Genes & Diseases (2024) 11, 101004

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/en/journals/genes-diseases

FULL LENGTH ARTICLE

Integrative transcriptome analysis identifies MYBL2 as a poor prognosis marker for osteosarcoma and a pan-cancer marker of immune infiltration

Xinzhu Qiu [a,b,c], Hongbo He [a,b], Hao Zeng [a,b], Xiaopeng Tong [a,b,c], Can Zhang [a,b], Yupeng Liu [a,b], Zhan Liao [a,b], Qing Liu [a,b,*]

[a] *Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China*
[b] *National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Changsha, Hunan 410008, China*
[c] *Department of Sports Medicine, Research Center of Sports Medicine, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China*

Received 30 November 2022; received in revised form 23 March 2023; accepted 29 April 2023
Available online 11 July 2023

KEYWORDS
Biomarker;
Immune checkpoint;
MYBL2;
Osteosarcoma;
Pan-cancer

Abstract MYBL2 (MYB proto-oncogene like 2) is an emerging prognostic marker for malignant tumors, and its potential role in osteosarcoma and its relationship with immune infiltration in pan-cancer is yet to be elucidated. We constructed a transcription factor activity profile of osteosarcoma using the single-cell regulatory network inference algorithm based on single-cell RNA sequencing data obtained from the Gene Expression Omnibus. Subsequently, we calculated the extent of MYBL2 activation in malignant proliferative osteoblasts. We also explored the association between MYBL2 and chemotherapy resistance in osteosarcoma. Furthermore, we systematically correlated MYBL2 with immunological signatures in the tumor microenvironment in pan-cancer, including immune cell infiltration, immune checkpoints, and tumor immunotherapy prognosis. Finally, we developed and validated a risk score (MRGS), derived an osteosarcoma risk score nomogram based on MRGS, and tested its ability to predict prognosis. MYBL2 and gene enrichment analyses in osteosarcoma and pan-cancer revealed that MYBL2 was positively correlated with cell proliferation and tumor immune pathways. MYBL2 expression positively correlated with SLC19A1 in pan-cancer and osteosarcoma cell lines. Pan-cancer immune infiltration analysis revealed that MYBL2 was correlated with myeloid-derived suppressor cells, Th2 cell infiltration, CD276, RELT gene expression, and tumor mutation burden.

* Corresponding author. Department of Orthopaedics, Xiangya Hospital, Central South University, Changsha, Hunan 410008, China.
E-mail address: 158112225@csu.edu.cn (Q. Liu).
Peer review under responsibility of Chongqing Medical University.

https://doi.org/10.1016/j.gendis.2023.04.035

In summary, MYBL2 regulates proliferation, progression, and immune infiltration in osteosarcoma and pan-cancer. Therefore, we found that MYBL2 could be used as a potential marker for predicting the osteosarcoma prognosis. Patients with osteosarcoma and high MYBL2 expression are theoretically more sensitive to methotrexate. An osteosarcoma prognostic nomogram can provide new ideas in the search for osteosarcoma prognostic markers.

Introduction

Osteosarcoma is the most common malignant primary bone tumor in children and adolescents. It primarily occurs in the metaphysis and diaphysis of long bones and is a malignant tumor characterized by forming an osteoid matrix and immature bone.[1] Approximately 15%—20% of patients with osteosarcoma have metastatic lesions on their first visit. More than 85% of metastatic lesions are located in the lungs, the most common site of distant metastasis from osteosarcoma.[2] Although current therapies can treat approximately 70% of patients with osteosarcoma, survival rates for patients with metastatic or recurrent lesions have remained largely unchanged over the past 30 years.[3]

Although the exact cellular origin of osteosarcoma is unknown, evidence suggests that it arises from mesenchymal stem cells or osteogenic precursor cells that cannot undergo terminal differentiation.[4] In these cells, copy number variations and chromosomal fusion sequences occur at high frequencies. These complex genetic scenarios have hindered the development of targeted therapeutic strategies.[5] Therefore, it is crucial to deepen our understanding of osteosarcoma. Tumor occurrence is a multifactor and multistep process. Osteosarcoma is aggressive, and the prognosis of patients with metastases is poor. The effects of molecular-targeted therapy in patients with osteosarcoma are also quite different.[6]

The MYB family plays an important role in regulating cell proliferation, differentiation, and angiogenesis, and is expressed in many tumors. Recently, researchers have found that MYB is associated with immunosuppression, and MYB, as a transcription factor, can down-regulate the effector ability of T cells and maintain the self-renewal ability of T exhausted cells. MYBL2 is a member of the MYB transcription factor family that plays an important role in cell cycle progression, survival, and differentiation.[7] MYBL2 is expressed in almost all proliferating cells.[8] MYBL2-knockout mice develop early embryonic death due to blocked colony formation within blastocysts.[8] Knockdown of MYBL2 at the cellular level can lead to abnormalities in the cell cycle's S, G2, and M phases, resulting in cell cycle arrest at the G2 phase.[9] Studies have demonstrated that MYBL2 is overexpressed in various tumors, including gallbladder,[10] colorectal,[11] bladder,[12] cervical,[13] lung,[14] and breast cancers.[15]

MYBL2 mRNA expression levels positively correlate with MKI67 expression levels in breast cancer.[15] Jin et al found that MYBL2 overexpression can promote the proliferation of non-small-cell lung cancer through the ERK signaling and Akt signaling pathways.[14] Liang et al found that MYBL2 improves the proliferation of gallbladder cancer cells by promoting the progression of the cell cycle S phase and G2/M phase in gallbladder cancer and suggested that MYBL2 is a factor for poor prognosis in patients with gallbladder cancer.[10] High MYBL2 expression is also associated with poor prognosis in neurofibroblastoma,[16] breast cancer,[17] and colorectal cancer.[11] Recently, Meng et al performed a pan-cancer analysis of MYBL2 in bladder cancer and confirmed that MYBL2 plays an immunomodulatory role in bladder cancer and may serve as an important prognostic biomarker.[18] However, the prognostic role of MYBL2 in osteosarcoma and pan-cancer is yet to be fully validated.

With the progress in sequencing technology, single-cell transcriptome technology has become necessary.[19] Single-cell sequencing is a technique for analyzing tissue transcription at the single-cell level, which can reveal changes in gene expression in every cell in the tissue and may even reveal new cell types.[20] Bioinformatic analysis of single-cell transcriptomes can more accurately demonstrate the cell state and gene expression regulation.[21]

Here, we first used osteosarcoma single-cell RNA-seq data to identify the transcription factor MYBL2 specifically expressed in proliferating osteosarcoma cells. Second, we validated the MYBL2 expression and function in osteosarcoma and pan-cancer transcriptome data and demonstrated that MYBL2 positively correlated with MDSC (myeloid-derived suppressor cell) and Th2 cells in pan-cancer. Finally, a clinical prediction model was developed to predict the prognosis of osteosarcoma based on MYBL2-related genes (Fig. 1A—D). This study provides valuable basic information for studying osteosarcoma biomarkers and prognostic factors with clinical relevance.

Materials and methods

Data sources

The sc-RNA-seq cohort GSE152048 was processed using the Seurat package (version 3.2.2) in R statistic software (version 4.0.3). We selected cells with 300—4000 genes and mitochondrial genes $<$ 10% of the total expressed genes. Subsequently, we removed the doublets using the DoubletFinder package in R. Eleven Seurat objects were merged. After normalization and removal of the batch effect using the Harmony package of R, we performed clustering analysis using the FindClusters function of the Seurat package. We identified differentially expressed genes

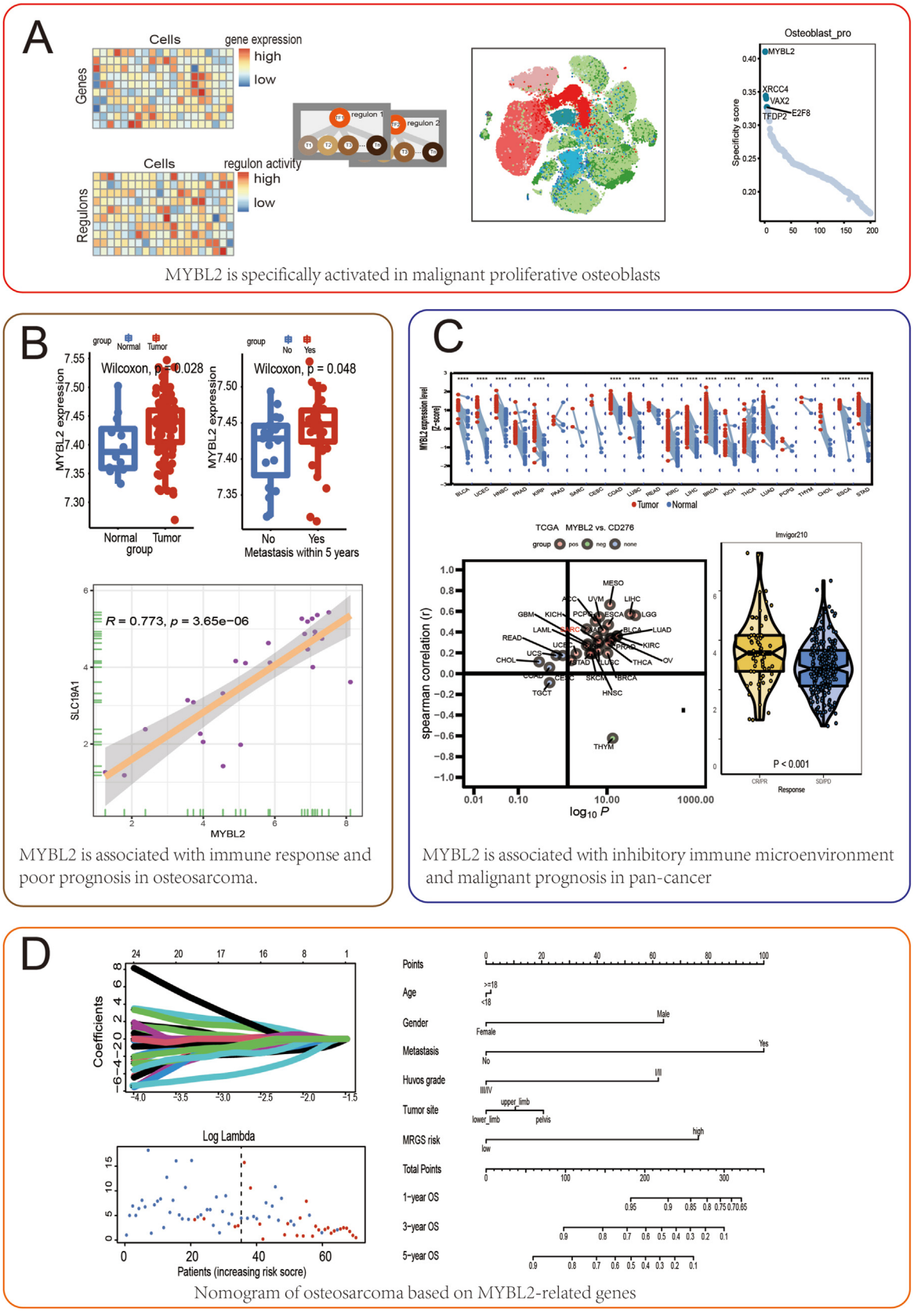

Figure 1 Overview of the study design. **(A)** MYBL2 was found to be a specific regulon in malignant osteoblasts through SCENIC. **(B)** MYBL2's expression level, prognostic effect, and function analysis in osteosarcoma. **(C)** MYBL2's expression level, prognostic effect, and function analysis in pan-cancer. **(D)** The establishment and validation of a nomogram for predicting overall survival based on MRGS.

