Ding *et al. Discover Oncology* (2026) 17:604
https://doi.org/10.1007/s12672-026-04761-3

Discover Oncology

ANALYSIS Open Access

Transcriptome combined with single-cell data to construct a prognostic model for glycosylation-related genes in osteosarcoma

Ning Ding[1], Jinlong Li[1], Songbo Shi[1], Jizu Wang[1], Yinliang Ding[1] and Qingshan Yang[1*]

*Correspondence:
Qingshan Yang
yangqingshangs@126.com
[1]Sports Medicine ward, Gansu Provincial Hospital, Lanzhou 730000, Gansu, China

Abstract

Background Osteosarcoma (OS) is a malignant primary bone tumor developing from primitive mesenchymal cells. Glycosylation is important in the adhesion, metastasis and transformation of cancer cells. Nevertheless, the investigation of glycosylation-related genes (GRGs) in OS has been infrequently described.

Methods The OS-related datasets (GSE36001, TARGET-OS and GSE21257), scRNA-seq dataset (GSE152048) and 582 GRGs were contained in this research. The identification of candidate genes was conducted by differential expression analysis and weighted gene co-expression network analysis (WGCNA). The prognostic genes filtering method was univariate cox regression, and risk score were generated by the implementation of LASSO. Samples comprising the TARGET-OS and GSE21257 datasets were categorized into high- and low-risk according to the risk score. Through applying various cox regression analysis, independent prognostic variables were identified. Furthermore, functional enrichment analysis and immune microenvironment analysis were performed. Finally, the key cells were identified, and the cell-cell communication and pseudo-time analyses of key cells were conducted.

Results The analysis of scRNA-seq data identified cell types in OS and detected the expression of prognostic genes. Seven glycosylation-related genes (*DCN*, *RENBP*, *UAP1*, *B4GALNT1*, *CCDC115*, *TUBA1A* and *B3GALT4*) were identified as prognostic genes. The low-risk group exhibited a comparatively elevated rate of survival and 1-year OS patients had the highest survival rate. Immune and apoptosis-related signaling pathways were activated in the high-risk group, indicating a worse prognosis. The analysis of scRNA-seq data revealed DCN and TUBA1A were high- expressed in osteoblastic cells. Thus, the osteoblastic cells were identified as key cells. Osteoblastic cells exhibited frequent interactions with other annotated cells. Pseudo-time analysis elucidated the differentiation direction of key cells from left to right.

Conclusions we identified seven prognostic genes for OS. The prognostic models were constructed, and underlying molecular mechanisms were explored based on diagnostic genes, which served as a benchmark for OS prognosis and treatment.

Keywords Osteosarcoma, Glycosylation-related genes, Prognostic model, Molecular mechanisms

1 Background

Osteosarcoma (OS) is a common bone cancer in both adults and adolescents [1], and it mostly occurs in long bones, especially around the knee [2]. Recently, a large amount of research has focused on chemotherapy and its combination regimens, but the therapeutic effects are still limited. As a new treatment method, targeted therapy provides a new treatment option for patients with osteosarcoma and has broad application prospects [3]. Currently, OS is dominantly thought to originate from mesenchymal stem cells (MSCs) or osteoprogenitor cells [4], however, the exact OS progenitor and generation are still unclear [5]. OS exhibits high genomic complexity and instability with easier rearrangement and copy number change, which is likely to impact immune system [6] as well as tumor microenvironment (TME) [7]. Such instability and mutability result in oncogenesis and bring challenges to diagnosing in the early stage and evaluating prognosis. Apparently, it is urgent to find some biological markers of OS. It has been revealed that TRIM21 plays an important role in the development of osteosarcoma by regulating the expression of the TXNIP/p21 axis to inhibit the aging of osteosarcoma cells [8]. In addition, miR-605-3p exerts tumor suppressive effects and enhances chemotherapy sensitivity of osteosarcoma cells by regulating the expression of RAF1. The study also suggests that utilizing stable nano vesicles to efficiently deliver miR-605-3p for tumor therapy may be a promising therapy for treating osteosarcoma [9].

Glycosylation is a post-translated modification pattern and is even more common than phosphorylation [10]. It is obtained through the formation of a linkage between protein and single sugar or glycan and is shown as several types which are N-, P-, C-, O-glycosylation, etc [11], and it is catalyzed by glycosyltransferases in the endoplasmic reticulum or Golgi apparatus normally [10]. Glycosylation is essential for function, structure, stability and solubility of protein [11]. However, abnormal glycosylation appears in various carcinomas, and functions in tumor progression, metastasis, cell junctions, as well as epithelial-mesenchymal transition (EMT) [12–15]. Proteins could be shown as having a low or enhanced level of glycan or some novel glycans [11]. And alternations in glycosylation related to tumor functions were first reported more than half century ago [16]. More and more studies reveal that glycosylation plays a vital role in tumorigenesis [17], developmental dysfunction, congenital diseases and immune disease [18]. Recently, altered glycosylation is regarded as a marker of cancer [17, 18]. The glycosylation of many cytokines, such as CD4, CD7 and CD38, is involved in regulating immune environment. CD44 glycosylation plays a role in pancreatic, breast cancers [19]. Furthermore, the changing glycosylation pattern of ST6GALNAC5 might indicate breast cancer and metastasis [20, 21]. Carcinoembryonic antigen (CEA) is highly glycosylated in colorectal cancer [22, 23]. EGFRs, as transmembrane proteins, are glycosylated and participate in physiological and pathological cellular regulatory pathways, including cell death and survival [24]. E-cadherin is a cell adhesion molecule that is widely located on the membrane and contains4 N-glycosylation sites [25].With the development of gastric carcinoma, E-cadherin becomes highly glycosylated [25–27]. PSA glycosylation plays an important role in prostate cancer [28, 29]. Moreover, several studies have revealed that glycosyltransferase are up-regulated in OS [30], and FUT8, one of the fucosyltransferases, promotes oncogenesis of OS [31]. Obviously, many of those hallmarks are glycoproteins or glycolipids [32]. Cytokines can be detected by antibodies, but they are difficult to measure in early stage [10].

Since its first application in 2009 [33], single-cell RNA sequencing (scRNA seq) technology has broken through the technical limitations of the biomedical field. It can deeply explore gene information in different tissues and diseases through single-cell expression profiling analysis, and capture the inherent heterogeneity of samples. With the continuous improvement of technology, it has enhanced high-throughput capability and sensitivity, promoting the development of full-length transcriptome analysis [34, 35]. By integrating transcriptome data with scRNA seq, researchers classified osteosarcoma samples into two molecular subtypes and significantly improved the accuracy of prognosis by combining immune cell composition and functional status [36]. In addition, scRNA seq further revealed the diversity and status changes of immune cells in osteosarcoma, identifying the role of key immune cell subsets in tumor immune regulation [36]. The application of scRNA seq in orthopedic research [37] provides a theoretical basis for precision medicine and personalized treatment, and is expected to accelerate clinical translation.

Studies on the regulation of OS glycosylation are limited. Therefore, differential glycosylation-related genes (GRGs) were screened, and their prognostic value was explored through public databases and bioinformatics analysis. In addition, the analysis of the immune microenvironment was combined to provide insights for immunotherapy of OS.

2 Methods

2.1 Data acquisition and processing

The OS-related datasets GSE36001, GSE21257 and TARGET-OS were collected from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) and the UCSC Xena (University of California Santa Cruz) website (http://xena.ucsc.edu/). In the GSE36001 datasets (Chip expression data), six normal samples (osteoblasts and bones) and 19 OS cell lines were preserved [38]. The GSE21257 (microarray expression) dataset included 53 patients with OS in the survival analysis to validate the risk model. The TARGET-OS (HTSeq-FPKM and HTSeq-Counts) retained 85 OS patients in the survival analysis, excluding patients with missing survival data, such as survival time and survival status. In this study, except for differential expression analysis which utilized the HTSeq-Counts expression matrix (for subsequent GSEA), all other analyses employed the log2-normalized HTSeq-FPKM expression matrix (log2(fpkm + 1)). Concurrently, gene symbols were derived from the Ensemble IDs in the expression matrix by referencing species-specific annotation files. In addition, OS patients with clinical data (age, gender, metastasis) were used to construct prognostic model. Single-cell RNA sequencing (scRNA-seq) dataset GSE152048 was retrieved from GEO website, containing 11 OS samples (Table 1). Additionally, a total of 582 GRGs were obtained through msigdbr package (version 7.5.1) [39].

2.2 Differential expression analysis between OS and normal groups

The differentially expressed genes (DEGs) in the GSE36001 dataset were selected adopting the limma package (version 3.54.0) [40] between the OS and normal groups, with threshold of p.adj < 0.05 and $|\log_2 FC| > 0.5$. Volcanic plots and heatmaps were employed to represent the DEGs, utilizing the ggplot2 (version 3.4.1) [41] and ComplexHeatmap (version 2.15.1) packages, respectively [42].

Table 1 Single cell sample information in the GSE152048 dataset

Sample	Pathological type	Type
BC2	Osteoblastic OS	Primary
BC3	Osteoblastic OS	Primary
BC5	Osteoblastic OS	Primary
BC6	Osteoblastic OS	Primary
BC10	Osteoblastic OS	Metastasis
BC11	Osteoblastic OS	Recurrent
BC16	Osteoblastic OS	Primary
BC17	Chondroblastic OS	Metastasis
BC20	Chondroblastic OS	Recurrent
BC21	Osteoblastic OS	Primary
BC22	Chondroblastic OS	Primary

2.3 Key modules and genes to be recognized

The co-expression network was constructed using the Weighted Gene Co-expression Network Analysis (WGCNA) package (version 1.70.3) [43] on the gene expression matrix of the GSE36001 dataset, with the OS group and the normal group serving as phenotypic traits. Firstly, the hierarchical cluster analysis was implemented and outlier samples were removed. Subsequently, the scale-free topology fit index of 0.9 was utilized to ascertain the most suitable soft threshold power (β). Next, constructing a modular dynamic shear tree was accomplished based on the dynamic tree cutting algorithm. Finally, in order to determine the key modules, the Pearson correlation coefficient was utilized to evaluate the relations within each trait and module. The genes comprising the key module were regarded as key module genes.

2.4 Identification and enrichment analysis of candidate genes

The positive-correlation intersection genes were identified by classifying overlapping genes in up-regulated DEGs and positive-correlation key module genes using the VennDiagram package (version 1.7.1) [44]. Similarly, down-regulated DEGs and negative-correlation key module genes were taken to intersect to obtain negative-correlation intersection genes. Then, the negative and positive-correlation intersection genes were combined, and intersection was taken with GRGs to obtain candidate genes.

Using the clusterProfiler package ($p.adj < 0.05$) (version 4.2.2) [45], Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses of the candidate genes were carried out to look into the functional items connected with the candidate genes. The establishment of a Protein-Protein Interaction (PPI) network was accomplished by utilizing the STRING database(http://string-db.org/) with a confid ence interval > 0.4 to uncover candidate genes interactions at the protein level [46].

2.5 The creation and verification of a risk model

Firstly, the univariate cox regression analysis was conducted deploying the survival package (version 3.3-1.3) [47] to filter the prognostic genes in training set (TARGET-OS) based on expression of candidate genes, with threshold of hazard ratio (HR) ≠ 1 and $p < 0.05$. Subsequently, risk model was built from prognostic genes applying the least absolute shrinkage and selection operator (LASSO) algorithm in glmnet (version 4.1-2.1) [48]. Then, the risk score was calculated using the risk model formula, and an optimal cutoff value was used to categorize the samples in the TARGET-OS dataset and the

validation set (GSE21257) into high- and low-risk groups. The risk model was validated using the GSE21257 dataset. Afterwards, the exploration of the survival status was conducted through the utilization of Kaplan–Meier (K–M) survival analyses ($p < 0.05$) and receiver operating characteristic (ROC) curves [49, 50]. Additionally, to explore differences in survival across clinical status, clinical information (age, gender, metastasis) of OS samples from the two groups was integrated into prognostic analysis. The heatmap was utilized to describe the expression patterns of prognostic genes across various clinical conditions.

2.6 Independent prognostic analysis

In the training set (TARGET-OS), the pathological indicators (age, gender, Non-metastatic) of OS and risk score were integrated into independent prognostic analyses. Univariate cox regression analyses ($p < 0.05$), proportional hazards (PH) assumption test and multivariate cox regression analyses were adopted to screened prognostic variables ($p < 0.05$). Thereby, nomogram of 1-, 3-, and 5-year survival rates for patients with OS were constructed based on independent prognostic variables. Each distinct value of a factor corresponds to a "point (individual score)" on the right-hand line segment. The "Total score" represents the sum of individual scores for independent prognostic factors. The "Pr" on the right vertical axis denotes the "probability of death". A higher total score correlates with a larger corresponding "Pr" value. In addition, calibration curves, were executed to indicate the differentiated validity of the nomogram for OS. Decision Curve Analysis (DCA), ROC analysis, and global sensitivity analysis were used to evaluate the performance of the nomogram.

2.7 Enrichment analysis

During data processing, genes with expression levels less than or equal to 1 across all samples were filtered out. The DEseq2 (version 1.26.0) [51] was then employed to conduct a differential expression analysis comparing the high- and low- risk groups within the training set (TARGET-OS) and $\log_2$FC were acquired. The $\log_2$FC was sorted from largest to smallest, followed by GSEA enrichment analysis utilizing the clusterProfiler package (version 4.2.2) [45]. The reference gene set was the "h: HALLMARK" from MSigDB database. The visualization of the signaling pathways was conducted with the enrichplot (version 1.18.3) [52] ($p.adj < 0.05$). Further, correlations between risk score and signaling pathways were elucidated in scatterplots.

2.8 Analysis of the immunological microenvironment in the context of OS

To explore the ratios of relevant immune cells, stromal cells and tumor cells in the OS microenvironment, the StromalScore, ImmuneScore and ESTIMATEScore of the biological samples was calculated by ESTIMATE algorithm to further assess OS prognosis [53]. Afterwards, the single sample gene set enrichment analysis (ssGSEA) algorithm from the GSVA (version 1.42.0) [54] was employed to obtain enrichment scores for 16 immune-infiltrating cells and 13 immunological functions within the two groups of the training set (TARGET-OS). The analysis of different immune cells in the two groups was conducted by employing the Wilcox test. ($p < 0.05$). Further, Spearman correlation analysis revealed correlations between prognostic genes, different immune cells and different immune functions.

2.9 The exposure of immune checkpoints and immunotherapy

Protein molecules that regulate the immune system were known as immune checkpoints, and inhibitors that target these immune checkpoints had emerged as a crucial approach in cancer immunotherapy. In this study, 48 immune checkpoints from the literature were extracted and intersected with genes in the TARGET-OS dataset to genes and expression matrix [55]. Differential expression analysis was conducted to identify the different immune checkpoints present in the two groups. The Spearman correlation analysis was employed to establish correlations between various immune checkpoints and the risk score. Meanwhile, the GSVA package was utilized to compute the ssGSEA score for the immunotherapy prediction pathway. The Spearman approach was employed to determine the relevance of the risk score to the enrichment score [56].

2.10 Sensitivity analysis of OS to chemotherapeutic drugs

As immunotherapy has become one of the key strategies for treating osteosarcoma, understanding patient sensitivity to chemotherapy drugs remains an important component of clinical treatment. For the purpose of assessing the susceptibility of patients with OS to chemotherapy drugs, we retrieved the drug sensitivity data of cancer cell lines (CCL) from the genomics of Drug Sensitivity in Cancer (GDSC2) database (https://www.cancerrxgene.org/), and then calculated the half maximal inhibitory concentration (IC_{50}) of each drug for the samples in the TARGET-OS dataset by utilizing acalcPhenotype function in the oncoPredictpackage (version 0.1) [57].Finally, the Wilcoxon test was employed for analyzing the disparity in IC_{50} values in two different groups to seek different response chemotherapy drugs.

2.11 Regulatory network analysis

To clarify the regulatory mechanisms of key prognostic genes and established OS driver genes-related non-coding RNAs (miRNAs and lncRNAs) in OS. Using multiMiR package to obtain mirnas associated with prognostic genes and MYC, TP53. Following this, starbase (http://starbase.sysu.edu.cn/) database was used to predict lncRNAs that int eracted with these key miRNAs (clipExpNum > 10). Integrating these relationships, a comprehensive lncRNA-miRNA-mRNA network was constructed and visualized using Cytoscape (v 3.10.0) [58].

