtypes based on Gene Expression. *BMC Bioinforma.* 15, 37. doi:10.1186/1471-2105-15-37
- Lu, Y., Zhang, J., Chen, Y., Kang, Y., Liao, Z., He, Y., et al. (2022). Novel immunotherapies for osteosarcoma. *Front. Oncol.* 12, 830546. doi:10.3389/fonc.2022.830546
- Maeser, D., Gruener, R. F., and Huang, R. S. (2021). OncoPredict: an R package for predicting *in vivo* or cancer patient drug response and biomarkers from cell line screening data. *Briefings Bioinforma.* 22, bbab260. doi:10.1093/bib/bbab260
- Mason, N. J. (1970). *Comparative immunology and immunotherapy of canine osteosarcoma*. SpringerLink. Available at: https://link.springer.com/chapter/10.1007/978-3-030-43085-6_14.
- Mens, M. M., and Ghanbari, M. (2018). Cell cycle regulation of stem cells by microRNAs. *Stem Cell Rev. Rep.* 14, 309–322. doi:10.1007/s12015-018-9808-y
- Mutsaers, A. J., and Walkley, C. R. (2014). Cells of origin in osteosarcoma: mesenchymal stem cells or osteoblast committed cells? *Bone* 62, 56–63. doi:10.1016/j.bone.2014.02.003
- Patil, V. S., Zhou, R., and Rana, T. M. (2013). Gene regulation by non-coding RNAs. *Crit. Rev. Biochem. Mol. Biol.* 49, 16–32. doi:10.3109/10409238.2013.844092
- Prudowsky, Z. D., and Yustein, J. T. (2020). Recent insights into therapy resistance in osteosarcoma. *Cancers* 13, 83. doi:10.3390/cancers13010083
- Sadykova, L. R., Ntekim, A. I., Muyangwa-Semenova, M., Rutland, C. S., Jeyapalan, J. N., Blatt, N., et al. (2020). Epidemiology and risk factors of osteosarcoma. *Cancer Investig.* 38, 259–269. doi:10.1080/07357907.2020.1768401
- Schiavone, K., Garnier, D., Heymann, M.-F., and Heymann, D. (2019). The heterogeneity of osteosarcoma: the role played by Cancer Stem Cells. *Stem Cells Heterogeneity Cancer* 1139, 187–200. doi:10.1007/978-3-030-14366-4_11
- Schoenfeld, J. D., Sibenaller, Z. A., Mapuskar, K. A., Wagner, B. A., Cramer-Morales, K. L., Furqan, M., et al. (2017). Clamo O₂- and H₂O₂-mediated disruption of Fe metabolism causes the differential susceptibility of NSCLC and GBM cancer cells to pharmacological ascorbate. *Cancer Cell* 31, 487–500. doi:10.1016/j.ccell.2017.02.018
- Scott, M. C., Temiz, N. A., Sarver, A. E., LaRue, R. S., Rathe, S. K., Varshney, J., et al. (2018). Comparative transcriptome analysis quantifies immune cell transcript levels, metastatic progression, and survival in osteosarcoma. *Cancer Res.* 78, 326–337. doi:10.1158/0008-5472.CAN-17-0576
- Selga, E., Oleaga, C., Ramirez, S., de Almagro, M. C., Noé, V., and Ciudad, C. J. (2009). Networking of differentially expressed genes in human cancer cells resistant to methotrexate. *Genome Med.* 1, 83. doi:10.1186/gm83
- Shi, Y., He, R., Zhuang, Z., Ren, J., Wang, Z., Liu, Y., et al. (2020). A risk signature-based on metastasis-associated genes to predict survival of patients with osteosarcoma. *J. Cell. Biochem.* 121, 3479–3490. doi:10.1002/jcb.29622
- Sun, Y., Zhang, C., Fang, Q., Zhang, W., and Liu, W. (2023). Abnormal signal pathways and tumor heterogeneity in osteosarcoma. *J. Transl. Med.* 21, 99. doi:10.1186/s12967-023-03961-7