(DEGs) using the FindAllMarkers function in the Seurat package to annotate the cell clusters. The cell groups were annotated based on the DEGs and well-known cellular markers downloaded from the original study.[3] The following three osteosarcoma gene expression omnibus (GEO) datasets with clinical data were downloaded: GSE87624 (GPL11154), GSE42352 (GPL10295), and GSE21257 (GPL10295).

The normalized RNA-seq data (FPKM: Fragments Per Kilobase of exon model per Million mapped fragments) in GSE87624 was transformed into TPM (Transcripts Per Kilobase of exon model per Million mapped reads). GSE42352 and GSE21257 are microarray datasets. Based on the GPL10295 platform information, the probe name was annotated as the gene name. Therapeutically applicable research to generate effective treatments (TARGET) RNA-seq and methylation cohorts were downloaded from the TARGET website (https://ocg.cancer.gov/programs/target). After filtering the data without clinical information, we obtained RNA-seq data from the 94 pre-chemotherapy biopsy samples. We converted the Ensembl gene ID into a gene symbol. Subsequently, the count was converted to log2 (TPM +1) to show the amount of gene expression. Pan-cancer RNA sequencing (RNA-seq) data (FPKM value), pan-tissue RNA sequencing (RNA-seq) data (FPKM value), survival information, and somatic mutation data for cancer were downloaded from the USCS Xena data portal.[22] Next, we annotated the gene names and transformed the gene expression from FPKM to log2 (TPM +1). VarScan2 was used to analyze somatic mutation data and calculate the tumor mutation burden (TMB). The osteosarcoma cell line data were downloaded from the CCLE (Cancer Cell Line Encyclopedia) database. We selected tumor cell lines of bone tissue origin, downloaded the gene expression matrix, and converted the expression level from FPKM into TPM.

Single-cell regulatory network inference (SCENIC)

We used the Python version of SCENIC pySCENIC to construct GRNs (gene regulatory networks) from single-cell RNA-seq data in Python 3.6.13. The SCENIC workflow contains three steps as follows: (i) determine the co-expression module between the transcription factor and the potential target gene, (ii) infer the potential target gene for each co-expression module based on the motif information of the corresponding transcription factor, and (iii) calculate the transcription factor activity matrix.

Calculating cell-type specificity score and validation

We used the method described by Shengbao[23] to determine cell-specific transcription factors. For quantification, an entropy-based method was adopted to calculate cell-type specifications.[23] The distribution of RAS in the cell population was represented as a vector $P^R=(p_1^R, p_2^R, \cdots p_n^R)$ (n is the number of all cells). Subsequently, we normalized the RAS such that $\sum_{i=1}^{n} p_i^R = 1$. A vector $P^C=(p_1^C, p_2^C, \cdots p_n^C)$ was created to show whether a cell belongs to a specific cell-type ($p_i^C = 1$) or not ($p_i^C = 0$). A similar normalization was performed such that $\sum_{i=1}^{n} p_i^C = 1$. Next, we quantified the Jensen–Shannon divergence (JSD), defined as $JSD(P^R, P^R) = H\left(\frac{P^R+P^C}{2}\right) - \frac{H(P^R)+H(P^C)}{2}$, where $H(P) = -\sum p_i \, log(p_i)$ represents the Shannon entropy of a probability distribution P. JSD ranges from 0 to 1, where 0 indicates the same distribution and one means an extreme difference. Lastly, the regulon specificity score (RSS) was defined as $RSS(R, C) = 1 - \sqrt{JSD(P^R, P^C)}$. Specific regulators were predicted based on their association with cell type-specific scores. For validation, we used the website Search-Based Exploration of Expression Compendium: SEEK (https://seek.princeton.edu/seek/) to validate whether the predicted regulons were authentically related to their associated cell types. We inputted the gene list of the MYBL2 regulon to the SEEK web server and searched for sarcoma-related keywords (such as "cancer", "sarcoma", and "proliferation" among others) from the complete dataset list ranked by query-co-expression score. Subsequently, we set a $P < 0.05$ cutoff to select significant datasets, and used Fisher's exact test to evaluate whether the selected datasets were significantly enriched in the top ranks.

InferCNV

We used the infercnv1.4.0 R package to analyze the InferCNV. Endothelial cells (benign cells) and proliferating osteoblasts from all patients in the osteosarcoma single-cell dataset were included in the analysis. The CNV score for each cell was obtained by dividing the sum of the squares of CNV values-1 of each gene in a single cell by the total number of genes.

Functional enrichment analysis

Gene ontology (GO) enrichment analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, and gene set enrichment analysis (GSEA) were performed using the clusterprofier R package.[24] The H gene set v7.2 collections and hallmarks were obtained from the molecular signature database as reference gene sets for GSEA. Single-sample GSEA (ssGSEA) was obtained using the R package gsva.

Immunohistochemical staining

Fifty-nine specimens of osteosarcoma with complete follow-up data were prepared for immunohistochemical staining. Representative tumor areas were obtained from formalin-fixed and paraffin-embedded primary osteosarcoma tissue specimens. Five high-power microscopic fields were randomly selected. The number of positive cells was calculated as a percentage of the total number of cells, presented as a positive expression value.

Pan-cancer immune infiltration analysis

The correlation between pan-cancer immune infiltration and MYBL2 was obtained using the Tumor Immune Estimation Resource database. Immune cell infiltration was

calculated using the QUANTISEQ, XCELL, CIBERSORT, CIBERSORT-ABS, and EPIC algorithms.

MYBL2-related gene score (MRGS) establishment and validation

We calculated the correlation coefficient between genes in the TARGET dataset and MYBL2. Subsequently, we filtered them with a standard correlation coefficient > 0.3 and $P < 0.05$ and obtained 570 genes. We merged the TARGET and GSE212557 datasets and used the sva package to remove batch effects and normalize the expression to obtain a sufficient and balanced sample. We randomly categorized the integrated osteosarcoma cohort into two groups: 75 and 72 participants in the training and control groups, respectively. The least absolute shrinkage and selector operation (LASSO) algorithm was used to further screen out genes that were optimal candidate MYBL2-related gene expression profiles with the best discriminative capability. Stepwise regression was applied to screen for the final genes. We developed a risk score model based on the MYBL2-related gene expression profiles, weighted using the multivariate Cox regression coefficient as follows:

$$MRGS = \sum \beta_i \times RNA_i$$

where β_i is the coefficient of RNA_i expression. We categorized the patients into high- and low-risk groups according to the median MRGS in both cohorts. The time-dependent receiver operating characteristic (ROC) curve was used to evaluate the predictive ability of MRGS for poor prognosis.

Statistical analysis

All statistical analyses and graphing were performed using R 4.0.3. Correlation analysis was performed using Spearman's method. The two groups were compared using the Wilcoxon rank-sum test. A random-effects model was used in the meta-analysis of the MYBL2 gene expression hazard ratio in pan-cancers. Kaplan—Meier plots and log-rank tests were used to present the survival analysis results and test the differences between the two groups for survival analysis, respectively. We used the cutoff finder to optimize the cutoff points for datasets that medians could not evaluate.[25] Statistical significance was set at $P < 0.05$, and all statistical tests were two-sided.

Results

Tumor heterogeneity in different patients is related to cell type composition and tumor regulon activity

We used GSE152048, a representative osteosarcoma single-cell dataset downloaded from GEO containing osteosarcoma tissues from 11 patients. These patients had osteogenetic, chondrogenic, metastatic, and primary osteosarcoma. After clustering and annotating the dataset (Fig. S1) and applying the SCENIC process, 304 activated regulons were identified (Table S1). The activity of each regulon in each cell type was calculated. Based on the activity level of these regulons, we performed t-distributed stochastic neighbor embedding on those cells to show the regulon patterns of different cells. Surprisingly, overall, the regulon patterns of patients varied (Fig. 2A). This difference also varied among cell types (Fig. 2B). Some cell types, including endothelial cells (Fig. 2C), had similar regulon patterns in different patients. These cells can be grouped into a dimensionality reduction cluster map based on the regulon activity. Some cell types, including malignant cells in osteosarcoma tissue, proliferative osteoblasts, and osteoblasts (Fig. 2D, E), and two types of cells in different patients did not aggregate well in the dimensionality reduction clustering based on the activity of regulons in cells. By comparing copy number variations, we found that proliferative osteoblasts were more malignant than other cells (Fig. S2A—E). We performed a KEGG analysis of the 334 transcription factors. Figure 2F shows the significant pathway.

Determination of the specific regulon in the main cell types of osteosarcoma

Although differences exist in the regulon activity patterns in different patients, the commonality of regulon activation status can still be found in specific cell types. For the main cells with high composition in osteosarcoma (osteoblastic proliferative cells, osteoblasts, chondroblasts, and osteoclasts), we calculated the activities of all regulons for them one by one, and based on the Jensen—Shannon divergence, calculated the RSS. We identified five regulons with the highest RSS in each cell and verified the highest RSS regulon function in the SEEK database.

In malignant osteoblastic proliferative osteosarcoma cells, the five most specific regulons were MYBL2, XRCC4, VAX2, E2F8, and TFDP2. The RSS of MYBL2 was higher than 0.4, which was the highest. We investigated the activation status of proliferative malignant osteoblasts and the MYBL2 regulon in different cells (green and gray represent activation and inactivation, respectively) and found that they overlapped substantially. SEEK analysis revealed that the co-expression level of MYBL2 and its target genes was significantly related to malignant proliferation ($P = 6.53e\text{-}09$) (Fig. 3D). The top five specific regulons in osteoblasts were DLX5, TCF7, RUNX2, CTCFL, and CREB3L1; the RSS of all five was higher than 0.55 (Fig. 3E). We found that the activation state distribution of the DLX5 regulon overlapped with that of osteoblasts (Fig. 3F, G), and the SEEK database verified that the degree of co-expression of the DLX5 regulon was correlated with osteogenic function ($P = 0.00371$) (Fig. 3H). The five most specific regulons in chondroblasts were FOXA3, SOX8, FOXC1, NKX2-2, and BHLHE40. The RSS of these regulons were all greater than 0.35. Among them, the highest was FOXA3, with an RSS of approximately 0.45 (Fig. 3I). The activation state distribution of the FOXA3 regulon was comparable to that of chondrocytes (Fig. 3J, K). The SEEK database verified that the FOXA3 regulon is correlated with the chondroblast function (Fig. 3L).

The five most specific regulons in osteoclasts included NFATC1, SP6, JDP2, VDR, and RELB. Among them, NFATC1 was the highest, with an RSS greater than 0.4 (Fig. 3M). The activation state distribution of the NFATC1 regulon was similar to that of the osteoclasts (Fig. 3M—P). The SEEK database verified that the NFATC1 regulon was correlated with osteoclast function (Fig. 3P).