2.12 Analysis of scRNA-seq data

The Seurat package (version 4.1.0) [59] was used to filter the data: (1) : cells in which the number of detected genes was 300-5,000 were retained; (2) : cells with a total number of detected molecules less than 25,000 were reserved; 3): cells with a proportion of mitochondrial genes less than 10% were included. For the purpose of to mitigate the influence of batch effects on the analysis and optimize gene expression information for the included cells, the IntegrateData function from the Seurat package was utilized. Following that, the data underwent normalization using the Normalize Data function, and the identification of highly variable genes was executed utilizing the FindVariableFeatures function (version 4.1.0) [59]. Subsequently, the process of dimensionality reduction using the uniform manifold approximation and projection (UMAP) algorithm was executed (resolution = 0.1). Further, the cell clusters obtained from UMAP dimensionality reduction clustering were annotated using the SingleR package (version 1.831) [60].

Marker gene was obtained from published literature [61], and the expression of prognostic genes in different cell clusters was detected by wilcoxon test. Additionally, a bar chart depicting the proportion distribution of different cell types in the single-cell sequencing data was generated. Notably, the cell cluster with the highest expression of prognostic genes were considered as key cell.

2.13 Enrichment analysis and cell-cell communication analysis

To explore the biological function of annotated cells, ReactomeGSA package (version 1.12.0) [62] was utilized. Besides, to investigate intercellular interactions among annotated cells, cell-cell communication analysis was conducted using CellChat package (version 1.6.1) [63].

2.14 Pseudo-time analysis

To examine the heterogeneity of key cell, dimensional clustering analysis was performed on the scRNA-seq data. The key cells were classified into different subtypes. Concurrently, marker genes for these subtypes were identified using the FindAllMarkers function (min.pct = 0.5, logfc.threshold = 0.25, test.use = "auc"). Following this, pseudo-time analysis was conducted using Monocle (version 2.30.0) [64] to explore the differentiation states and trajectories of key cell. Additionally, the expression of prognostic genes was assessed at different stages of differentiation in the key cell.

2.15 Prognostic gene expression in the OS and Normal groups

Differential expression analysis was carried out on the dataset GSE36001 to investigate the expression patterns of prognostic genes in the OS group compared to the normal group, and the ggplot2 package(version 3.4.1) [41] was utilized for visualization of the results.

2.16 Statistical analysis

The statistical analysis was conducted using R programming language version 4.1.0. The Wilcoxon test was utilized to conduct a comparison of data across different categories. Without additional specification, statistical significance was defined as a p-value less than 0.05.

3 Results

3.1 Identification of DEGs and key modules

Analysis of the GSE36001 dataset (OS vs. Normal) identified a total of 2,345 DEGs, including 1,149 up-regulated and 1,196 down-regulated genes (Fig. 1a-b). The samples obtained from the GSE36001 dataset were subjected to clustering analysis, which no outlier samples were excluded (Fig. 1c). Afterwards, in order to ensure a biologically meaningful scale-free topology, a optimal value of 8 was selected for β, taking into consideration a scale independence criterion of >0.9. (Fig. 1d). A total of 13 co-expression modules were identified. Among them, the MEturquoise and MEblue modules showed the strongest (negative) correlations with OS (Cor = −0.81 and −0.88, respectively). These two modules contained 9,677 and 7,512 key module genes, respectively (Fig. 1e-f).

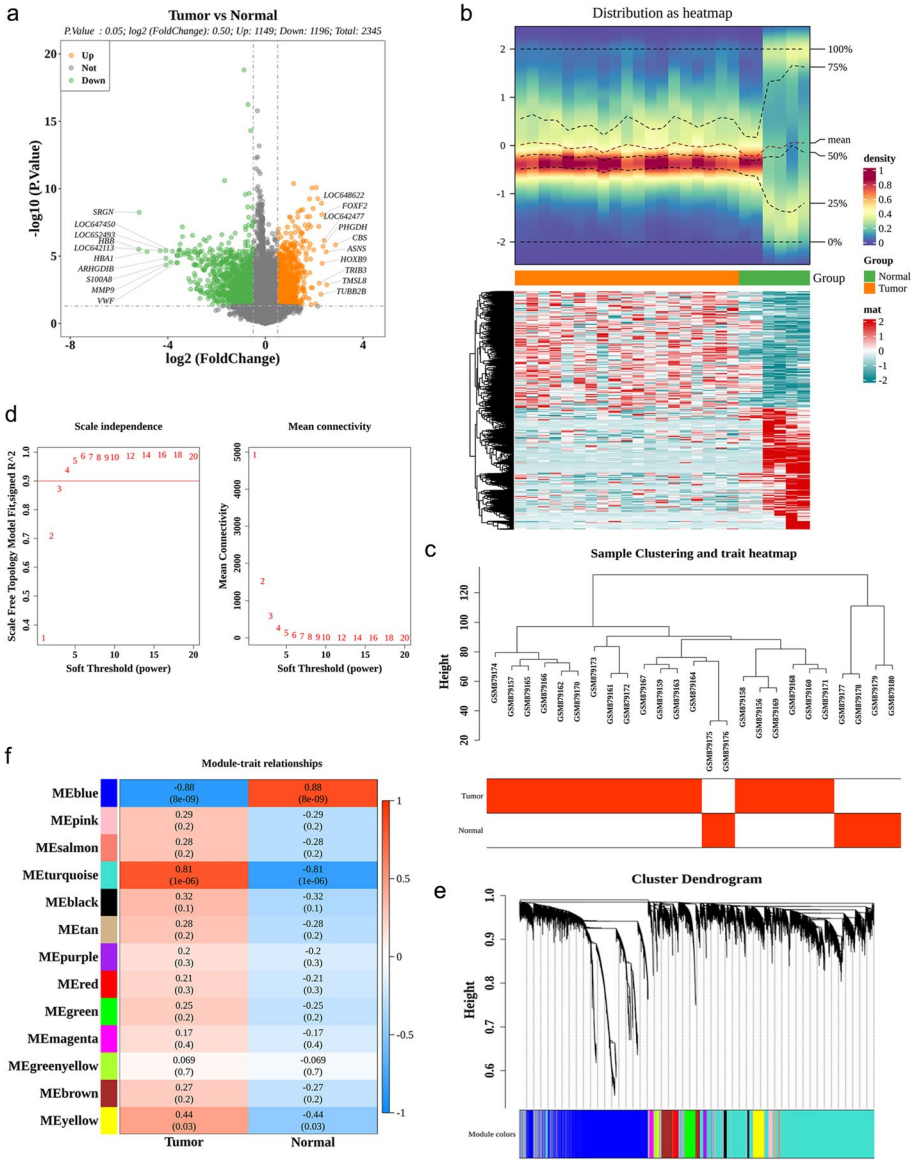

Fig. 1 Identification of differential expression genes (DEGs) and construction of weighted gene co-expression network analysis (WGCNA) in GSE36001. **a** Volcano plot of DEGs. (Green points represent down-regulated genes, and orange points represent up-regulated genes. Top 10 upregulated and downregulated genes are labeled by names). **b** Heatmap of DEGs with normal in green and tumor in orange. **c** Samples clustering diagram. **d** Determination of soft-threshold power. **e** Analysis of Cluster Dendrogram. **f** Correlation between module and trait in both tumor and normal group

3.2 Identification and functional exploration of candidate genes

A total of 59 candidate genes were identified by VennDiagram (Fig. 2a-c), which were enriched in 144 biological processes such as glycosylation, protein glycosylation, macromolecular glycosylation and protein N-linked glycosylation (Fig. 2d) ($p.adj < 0.05$). Moreover, the signaling pathways primarily involved in candidate genes were salmonella infection, amino sugar and nucleotide sugar metabolism and N-Glycan biosynthesis (Fig. 2e) ($p.adj < 0.05$). The PPI network consisted of 55 nodes (red circles indicating upregulated proteins and blue circles indicating downregulated proteins) and 119 edges. Notably, four gene-encoded proteins were excluded from the network as they could not be mapped to the STRING database at the confidence score threshold of > 0.4 and thus

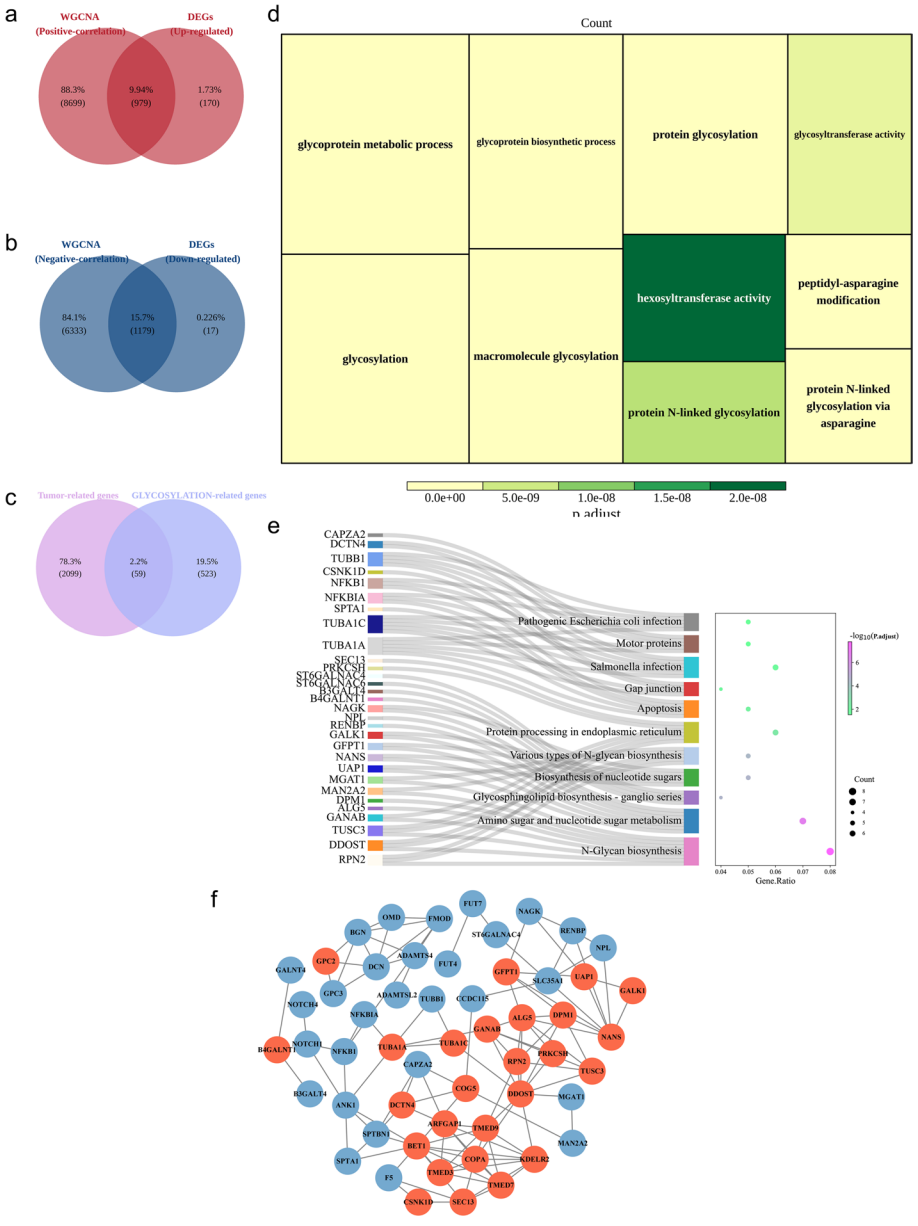

Fig. 2 Screening candidate genes and enrichment analysis. **a**, **b** Venn diagram of the tumor-related genes from WGCNA and DEGs. **c** Venn diagram of tumor-related genes and glycosylation-related genes. **d** Gene Ontology (GO) analysis of candidate genes. **e** Sankey diagram of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway of candidate genes. **f** Protein-protein interaction (PPI) network of candidate genes with 55 nodes. Red circles indicate upregulated proteins; blue circles indicate downregulated proteins

lacked qualified protein-protein interaction information. The network suggests a potential interaction between COPA and the TMED family (Fig. 2f). There is currently no direct evidence indicating a functional association between these two proteins and the glycosylation process or the prognosis of osteosarcoma. Therefore, further research is needed to explore this possibility.

3.3 Risk model based on seven prognostic genes well predicted survival in OS patients

First, seven prognostic gene were included in the analysis by univariate Cox regression, of which *UAP1* and *B4GALNT1* exhibited a high-risk profile (HR > 1), while

B3GALT4, *TUBA1A*, *CCDC115*, *RENBP*, and *DCN* had a low-risk profile (HR < 1) (Fig. 3a). The ideal value of lambda in the Lasso regression was found to be 0.020 by the use of 10-fold cross-validation. At this lambda value, the number of genes selected was 7. (Fig. 3b, c), and the risk model formula was "RiskScore=−0.426*DCN-0.175*RENBP + 0.136*UAP1 + 0.242*B4GALNT1-0.259*CCDC115-0.469*TUBA1A-0.639*B3GALT4". Then, the samples from the TARGET-OS dataset were grouped into

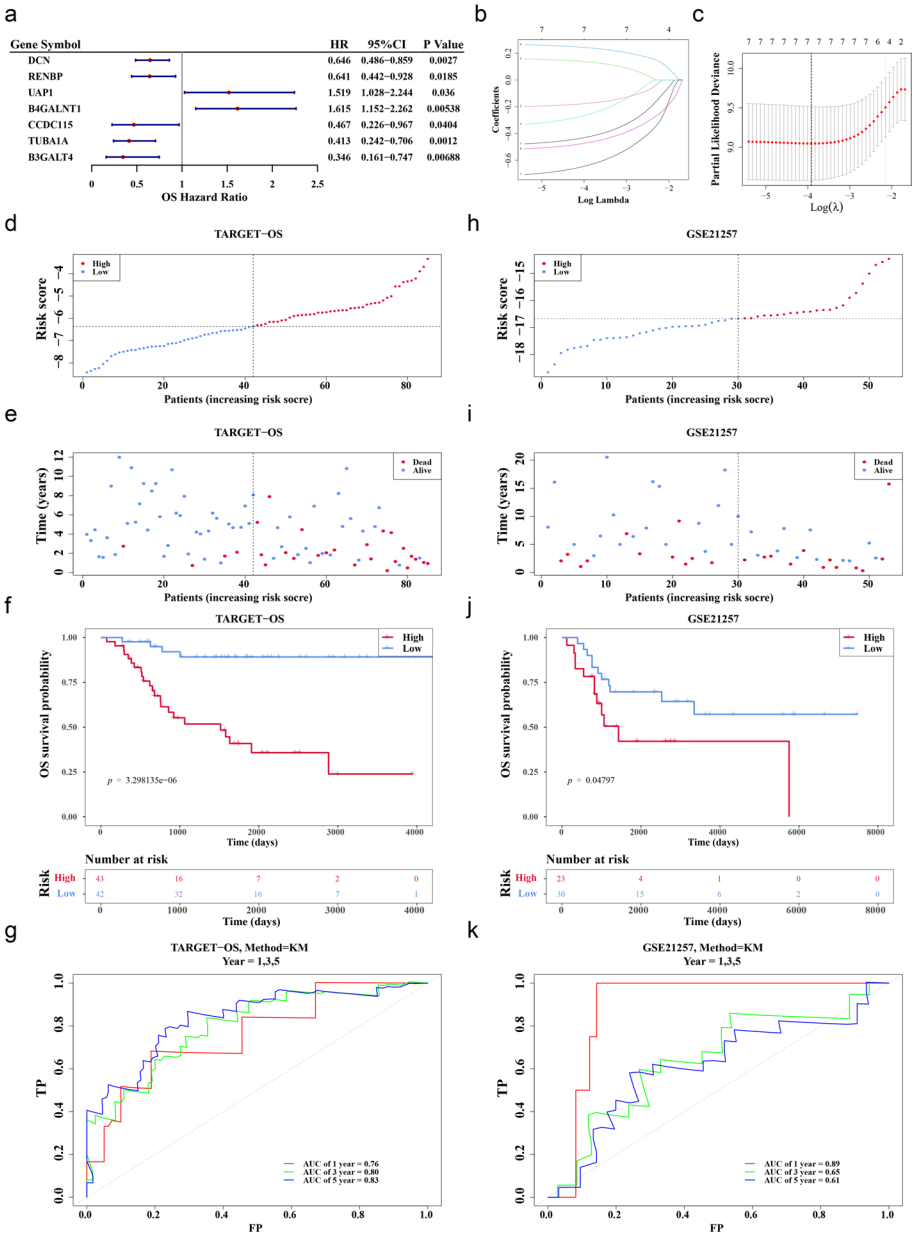

Fig. 3 Construction and identification of risk prognostic model in TARGET-OS. **a** Univariate cox regression analysis exhibiting 2 high-risk genes and 5 low-risk genes. **b** Least absolute shrinkage and selection operator (LASSO) regression analysis. **c** Analysis of optimal penalty parameter by LASSO regression. **d** The determination of high-and low-risk groups. **e** Survival status map. **f** Survival curve. **g** Receiver operating characteristic (ROC) curves. Test group from GSE21257 (**h**–**k**). **h** The determination of risk score. **i** Survival status map: The red dots represent the deceased samples, and the blue dots represent the surviving samples. **j** Survival curve: The red curve represents the high-risk group, and the blue curve represents the low-risk group. A statistically significant difference was observed ($p < 0.05$). **k** The Receiver Operating Characteristic (ROC) curve: The red, green, and blue curves correspond to 1-year, 3-year, and 5-year survival, respectively

two distinct groups, namely high risk and low risk, utilizing the optimal cutoff value (−6.370) for the Risk score. (Fig. 3d). Figure 3e displayed a higher mortality rate in the high-risk groups based on the distributions of risk score and survival status. The survival rate noticed in the high-risk group exhibited a meaningful decrease ($p < 0.05$), according to the K-M's findings. Furthermore, the AUC values for 1-, 3-, and 5-year survival in patients with OS exceeded above 0.70, indicating that the risk model accurately predicted survival in OS (Fig. 3f-g). Meanwhile, similar results were obtained in the GSE21257 dataset. The samples comprising GSE21257 were classified into high-risk and low-risk groups according to the risk score's optimal cutoff value (−16.674) (Fig. 3h), The distribution of risk score, survival status and K-M curve ($p = 0.048$) indicated higher mortality in the high-risk group. Furthermore, the AUC values for 1-, 3-, and 5-year survival in patients with OS were all greater than 0.60, indicating moderate predictive performance (Fig. 3i-k).