- Torti, S. V., and Torti, F. M. (2013). Iron and cancer: more ore to be mined. *Nat. News* 13, 342–355. doi:10.1038/nrc3495
- Tuohy, J. L., Lascelles, B. D. X., Griffith, J. H., and Fogle, J. E. (2016). Association of canine osteosarcoma and monocyte phenotype and chemotactic function. *J. Veterinary Intern. Med.* Available at: <https://pubmed.ncbi.nlm.nih.gov/27338235/>. doi:10.1111/jvim.13983
- Wang, F., Qin, G., Liu, J., Wang, X., and Ye, B. Integrated genome-wide methylation and expression analyses reveal key regulators in osteosarcoma. *Comput. Math. Methods Med.* 2020(2020) 7067649–11. doi:10.1155/2020/7067649
- Wang, P., Gao, L., Hu, Y., and Li, F. (2018). Feature related multi-view nonnegative matrix factorization for identifying conserved functional modules in multiple biological networks. *BMC Bioinforma.* 19, 394. doi:10.1186/s12859-018-2434-5
- Zhang, B., Liu, J., Li, H., Huang, B., Zhang, B., Song, B., et al. (2023b). Integrated multi-omics identified the novel intratumor microbiome-derived subtypes and signature to predict the outcome, tumor microenvironment heterogeneity, and immunotherapy response for pancreatic cancer patients. *Frontiers* 14, 1244752. doi:10.3389/fphar.2023.1244752
- Zhang, C., Zheng, J.-H., Lin, Z.-H., Lv, H.-Y., Ye, Z.-M., Chen, Y.-P., et al. (2020). Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma. *Aging* 12, 3486–3501. doi:10.18632/aging.102824
- Zhang, J., He, Y., Yu, Y., Chen, X., Cui, G., Wang, W., et al. (2018). Upregulation of mir-374a promotes tumor metastasis and progression by downregulating LACTB and predicts unfavorable prognosis in breast cancer. *Cancer Med.* 7, 3351–3362. doi:10.1002/cam4.1576
- Zhang, J., Wang, S., Bai, Y., Ali, A. M., Deng, J., Chen, Y., et al. (2023a). Mir-197-3p promotes osteosarcoma stemness and chemoresistance by inhibiting SPOPL. *J. Clin. Med.* 12, 1177. doi:10.3390/jcm12031177
- Zhang, Q., Wu, J., Zhang, X., Cao, L., Wu, Y., and Miao, X. (2021). Transcription factor ELK1 accelerates aerobic glycolysis to enhance osteosarcoma chemoresistance through mir-134/PTBP1 signaling Cascade. *Aging* 13, 6804–6819. doi:10.18632/aging.202538
- Zhang, S., Liu, C.-C., Li, W., Shen, H., Laird, P. W., and Zhou, X. J. (2012). Discovery of multi-dimensional modules by integrative analysis of cancer genomic data. *Nucleic Acids Res.* 40, 9379–9391. doi:10.1093/nar/gks725
- Zhou, L., Zhang, L., Wang, S., Zhao, B., Lv, H., and Shang, P. (2020). Labile iron affects pharmacological ascorbate-induced toxicity in osteosarcoma cell lines. *Free Radic. Res.* 54, 385–396. doi:10.1080/10715762.2020.1744577
- Zhu, Y.-L., Gao, Y.-L., Liu, J.-X., Zhu, R., and Kong, X.-Z. (2021). Ensemble adaptive total variation graph regularized NMF for Singlecell RNA-seq data analysis. *Curr. Bioinforma.* 16, 1014–1023. doi:10.2174/1574893616666210528164302
- Zhuang, J., Tian, J., Xiong, X., Li, T., Chen, Z., Chen, R., et al. (2023). Associating brain imaging phenotypes and genetic risk factors via a hypergraph based netNMF method. *Front. Aging Neurosci.* 15, 1052783. doi:10.3389/fnagi.2023.1052783