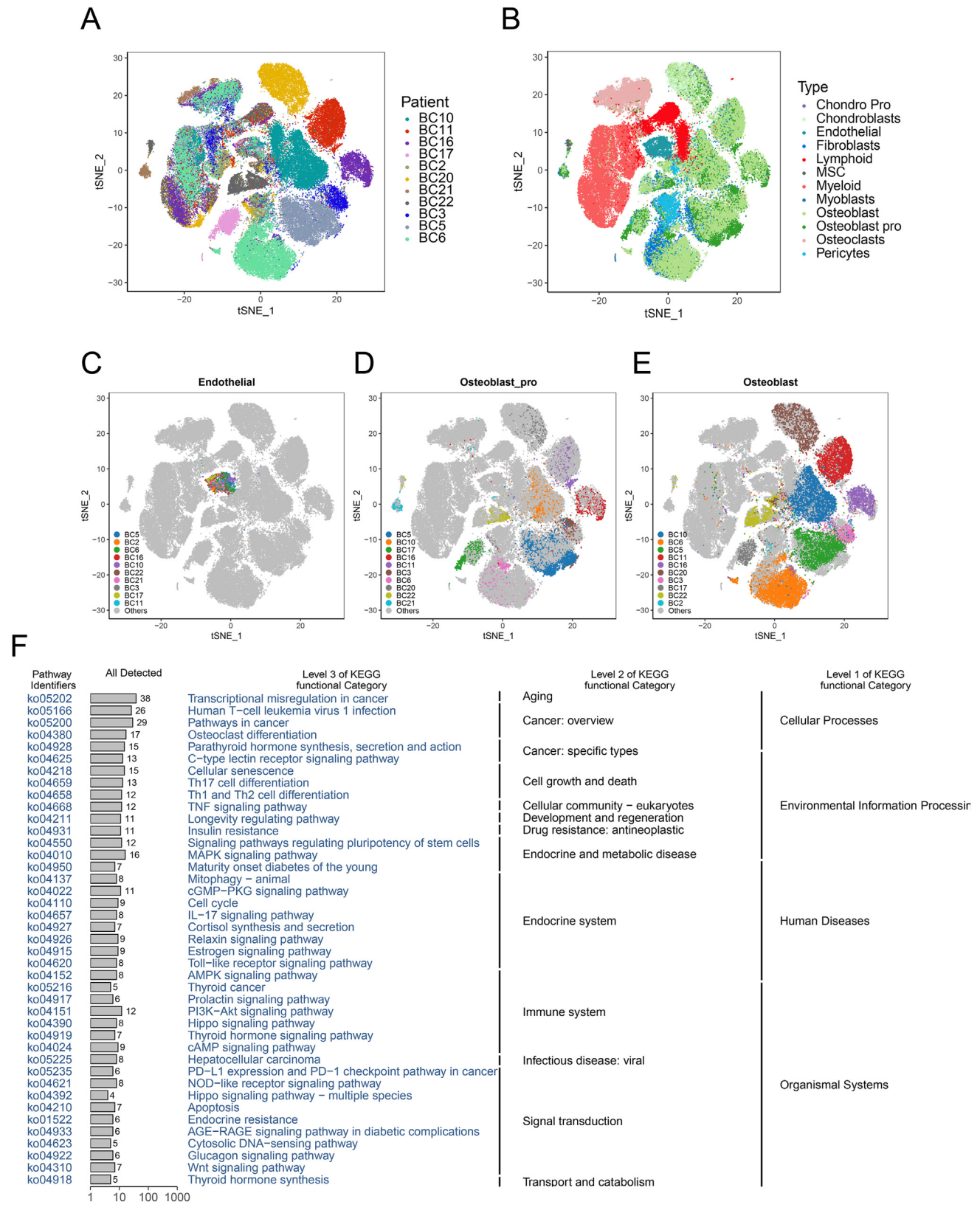

Figure 2 Osteosarcoma regulon map obtained from SCENIC. **(A)** Display of cell clustering map based on regulon activity on different patients. **(B)** Display of cell clustering map based on regulon activity on different cells in osteosarcoma. **(C)** The patterns of regulon activity in different patients are similar in endothelial cells. **(D)** Different patients have different regulon activity patterns in malignant osteoblasts with proliferative activity. **(E)** Regulon activity pattern in malignant osteoblasts varies in different patients. **(F)** KEGG analysis of activated transcription factors in osteosarcoma.

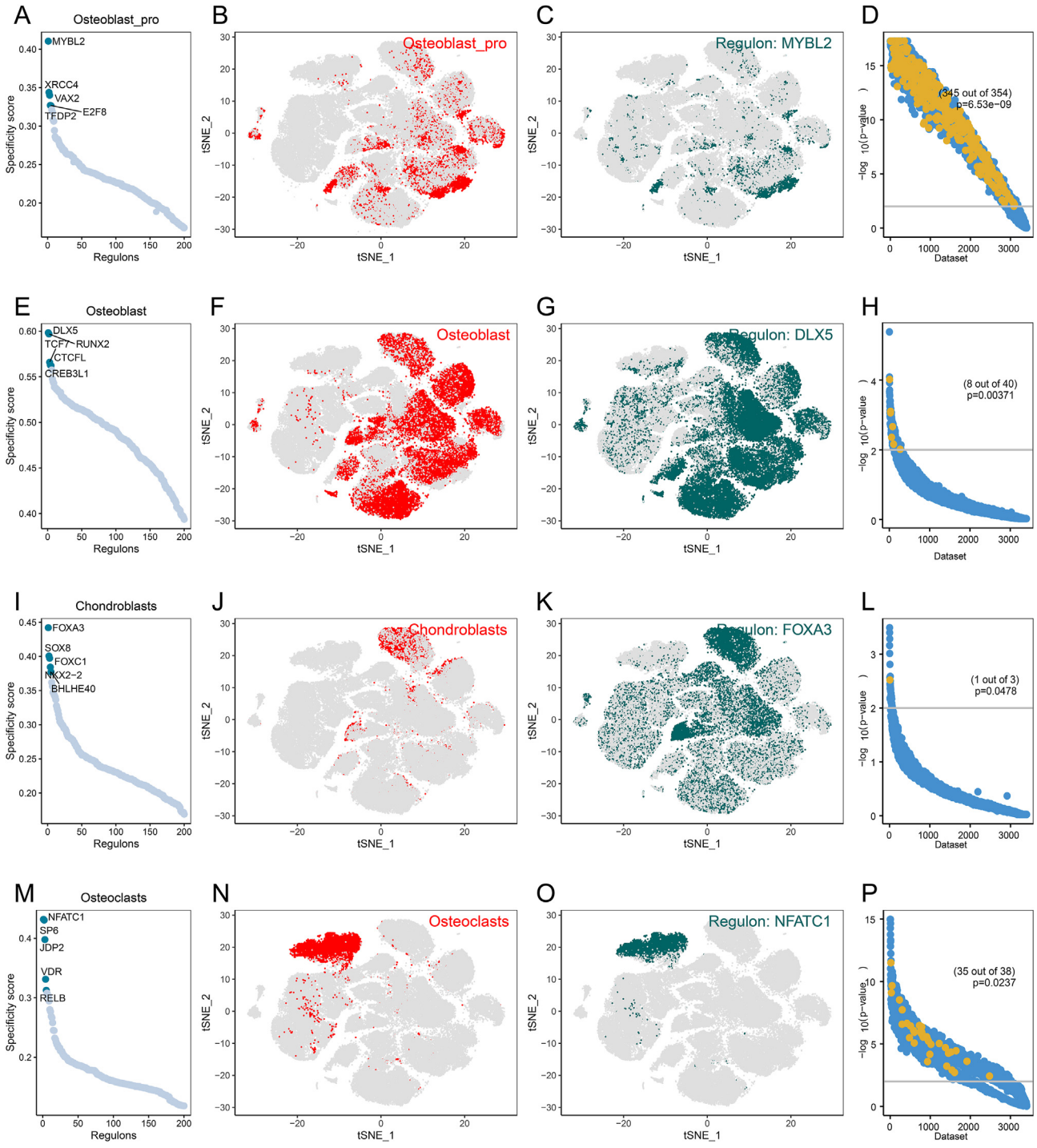

Figure 3 Cell-type-specific regulon activity analysis. **(A—D)** Proliferative osteoblasts. (A) Rank for regulons in proliferative osteoblasts based on regulon specificity score (RSS). (B) Proliferative osteoblasts are highlighted in the t-SNE map (red dots). (C) Binarized regulon activity scores (RAS) for top regulon MYBL2 on t-SNE map (dark green dots). (D) SEEK co-expression results for target genes of top regulon MYBL2 in different GEO datasets. The *x* axis represents different datasets, and the *y* axis represents the co-expression significance of target genes in each dataset. Proliferative osteoblasts related datasets with significant correlation ($P < 0.05$) are highlighted by yellow dots. **(E—H)** Same as (A—D) but for osteoblasts. **(I—L)** Same as (A—D) but for chondroblasts. **(M—P)** Same as (A—D) but for osteoclasts.

Expression and function of MYBL2 in osteosarcoma

MYBL2 was identified as the most specific regulator of proliferative osteoblasts of osteosarcoma in our previous study. To clarify the function of the transcription factor MYBL2, we analyzed it using the available osteosarcoma dataset.

In the GSE42352 dataset, 84 pre-chemotherapy biopsy samples and 19 osteosarcoma cell lines were collected as osteosarcoma samples, and 12 mesenchymal stem cell lines were set as normal samples. We compared the MYBL2 expression levels between the two groups. We found that the expression of MYBL2 in the tumor group was significantly higher than that in the normal group (Wilcoxon test $P = 0.028$) (Fig. 4A, 5).

In the GSE42352 dataset, we compared MYBL2 expression between 34 and 19 patients with and without metastasis within 5 years, respectively. We observed significantly higher expression in the metastasis group (Wilcoxon test, $P = 0.048$) (Fig. 4B).

We compared the overall survival analysis of GSE21257 and the TARGET dataset according to the level of MYBL2 expression to further explore the correlation between MYBL2 and the prognosis of patients with osteosarcoma. Patients with high MYBL2 expression sets showed worse overall survival in both datasets (GSE21257, log-rank test $P = 0.034$; TARGET, log-rank test $P = 0.044$) (Fig. 4C, D).

We selected the largest sample size dataset, TARGET, performed pathway enrichment analysis to obtain the enrichment score of each pathway, and subsequently calculated the Spearman correlation between the expression of MYBL2 and the enrichment score. We found that MYBL2 was related to cell proliferation pathways (including various aspects of the cell cycle) and immune response pathways in osteosarcoma (Fig. 4E). According to the association theory, co-expressed genes can reflect a specific gene's function. We used GSEA to identify enriched features and functional differences based on the Spearman correlation coefficients of all genes and MYBL2. We found that the MYBL2-related signaling pathways in osteosarcoma were almost all related to tumor progression or occurrence, including activation of tumor proliferation, tumor metabolism, immunosuppression, inflammatory immune responses, and inhibition of oxidative phosphorylation and normal DNA repair (Fig. 4F), indicating that MYBL2 may promote tumorigenesis and progression through multiple channels.

We performed GO and KEGG enrichment analyses on the MYBL2 target genes (189) calculated by SCENIC using clusterprofier. These genes were almost all enriched in the tumor.