3.4 A worse prognosis appeared in patients in the high-risk group with varying clinical status

The K-M curve, which incorporated age, gender, and metastasis for both groups, revealed that the prognosis was significantly worse for all subgroups within the high-risk group (Fig. 4a–c). Analysis of the expression patterns of prognostic genes in different clinical states showed that *B3GALT4*, *RENBP*, *CCDC115*, *DCN* and *TUBA1A* expression was significantly elevated in the high-risk + non-metastatic group (Fig. 4d).

3.5 Construction of nomograms for OS based on independent prognostic variables

Two independent prognostic variables (risk score and Non-metastatic) were screened by univariate cox regression analyses, PH hypothesis test and multivariate cox regression analyses (Fig. 5a, b). Based on the nomogram, the anticipated mortality rates for patients with osteosarcoma (OS) at 1 year, 3 years, and 5 years were 2.6%, 27.6%, and 35.7%, respectively. (Fig. 5c). The great precision of nomogram in forecasting the likelihood of survival for OS patients was demonstrated by the calibration curves (Fig. 5d). The Decision Curve Analysis (DCA) curve consistently lies above the other curves (Fig. 5e). In the ROC analysis, the AUC values for 1-year, 3-year, and 5-year survival were all greater than 0.8 (Fig. 5f). In the sensitivity analysis results of both the training set (Supplementary Table 1) and the validation set (Supplementary Table 2), the sigma values for each variable were close to zero, indicating stable and predictable model performance.

3.6 A total of nine signaling pathways were activated in high-risk group for OS

This study confirmed that both "risk score" and "non-metastatic status" were independent prognostic variables for OS, but subsequent pathway analysis focused on the "risk score" because it better reflected the expression pattern of GRGs and was closely related to molecular-level analyses such as pathway activation and immune microenvironment, supporting the core logic of GRGs-pathway-prognosis. From the GSEA results, it could be observed that 9 signaling pathways were activated in the high-risk group of OS, namely, interferon γ response, inflammatory response, allograft rejection, interferon α response, IL6-JAK-STAT3 signaling, epithelial mesenchymal transition, complement, apoptosis, and TNF-α signaling via NF-κB (Fig. 6). Moreover, risk score was negatively correlated with nine signaling pathways (Fig. 7).

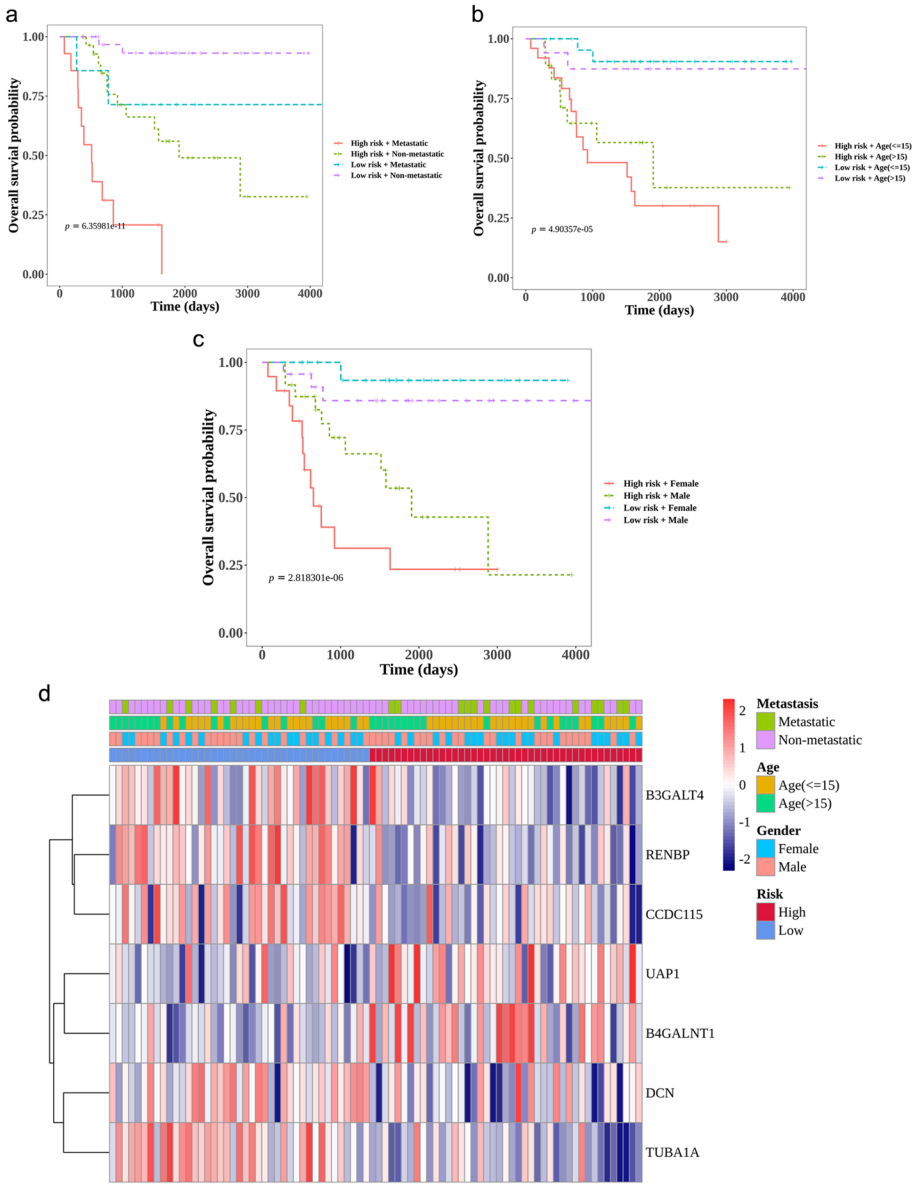

Fig. 4 Survival analysis of TARGET-OS. **a** Survival cure related to metastasis. **b** Survival cure related to age. **c** Survival cure related to gender. **d** Risk heatmap of prognostic genes

3.7 Immune cells and prognostic genes played important roles in OS

The high-risk group exhibited low StromalScore, ImmuneScore, and ESTIMATEScore according to the ESTIMATE calculation, meaning that OS patients at high-risk group had an unfavorable prognosis (Fig. 8a). Then, we observed that 11 categories of immune cells, containing macrophages, neutrophils, T helper cells, and TIL, were extensively infiltrated in the low-risk group. In addition, 13 immune functions, which comprised T cell co-inhibition, parainflammation, T cell co-stimulation, type I IFN Reponse, type II IFN Reponse, and others, exhibited higher scores in the low-risk group (Fig. 8b, c). The Correlation analysis between different immune cells exhibited significant positive correlations for multiple pairs of cells, for instance, macrophages-T helper cells ($r = 0.861$), neutrophils-TIL ($r = 0.765$), TIL-pDCs ($r = 0.864$). Correlation analysis between different immune function showed significant positive correlation between Check-point and CCR

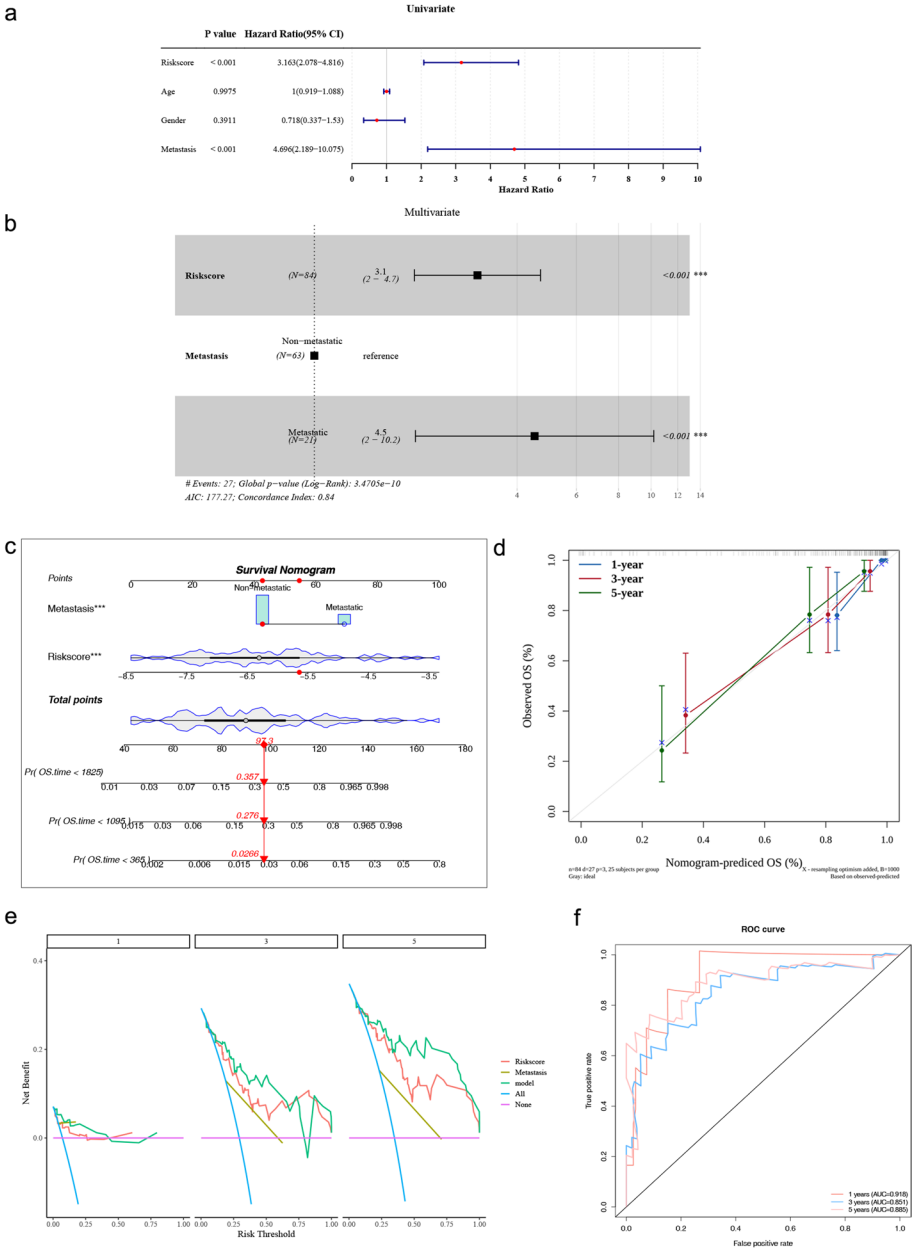

Fig. 5 A nomogram associated with multiple clinical factors. **a** Univariate cox regression analysis. **b** Multivariate cox regression analysis containing riskscore and Non-metastatic. **c** Survival nomogram for evaluating the 1-, 3- and 5-year overall survival of OS patients. **d** Nomogram calibration for OS patients. **e** Nomogram DCA curve. **f** Nomogram ROC curve

($r = 0.904$), T cell co-inhibition ($r = 0.952$), parainflammation ($r = 0.880$), HLA ($r = 0.807$), and inflammation-promoting ($r = 0.866$) (Fig. 8d-e). Further, *RENBP* exhibited the highest correlation with T cell co inhibition ($r = 0.501$) and T helper cells ($r = 0.596$) (Supplementary Fig. 1a-b). A sum of 19 immune checkpoints exhibited differential expression in two groups. Notably, the low-risk group possessed high-level *HAVCR2*, *LAIR1*, *LGALS9*, and *SELPG*, whereas the high-risk group proved high-level of *CD27* and *PDCD1LG2* (Supplementary Fig. 1c). Moreover, there was a substantial negative correlation across risk score and *HAVCR2* ($r = -0.433$, $p < 0.001$) and *SELPG* ($r = -0.414$,

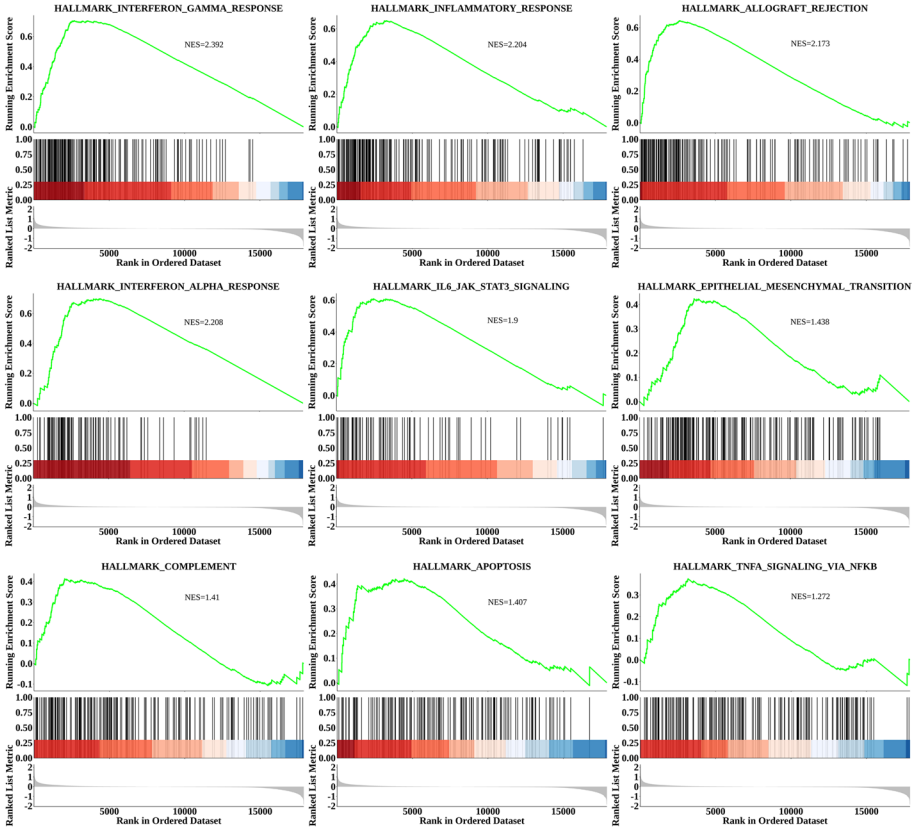

Fig. 6 Enrichment plots of high-risk group from gene set enrichment analysis (GSEA)

$p < 0.001$) (Supplementary Fig. 1d), were significantly linked with IFN-γ-signature ($r < -0.4$, $p < 0.001$) and pyrimidine metabolism ($r > 0.3$, $p < 0.01$) (Supplementary Fig. 1e).

3.8 A total of 24 chemotherapeutic drugs with high efficacy in high-risk group were identified

In order to investigate the sensitivity of patients with OS to chemotherapeutic drugs, a total of 198 data of drug sensitivity were extracted from the GDSC2 database. Among these, 33 of drugs displayed IC_{50} differences in the two groups ($p < 0.05$). The IC50 values for 24 drugs, including BI.2536 – 01086, Daporinad-1248, and Linsitinib-1510, were considerably greater among the low-risk group. This finding suggested that the high-risk group exhibited superior efficacy for 24 of the drugs (Fig. 9a). Further analysis of the correlation unveiled a substantial negative association with the drugs and risk score (Supplementary Fig. 2).

