

BCL7B is a potential novel diagnosis and prognosis biomarker for sarcomas using bioinformatics analysis

Dinglong Yang, MM^a, Xiaodong Gu, MD^b, Chunjiang Li, MM^b, Junjun Shi, MD^b, Yujing Chen, MM^c, Mingjie Dong, MM^a, Zhiqiang Zhang, MD^{b,*}

Abstract

BCL7B plays a potential role in the progression of various cancers, while its role in sarcomas is unknown. We aimed to evaluate BCL7B's diagnostic and prognostic value, and potential BCL7B-related mechanisms in sarcomas based on The Cancer Genome Atlas (TCGA) database. We collected patients with sarcoma from TCGA. Wilcoxon rank sum test was used to compare the expression of BCL7B in sarcoma samples with different clinical-pathologic features. Univariate Cox regression analysis and multivariate Cox regression analysis were used to evaluate prognosis factors for sarcoma. Gene set enrichment analysis (GSEA) was conducted to elucidate the significant functions and pathways associated with BCL7B. BCL7B was a potential biomarker for distinguishing normal and tumor tissues with the analysis of ROC curve (AUC=0.588). Low BCL7B expression was significantly correlated with tumor multifocal (OR=0.39 for yes vs no), larger residual tumor (OR=0.40 for R1,R2 vs RO), male gender (OR=0.48 for male vs female) and White race (OR=0.29 for White vs Asian, Black or African American). High BCL7B expression was correlated with leiomyosarcoma histological type (OR=6.08 for leiomyosarcoma vs dedifferentiated liposarcoma, pleomorphic sarcoma). Univariate and multivariate Cox regression analysis showed that low BCL7B expression was independently associated with poor overall survival ($P=.008$). GSEA showed that GPCR (G protein-coupled receptors) ligand binding, secreted factors, class A1 rhodopsin-like receptors, extracellular matrix organization, core matrisome, Fc epsilon receptor I mediated NF- κ B activation, and WNT signaling pathway were differentially enriched in BCL7B low expression phenotype ($|NES| > 1$, adjusted P value $< .05$, and FDR value < 0.25). BCL7B may play an important role in sarcoma progression and may be a potential biomarker for prognosis and diagnosis in sarcomas.

Abbreviations: ACC = adrenocortical carcinoma, BCL7B = BAF chromatin remodeling complex subunit BCL7B, BCR = B cell receptor, BLCA = bladder urothelial carcinoma, BRCA = breast invasive carcinoma, CESC = cervical squamous cell carcinoma and endocervical adenocarcinoma, CHOL = cholangiocarcinoma, COAD = colon adenocarcinoma, CREB = cAMP response element-binding protein, DLBC = lymphoid neoplasm diffuse large B-cell lymphoma, DSS = disease-specific survival, ECM = extracellular matrix, ESCA = esophageal carcinoma, FcεRI = Fc epsilon receptor I, FDR = false discovery rate, GBM = glioblastoma multiforme, GPCR = G protein-coupled receptors, GSEA = gene set enrichment analysis, HNSC = head and neck squamous cell carcinoma, KICH = kidney chromophobe, KIRC = kidney renal clear cell carcinoma, KIRP = kidney renal papillary cell carcinoma, LAML = acute myeloid leukemia, LGG = brain lower grade glioma, LIHC = liver hepatocellular carcinoma, LUAD = lung adenocarcinoma, LUSC = lung squamous cell carcinoma, MESO = mesothelioma, NES = normalized enrichment score, OS = overall survival, OV = ovarian serous cystadenocarcinoma, PAAD = pancreatic adenocarcinoma, PCPG = pheochromocytoma and paraganglioma, PFI = progression-free interval, PRAD = prostate adenocarcinoma, READ = rectum adenocarcinoma, rRNA = ribosomal RNA, SARC = sarcoma, SKCM = skin cutaneous melanoma, STAD = stomach adenocarcinoma, STES = stomach and esophageal carcinoma, TCGA = The Cancer Genome Atlas, TGCT = testicular germ cell tumors, THCA = thyroid carcinoma, THYM = thymoma, UCEC = uterine corpus endometrial carcinoma, UCS = uterine carcinosarcoma, UVM = uveal melanoma.

Keywords: BCL7B, GSEA, prognosis, sarcoma, TCGA

Editor: Milind Chalishazar.

The authors have no funding and conflicts of interest to disclose.

All datasets generated and analyzed during the current study are available in the TCGA repository (<https://portal.gdc.cancer.gov/>): search term TCGA-SARC.