Methylation and RNA-seq expression data from the TARGET database were used for this analysis. We found seven methylation sites in the MYBL2 gene region (Fig. S4A). They included cg18266010, cg23843505, cg01086075, cg16487280, cg09604203, ch.20.905431R, and ch.20.9.4058R. We analyzed the correlation between the methylation rate of all methylation sites and the expression level of MYBL2 to identify the methylation sites that may regulate the expression of MYBL2. We calculated the correlation between the methylation rate of these methylation sites and the expression level of MYBL2 and found no statistically significant results (Fig. S4B–H).

We identified 24 methylation sites with an absolute value of correlation coefficient greater than 0.4 ($P < 0.05$) (Fig. S4I). These sites' methylation levels were significantly related to MYBL2 and may directly or indirectly affect the expression of MYBL2.

Expression and function of MYBL2 in pan-cancer

Expression, prognostic relationship, function, and regulation of MYBL2 in pan-cancer were investigated by comparing the tumor tissues from TCGA (The Cancer Genome Atlas) with the corresponding GTEx normal tissues, and we found that MYBL2 was overexpressed in many types of human tumors, despite the different tissue sources (Fig. 6A). This indicates that MYBL2-mediated tumor occurrence and development is a general mechanism adopted by solid tumors. We reached similar conclusions in the TCGA matching organization (Fig. 6B). We compared the enrichment of the MYBL2 regulon in TCGA tumor and normal tissue-paired samples. The results obtained were identical to the results obtained from the expression of the *MYBL2* gene (Fig. 6C), indicating that the MYBL2 regulon may function synergistically as a whole in pan-tumors.

We calculated the correlation between MYBL2 expression in each tumor and MYBL2 target gene expression to verify whether the MYBL2 target gene and MYBL2 regulatory relationship calculated in the osteosarcoma single-cell data were consistent with other tumors. We found that the expression level of MYBL2 in pan-cancer was positively correlated with most target genes' expression levels, indicating that MYBL2 may similarly promote pan-cancer tumor progression. WDFY3, TRPM8, SYT2, and SYCP1 had different correlations with MYBL2 expression levels in various tumors. This indicated that the transcriptional regulation of MYBL2 was also affected by tissue sources (Fig. S5A).

A univariate Cox regression analysis was performed based on the MYBL2 expression level in pan-cancer to determine the relationship between the MYBL2 expression level and the overall survival of pan-cancers. It was found that in 14 of 33 (42%) types of tumors, MYBL2's expression was a risk factor ($HR > 1$, $P < 0.05$) (Fig. S5B). Based on the clinical data from TCGA, we performed a univariate Cox regression analysis of MYBL2 expression levels for DSS (disease free survival), DFI (disease free interval), and PFI (progression free interval). We found that MYBL2 expression levels were associated with worse DSS in ACC (adrenocortical carcinoma), KICH (kidney chromophobe), KIRC (kidney renal clear cell carcinoma), KIRP (kidney renal papillary cell carcinoma), LGG (brain lower grade glioma), LIHC (liver hepatocellular carcinoma), LUAD (lung adenocarcinoma), MESO (mesothelioma), PAAD (pancreatic adenocarcinoma), PCPG (pheochromocytoma and paraganglioma), PRAD (prostate adenocarcinoma), SARC (sarcoma), SKCM (skin cutaneous melanoma), and THCA (thyroid carcinoma); worse DFI in BRCA (breast invasive carcinoma), KIRP (kidney renal papillary cell carcinoma), LUAD (lung adenocarcinoma), MESO, PAAD (pancreatic adenocarcinoma), PRAD (prostate adenocarcinoma), and THCA (thyroid carcinoma); and worse PFI in ACC, KICH,

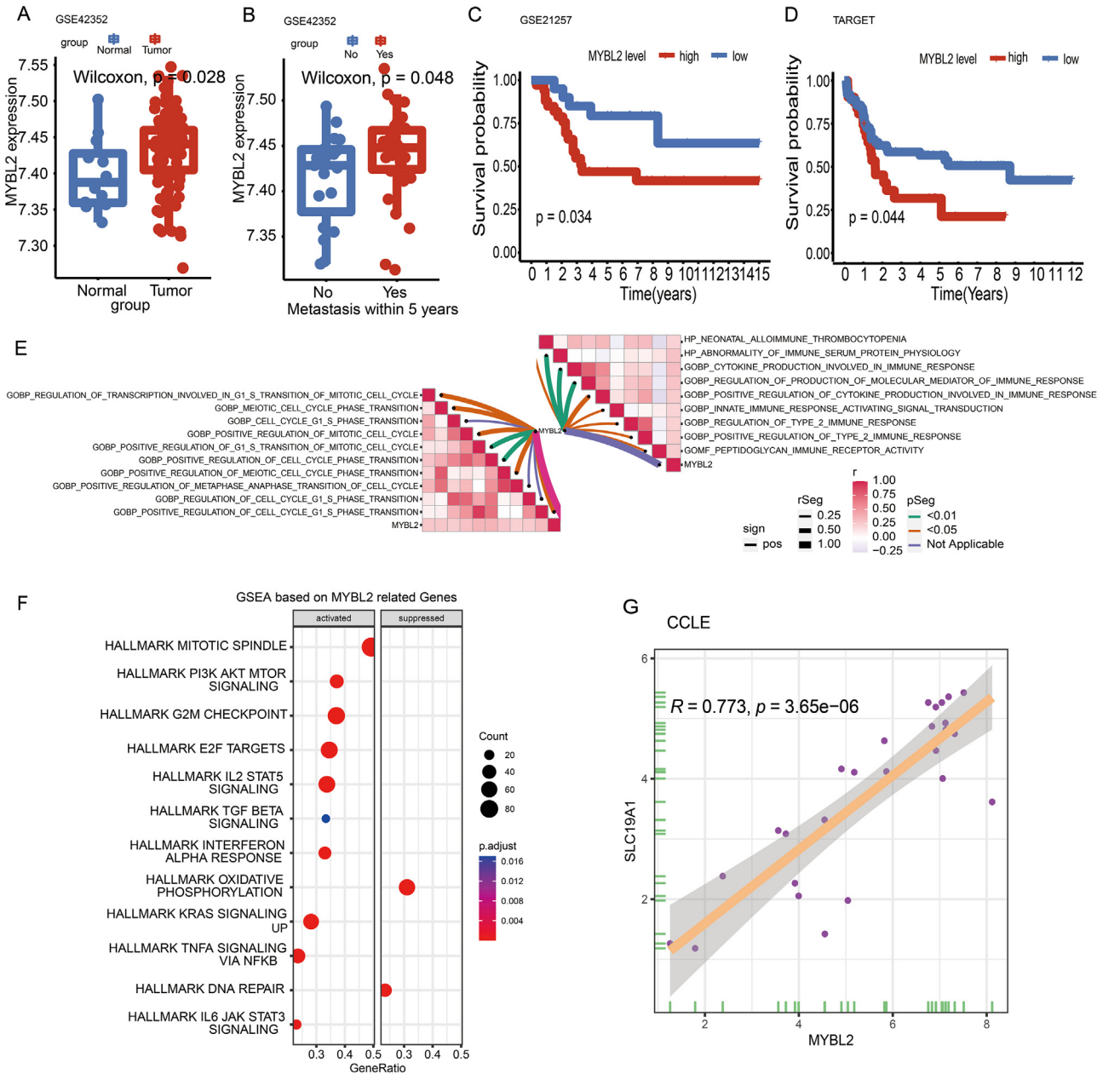

Figure 4 Expression and function of MYBL2 in osteosarcoma. **(A)** MYBL2 expressed higher in osteosarcoma tissues than in normal tissues. **(B)** MYBL2 expressed higher in patients with metastasis than in patients without metastasis in osteosarcoma. **(C, D)** Patients with higher MYBL2 expression have worse survival. **(E)** MYBL2 expression is related to cell proliferation and immunity in osteosarcoma tissue samples. **(F)** GSEA based on MYBL2 related genes indicates that MYBL2 may promote tumor progression in a variety of ways. **(G)** MYBL2 is positively correlated with SLC19A1 in malignant cells in bone.

KIRC, KIRP, LGG, LIHC, LUAD, MESO, PAAD, PCPG, PRAD, SARC, SKCM, and UVM (uveal melanoma) (HR $>$ 1, $P < 0.05$) (Fig. S5C—E).

We also meta-analyzed the HRs of MYBL2 to osteosarcoma, DSS, DFI, and PFI in different tumors. We adopted a random-effects model because the I2 value was over 50%. The HRs of the MYBL2 were 1.48 (95% CI [1.27, 1.71], OS), 1.38 (95% CI [1.19, 1.61], DSS), 1.21 (95% CI [1.08, 1.38], DFI), and 1.32 (95% CI [1.20, 1.45], PFI) respectively (Fig. S6A—D). Due to the large heterogeneity between different tumors, more experiments are needed to prove the prognostic role of MYBL2 in some tumors, including DLBC (lymphoid neoplasm diffuse large B-cell lymphoma) and TGCT (testicular germ cell tumors).

We verified the correlation between MYBL2 and SLC19A1 expression in pan-cancer. We found that in 29/33 tumors, MYBL2 expression was positively correlated with SLC19A1 expression (Fig. 6D). These findings indicate that the correlation between MYBL2 and SLC19A1 may not be affected by the source of the tumor tissue.