3.9 The regulation network was helpful for exploring the mechanism of OS

In the CeRNA network regulatory analysis of prognostic genes and OS, 154 miRNAs and 37 lncRNAs were identified. A network comprising 151 nodes and 235 edges was constructed (Figure 9b). Notable pathways within this network include 'TP53-hsa-miR-4284-PCBP2-OT1', 'MYC-hsa-miR-1294-AC007161.3', and 'TUBA1A-hsa-miR-222-3p-TUG1', among others. The Gene-gene Interaction 𝔊GI𝔄analysis revealed that key genes and other related genes were enriched in pathways including ubiquitin-like protein ligase binding, cellular response to oxygen levels, negative regulation of cell cycle G1/S

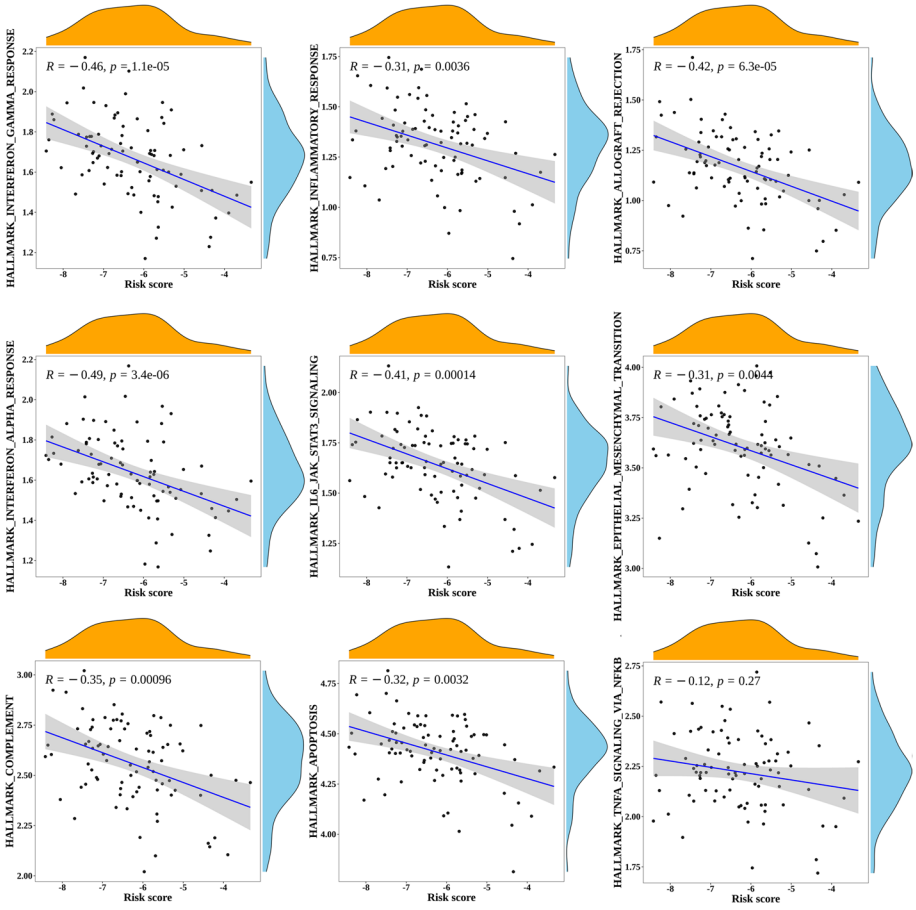

Fig. 7 Scatter plot visualizing correlation analysis between riskscore and 9 enrichment pathways

phase transition, response to oxygen levels, cell cycle G1/S phase transition, regulation of G1/S transition of mitotic cell cycle, and intrinsic apoptotic signaling pathway by p53 class mediator (Fig. 9c).

3.10 Cellular distribution and characteristics in OS

Following scRNA-seq data normalization and quality control, 26,145 genes and 102,589 cells were preserved, and sample batch effects were effectively eliminated (Supplementary Fig. 3). The cells in the GSE152048 dataset were effectively categorized into eight distinct clusters, whether in metastasis, primary and recurrent lesions, or in the two types of conventional OS (chondroblastic OS and osteoblastic OS) using UMAP dimensionality reduction clustering (Fig. 10a, b). According to the expression patterns of the marker genes, eight cell clusters were annotated, which uncovered 7 types of cells, namely osteoblastic cell, myeloid cells, myoblasts, TILs, osteoclasts, endothelial cells and fibroblasts (Fig. 10c–f). In the single-cell sequencing data, osteoblastic cells accounted for the largest proportion of OS tissue (46.1%), followed by myeloid cells (23.93%). TILs made up 6.46%, and these data supported the results of immune infiltration (Supplementary Fig. 4). For instance, the abundance of osteoblastic cells directly impacted the StromalScore, with the low StromalScore in the high-risk group possibly indicating a reduction in osteoblastic cells, which could further exacerbate immune suppression (e.g., by decreasing the secretion of osteoprotegerin and TGF-β). Meanwhile, the

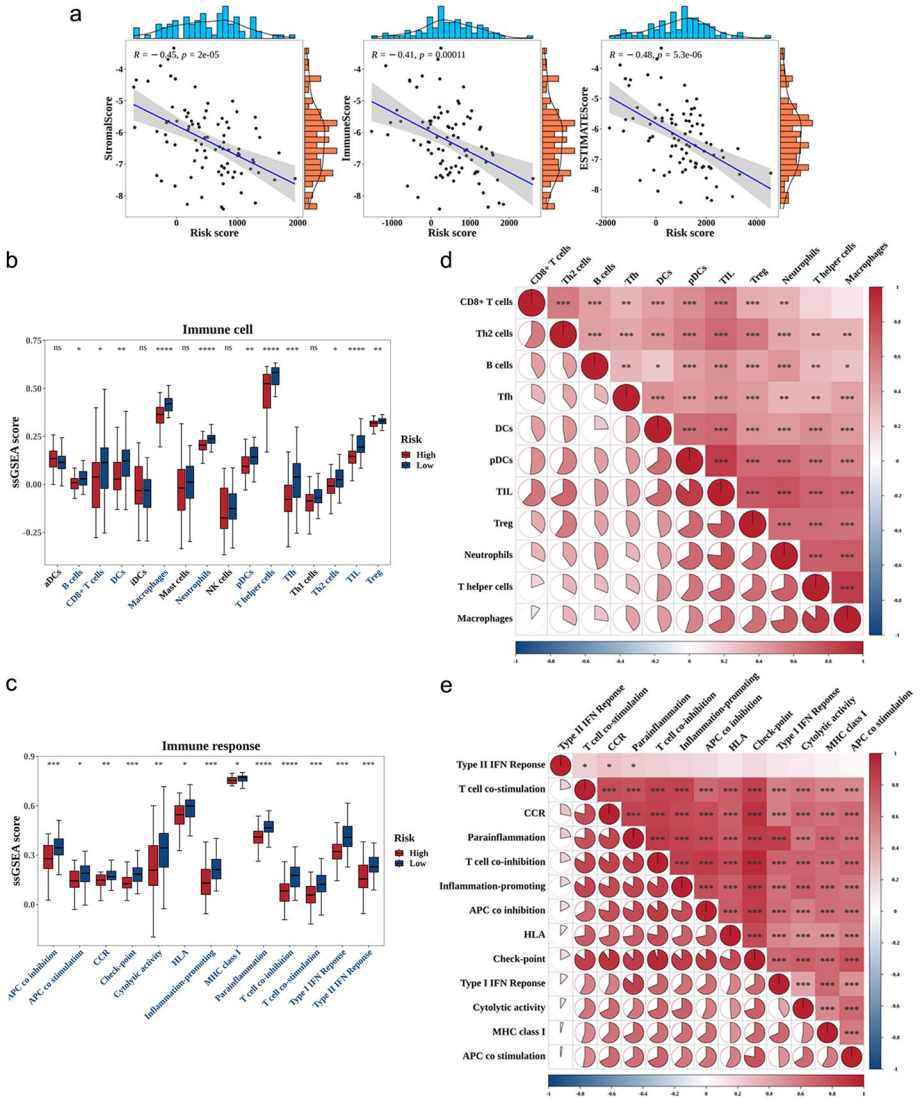

Fig. 8 Immune correlation analysis of prognostic genes. **a** Scatter plot visualizing correlation analysis between-Riskscore and StromalScore, Immunescore and ESTIMATEScore in TARGET-OS. **b** Immune cells infiltration comparing high and low-risk group. **c** Immune responses comparing high and low risk group. **d** Correlation analysis of immune cells. **e** Correlation analysis of immune response

enrichment of TILs in the low-risk group suggested that they were an important cell population in immune regulation, supporting the conclusion that the low-risk group relied on TILs to enhance anti-tumor immunity. However, this mechanism requires subsequent validation through cell co-culture experiments. Furthermore, it was observed that prognostic genes exhibited distinct expression patterns across various cell types. Specifically, the *DCN* and *TUBA1A* displayed considerably elevated expression levels in osteoblastic cells and fibroblasts (Fig. 10g). The osteoblastic cells were identified as key cells. Moreover, function enrichment analysis showed annotated cells were enriched in "ficolins bind to repetitive carbohydrate structures on the target cell surface", "metabolism of ingested MeSeO2H into MeSeH", "proton-coupled neutral amino acid transporters", etc. (Fig. 10h).

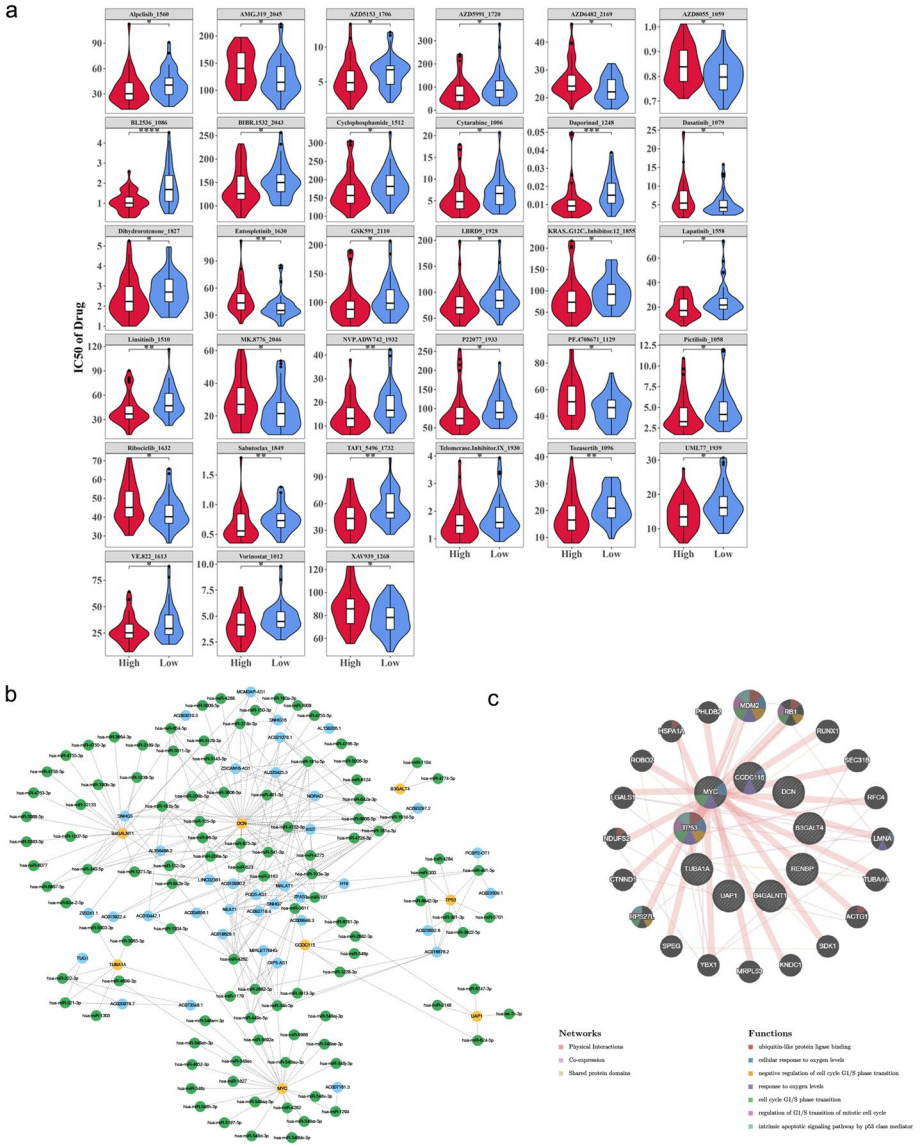

Fig. 9 Sensitivity analysis of chemotherapeutic drugs and lncRNA-miRNA-mRNA network associated with prognostic genes. **a** The comparison of sensitive drugs between low and high risk group. **b** The prediction miRNAs targeting prognostic genes. **c** lncRNA-miRNA-mRNA network. (*$p < 0.05$;**$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns, $p > 0.05$)

3.11 The communication and differentiation state of key cells were found

Cell-cell communication analysis revealed distinct interaction patterns in the OS samples across primary, recurrent, and metastatic groups. In primary samples, osteoblastic cells and osteoclasts exhibited frequent interactions (Fig. 11a), while myoblasts and osteoclasts interacted via SPP1-(ITGAV+ITGB1) (Supplementary Fig. 5a). In recurrent samples, myeloid cells and fibroblasts had frequent interactions (Fig. 11b), and myoblasts and tumor-infiltrating lymphocytes (TILs) interacted via MIF-(CD74 + CXCR4) (Supplementary Fig. 5b). In metastatic samples, osteoblastic cells and fibroblasts showed frequent interactions (Fig. 11c), and myoblasts and myeloid cells interacted via MIF-(CD74 + CXCR4) (Supplementary Fig. 5c). Pseudo-time analysis revealed a progressive differentiation of osteoblastic cells from left to right (Fig. 11d). Additionally, osteoblastic cells were classified into 14 subtypes (Fig. 12a). The expression of prognostic genes

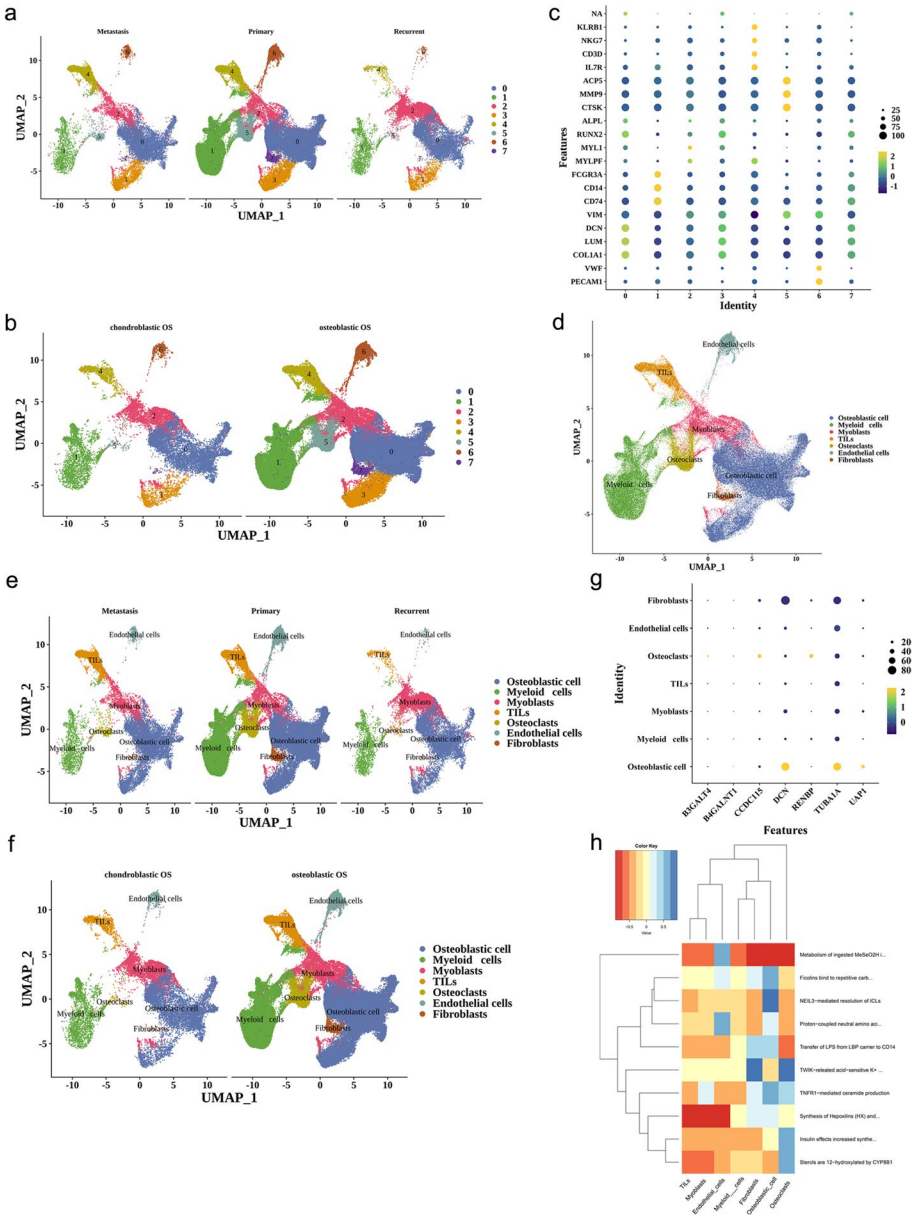

Fig. 10 Uniform manifold approximation and projection (UMAP) analysis. **a** Cell clusters by UMAP dimensionality reduction clustering in metastasis, primary and recurrent lesions. **b** Cell clusters analysis in chondroblastic and osteoblastic OS. **c** Marker genes annotation. **d**–**f** Identification of 7 major Cell clusters. **g** Expressions of prognostic genes in different cells. **h** Enrichment analysis of annotated cells

indicated that *CCDC115* remained stable and was upregulated in the late stages. *DCN* was upregulated in the early, the downregulated in middle, and then upregulated in late stages, while *TUBA1A* showed increased expression in the middle stage. UAP1 was upregulated in the middle and late stages. In contrast, the expression of *RENBP*, *B4GALNT1*, and *B3GALT4* did not significantly change during osteoblastic cell differentiation (Fig. 12b).