The datasets generated during and/or analyzed during the current study are publicly available.

^a Second Clinical Medical College, Shanxi Medical University, China, ^b Department of Orthopedics, Second Hospital of Shanxi Medical University, Taiyuan, China,

^c School of Public Health, Xi'an Jiaotong University, Xian, Shaanxi, China.

* Correspondence: Zhiqiang Zhang, Department of Orthopedics, Second Hospital of Shanxi Medical University, Taiyuan, Shanxi 030001, China (e-mail: 13393430930@163.com).

Copyright © 2021 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Yang D, Gu X, Li C, Shi J, Chen Y, Dong M, Zhang Z. BCL7B is a potential novel diagnosis and prognosis biomarker for sarcomas using bioinformatics analysis. *Medicine* 2021;100:28(e26632).

Received: 13 November 2020 / Received in final form: 24 March 2021 / Accepted: 23 June 2021

<http://dx.doi.org/10.1097/MD.00000000000026632>

1. Introduction

Sarcomas are a rare (12% of all human cancers), and heterogeneous group of malignant mesenchymal tumors of soft tissue or bone. The mortality of patients with rare cancers is higher than those with common cancers, because rare cancers can't be diagnosed in time and accurately, and their subsequent treatment is suboptimal and inadequate.^[1] Sarcomas are difficult to diagnose early, and different subtypes have considerable heterogeneity in terms of age, anatomical location, rate of progression, and prognosis.^[2] So far, the therapies of sarcomas are still mainly surgery, followed by radiotherapy and chemotherapy. But due to cell recurrences and drug resistance, the mortality of sarcomas remains unchanged.^[3] Therefore, it is necessary to identify reliable predictor for early detection and provide targets for treatment and prognosis evaluation of sarcomas. Molecules including DNA, RNA and protein are increasingly used as diagnostic and prognostic biomarkers and molecular targets in therapy.^[4]

The B-cell CLL/lymphoma 7 protein family, including BCL7A, BCL7B and BCL7C, was first discovered when BCL7A gene was involved in a complex translocation of a Burkitt lymphoma cell line.^[5] BCL7 family members play an important role in the development and progression of several cancers. For example, reduced expression of BCL7A may increase the risk of Burkitt lymphoma,^[5] mycosis fungoides,^[6] astrocytoma,^[7] non-Hodgkin's lymphoma^[8] and cutaneous T cell lymphoma.^[9] BCL7B, also known as BCL tumor suppressor 7B, is a protein encoding-gene located on 7q11.23, which is thought to be deleted in the Williams-Beuren syndrome.^[10] A study had shown a hemizygous loss of 7q11.23 with the reduced expression of BCL7B in pilocytic astrocytomas.^[7] BCL7B expressed in tibial artery, gastrocnemius, bone marrow, etc with low tissue specificity and cell specificity (The Human Protein Atlas: <https://www.proteinatlas.org/ENSG00000106635-BCL7B>). BCL7B knock-down suppressed cell death, promoted the multinuclei phenotype and induced nuclear enlargement in KATOIII human gastric cancer cells. The expression of β -catenin and high-mobility group A1, which are components of WNT signaling pathway, was significantly upregulated in BCL7B-knockdown KATOIII cells.^[11] These results indicated that BCL7B may be a negative regulator of WNT signaling pathways, which was consistent with our study. Besides, BCL7B may play a role in the development and progression of mouse lung tumors.^[12] Apart from cancers, BCL7B of SWI/SNF (SWItching defective/Sucrose Non-Fermenting) complex was found to contribute to alcohol dependence.^[13] The role of BCL7B in tumors has been only partly explored, and its effects in the progression of sarcomas remain unclear. Hence, whether BCL7B can be used as a biomarker for diagnosis and treatment of sarcomas is of vital importance.

In order to explore the correlation between BCL7B and sarcomas and analyze the prognostic correlation, we analyzed the expression level of BCL7B in sarcomas and normal tissues by using RNA-seq data in The Cancer Genome Atlas (TCGA) database. Furthermore, Gene set enrichment analysis (GSEA)^[14] was performed on high and low expression groups to detect BCL7B related biological pathways involved in sarcomas. Immune infiltration analysis can estimate the relative abundance of a certain cell type by using an expression matrix. So, we explored the possible development mechanism of sarcomas by analyzing the immune infiltration related to BCL7B expression.

Our findings may reveal potential therapeutic targets and molecular mechanisms of sarcomas.