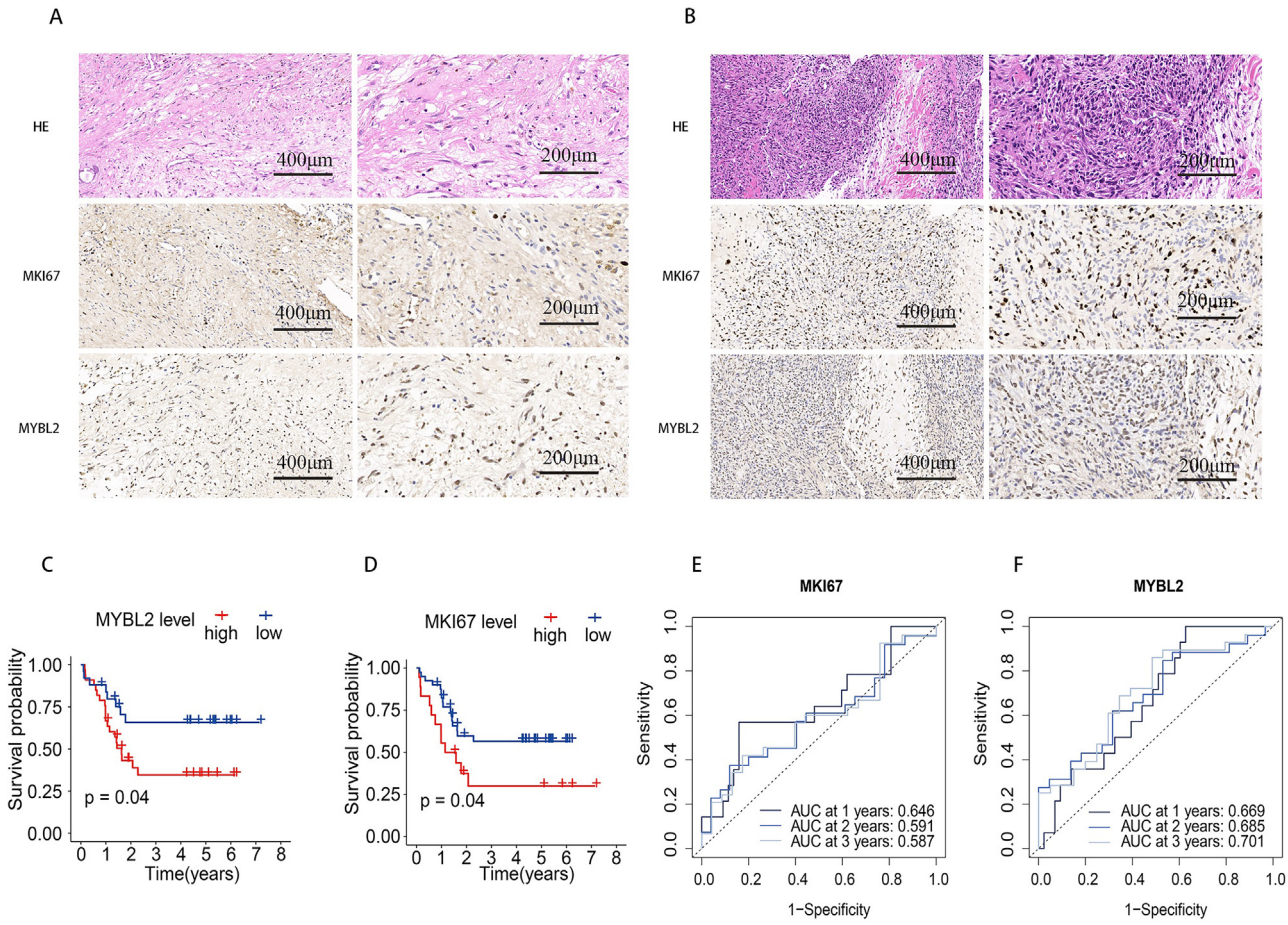

Figure 5 Validation of the predictive effect of MYBL2 on prognosis in the Xiangya cohort. **(A, B)** Both MYBL2 expression levels and MKI67 expression levels were associated with poor patient prognosis. **(C)** Patients with higher MYBL2 expression levels in the Xiangya osteosarcoma cohort had shorter survival times. **(D)** Patients with higher expression levels of MKI67 in the Xiangya osteosarcoma cohort had a shorter survival time. **(E)** The time-dependent ROC curve showed that the predictive accuracy of MKI67 on patient survival at 1, 2, and 3 years after the diagnosis of osteosarcoma was 0.646, 0.591, and 0.587, respectively. **(F)** The time-dependent ROC curve showed that the predictive accuracy of MYBL2 on patient survival at 1, 2, and 3 years after diagnosis of osteosarcoma was 0.669, 0.0.685, and 0.701, respectively.

MYBL2 is a potential immune checkpoint

From the heat map of the correlation between the infiltration of various immune cells and MYBL2 expression, we found that the expression level of MYBL2 in some tumors (including THCA and THYM (thymoma)) was associated with various immune cell infiltrations. However, this correlation was inconsistent in other tumors, which may be due to the specificity of tumor tissue. Interestingly, we found that a positive correlation between Th2, MDSC infiltration, and MYBL2 expression remained consistent in almost all types of tumors (Fig. S7).

We focused on Th2 cells and MDSC and found that MYBL2 was positively correlated with Th2 cell infiltration in almost all tumors except UCS, CESC, and COAD (Fig. 7A). Similarly, the positive correlation between MYBL2 and MDSC infiltration was consistent in approximately all tumors except BRCA-Her2, DLBC, TGCT, UCS, and HNSC-HPV+ (head and neck squamous cell carcinoma-HPV+) (Fig. 7B).

In the exploration of several typical tumors, we found that higher Th2 and MDSC infiltration in these tumors was associated with poorer overall survival (HR $>$ 1, $P < 0.05$) (Fig. S8A—P).

MYBL2 expression level in 28 tumors was significantly positively correlated with the expression of immune checkpoint CD276 (B7—B3) (with correlation coefficient > 0 and $P < 0.05$) (Fig. 7D). Among these tumors, CD276 was correlated with MYBL2, with the highest degree in MESO.

Additionally, MYBL2 expression was still present in 28 types of tumors. It was correlated with RELT, a member of the tumor necrosis factor receptor superfamily (Fig. 7E). Among these, RELT was associated with MYBL2, which had the highest ACC.

We determined the correlation between the expression of MYBL2 and TMB in pan-cancer. In 17 tumors, MYBL2 expression was positively correlated with TMB (Spearman correlation > 0, $P < 0.05$) (Fig. 7C).

In the classic immunotherapy cohort imvigor210 for urothelial cell tumors, we compared the expression of

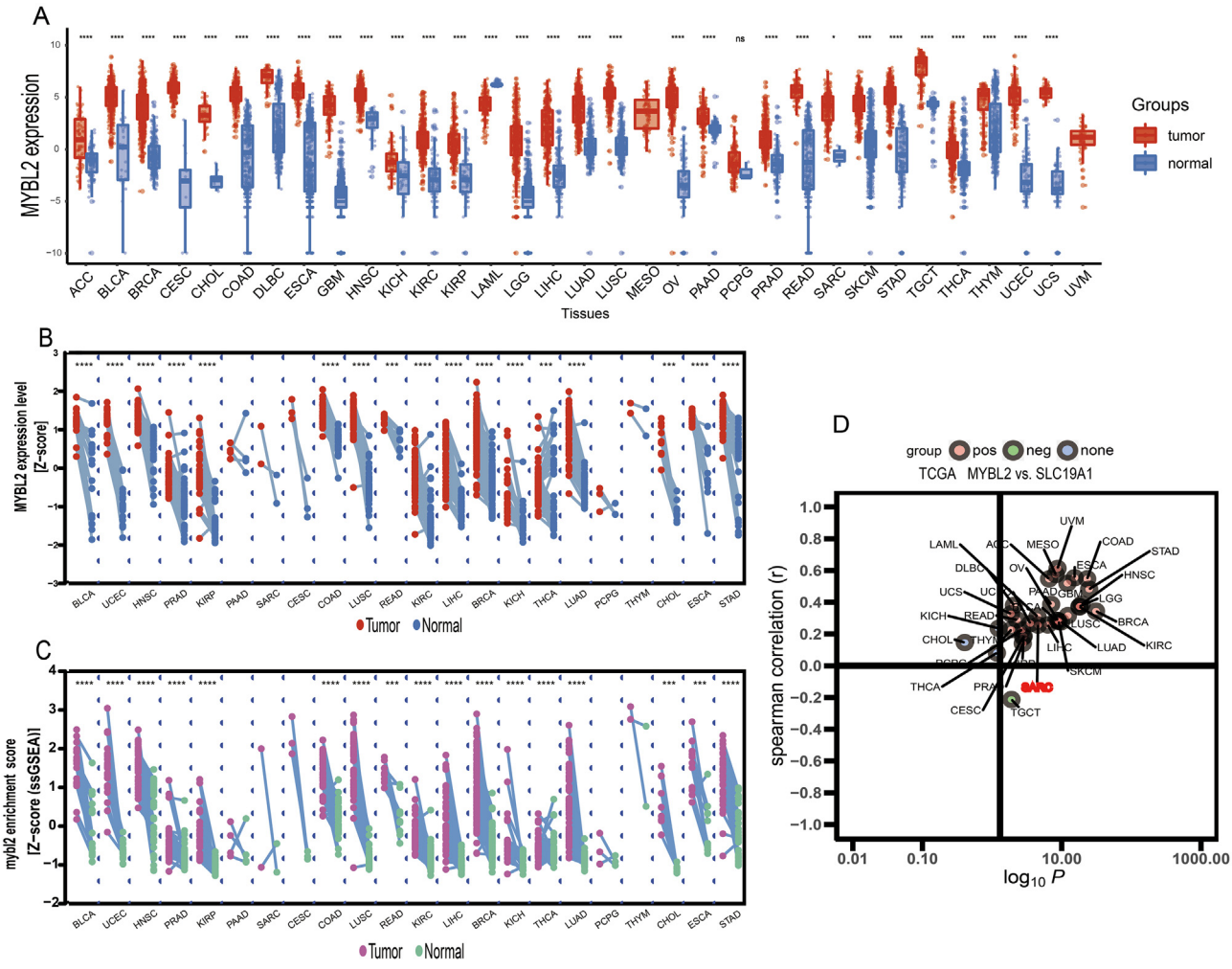

Figure 6 Expression and function of MYBL2 in pan-cancer. **(A)** The MYBL2 gene in TCGA and GTEX databases has higher expression in tumor tissues than in normal tissues. **(B)** The expression of the MBYL2 gene was higher in tumor tissues than in matched adjacent tissues in TCGA-paired samples. **(C)** The enrichment scores of MYBL2 and its target genes were higher in TCGA-paired tumor tissues than in normal tissues. **(D)** MYBL2 gene positively correlates with the expression of tumor drug sensitivity-related gene SLC19A1 in pan-cancer.

MYBL2 between the CR (complete response)/PR (partial response) and SD (stable disease)/PD (progressive disease) groups. We found that the MYBL2 expression was higher in the group with better immunotherapy effects ($P < 0.001$) (Fig. 7F).

Patients were categorized into high- and low-expression groups according to the median value of MYBL2 expression. Patients with high MYBL2 expression had better immunotherapy responses (OR = 0.423; 95A% CI [0.228—0.767]; $P = 0.004$). The baseline information of patients in the imvigor210 dataset is presented in Table S2. We validated MYBL2 expression in rectal cancer and esophageal adenocarcinoma and found that the mean MYBL2 level in the better-immunotherapy-response group was higher than that in the worse-immunotherapy-response group (Fig. S9).

Construction and validation of MRGS

The osteosarcoma dataset with the largest sample size was selected, and the expression of MYBL2 and the expression of each gene were analyzed for correlation. After screening according to the established rules (correlation coefficient $>$ 0.3, $P < 0.05$), 570 genes that cooperated with MYBL2 were obtained.