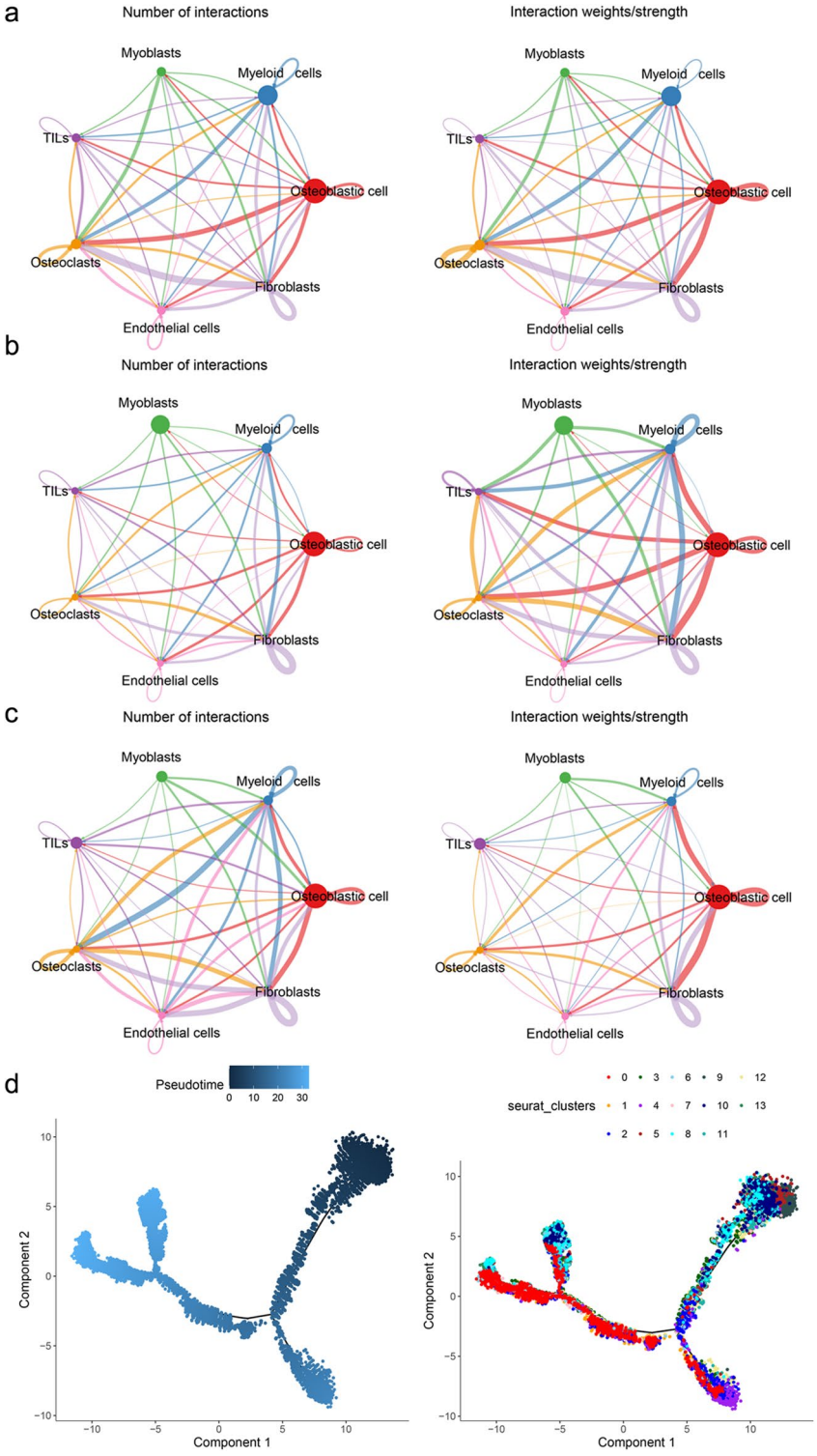

Fig. 11 The cell-cell communications and pseudo-time analysis. **a** The communications in primary samples. **b** The communications in recurrent samples. **c** The communications in metastatic samples. **d** Pseudo-time analysis

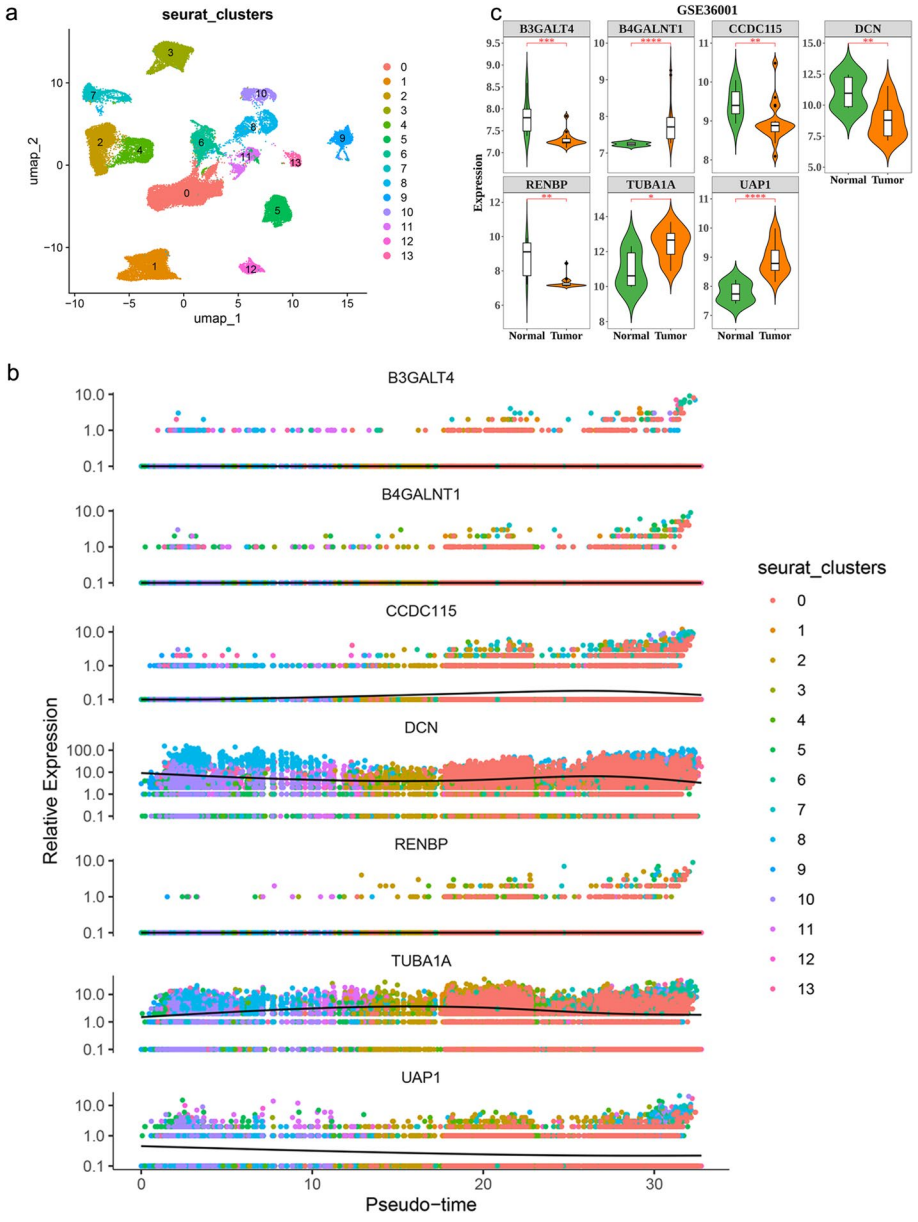

Fig. 12 The analysis of subtypes and prognostic genes. **a** osteoblastic cells were classified into 14 subtypes. **b** The pseudo-time analysis of prognostic genes. **c** Prognostic genes expressions in GSE36001 between normal and OS group. (*$p < 0.05$;**$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$)

3.12 Levels of prognostic gene expression in OS and normal samples

The GSE36001 dataset revealed that OS patients presented high levels of expression for *B4GALNT1*, *TUBA1A*, and *UAP1*. Conversely, Normal samples demonstrated elevated expression levels for *B3GALT4*, *CCDC115*, *DCN*, and *RENBP* (Fig. 12c).

4 Discussion

OS is a common primary bone cancer, and the 5-year survival rate of OS is no more than 30% [65–68]. Early detection and effective therapy are urgent. It has been revealed that glycosylation, as a widespread post-translation modification, occurs in certain cancers [69], such as hepatocellular carcinoma, lung adenocarcinoma [70], and prostatic cancer

[29], and regulates various biological programmes. Currently, glycosylation in OS is receiving increasing attention.

In this study, 59 candidate genes were identified from GSE36001, most of the enrichment fields of which were related to glycosylation, such as hexosyltransferase activity, N-linked glycosylation, glycosyltransferase activity, protein glycosylation, and so on. N-linked glycosylation occurred through the attachment of glycans to nitrogen sides on proteins, which mainly took place in the endoplasmic reticulum (ER) [71, 72]. Protein glycosylation occurred within the ER and whether they were transported to the Golgi was determined by the correctness of folding [73]. Further modifications by glycosyltransferases and mannosidases were completed in Golgi [74]. In terms of functionality, N-glycosylation contributes to protein stabilization [75–77], cell adhesion [78], peptide folding, subcellular localization of protein, and inhibiting protein aggregation [79, 80], as well as serving as markers of activation, diseases and cancers [29, 69, 81, 82]. Abnormal N-glycosylation reflected cancer development and metastasis to some extent [16, 83, 84]. From the previous studies, we can find that N-glycosylation was critical for various cancers, which suggested it may also play a role in OS development, progression, and metastasis. Our functional enrichment analysis results provided strong evidence for this.

Further, seven prognostic genes were identified among the 59 candidate genes, including *UAP1*, *B4GALNT1*, *B3GALT4*, *TUBA1A*, *CCDC115*, *RENBP* and *DCN*. *UAP1* as a positive regulator of innate immune response is pyrophosphorylated at s386 by targeting IRF3 [85]. UAP1 and B4GALNT1 were associated with a high risk for OS, while the rest posed a low risk. DCN (Decorin) had previously been considered a biomarker for lung adenocarcinoma [70]. As a proteoglycan, glycosylation modification of DCN may play an important role in regulating immune responses, particularly in triggering inflammatory processes [86]. OS cells induced bone mesenchymal stem cells (BMSCs) to differentiate into cancer-associated fibroblasts (CAFs), which generated DCN [87]. During iron death, DCN is released by dying cells, triggering innate and adaptive immune responses. It binds to the specific receptor ager on macrophages, thereby inducing the production of proinflammatory cytokines [86]. B3GALT4 has been validated as a diagnostic marker for gynecological cancer and carcinoma of colon and rectum (CRC) [88, 89]. CCDC115 (Coiled-coil domain-containing protein 115) was overexpressed in liver hepatocellular carcinoma (LIHC) and promoted infiltration of immunosuppressive cells [90]. The absence of CCDC115 is closely related to abnormal glycosylation processes, especially in the Golgi apparatus [91]. Previous studies suggested that it might be a hallmark of liver and prostatic cancer [90, 92]. CCDC115 can promote the infiltration of immunosuppressive cells [91, 93]. Tubulin alpha 1a (TUBA1A) encoded the tubulin-1 A chain and was a member of the tubulin superfamily. Its aberrant expression results in a variety of organ malformations [94]. Although there is no direct research elucidating the specific role of TUBA1A in glycosylation, its expression level is associated with post-translational modifications of O-GlcNAc glycosylation [95]. In addition, TUBA1A was overexpressed in breast cancer [62] and gastric cancer [63], facilitating massive infiltration of M2 macrophages in TAMs [96]. And TUBA1A indirectly regulates immune cell function by altering microtubule dynamics [97]. However, limited research has been conducted on RENBP (renin-binding protein).

The nomogram of this study integrates the risk scores of seven prognosis related GRGs (such as DCN, RENBP, UAP1, etc.) with clinical key variables (such as metastatic

status), showing excellent predictive ability for 1 -, 3 -, and 5-year survival rates of patients (AUC > 0.85). Compared with the existing prediction tools [98–100], this study does not rely on the traditional clinicopathological characteristics (such as tumor size, metastasis status, histological grading), and realizes the dual evaluation of "molecular characteristics+clinical phenotype". In addition, the traditional staging system (such as TNM) has certain limitations in the risk stratification of early (no metastasis) osteosarcoma, and often cannot effectively distinguish patients who "appear to be low-risk but are actually prone to recurrence". The nomogram of this study can accurately identify high-risk subgroups in the early clinical stage (without metastasis) through the molecular characteristics of GRGs, especially for newly diagnosed osteosarcoma patients (such as adolescents, those without definite metastasis) and postoperative recurrence risk assessors. In clinical practice, the expression of seven GRGs was detected by RNA sequencing of tumor tissue, combined with the clinical metastasis status, and then the total score was calculated by nomogram, which could estimate the 1 -, 3 -, and 5-year survival probability. When formulating treatment strategies, the high-risk group should give priority to intensive chemotherapy + immune checkpoint inhibitors. The low-risk group can simplify the chemotherapy regimen, while the metastatic patients can choose precision drugs according to drug sensitivity. However, the nomogram model of this study also has some limitations. For example, primary hospitals usually lack high-throughput sequencing platforms, and the cost of single RNA sequencing is high, which is much higher than that of traditional pathological detection (such as TNM staging, only imaging + pathological section is required). In general, although there are some challenges, the nomogram proposed in this study provides a new idea for the individualized risk assessment of osteosarcoma patients and has important clinical application potential.

UDP-N-acetylglucosaminepyrophosphorylase 1 (UAP1) is a key rate limiting enzyme responsible for the generation of Uridine Diphospho-N-acetylglucosamine(UDP-GlcNAc), it is an important substrate for glycosylation reactions [101]. Studies have shown that UDP GlcNAc in OS may promote tumor growth and metastasis by regulating key signaling pathways (such as PI3K/AKT and WNT/β -catenin) [102]. Therefore, UAP1, as a key enzyme regulating the generation of UDP GlcNAc, plays an important role in the occurrence and development of OS, and may serve as an important biomarker and therapeutic target of OS. In addition, UAP1 is highly expressed in prostate cancer, protecting cancer cells from endoplasmic reticulum stress, thereby promoting cancer cell growth [103].

Beta-1,3-galactosyltransferase-4 (B3GALT4), β3GalT gene family's members, promoted glycosylation. Studies have found that the mutation of B4GALNT1 (β −1,4-n-acetylgalactosamine transferase 1) directly affects the biological processes related to glycosylation, especially in the nervous system [104]. In addition, B4GALNT1 promotes the development of hepatocellular carcinoma (HCC) [105], and participates in the metastasis of lung adenocarcinoma through JNK/c-Jun/slug pathway [106]. In recent years, studies have found that B4GALNT1 expression is significantly elevated in a variety of cancer cells [107–109]. Especially in OS, the expression level of B4GALNT1 is closely related to the occurrence and development of tumors, which may promote tumor progression by regulating cell proliferation, migration and invasion. Therefore, B4GALNT1

is considered as a potential molecular target of OS, which may provide a new strategy for the treatment of it [110].

These results supported the conclusion that UAP1, B4GALNT1, B3GALT4, TUBA1A, CCDC115, RENBP and DCN were associated with OS and other cancers and had the potential to serve as prognostic predictors of OS. Independent prognostic analysis showed risk score and metastasis correlate with prognosis of OS. In case of high risk and metastasis, the survival rate decreases more sharply. The occurrence of metastasis implied more mortal.

Furthermore, pathways were identified in the high-risk OS, including INF-γ, inflammatory, allograft, INF-α, IL-6/JAK/STAT3 signaling, epithelial mesenchymal transition (EMT), complement and apoptosis, which were correlated to riskscore. EMT represented the transformation of cells from an epithelial phenotype to a mesenchymal phenotype. When EMT occurred, cancer cells lost tight cell junction and obtained high probability of migration, invasion and metastasis [111]. JAK/STAT, PI3K/AKT pathway, as well as Tim-3 (T-cell immunoglobulin domain and mucin domain-3) and LAIR1 (Leukocyte associated immunoglobulin like receptor 1, or CD305), were associated with EMT [112–114]. B4GALNT1 and B3GALT4 promoted EMT of HCC [105] and breast cancer [115] respectively, finally propelling the development of cancer. However, research on EMT of OS was rare. Certain cytokines regulating OS were not clear. In addition, many other pathways were related to the immune system and inflammation. Immune cells and cytokines played a crucial role. Understanding the OS immune microenvironment contributed to the immunotherapy of OS. In this study, the high-risk group exhibited decreased infiltration of TILs, neutrophils, T helper cells and macrophages, along with lower immune function scores and decreased expression of immune checkpoints. This suggested a poorer prognosis.

Moreover, 19 differentially expressed check points significantly were identified and illustrated. All of them were down regulated in high risk group, and their expressions were negatively correlated with rickscore, especially SELPLG, LGALS9, HAVCR2 and LAIR1.SELPLG (Selectin P ligand gene), also regarded as CD162 and PSGL-1 [116], was generated from immune cells and affected inflammatory response [117] and myeloid cells differentiation [118]. Previous study suggested that low level of SELPLG were correlated to metastasis and a poor outcome in OS [119]. LGALS9 functions as an inhibitor of immune response [120, 121], and was regarded as a hallmark in breast cancer [122], pancreatic cancer [123] and pancreatic ductal adenocarcinoma [124]. Additionally, LGALS9 was a ligand of Tim-3 [125] which was encoded by HAVCR2 [126]. The interaction between TIM-3 and LGALS9 can negatively regulate tumor immune response [125]. Previous study suggested HAVCR2 was related to the expression of CD4 which might become a predictor of OS outcome [127]. Apparently, there existed a close interaction between LGALS9 and HAVCR2, making them potential therapy targets [126, 128–131]. Another checkpoint, LAIR1, was widely expressed in immune cells such as NK cells, B cells, T cells, Monocytes, etc [132],. LAIR1actedas an immune inhibitory receptor to suppress T cell activation [133], monocyte differentiation [134], M1 macrophage polarization [135] and other immune reactions. In OS, High expressed LAIR1 represses cell migration instead of proliferation [113]. Apparently, various cytokines generated from different cells in TME which were important for regulating cancer cell development, tumorigenesis, metastasis and mortality. The levels of checkpoints in high risk group

were down-regulated, and similarly, there were fewer immune cells in high risk OS group. According to existing researches, various immune checkpoints had potential role in suppressing cancer cells by regulating immune cells and immune reactions. Immune therapy originating from immune checkpoints was receiving increasing attention and might have a satisfactory prospect, just like the application of PD-1/PD-L1 inhibitors.

Through database analysis, more than 20 drugs targeting OS have been identified as inhibitors of enzymes or signaling pathways with potential effects. For example, BI.2536, an inhibitor of PLK1, suppressed cancer cell proliferation and induced pyroptosis [136, 137]. Also, BI.2536 had an effect on diabetic kidney disease via NF-κB and Smad3 pathway [138]. This shows its potential in diabetes related diseases and provides a new idea for clinical treatment Importantly, Daporinad (FK866), an inhibitor of nicotinamide-phosphate ribosyltransferase (NAMPT) that promoted apoptosis in tumor cells, has been tested in phase 2 clinical trials and had the potential to serve as a potential drug for OS [139]. As an IGF-1R inhibitor, linsitinib inhibited the proliferation of cancer cells and induced apoptosis [140, 141]. The drug has been evaluated in a number of clinical trials for the treatment of cancer [141–143](PMID: 39596005, PMID: 35197754, PMID: 33509009), indicating its potential clinical application in malignant tumors such as osteosarcoma. Additionally, NVP.ADW742, another IGF-1R inhibitor, regulated adrenocortical, prostate, and pulmonary malignancies [144–146]. It has been revealed sabutoclax (BI-97C1) had significant resistance effect on prostate cancer [147], pancreatic cancer [148], colorectal cancer [149] and breast cancer [150]. Also, sabutoclax can inhibit IL-6/STAT3 pathway and promoted apoptosis in cancer cells [150]. Lapatinib was an inhibitor of EGFR and HER2 [151]. Lapatinib was used in the treatment of breast cancer [152]. Vorinostat was an inhibitor of HDAC (histone deacetylase) which acted as an anticancer drug in various cancer cells, as it can cause apoptosis [153], autophagy of cancer [154] and cell cycle arrest [155], as well as regulate macrophage polarization [156]. The involvement of these drugs in diverse signaling pathways and their ability to modulate specific cancers implied that they function within the OS for a rational reason.

miRNAs as critical cytokines are involved in numerous regulation pathways. Through databases analysis, 5 miRNAs are selected which are hsa-miR-3163, hsa-miR-1179, hsa-miR-624-5p, hsa-miR-222-3p and hsa-miR-221-3p. Limited studies revealed that those miRNAs played important role in tumerogenesis, including breast cancer [157, 158], pancreatic ductal adenocarcinoma [159] and lung cancer [160]. And among of which, hsa-miR-624-5p increased in OS and might be a risk factor of metastasis [161]. And circRNA acts as sponges for miRNA to regulate functions of miRNA [162]. hsa_circ_0000652 can bound with hsa-miR-1179 in macrophages to regulate immune response [163]. RNA networks composed by miRNA, mRNA and lncRNA influence whole signal pathways. However, a large number of RNAs' functions are unclear, the exact functions need further exploration.

In OS samples, the interaction between osteoblastic cells and osteoclasts is more frequent, while myeloid cells and fibroblasts have more interactions in recurrent samples. Those different contents might result from different OS sample functions. Recurrent OS samples are more characterized by generation and proliferation. In primary sits, OS tissues give priority to their own functions. No matter in primary, recurrent and metastatic samples, it is clear that osteoblast cells are dominant. In the scRNA-seq analysis from GSE152048, 7 types of cells were identified, including osteoblast, osteoclast, myeloid

cells, etc. According to cell cluster studies of OS, Osteoblastic OS was one of the major subtypes of OS. The osteoblastic cell makers included*COL1A1*, *COL3A1A* and *RUNX2*. Compared to chondroblastic cells, osteoblastic OS has potential role in aggressive activities and metastasis of OS [164].The expression levels of *B3GALT4*, *CCDC115*, *DCN* and *RENBP* decreased in OS, while *UAP1*, *TUBA1A* and *B4GALNT1*were highly expressed in OS. Apparently, the risk model mentioned before was quite accurate. However, the trend of TUBA1A was different. This might be resulted from different tumor specimens acquired in different experimental conditions.

This study mainly relies on the analysis of public databases, and lacks experimental data for the functional validation of these genes in cell lines or animal models. In the future, it is necessary to explore how these genes affect the proliferation, migration, invasion and apoptosis of OS cells through gene knockout or overexpression experiments. Second, sample heterogeneity and insufficient mechanism research may also affect the results, especially the specific molecular mechanisms of these genes regulating OS progression through glycosylation has not been fully elucidated. In addition, the complexity of the immune microenvironment needs to be further analyzed in combination with single-cell sequencing technology. Moreover, circulating tumor cells (CTC) as an important biomarker for the assessment of metastasis potential of OS, its association with the seven prognostic GRGs screened in this study has not been explored. At the same time, PD-L1 is a key marker of immunotherapy for OS, and its association with the prognosis model of GRGs constructed in this study is also worth further exploration. Finally, this study adopts a retrospective, single center design, which may introduce sample selection and regional bias, limiting the generalizability of the conclusions. The independent validation results of the model further highlighted its limitations: in GSE21257, the model not only showed a decrease in long-term predictive performance (3/5-year AUC), but also failed to maintain independent prognostic value in risk scores, resulting in the inability to construct and validate column charts. These results may be due to technical differences between RNA seq and microarray platforms, interference from confounding factors such as treatment and recurrence in long-term survival outcomes, and limited sample size in the validation cohort, collectively suggesting that the model is sensitive to data platforms and cohort characteristics. Given the limitations mentioned above, future research should focus on functional validation of key genes, exploration of molecular mechanisms, and multi omics integration analysis. The model should be optimized and re-evaluated based on a larger scale, multi center, technically homogeneous prospective cohort, and ultimately promoted through clinical trials to transform it into an effective treatment strategy for OS.

5 Conclusions

On the basis of public databases and bioinformatics analysis, the prognostic significance of glycosylation-related genes, including DCN, RENBP, UAP1, B4GALNT1, CCDC115, TUBA1A, and B3GALT4, for OS was identified. In addition, immunoscape analysis, immune response analysis and drug prediction, provided a foundation for the treatment of OS. The results of this study not only have important value in prognosis evaluation, but also may provide a new direction for the diagnosis and treatment of osteosarcoma. The expression level of prognostic related genes can be used as potential biomarkers for early diagnosis and typing of osteosarcoma, and become a potential target for

osteosarcoma treatment. The dynamic changes of gene expression may be used to monitor the treatment effect and disease progression, and provide the basis for adjusting the treatment plan. Future studies should further verify the expression of these genes in clinical samples and their practical application value in treatment.

Abbreviations

OS	Osteosarcoma
GRGs	Glycosylation-related genes
scRNA-seq	Single-cell RNA sequencing
WGCNA	Weighted gene co-expression network analysis
LASSO	Least absolute shrinkage and selection operator
PH	Proportional hazards
MSCs	Mesenchymal stem cells
TME	Tumor microenvironment
EMT	Epithelial–mesenchymal transition
CEA	Carcinoembryonic antigen
GEO	Gene Expression Omnibus
DEGs	Differentially expressed genes
GO	Gene Ontology
KEGG	Kyoto Encyclopedia of Genes and Genomes
PPI	Protein–Protein Interaction
K–M	Kaplan–Meier
HR	Hazard ratio
ROC	Receiver operating characteristic
ssGSEA	Single sample gene set enrichment analysis
CCL	Cancer cell lines
GDSC2	Genomics of Drug Sensitivity in Cancer 2
IC_{50}	The half maximal inhibitory concentration
UMAP	Uniform manifold approximation and projection
ER	Endoplasmic reticulum
DCN	Decorin
BMSCs	Bone mesenchymal stem cells
CAFs	Cancer-associated fibroblasts
UAP1	UDP-N-acetylglucosaminepyrophosphorylase 1
B3GALT4	β-1,3-galactosyltransferase-4
B4GALNT1	β-1,4-N-Acetyl galactosaminyltransferase 1
UDP-GlcNAc	Uridine Diphospho-N-acetylglucosamine
HCC	Hepatocellular carcinoma
CCDC115	Coiled-coil domain-containing protein 115
LIHC	Liver hepatocellular carcinoma
TUBA1A	Tubulin alpha 1a
RENBP	Renin-binding protein
EMT	Epithelial mesenchymal transition
Tim-3	T-cell immunoglobulin domain and mucin domain-3
LAIR1	Leukocyte associated immunoglobulin like receptor 1, or CD305
SELPLG	Selectin P ligand gene
GGI	The Gene–gene Interaction

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s12672-026-04761-3.

Supplementary Material 1: Figure S1. Correlation analysis related to prognostic genes and immun. (a) Correlation between prognostic genes and immune cells. (b) Correlation between prognostic genes and immune responses. (c) Different expression of checkpoints between high and low risk group. (d) Correlation analysis between riskscore and checkpoints. (e) Correlation analysis between riskscore and immune responses. (*$p < 0.05$;**$p < 0.01$; ***$p < 0.001$; ****$p < 0.0001$; ns, $p > 0.05$).

Supplementary Material 2: Figure S2. Correlation analysis between riskscore and ICD10 of chemotherapeutic drugs.

Supplementary Material 3: Figure S3. Quality control and data processing of scRNA-seq analysis.

Supplementary Material 4: Figure S4. Proportional distribution chart of 7 cell types.

Supplementary Material 5: Figure S5. The interactions of different cells in tumor samples. (a) In primary samples. (b) In recurrent samples. (c) In metastatic samples.

Supplementary Material 6.

Supplementary Material 7.

Acknowledgements
Not applicable.

Author contributions
Ning Ding performed the research; Ning Ding, Songbo Shi and Qingshan Yang designed the research study; Yinliang Ding, Jinlong Li and Jizu Wang analysed the data; Ning Ding and Qingshan Yang wrote the paper.

Funding
This research was supported by Lanzhou Science and Technology Development Plan Project (2024-9-23).

Data availability
The dataset(s) supporting the conclusions of this article are available in the GEO at [https://www.ncbi.nlm.nih.gov/geo/](https:/www.ncbi.nlm.nih.gov/geo) and UCSC Xena at [http://xena.ucsc.edu/](http:/xena.ucsc.edu). Cancer cell lines (CCL) from database at https://www.cancerrxgene.org/.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 11 June 2025 / Accepted: 2 March 2026
Published online: 12 March 2026

References

1. Arndt CA, Rose PS, Folpe AL, Laack NN. Common musculoskeletal tumors of childhood and adolescence. Mayo Clin Proc. 2012;87(5):475–87.
2. Ritter J, Bielack SS. Osteosarcoma. Ann Oncol. 2010;21(Suppl 7):vii320-5.
3. Wang S, Ren Q, Li G, Zhao X, Zhao X, Zhang Z. The targeted therapies for osteosarcoma via six major pathways. Curr Mol Pharmacol. 2024;17:e210823220109.
4. Mohseny AB, Szuhai K, Romeo S, Buddingh EP, de Briaire- Bruijn I, de Jong D, et al. Osteosarcoma originates from mesenchymal stem cells in consequence of aneuploidization and genomic loss of Cdkn2. J Pathol. 2009;219(3):294–305.
5. Yan GN, Lv YF, Guo QN. Advances in osteosarcoma stem cell research and opportunities for novel therapeutic targets. Cancer Lett. 2016;370(2):268–74.
6. Ock CY, Hwang JE, Keam B, Kim SB, Shim JJ, Jang HJ, et al. Genomic landscape associated with potential response to anti-CTLA-4 treatment in cancers. Nat Commun. 2017;8(1):1050.
7. Wu CC, Livingston JA. Genomics and the immune landscape of osteosarcoma. Adv Exp Med Biol. 2020;1258:21–36.
8. Xu GS, Lin YN, Zeng Q, Li ZP, Xiao T, Ye YS, et al. HSP90-regulated CHIP/TRIM21/p21 axis involves in the senescence of osteosarcoma cells. Protein Pept Lett. 2023;30(6):513–9.
9. Wang M, Li W, Han G, Bai X, Xie J. MicroRNA-605-3p inhibited the growth and chemoresistance of osteosarcoma cells via negatively modulating RAF1. Protein Pept Lett. 2024;31(7):559–68.
10. Taniguchi N, Kizuka Y. Glycans and cancer: role of N-glycans in cancer biomarker, progression and metastasis, and therapeutics. Adv Cancer Res. 2015;126:11–51.
11. Eichler J. Protein glycosylation. Curr Biol. 2019;29(7):R229–31.
12. Chen CY, Jan YH, Juan YH, Yang CJ, Huang MS, Yu CJ, et al. Fucosyltransferase 8 as a functional regulator of nonsmall cell lung cancer. Proc Natl Acad Sci USA. 2013;110(2):630–5.
13. Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009;119(6):1420–8.
14. Li N, Xu H, Fan K, Liu X, Qi J, Zhao C, et al. Altered β1,6-GlcNAc branched N-glycans impair TGF-β-mediated epithelial-to-mesenchymal transition through Smad signalling pathway in human lung cancer. J Cell Mol Med. 2014;18(10):1975–91.
15. Pinho SS, Oliveira P, Cabral J, Carvalho S, Huntsman D, Gärtner F, et al. Loss and recovery of Mgat3 and GnT-III mediated E-cadherin N-glycosylation is a mechanism involved in epithelial-mesenchymal-epithelial transitions. PLoS One. 2012;7(3):e33191.
16. Pinho SS, Reis CA. Glycosylation in cancer: mechanisms and clinical implications. Nat Rev Cancer. 2015;15(9):540–55.
17. Hakomori S. Glycosylation defining cancer malignancy: new wine in an old bottle. Proc Natl Acad Sci U S A. 2002;99(16):10231–3.
18. Kudelka MR, Ju T, Heimburg-Molinaro J, Cummings RD. Simple sugars to complex disease–mucin-type O-glycans in cancer. Adv Cancer Res. 2015;126:53–135.
19. Mallard BW, Tiralongo J. Cancer stem cell marker glycosylation: nature, function and significance. Glycoconj J. 2017;34(4):441–52.
20. Bos PD, Zhang XH, Nadal C, Shu W, Gomis RR, Nguyen DX, et al. Genes that mediate breast cancer metastasis to the brain. Nature. 2009;459(7249):1005–9.
21. Kölbl AC, Andergassen U, Jeschke U. The role of glycosylation in breast cancer metastasis and cancer control. Front Oncol. 2015;5:219.
22. Duffy MJ. Carcinoembryonic antigen as a marker for colorectal cancer: is it clinically useful? Clin Chem. 2001;47(4):624–30.
23. Goldstein MJ, Mitchell EP. Carcinoembryonic antigen in the staging and follow-up of patients with colorectal cancer. Cancer Invest. 2005;23(4):338–51.