The GSE21257 and TARGET datasets, with detailed clinical data, were selected to construct the prognostic model. After excluding the batch effect, we normalized gene expression. To train for prognostic factors, 147 patients were randomly categorized into two groups: 75 and 72 in the training and test groups, respectively. Overall, 37 MYBL2-related genes affected the prognosis based on univariate Cox analysis (Table S3). Using the LASSO algorithm, we identified 14 genes with minimal lambda (0.0744) (Fig. 8A and B). After stepwise multivariate Cox regression, seven final candidate genes were screened to construct MRGS. Multivariate Cox regression analysis demonstrated that the expression of *CCDC58*, *NDUFB4*, and *RAB15* were independent risk factors for osteosarcoma, while *CTNBIP1*, *PLEKHG1*, and *SGPL1* were protective factors after osteosarcoma (Fig. 8C and Table S4). The risk score was

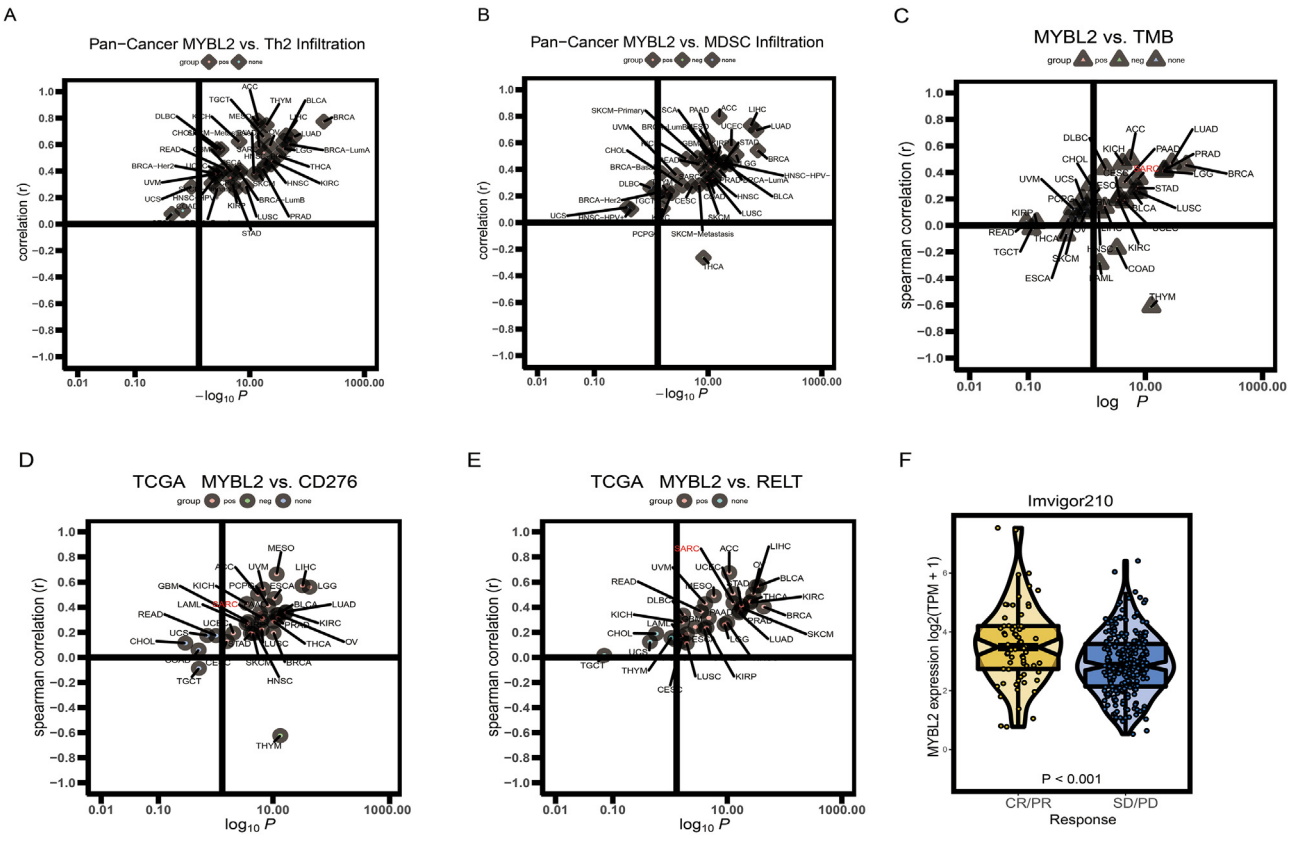

Figure 7 MYBL2 is a potential immune checkpoint. **(A)** MYBL2 expression level in pan-cancer is positively correlated with Th2 immune cell infiltration. **(B)** MYBL2 expression level in pan-cancer is positively correlated with MDSC immune cell infiltration level. **(C)** MYBL2 is positively correlated with tumor mutation burden in pan-cancer. **(D)** MYBL2 expression level in pan-cancer is positively correlated with immune checkpoint CD276. **(E)** The expression level of MYBL2 in pan-cancer is positively correlated with tumor necrosis factor receptor RELT. **(F)** Imvigor210 immune patients in the treatment cohort with better immunotherapy response had higher levels of MYBL2 expression.

calculated using the following equation: risk score = 2.690 × AUNIP + 4.434 × CCDC58 − 3.383 × CTNNBIP1+ 2.398 × NDUFB4 − 7.068 × PLEKHG1 + 7.765 × RAB15 + 6.188 × SGPL1.

We calculated the risk score of each patient and categorized them into high- and low-risk groups in the training and test groups, respectively, according to the median risk score. The distribution of the risk score, survival status, and corresponding heatmap of the expression levels of MRGs in patients in the training and testing sets are displayed (Fig. 8D, E). To evaluate the impact of high-risk and low-risk scores on prognosis, we evaluated the overall survival of the training and validation groups using the Kaplan−Meier curve (Fig. 8G, I). To further clarify the accuracy of the MRGS model in predicting the overall survival of patients with osteosarcoma, we analyzed the time-dependent ROC curves. In the training set, the AUCs of the prognostic signature were 0.841, 0.885, and 0.936 at 1, 2, and 3 years, respectively. Similarly, in the validation set, the AUCs were 0.812, 0.778, and 0.759, respectively (Fig. 8H, J). These results indicated that the MRGS model accurately predicted osteosarcoma development.

The risk of all patients after integration was classified into high and low groups according to the median risk score. After excluding patients with missing clinical information, we obtained data from 98 patients. Ninety-eight patients were randomly categorized into training and validation groups at a 1:1 ratio. The baseline information is presented in Table S4. In the training group, we performed multivariate regression analysis on age, sex, Huevos grade, tumor site, metastasis at diagnosis, and MRGS. We verified that metastasis at diagnosis and MRGS were independent prognostic factors (Fig. 9A). We constructed a nomogram based on the multivariate Cox regression model (Fig. 9B). Metastasis at diagnosis had the most significant effect on osteosarcoma, followed by MRGS. Based on this nomogram, by summarizing the characteristics of each predictor and measuring the total osteosarcoma score, the individual's survival probability can be easily calculated. Calibration plots were used to visualize the performance of the nomograms in the training and validation sets. The 45° line represents the best prediction. The calibration plots showed that the nomogram performed well in both the training and control groups (Fig. 9C−E; Figure 9F−H in the training and validation sets, respectively). The concordance index (C-index) of the training, validation, and whole sets were 0.805, 95% CI [0.636, 0.975], 0.764, 95% CI [0.598, 0.930], and 0.787, 95% CI [0.598, 0.930], respectively (Table S6). Generally, the prediction nomogram based on MRGS has good predictive ability in the short and long terms.

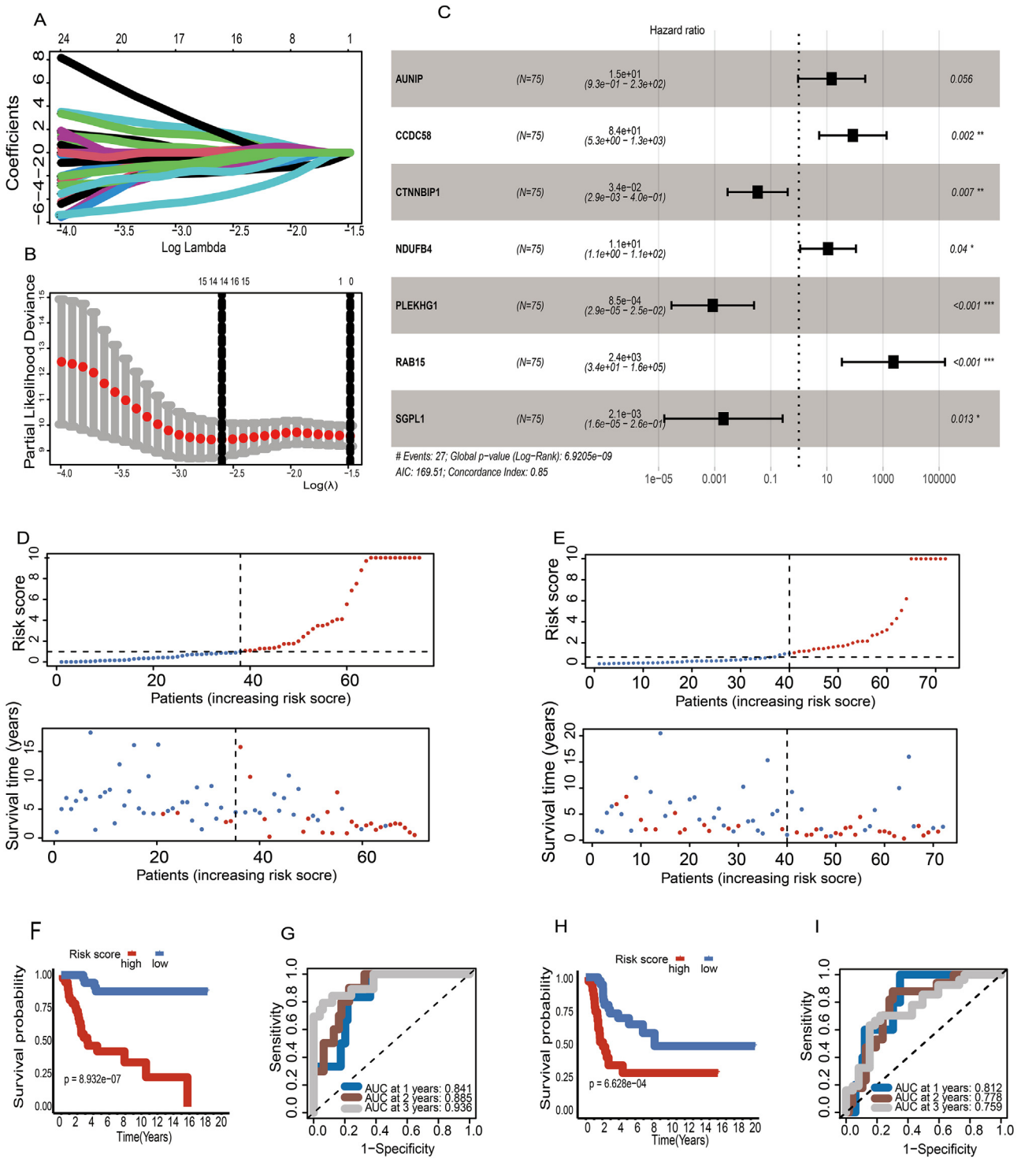

Figure 8 The construction and validation of the MRGS. **(A, B)** LASSO regression analysis of the MYBL2-related genes associated with osteosarcoma prognosis. **(C)** Multivariate Cox regression analysis of the 7 genes (AUNIP, CCDC58, CTNNBIP1, NDUFB4, PLEKHG1, RAB15, and SGPL1) and establishment of the regression equation. **(D, E)** Patients are ranked by risk score and corresponding survival status of the training and validation sets. **(F, H)** Kaplan—Meier survival curve of overall survival of OS patients in the training set and the validation set according to the median cutoff value. **(G, I)** ROC curves at 1, 2, and 3 years in the training and validation sets.