24. Gao Y, Luan X, Melamed J, Brockhausen I. Role of glycans on key cell surface receptors that regulate cell proliferation and cell death. Cells. 2021. https://doi.org/10.3390/cells10051252.
25. Pinho SS, Seruca R, Gärtner F, Yamaguchi Y, Gu J, Taniguchi N, et al. Modulation of E-cadherin function and dysfunction by N-glycosylation. Cell Mol Life Sci. 2011;68(6):1011–20.
26. Pinho SS, Osório H, Nita-Lazar M, Gomes J, Lopes C, Gärtner F, et al. Role of E-cadherin N-glycosylation profile in a mammary tumor model. Biochem Biophys Res Commun. 2009;379(4):1091–6.
27. Paredes J, Figueiredo J, Albergaria A, Oliveira P, Carvalho J, Ribeiro AS, et al. Epithelial E- and P-cadherins: role and clinical significance in cancer. Biochim Biophys Acta. 2012;1826(2):297–311.
28. Drake RR, Jones EE, Powers TW, Nyalwidhe JO. Altered glycosylation in prostate cancer. Adv Cancer Res. 2015;126:345–82.
29. Gilgunn S, Conroy PJ, Saldova R, Rudd PM, O'Kennedy RJ. Aberrant PSA glycosylation–a sweet predictor of prostate cancer. Nat Rev Urol. 2013;10(2):99–107.
30. Zhang L, Lv B, Shi X, Gao G. High expression of N-acetylgalactosaminyl-transferase 1 (GALNT1) associated with invasion, metastasis, and proliferation in osteosarcoma. Med Sci Monit. 2020;26:e927837.
31. Lin S, Zhou L, Dong Y, Yang Q, Yang Q, Jin H, et al. Alpha-(1,6)-fucosyltransferase (FUT8) affects the survival strategy of osteosarcoma by remodeling TNF/NF-κB2 signaling. Cell Death Dis. 2021;12(12):1124.
32. Packer NH, von der Lieth CW, Aoki-Kinoshita KF, Lebrilla CB, Paulson JC, Raman R, et al. Frontiers in glycomics: bioinformatics and biomarkers in disease. An NIH white paper prepared from discussions by the focus groups at a workshop on the NIH campus, Bethesda MD (September 11–13, 2006). Proteomics. 2008;8(1):8–20.
33. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. Nat Methods. 2009;6(5):377–82.
34. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretsky I, et al. Massively parallel single-cell RNA-seq for marker-free decomposition of tissues into cell types. Science. 2014;343(6172):776–9.
35. Hashimshony T, Wagner F, Sher N, Yanai I. CEL-Seq: single-cell RNA-Seq by multiplexed linear amplification. Cell Rep. 2012;2(3):666–73.
36. Long C, Ma B, Zhong X, Zou M, Li K, Liu S. Comprehensive multi-omics analysis reveals the prognostic and immune regulatory characteristics of the PTPN family in osteosarcoma. PLoS One. 2025;20(6):e0326872.
37. Wang Q, Huang M, Guo JJ. From cells to clinic: single-cell transcriptomics shaping the future of orthopedics. J Orthop Translat. 2025;53:1–11.
38. Kresse SH, Rydbeck H, Skårn M, Namløs HM, Barragan-Polania AH, Cleton-Jansen AM, et al. Integrative analysis reveals relationships of genetic and epigenetic alterations in osteosarcoma. PLoS One. 2012;7(11):e48262.
39. Chen W, Yu X, Li H, Yuan S, Fu Y, Hu H, et al. Single-cell RNA-seq reveals MIF-(CD74 + CXCR4) dependent inhibition of macrophages in metastatic papillary thyroid carcinoma. Oral Oncol. 2024;148:106654.
40. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43(7):e47.
41. Cao T, Li Q, Huang Y, Li A. plotnineSeqSuite: a Python package for visualizing sequence data using ggplot2 style. BMC Genomics. 2023;24(1):585.
42. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847–9.
43. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
44. Chen H, Boutros PC. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinformatics. 2011;12:35.
45. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation (Camb). 2021;2(3):100141.
46. Zhou J, Xiong W, Wang Y, Guan J. Protein function prediction based on PPI networks: network reconstruction vs edge enrichment. Front Genet. 2021;12:758131.
47. Bommert A, Welchowski T, Schmid M, Rahnenführer J. Benchmark of filter methods for feature selection in high-dimensional gene expression survival data. Brief Bioinform. 2022. https://doi.org/10.1093/bib/bbab354.
48. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J Stat Softw. 2010;33(1):1–22.
49. Kim J. Drawing guideline for JKMS manuscript (01) Kaplan-Meier curve and survival analysis. J Korean Med Sci. 2019;34(8):e35.
50. Song Y, Xiang Z, Lu Z, Su R, Shu W, Sui M, et al. Identification of a brand intratumor microbiome signature for predicting prognosis of hepatocellular carcinoma. J Cancer Res Clin Oncol. 2023;149(13):11319–32.
51. Deryabin PI, Shatrova AN, Borodkina AV. Apoptosis resistance of senescent cells is an intrinsic barrier for senolysis induced by cardiac glycosides. Cell Mol Life Sci. 2021;78(23):7757–76.
52. Zhang C, Zheng Y, Li X, Hu X, Qi F, Luo J. Genome-wide mutation profiling and related risk signature for prognosis of papillary renal cell carcinoma. Ann Transl Med. 2019;7(18):427.
53. Díez-Villanueva A, Sanz-Pamplona R, Solé X, Cordero D, Crous-Bou M, Guinó E, et al. COLONOMICS - integrative omics data of one hundred paired normal-tumoral samples from colon cancer patients. Sci Data. 2022;9(1):595.
54. Hänzelmann S, Castelo R, Guinney J. GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics. 2013;14:7.
55. Wu J, Li L, Zhang H, Zhao Y, Zhang H, Wu S, et al. A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients with lung adenocarcinoma. Oncogene. 2021;40(26):4413–24.
56. Liu Z, Tang Q, Qi T, Othmane B, Yang Z, Chen J, et al. A robust hypoxia risk score predicts the clinical outcomes and tumor microenvironment immune characters in bladder cancer. Front Immunol. 2021;12:725223.
57. Maeser D, Gruener RF, Huang RS. oncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. Brief Bioinform. 2021. https://doi.org/10.1093/bib/bbab260.
58. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: network analysis and visualization of proteomics data. J Proteome Res. 2019;18(2):623–32.
59. Hao Y, Hao S, Andersen-Nissen E, Mauck WM, Zheng S, Butler A, et al. Integrated analysis of multimodal single-cell data. Cell. 2021;184(13):3573-e8729.

60. Huang Q, Liu Y, Du Y, Garmire LX. Evaluation of cell type annotation R packages on single-cell RNA-seq data. Genomics Proteomics Bioinformatics. 2021;19(2):267–81.
61. Liao Z, Li M, Wen G, Wang K, Yao D, Chen E, et al. Comprehensive analysis of angiogenesis pattern and related immune landscape for individual treatment in osteosarcoma. NPJ Precis Oncol. 2023;7(1):62.
62. Griss J, Viteri G, Sidiropoulos K, Nguyen V, Fabregat A, Hermjakob H. ReactomeGSA - efficient multi-omics comparative pathway analysis. Mol Cell Proteom. 2020;19(12):2115–25.
63. Luo J, Deng M, Zhang X, Sun X. ESICCC as a systematic computational framework for evaluation, selection, and integration of cell-cell communication inference methods. Genome Res. 2023;33(10):1788–805.
64. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, et al. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods. 2017;14(10):979–82.
65. Vesely MD, Kershaw MH, Schreiber RD, Smyth MJ. Natural innate and adaptive immunity to cancer. Annu Rev Immunol. 2011;29:235–71.
66. Ding L, Sun R, Yan Q, Wang C, Han X, Cui Y, et al. MiR-506 exerts antineoplastic effects on osteosarcoma cells via inhibition of the Skp2 oncoprotein. Aging Albany NY. 2021;13(5):6724–39.
67. Wang B, Hong Z, Zhao C, Bi Q, Yuan J, Chen J, et al. The effects of MEX3A knockdown on proliferation, apoptosis and migration of osteosarcoma cells. Cancer Cell Int. 2021;21(1):197.
68. Huang G, Hua S, Liu H, Zhou H, Chen X, Wang Z, et al. Efficacy of ifosfamide combined with liposome doxorubicin on osteosarcoma and its effects on serum IL-10, TNF-α, and IFN-γ in patients with osteosarcoma. Am J Transl Res. 2022;14(2):1288–96.
69. Cadena AP, Cushman TR, Welsh JW. Glycosylation and antitumor immunity. Int Rev Cell Mol Biol. 2019;343:111–27.
70. Yan Y, Xu Z, Qian L, Zeng S, Zhou Y, Chen X, et al. Identification of CAV1 and DCN as potential predictive biomarkers for lung adenocarcinoma. Am J Physiol Lung Cell Mol Physiol. 2019;316(4):L630-l43.
71. Breitling J, Aebi M. N-linked protein glycosylation in the endoplasmic reticulum. Cold Spring Harb Perspect Biol. 2013;5(8):a013359.
72. Qin JZ, Xin H, Nickoloff BJ. 2-deoxyglucose sensitizes melanoma cells to TRAIL-induced apoptosis which is reduced by mannose. Biochem Biophys Res Commun. 2010;401(2):293–9.
73. Braakman I, Hebert DN. Protein folding in the endoplasmic reticulum. Cold Spring Harb Perspect Biol. 2013;5(5):a013201.
74. Stanley P. Golgi glycosylation. Cold Spring Harb Perspect Biol. 2011;3(4):a005199.
75. Vagin O, Kraut JA, Sachs G. Role of N-glycosylation in trafficking of apical membrane proteins in epithelia. Am J Physiol Ren Physiol. 2009;296(3):F459–69.
76. Esmail S, Yao Y, Kartner N, Li J, Reithmeier RA, Manolson MF. N-linked glycosylation is required for vacuolar H(+) -ATPase (V-ATPase) a4 subunit stability, assembly, and cell surface expression. J Cell Biochem. 2016;117(12):2757–68.
77. Lee HS, Qi Y, Im W. Effects of N-glycosylation on protein conformation and dynamics: protein data bank analysis and molecular dynamics simulation study. Sci Rep. 2015;5:8926.
78. Crean SM, Meneski JP, Hullinger TG, Reilly MJ, DeBoever EH, Taichman RS. N-linked sialyated sugar receptors support haematopoietic cell-osteoblast adhesions. Br J Haematol. 2004;124(4):534–46.
79. Ryšlavá H, Doubnerová V, Kavan D, Vaněk O. Effect of posttranslational modifications on enzyme function and assembly. J Proteome. 2013;92:80–109.
80. Mikolajczyk K, Kaczmarek R, Czerwinski M. How glycosylation affects glycosylation: the role of N-glycans in glycosyltransferase activity. Glycobiology. 2020;30(12):941–69.
81. Vasconcelos-Dos-Santos A, Oliveira IA, Lucena MC, Mantuano NR, Whelan SA, Dias WB, et al. Biosynthetic Machinery Involved in Aberrant Glycosylation: Promising Targets for Developing of Drugs Against Cancer. Front Oncol. 2015;5:138.
82. Li D, Li Y, Wu X, Li Q, Yu J, Gen J, et al. Knockdown of Mgat5 inhibits breast cancer cell growth with activation of CD4 + T cells and macrophages. J Immunol. 2008;180(5):3158–65.
83. Fernandes B, Sagman U, Auger M, Demetrio M, Dennis JW. Beta 1–6 branched oligosaccharides as a marker of tumor progression in human breast and colon neoplasia. Cancer Res. 1991;51(2):718–23.
84. Lange T, Ullrich S, Müller I, Nentwich MF, Stübke K, Feldhaus S, et al. Human prostate cancer in a clinically relevant xenograft mouse model: identification of β(1,6)-branched oligosaccharides as a marker of tumor progression. Clin Cancer Res. 2012;18(5):1364–73.
85. Yang S, Jin S, Xian H, Zhao Z, Wang L, Wu Y, et al. Metabolic enzyme UAP1 mediates IRF3 pyrophosphorylation to facilitate innate immune response. Mol Cell. 2023;83(2):298-313.e8.
86. Liu J, Zhu S, Zeng L, Li J, Klionsky DJ, Kroemer G, et al. DCN released from ferroptotic cells ignites AGER-dependent immune responses. Autophagy. 2022;18(9):2036–49.
87. Cortini M, Avnet S, Baldini N. Mesenchymal stroma: role in osteosarcoma progression. Cancer Lett. 2017;405:90–9.
88. Zhang T, Wang F, Wu JY, Qiu ZC, Wang Y, Liu F, et al. Clinical correlation of B7-H3 and B3GALT4 with the prognosis of colorectal cancer. World J Gastroenterol. 2018;24(31):3538–46.
89. Zhang M, Liu Y, Kong D. Identifying biomolecules and constructing a prognostic risk prediction model for recurrence in osteosarcoma. J Bone Oncol. 2021;26:100331.
90. Su C, Yang JC, Rong Z, Li F, Luo LX, Liu G, et al. Identification of CCDC115 as an adverse prognostic biomarker in liver cancer based on bioinformatics and experimental analyses. Heliyon. 2023;9(9):e19233.
91. Jansen JC, Cirak S, van Scherpenzeel M, Timal S, Reunert J, Rust S, et al. CCDC115 deficiency causes a disorder of Golgi homeostasis with abnormal protein glycosylation. Am J Hum Genet. 2016;98(2):310–21.
92. Winter JM, Curry NL, Gildea DM, Williams KA, Lee M, Hu Y, et al. Modifier locus mapping of a transgenic F2 mouse population identifies CCDC115 as a novel aggressive prostate cancer modifier gene in humans. BMC Genomics. 2018;19(1):450.
93. Girard M, Poujois A, Fabre M, Lacaille F, Debray D, Rio M, et al. CCDC115-CDG: A new rare and misleading inherited cause of liver disease. Mol Genet Metab. 2018;124(3):228–35.
94. Xie L, Huang J, Dai L, Luo J, Zhang J, Peng Q, et al. Loss-of-function plays a major role in early neurogenesis of tubulin α-1 A (TUBA1A) mutation-related brain malformations. Mol Neurobiol. 2021;58(4):1291–302.
95. Wex B, Safi RM, Antonios G, Zgheib PZ, Awad DB, Kobeissy FH, et al. SLC35B4, an inhibitor of gluconeogenesis, responds to glucose stimulation and downregulates Hsp60 among other proteins in HepG2 liver cell lines. Molecules. 2018. https://doi.org/10.3390/molecules23061350.