Discussion

Based on the SCENIC algorithm, we found a critical role for MYBL2 in proliferating osteosarcoma cells. Multiple datasets have collectively demonstrated that MYBL2 plays a crucial predictive role in osteosarcoma prognosis. Based on the evidence that MYBL2 is positively correlated with a suppressive immune status in pan-cancer, the inhibitory

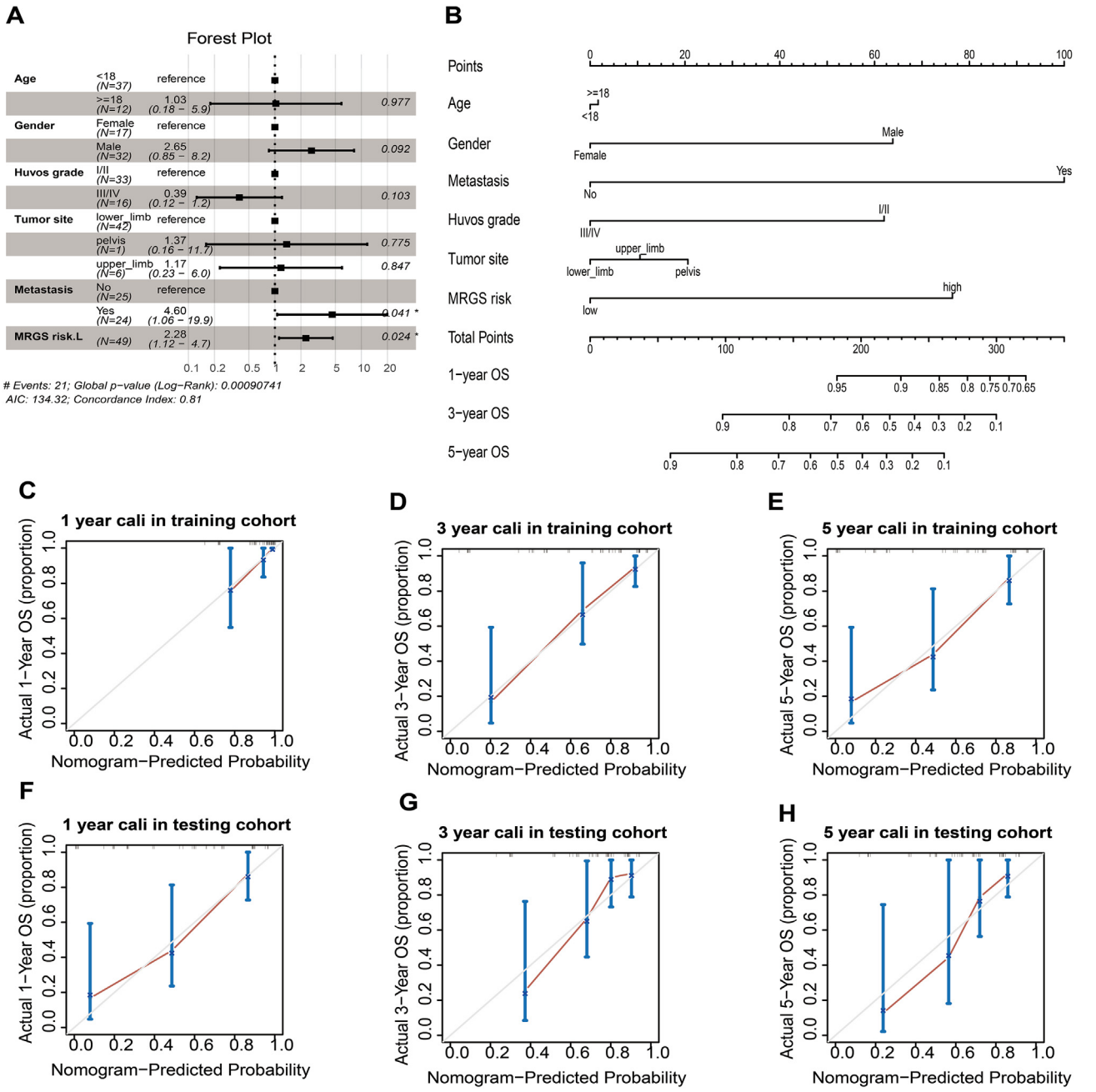

Figure 9 Establishment and validation of a nomogram based on MRGS. **(A)** Multivariate regression analysis based on age, gender, Huovs grade, metastasis at diagnosis, and MRGS in the training group. **(B)** Nomogram for predicting the probability of 1-, 3-, and 5-year overall survival for patients with osteosarcoma. **(C—E)** Calibration curves of the nomogram for predicting the overall survival at 1, 3, and 5 years in the training cohort. **(F—H)** Calibration curves of the nomogram for predicting the overall survival at 1, 3, and 5 years in the validation cohort.

effect of MYBL2 on tumor immunity was further confirmed. Finally, MRGS was developed to predict the osteosarcoma prognosis.

Osteosarcoma is a highly aggressive primary malignant bone tumor that occurs more frequently in adolescents and elderly individuals. The incidence of osteosarcoma is five per million people, and although the incidence is low, the prognosis of patients with osteosarcoma is poor.[26,27]

Tumor heterogeneity is a feature of malignancy and can cause differences in tumor growth rate, invasion, metastasis, drug sensitivity, and prognosis. Intertumoral and intratumor heterogeneities are closely related to tumor progression and metastasis, affecting the internal biology of tumors.[28,29] These features are closely related to tumor diagnosis, response to targeted therapy, and patient survival. Through SCENIC analysis, we found that the regulon activity of malignant tumor cells showed greater variability between patients with different tumors. In contrast, non-tumor cells (*e.g.*, endothelial cells) were less abundant.

Here, by comparing the expression of MYBL2 between osteosarcoma tissues and non-tumor controls in the GSE42352 and GSE87624 datasets, we demonstrated that MYBL2 is highly expressed in osteosarcoma tissues. In the GSE42352 dataset, we confirmed that MYBL2 was significantly associated with the malignant biological behavior of distant metastasis in osteosarcoma. In the GSE21257 and TARGET datasets, patients with high MYBL2 expression had shorter survival times. These results imply that MYBL2 plays a vital role in the development of osteosarcoma. Similar results have been reported for many other tumors. Liang et al detected the protein expression of MYBL2 in 65 samples each from patients with primary gallbladder cancer and cholecystitis.[14] They found that the expression rate of MYBL2 in gallbladder cancer samples was significantly higher than that in cholecystitis samples. It was also found that a high expression level of MYBL2 was significantly correlated with lymph node metastasis and tumor-node-metastasis stage in patients with gallbladder cancer.[14] Jin et al found that the MYBL2 expression rate in non-small-cell lung cancer was significantly higher than that in normal lung tissue and was associated with tumor lymph node metastasis.[18] Ren et al used immunohistochemistry to detect MYBL2 protein expression in 97 colorectal cancer specimens and 104 normal intestinal tissues. They found that the MYBL2 expression rate in cancer tissues was higher than that in normal intestinal epithelial tissues, and MYBL2 was involved in recurrence and metastasis. The high expression rate of the group was also significantly higher than that of the non-recurrence and metastasis group.[15] To further explore the prognostic role of MYBL2 in osteosarcoma, we detected the expression of MYBL2 and MKI67, reliable markers of tumor proliferation activity, in osteosarcoma specimens from the Xiangya cohort (Fig. 5). The predictive effect on prognosis was better than that of MKI67.

Through GSEA and ssGSEA analyses based on MYBL2 correlation, this study found that in osteosarcoma, the MYBL2 expression is closely related to the pathways of tumorigenesis and progression, including tumor proliferation, metabolism, immunosuppression, immune-inflammatory response, inhibition of apoptosis, and DNA repair. The relationship between MYBL2 and cell proliferation has been reported in many studies, which have demonstrated that MYBL2 plays a vital role in G1/S progression.[30–32] Moreover, the roles of MYBL2 in regulating apoptosis,[33–35] inhibiting cell senescence,[36–38] and maintaining the undifferentiated nature of stem cells[36,39] have been reported in other tumors. These studies have shown that MYBL2 is closely associated with tumor progression and proliferation. However, few reports on the relationship between MYBL2 and the immune microenvironment in malignant tumors are available.

Although immunotherapy has achieved remarkable results in various malignant tumors, some patients are still insensitive to immunotherapy. Technological advancements have gradually recognized the complexity and diversity of the tumor microenvironment and its impact on immunotherapy. The degree of infiltration of these two cell types is also associated with a worse prognosis in various tumors. Meng et al found that the MYBL2 expression level correlated with the MDSC infiltration level and CD276 immune checkpoint expression in prostate cancer. However, this finding was confirmed only in prostate cancer and not in other tumors.[40] MDSCs play a crucial role in tumor immunity and are involved in various aspects of immune regulation. MDSCs primarily induce immunosuppression via T cells.[41] Major factors in MDSC-mediated immunosuppression include arginase, inducible nitric oxide synthase, tumor growth factor, interleukin 10, COX2, indoleamine 2,3 dioxygenase immobilization of cysteine, and decreased expression of L-selectin by T cells, among others.[41] Th2 cells differentiate from $CD4^+$ T-cells. Studies have shown that Th2 cells activate tumor-associated M2 macrophages by secreting the cytokines IL-4, IL-6, IL-10, and IL-13, thereby promoting tumor progression.[42] Here, Le et al found that in a subset of ER-positive breast cancers, high Th2 expression was associated with Nottingham pathological grade, MKI67, and proliferation score.[43] This study also found that MYBL2 is associated with the immune checkpoints CD276 and RELT. Beom et al found that RELT might play an essential role in tumor immune escape as a negative regulator of the early stage of T cell activation by promoting T cell apoptosis.[44] Recently, CD276 has been a newly discovered immune checkpoint,[45] and researchers have found that the expression of CD276 is significantly negatively correlated with $CD8^+$ T cell infiltration. Blockade of CD276 can significantly increase the infiltration of $CD8^+$ cells, and depletion of $CD8^+$ T cells can significantly reduce the antitumor immune function blocked by CD276.[46]

Resistance to chemotherapy is an essential factor in determining tumor progression. Previous studies have demonstrated that MYBL2 increases resistance to chemotherapy by inhibiting apoptosis. In neuroblastoma, MYBL2 can regulate clusterin expression and resist apoptosis induced by dolephomycin.[47] In IL-2-dependent T cells, overexpression of MYBL2 can promote the expression of BCL-2, a gene that inhibits apoptosis, thereby increasing resistance to drugs, including ceramide and doxorubicin.[34] However, in this study, it was found that the expression of MYBL2 did not show significant abnormalities in the chemotherapy and drug-resistant groups in osteosarcoma.

The protein encoded by *SLC19A1* can assist in transporting methotrexate and enhancing the inhibitory effect of methotrexate on dihydrofolate reductase in cells.[48] A multidrug combination is usually used because of the complex chemotherapy regimen for osteosarcoma. Finally, we developed and validated an MRGS expression profile-based predictive model to predict nomograms for osteosarcoma prognosis.

This study had some limitations. First, the low incidence of osteosarcoma makes it difficult for a single center to have a large cohort; therefore, our cohort and the cohort of public databases have limited sample sizes. Therefore, more reliable conclusions should be drawn after more results from other osteosarcoma centers are combined. Second, because of the limited corresponding resources in the public database, this study did not verify the predictive ability of MYBL2 for immunotherapy in osteosarcoma, and it is necessary to further verify the role of MYBL2 in the osteosarcoma immunotherapy cohort at a later stage. Finally, most of the results of this study were based on transcriptomic sequencing data, and more reliable conclusions may require further verification by cell and animal experiments.

Author contributions

Xinzhu Qiu: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Supervision; Validation; Visualization; Roles/Writing - original draft.
Hongbo He: Conceptualization; Investigation; Writing - review & editing.
Hao Zeng: Conceptualization; Investigation; Writing - review & editing.
Xiaopeng Tong: Investigation; Writing - review & editing.
Can Zhang: Conceptualization; Investigation; Writing - review & editing.
Yupeng Liu: Conceptualization; Investigation; Writing - review & editing.
Zhan liao: Conceptualization; Supervision; Validation; Writing - review & editing.
Qing Liu: Conceptualization; Funding acquisition; Investigation; Methodology; Resources; Supervision; Validation; Visualization; Roles/Writing - original draft.

Ethics declaration

This study has been approved by the clinical medical research ethics committee of Xiangya Hospital of Central South University.

Conflict of interests

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

Funding

This work was supported by Changsha Natural Science Foundation, Hunan, China (No. kq2202382), the Natural Science Foundation of Hunan, China (No. 2022JJ40802, 2022JJ30928), and the Project funded by China Postdoctoral Science Foundation (No. 2022M713525).

Data availability

The data generated by this study are publicly available through the Gene Expression Omnibus (GEO: https://www.ncbi.nlm.nih.gov/gds), The Cancer Genome Atlas (TCGA: https://portal.gdc.cancer.gov), Cancer Cell Line Encyclopedia (CCLE: https://depmap.org/portal), Therapeutically Applicable Research To Generate Effective Treatments (TARGET: https://ocg.cancer.gov/programs/target), and Genotype-Tissue Expression Project (GTEx: https://gtexportal.org/home/datasets). The data of the Xiangya cohort can be obtained from the medical record information system of Xiangya Hospital, Central South University. Any questions or inquiries regarding the present study can be directed to Qing Liu, MD, Ph. D (158112225@csu.edu.cn), as corresponding author.

Acknowledgements

We would like to give our sincere thanks to Dr. Wei Luo, Dr. Jun Wan, Dr. Feng Long, Dr. Jian Tian, and Dr. Xiang Ding for their help in data processing and collection of clinical sample case data. Besides, we would like to thank Editage (www.editage.cn) for English language editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gendis.2023.04.035.

References

1. Yoshida A. Osteosarcoma: old and new challenges. *Surg Pathol Clin*. 2021;14(4):567—583.
2. Zhao J, Zhao Y, Ma X, Zhang B, Feng H. Targeting ferroptosis in osteosarcoma. *J Bone Oncol*. 2021;30:100380.
3. Zhou Y, Yang D, Yang Q, et al. Single-cell RNA landscape of intratumoral heterogeneity and immunosuppressive microenvironment in advanced osteosarcoma. *Nat Commun*. 2020;11(1):6322.
4. Kelley LM, Schlegel M, Hecker-Nolting S, et al. Pathological fracture and prognosis of high-grade osteosarcoma of the extremities: an analysis of 2, 847 consecutive cooperative osteosarcoma study group (COSS) patients. *J Clin Oncol*. 2020;38(8):823—833.
5. Zheng W, Huang Y, Chen H, et al. Nomogram application to predict overall and cancer-specific survival in osteosarcoma. *Cancer Manag Res*. 2018;10:5439—5450.
6. Zhang Y, Yan S, Lu D, Dong F, Lian Y. Screening feature genes of osteosarcoma with DNA microarray: a bioinformatic analysis. *Int J Clin Exp Med*. 2015;8(5):7134—7142.
7. 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.
8. Mutsaers AJ, Walkley CR. Cells of origin in osteosarcoma: mesenchymal stem cells or osteoblast committed cells? *Bone*. 2014;62:56—63.
9. Kuijjer ML, Rydbeck H, Kresse SH, et al. Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. *Genes Chromosomes Cancer*. 2012;51(7):696—706.
10. Zhu J, Simayi N, Wan R, Huang W. CAR T targets and microenvironmental barriers of osteosarcoma. *Cytotherapy*. 2022;24(6):567—576.
11. Musa J, Aynaud MM, Mirabeau O, Delattre O, Grünewald TG. MYBL2 (B-Myb): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis. *Cell Death Dis*. 2017;8(6):e2895.
12. Cicirò Y, Sala A. MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. *Oncogenesis*. 2021;10(2):19.
13. Zhan M, Riordon DR, Yan B, et al. The B-MYB transcriptional network guides cell cycle progression and fate decisions to sustain self-renewal and the identity of pluripotent stem cells. *PLoS One*. 2012;7(8):e42350.
14. Liang HB, Cao Y, Ma Q, et al. MYBL2 is a potential prognostic marker that promotes cell proliferation in gallbladder cancer. *Cell Physiol Biochem*. 2017;41(5):2117—2131.
15. Ren F, Wang L, Shen X, et al. MYBL2 is an independent prognostic marker that has tumor-promoting functions in colorectal cancer. *Am J Cancer Res*. 2015;5(4):1542—1552.

16. Nord H, Segersten U, Sandgren J, et al. Focal amplifications are associated with high grade and recurrences in stage Ta bladder carcinoma. *Int J Cancer*. 2010;126(6):1390—1402.
17. Astbury K, McEvoy L, Brian H, et al. MYBL2 (B-MYB) in cervical cancer: putative biomarker. *Int J Gynecol Cancer*. 2011;21(2): 206—212.
18. Jin Y, Zhu H, Cai W, et al. B-myb is up-regulated and promotes cell growth and motility in non-small cell lung cancer. *Int J Mol Sci*. 2017;18(6):860.
19. Thomas C, Robinson C, Dessauvagie B, et al. Expression of proliferation genes in formalin-fixed paraffin-embedded (FFPE) tissue from breast carcinomas. Feasibility and relevance for a routine histopathology laboratory. *J Clin Pathol*. 2017;70(1): 25—32.
20. Raschellà G, Cesi V, Amendola R, et al. Expression of B-myb in neuroblastoma tumors is a poor prognostic factor independent from MYCN amplification. *Cancer Res*. 1999; 59(14):3365—3368.
21. Thorner AR, Hoadley KA, Parker JS, Winkel S, Millikan RC, Perou CM. *In vitro* and *in vivo* analysis of B-Myb in basal-like breast cancer. *Oncogene*. 2009;28(5):742—751.
22. Goldman MJ, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data via the *Xena* platform. *Nat Biotechnol*. 2020;38(6):675—678.
23. Suo S, Zhu Q, Saadatpour A, Fei L, Guo G, Yuan GC. Revealing the critical regulators of cell identity in the mouse cell atlas. *Cell Rep*. 2018;25(6):1436—1445.e3.
24. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284—287.
25. Budczies J, Klauschen F, Sinn BV, et al. Cutoff Finder: a comprehensive and straightforward Web application enabling rapid biomarker cutoff optimization. *PLoS One*. 2012;7(12): e51862.
26. Sampo M, Koivikko M, Taskinen M, et al. Incidence, epidemiology and treatment results of osteosarcoma in Finland - a nationwide population-based study. *Acta Oncol*. 2011;50(8): 1206—1214.
27. Ottaviani G, Jaffe N. The epidemiology of osteosarcoma. *Cancer Treat Res*. 2009;152:3—13.
28. Shen WC, Lai YC, Li LH, et al. Methylation and PTEN activation in dental pulp mesenchymal stem cells promotes osteogenesis and reduces oncogenesis. *Nat Commun*. 2019;10(1):2226.
29. Gibbs CP, Kukekov VG, Reith JD, et al. Stem-like cells in bone sarcomas: implications for tumorigenesis. *Neoplasia*. 2005; 7(11):967—976.
30. Iwai N, Kitajima K, Sakai K, Kimura T, Nakano T. Alteration of cell adhesion and cell cycle properties of ES cells by an inducible dominant interfering Myb mutant. *Oncogene*. 2001; 20(12):1425—1434.
31. Lin D, Fiscella M, O'Connor PM, et al. Constitutive expression of B-myb can bypass p53-induced Waf1/Cip1-mediated G1 arrest. *Proc Natl Acad Sci U S A*. 1994;91(21):10079—10083.
32. Arsura M, Introna M, Passerini F, Mantovani A, Golay J. B-myb antisense oligonucleotides inhibit proliferation of human hematopoietic cell lines. *Blood*. 1992;79(10):2708—2716.
33. Lang G, Gombert WM, Gould HJ. A transcriptional regulatory element in the coding sequence of the human Bcl-2 gene. *Immunology*. 2005;114(1):25—36.
34. Grassilli E, Salomoni P, Perrotti D, Franceschi C, Calabretta B. Resistance to apoptosis in CTLL-2 cells overexpressing B-Myb is associated with B-Myb-dependent bcl-2 induction. *Cancer Res*. 1999;59(10):2451—2456.
35. Yuan J, Zhang Y, Sheng Y, Fu X, Cheng H, Zhou R. MYBL2 guides autophagy suppressor VDAC2 in the developing ovary to inhibit autophagy through a complex of VDAC2-BECN1-BCL2L1 in mammals. *Autophagy*. 2015;11(7):1081—1098.
36. Huang Y, Wu J, Li R, et al. B-MYB delays cell aging by repressing $p16^{INK4\alpha}$ transcription. *Cell Mol Life Sci*. 2011;68(5):893—901.
37. Johung K, Goodwin EC, DiMaio D. Human papillomavirus E7 repression in cervical carcinoma cells initiates a transcriptional cascade driven by the retinoblastoma family, resulting in senescence. *J Virol*. 2007;81(5):2102—2116.
38. Sala A, Calabretta B. Regulation of BALB/c 3T3 fibroblast proliferation by B-myb is accompanied by selective activation of cdc2 and cyclin D1 expression. *Proc Natl Acad Sci U S A*. 1992;89(21):10415—10419.
39. Bies J, Hoffman B, Amanullah A, Giese T, Wolff L. B-Myb prevents growth arrest associated with terminal differentiation of monocytic cells. *Oncogene*. 1996;12(2):355—363.
40. Jiao M, Zhang F, Teng W, Zhou C. MYBL2 is a novel independent prognostic biomarker and correlated with immune infiltrates in prostate cancer. *Int J Gen Med*. 2022;15:3003—3030.
41. Gabrilovich DI. Myeloid-derived suppressor cells. *Cancer Immunol Res*. 2017;5(1):3—8.
42. DeNardo DG, Coussens LM. Inflammation and breast cancer. Balancing immune response: crosstalk between adaptive and innate immune cells during breast cancer progression. *Breast Cancer Res*. 2007;9(4):212.
43. Le L, Tokumaru Y, Oshi M, et al. Th2 cell infiltrations predict neoadjuvant chemotherapy response of estrogen receptor-positive breast cancer. *Gland Surg*. 2021;10(1):154—165.
44. Choi BK, Kim SH, Kim YH, et al. RELT negatively regulates the early phase of the T-cell response in mice. *Eur J Immunol*. 2018;48(10):1739—1749.
45. Zhou WT, Jin WL. B7-H3/CD276: an emerging cancer immunotherapy. *Front Immunol*. 2021;12:701006.
46. Wang C, Li Y, Jia L, et al. CD276 expression enables squamous cell carcinoma stem cells to evade immune surveillance. *Cell Stem Cell*. 2021;28(9):1597—1613.e7.
47. Cervellera M, Raschella G, Santilli G, et al. Direct transactivation of the anti-apoptotic gene apolipoprotein J (clusterin) by B-MYB. *J Biol Chem*. 2000;275(28):21055—21060.
48. Lopez-Lopez E, Autry RJ, Smith C, et al. Pharmacogenomics of intracellular methotrexate polyglutamates in patients' leukemia cells *in vivo*. *J Clin Invest*. 2020;130(12):6600—6615.