96. Wang D, Jiao Z, Ji Y, Zhang S. Elevated TUBA1A might indicate the clinical outcomes of patients with gastric cancer, being associated with the infiltration of macrophages in the tumor immune microenvironment. J Gastrointestin Liver Dis. 2020;29(4):509–22.
97. Liang X, Gong M, Wang Z, Wang J, Guo W, Cai A, et al. LncRNA TubAR complexes with TUBB4A and TUBA1A to promote microtubule assembly and maintain myelination. Cell Discovery. 2024;10(1):54.
98. Yu Y, Wang S, Liu J, Ge J, Guan H. Development and validation of a nomogram to predict long-term cancer-specific survival for patients with osteosarcoma. Sci Rep. 2023;13(1):10230.
99. Lu S, Wang Y, Liu G, Wang L, Wu P, Li Y, et al. Construction and validation of nomogram to predict distant metastasis in osteosarcoma: a retrospective study. J Orthop Surg Res. 2021;16(1):231.
100. Xue W, Zhang Z, Yu H, Li C, Sun Y, An J, et al. Development of nomogram and discussion of radiotherapy effect for osteosarcoma survival. Sci Rep. 2023;13(1):223.
101. Su Z, Wang C, Pan R, Li H, Chen J, Tan J, et al. The hexosamine biosynthesis pathway-related gene signature correlates with immune infiltration and predicts prognosis of patients with osteosarcoma. Front Immunol. 2022;13:1028263.
102. Wei G, Jia H, Zhang Z, Qin J, Ao J, Qian H. O-GlcNAcylation: sagacious orchestrator of bone-, joint-, and spine-related diseases. J Proteome Res. 2025;24(3):981–94.
103. Itkonen HM, Engedal N, Babaie E, Luhr M, Guldvik IJ, Minner S, et al. UAP1 is overexpressed in prostate cancer and is protective against inhibitors of N-linked glycosylation. Oncogene. 2015;34(28):3744–50.
104. Trinchera M, Parini R, Indellicato R, Domenighini R, dall'Olio F. Diseases of ganglioside biosynthesis: an expanding group of congenital disorders of glycosylation. Mol Genet Metab. 2018;124(4):230–7.
105. Liu W, Chen Y, Yang J, Guo M, Wang L. B4GALNT1 promotes carcinogenesis by regulating epithelial-mesenchymal transition in hepatocellular carcinoma based on pan-cancer analysis. J Gene Med. 2023;25(11):e3552.
106. Jiang T, Wu H, Lin M, Yin J, Tan L, Ruan Y, et al. B4GALNT1 promotes progression and metastasis in lung adenocarcinoma through JNK/c-Jun/Slug pathway. Carcinogenesis. 2021;42(4):621–30.
107. Yoshida H, Koodie L, Jacobsen K, Hanzawa K, Miyamoto Y, Yamamoto M. B4GALNT1 induces angiogenesis, anchorage independence growth and motility, and promotes tumorigenesis in melanoma by induction of ganglioside GM2/GD2. Sci Rep. 2020;10(1):1199.
108. Yamaguchi T, Yamauchi Y, Furukawa K, Ohmi Y, Ohkawa Y, Zhang Q, et al. Expression of B4GALNT1, an essential glycosyltransferase for the synthesis of complex gangliosides, suppresses BACE1 degradation and modulates APP processing. Sci Rep. 2016;6:34505.
109. Li TA, Schnaar RL. Congenital disorders of ganglioside biosynthesis. Prog Mol Biol Transl Sci. 2018;156:63–82.
110. Li S, Wang B, Chen X, Tu C, Peng D, Dai Y. Downregulation of B4GALNT1 inhibits proliferation and metastasis of osteosarcoma cells. Zhong Nan Da Xue Xue Bao Yi Xue Ban. 2024;49(6):870–7.
111. Brabletz S, Schuhwerk H, Brabletz T, Stemmler MP. Dynamic EMT: a multi-tool for tumor progression. EMBO J. 2021;40(18):e108647.
112. Feng ZM, Guo SM. Tim-3 facilitates osteosarcoma proliferation and metastasis through the NF-κB pathway and epithelial-mesenchymal transition. Genet Mol Res. 2016;15(3):1.
113. Zhang J, Zhang Y, Cheng S, Mu Y, Liu Y, Yi X, et al. LAIR-1 overexpression inhibits epithelial-mesenchymal transition in osteosarcoma via GLUT1-related energy metabolism. World J Surg Oncol. 2020;18(1):136.
114. Hinton K, Kirk A, Paul P, Persad S. Regulation of the epithelial to mesenchymal transition in osteosarcoma. Biomolecules. 2023. https://doi.org/10.3390/biom13020398.
115. Ma Q, Zhuo D, Guan F, Li X, Yang X, Tan Z. Vesicular ganglioside GM1 from breast tumor cells stimulated epithelial-to-mesenchymal transition of recipient MCF-10A cells. Front Oncol. 2022;12:837930.
116. Bime C, Pouladi N, Sammani S, Batai K, Casanova N, Zhou T, et al. Genome-wide association study in African Americans with acute respiratory distress syndrome identifies the selectin P ligand gene as a risk factor. Am J Respir Crit Care Med. 2018;197(11):1421–32.
117. Tinoco R, Otero DC, Takahashi AA, Bradley LM. PSGL-1: A New Player in the Immune Checkpoint Landscape. Trends Immunol. 2017;38(5):323–35.
118. Sultana DA, Zhang SL, Todd SP, Bhandoola A. Expression of functional P-selectin glycoprotein ligand 1 on hematopoietic progenitors is developmentally regulated. J Immunol. 2012;188(9):4385–93.
119. Wang B, Sun Y. SELPLG expression was potentially correlated with metastasis and prognosis of osteosarcoma. Pathol Oncol Res. 2022;28:1610047.
120. Zhu C, Anderson AC, Schubart A, Xiong H, Imitola J, Khoury SJ, et al. The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. Nat Immunol. 2005;6(12):1245–52.
121. Seki M, Oomizu S, Sakata KM, Sakata A, Arikawa T, Watanabe K, et al. Galectin-9 suppresses the generation of Th17, promotes the induction of regulatory T cells, and regulates experimental autoimmune arthritis. Clin Immunol. 2008;127(1):78–88.
122. Fang J, Chen F, Liu D, Gu F, Chen Z, Wang Y. Prognostic value of immune checkpoint molecules in breast cancer. Biosci Rep. 2020. https://doi.org/10.1042/BSR20201054.
123. Fan Y, Li T, Xu L, Kuang T. Comprehensive analysis of immunoinhibitors identifies LGALS9 and TGFBR1 as potential prognostic biomarkers for pancreatic cancer. Comput Math Methods Med. 2020;2020:6138039.
124. Seifert AM, Reiche C, Heiduk M, Tannert A, Meinecke AC, Baier S, et al. Detection of pancreatic ductal adenocarcinoma with galectin-9 serum levels. Oncogene. 2020;39(15):3102–13.
125. Nishino M, Ramaiya NH, Hatabu H, Hodi FS. Monitoring immune-checkpoint blockade: response evaluation and biomarker development. Nat Rev Clin Oncol. 2017;14(11):655–68.
126. Holderried TAW, de Vos L, Bawden EG, Vogt TJ, Dietrich J, Zarbl R, et al. Molecular and immune correlates of TIM-3 (HAVCR2) and galectin 9 (LGALS9) mRNA expression and DNA methylation in melanoma. Clin Epigenetics. 2019;11(1):161.
127. Song YJ, Xu Y, Deng C, Zhu X, Fu J, Chen H, et al. Gene expression classifier reveals prognostic osteosarcoma microenvironment molecular subtypes. Front Immunol. 2021;12:623762.
128. Lv Y, Ma X, Ma Y, Du Y, Feng J. A new emerging target in cancer immunotherapy: Galectin-9 (LGALS9). Genes Dis. 2023;10(6):2366–82.
129. Li H, Yang D, Hao M, Liu H. Differential expression of HAVCR2 gene in pan-cancer: a potential biomarker for survival and immunotherapy. Front Genet. 2022;13:972664.

130. Keir ME, Butte MJ, Freeman GJ, Sharpe AH. PD-1 and its ligands in tolerance and immunity. Annu Rev Immunol. 2008;26:677–704.
131. Qian C, Wu D, Du J. RIPK3 modulates sarcoma through immune checkpoint HAVCR2. Oncol Lett. 2022;24(5):381.
132. Van Laethem F, Donaty L, Tchernonog E, Lacheretz-Szablewski V, Russello J, Buthiau D, et al. LAIR1, an ITIM-containing receptor involved in immune disorders and in hematological neoplasms. Int J Mol Sci. 2022. https://doi.org/10.3390/ijms232416136.
133. Xie J, Gui X, Deng M, Chen H, Chen Y, Liu X, et al. Blocking LAIR1 signaling in immune cells inhibits tumor development. Front Immunol. 2022;13:996026.
134. Poggi A, Tomasello E, Ferrero E, Zocchi MR, Moretta L. P40/LAIR-1 regulates the differentiation of peripheral blood precursors to dendritic cells induced by granulocyte-monocyte colony-stimulating factor. Eur J Immunol. 1998;28(7):2086–91.
135. Jin J, Wang Y, Ma Q, Wang N, Guo W, Jin B, et al. LAIR-1 activation inhibits inflammatory macrophage phenotype in vitro. Cell Immunol. 2018;331:78–84.
136. Steegmaier M, Hoffmann M, Baum A, Lénárt P, Petronczki M, Krssák M, et al. Bi 2536, a potent and selective inhibitor of polo-like kinase 1, inhibits tumor growth in vivo. Curr Biol. 2007;17(4):316–22.
137. Huo J, Shen Y, Zhang Y, Shen L. BI 2536 induces gasdermin E-dependent pyroptosis in ovarian cancer. Front Oncol. 2022;12:963928.
138. Zhang L, Wang Z, Liu R, Li Z, Lin J, Wojciechowicz ML, et al. Connectivity mapping identifies BI-2536 as a potential drug to treat diabetic kidney disease. Diabetes. 2021;70(2):589–602.
139. Park M, Lee BI, Choi J, Park Y, Park SJ, Lim JH, et al. Quantitative analysis of daporinad (FK866) and its in vitro and in vivo metabolite identification using liquid chromatography-quadrupole-time-of-flight mass spectrometry. Molecules. 2022. https://doi.org/10.3390/molecules27062011.
140. von Mehren M, George S, Heinrich MC, Schuetze SM, Yap JT, Yu JQ, et al. Linsitinib (OSI-906) for the treatment of adult and pediatric wild-type gastrointestinal stromal tumors, a SARC phase II study. Clin Cancer Res. 2020;26(8):1837–45.
141. Guo Y, Mehrabi Nasab E, Hassanpour F, Athari SS. Linsitinib and aspirin as the IGF1-R antagonists, inhibit regorafenib-resistant chemotherapy in colon cancer. Saudi J Biol Sci. 2022;29(2):872–7.
142. Almeida-Nunes DL, Silva JPN, Nunes M, Silva PMA, Silvestre R, Dinis-Oliveira RJ, et al. Metformin impairs linsitinib anti-tumor effect on ovarian cancer cell lines. Int J Mol Sci. 2024. https://doi.org/10.3390/ijms252211935.
143. Khan S, LeBlanc R, Gyger M, White D, Kaufman J, Jazubowiak A, et al. A phase-1 trial of linsitinib (OSI-906) in combination with bortezomib and dexamethasone for the treatment of relapsed/refractory multiple myeloma. Leuk Lymphoma. 2021;62(7):1721–9.
144. Warshamana-Greene GS, Litz J, Buchdunger E, Hofmann F, García-Echeverría C, Krystal GW. The insulin-like growth factor-I (IGF-I) receptor kinase inhibitor NVP-ADW742, in combination with STI571, delineates a spectrum of dependence of small cell lung cancer on IGF-I and stem cell factor signaling. Mol Cancer Ther. 2004;3(5):527–35.
145. Ozkan EE. Plasma and tissue insulin-like growth factor-I receptor (IGF-IR) as a prognostic marker for prostate cancer and anti-IGF-IR agents as novel therapeutic strategy for refractory cases: a review. Mol Cell Endocrinol. 2011;344(1–2):1–24.
146. Catalano R, Giardino E, Treppiedi D, Mangili F, Morelli V, Elli FM, et al. The cytoskeleton actin binding protein filamin A impairs both IGF2 mitogenic effects and the efficacy of IGF1R inhibitors in adrenocortical cancer cells. Cancer Lett. 2021;497:77–88.
147. Dash R, Azab B, Quinn BA, Shen X, Wang XY, Das SK, et al. Apogossypol derivative BI-97C1 (Sabutoclax) targeting Mcl-1 sensitizes prostate cancer cells to mda-7/IL-24-mediated toxicity. Proc Natl Acad Sci U S A. 2011;108(21):8785–90.
148. Quinn BA, Dash R, Sarkar S, Azab B, Bhoopathi P, Das SK, et al. Pancreatic cancer combination therapy using a BH3 mimetic and a synthetic tetracycline. Cancer Res. 2015;75(11):2305–15.
149. Azab B, Dash R, Das SK, Bhutia SK, Shen XN, Quinn BA, et al. Enhanced delivery of mda-7/IL-24 using a serotype chimeric adenovirus (Ad.5/3) in combination with the Apogossypol derivative BI-97C1 (Sabutoclax) improves therapeutic efficacy in low CAR colorectal cancer cells. J Cell Physiol. 2012;227(5):2145–53.
150. Hu Y, Yagüe E, Zhao J, Wang L, Bai J, Yang Q, et al. Sabutoclax, pan-active BCL-2 protein family antagonist, overcomes drug resistance and eliminates cancer stem cells in breast cancer. Cancer Lett. 2018;423:47–59.
151. Maennling AE, Tur MK, Niebert M, Klockenbring T, Zeppernick F, Gattenlöhner S, et al. Molecular targeting therapy against EGFR family in breast cancer: progress and future potentials. Cancers (Basel). 2019. https://doi.org/10.3390/cancers11121826.
152. Higa GM, Abraham J. Lapatinib in the treatment of breast cancer. Expert Rev Anticancer Ther. 2007;7(9):1183–92.
153. Richon VM. Cancer biology: mechanism of antitumour action of vorinostat (suberoylanilide hydroxamic acid), a novel histone deacetylase inhibitor. Br J Cancer. 2006. https://doi.org/10.1038/sj.bjc.6603463.
154. Patra S, Praharaj PP, Klionsky DJ, Bhutia SK. Vorinostat in autophagic cell death: a critical insight into autophagy-mediated, -associated and -dependent cell death for cancer prevention. Drug Discov Today. 2022;27(1):269–79.
155. Alqosaibi AI, Abdel-Ghany S. Vorinostat induces G2/M cell cycle arrest in breast cancer cells via upregulation of PTEN. Eur Rev Med Pharmacol Sci. 2023;27(4):1503–11.
156. Hsieh TH, Hsu CY, Wu CW, Wang SH, Yeh CH, Cheng KH, et al. Vorinostat decrease M2 macrophage polarization through ARID1A(6488delG)/HDAC6/IL-10 signaling pathway in endometriosis-associated ovarian carcinoma. Biomed Pharmacother. 2023;161:114500.
157. Qiu P, Guo Q, Yao Q, Chen J, Lin J. Hsa-mir-3163 and CCNB1 may be potential biomarkers and therapeutic targets for androgen receptor positive triple-negative breast cancer. PLoS One. 2021;16(11):e0254283.
158. Shao X, Zheng Y, Huang Y, Li G, Zou W, Shi L. Hsa-miR-221-3p promotes proliferation and migration in HER2-positive breast cancer cells by targeting LASS2 and MBD2. Histol Histopathol. 2022;37(11):1099–112.
159. Xin W, Fu Y, Chen L, Ding H, Feng T, Sun J, et al. Identification and validation of the microRNAs and hub genes for pancreatic ductal adenocarcinoma by an integrated bioinformatic analysis. J Gastrointest Oncol. 2023;14(2):719–32.
160. Tepebaşı MY, Öztürk Ö. MiR-21, miR-221, and miR-222 upregulation in lung cancer promotes metastasis by reducing oxidative stress and apoptosis. Rev Assoc Méd Brasil. 2023;69(6): e20221688.
161. Luo Y, Liu W, Tang P, Jiang D, Gu C, Huang Y, et al. MiR-624-5p promoted tumorigenesis and metastasis by suppressing hippo signaling through targeting PTPRB in osteosarcoma cells. J Exp Clin Cancer Res. 2019;38(1):488.
162. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013;495(7441):333–8.

163. Song M, Gao J, Yan T, Bi E, An T, Wang X, et al. Hsa_circ_0000652 aggravates inflammation by activation of macrophages and enhancement of OX40/OX40L interaction in ankylosing spondylitis. Front Cell Dev Biol. 2021;9:737599.
164. Zhou Y, Yang D, Yang Q, Lv X, Huang W, Zhou Z, et al. Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma. Nat Commun. 2020;11(1):6322.

Publisher's Note