Our results indicated that BCL7B was a potential treatment and diagnosis biomarker of sarcomas. In addition, GSEA showed that several functions and signaling pathways were related to low BCL7B expression, such as GPCR ligand binding, secreted factors, class A1 rhodopsin-like receptors, extracellular matrix organization, core matrisome, cytokine-cytokine receptor interaction, and WNT signaling pathway. Our findings suggest that a low level of BCL7B predicts poor prognosis and may contribute to the early detection in patients with sarcomas.

2. Materials and methods

2.1. RNA-sequencing data and bioinformatics analysis

The TCGA database (<https://tcgadata.nci.nih.gov/tcga/>) collected cases from sarcoma projects. The RNA-seq data without clinical information were excluded. The RNA-seq data (level 3 HTSeq-FPKM) and clinical features of 259 cases were extracted for further analysis. Then level 3 HTSeq-FPKM was transformed into TPM (transcripts per million reads) for the following analysis. Unavailable or unknown clinical features were considered as missing values. In the end, according to the level of BCL7B expression (the median as cut-off), the sarcoma samples were divided into high and low expression groups. Our study was in accordance with the publication guidelines provided by TCGA.

2.2. Gene set enrichment analysis

GSEA, according to the gene expression matrix, can predict gene-related phenotypes and signaling pathways by analyzing the enrichment difference of pathways between high and low expression group.^[14] In this study, to elucidate the significant differences of pathways and functions between high and low expression BCL7B groups, GSEA was conducted using R package clusterProfiler (3.14.3).^[15] The gene set permutations were performed 1000 times for each analysis. And the expression level of BCL7B was regarded as a phenotype label. The adjusted *P* value ($< .05$), FDR *q* value (< 0.25) and normalized enrichment score ($|NES| > 1$) were used to classify the pathways enriched in each phenotype.

2.3. Immune infiltration analysis by ssGSEA

In this study, we used the ssGSEA (single-sample Gene Set Enrichment Analysis) method of GSVA package (<http://www.bioconductor.org/packages/release/bioc/html/gsva.htm>) to analyze the immune infiltration of 24 immunocytes in tumor samples. The following immunocytes were obtained: CD4+ T cells, CD8+ T cells, B cells, natural killer (NK) cells, CD56bright natural killer cells (CD56+ NK), CD56dim natural killer cells (CD56- NK), dendritic cells (DCs), central memory T cells (Tcm), eosinophils, effector memory T cells (Tem), activated dendritic cells (aDCs), gamma delta T cells ($\gamma\delta$ T), immature dendritic cells (iDCs), mast cells, neutrophils, macrophages, plasmacytoid dendritic cells (pDCs), cytotoxic cells, T helper cells, regulatory T cells (Tregs), type-1T helper cells (Th1), follicular helper T cells (Tfh), type-2T helper cells (Th2), and type-17T helper cells (Th17). The relative infiltration levels of immunocytes in tumor were quantified based on the genes of the 24 immunocytes in published signature gene lists.^[16] The correlation between BCL7B and the enrichment

Table 1
Characteristics of patients with sarcomas based on TCGA.

Characters	Level	Low expression of BCL7B	High expression of BCL7B	P	Test
n		130	129		
Tumor depth (%)	Deep	99 (93.4%)	86 (86.0%)	.128	
	Superficial	7 (6.6%)	14 (14.0%)		
Tumor multifocal (%)	No	94 (77.0%)	103 (89.6%)	.017	
	Yes	28 (23.0%)	12 (10.4%)		
Residual tumor (%)	R0	66 (56.4%)	88 (76.5%)	.003	Exact
	R1	44 (37.6%)	25 (21.7%)		
	R2	7 (6.0%)	2 (1.7%)		
Metastasis (%)	No	57 (71.2%)	63 (65.6%)	.525	
	Yes	23 (28.7%)	33 (34.4%)		
Margin status (%)	Negative	61 (61.6%)	75 (68.2%)	.396	
	Positive	38 (38.4%)	35 (31.8%)		
Radiation therapy (%)	No	94 (73.4%)	83 (66.4%)	.279	
	Yes	34 (26.6%)	42 (33.6%)		
Gender (%)	Female	59 (45.4%)	82 (63.6%)	.005	
	Male	71 (54.6%)	47 (36.4%)		
Race (%)	Asian	0 (0.0%)	6 (4.8%)	.009	Exact
	Black or African American	6 (4.8%)	12 (9.7%)		
	White	120 (95.2%)	106 (85.5%)		
Age (median [IQR])		61.00[50.25,73.00]	61.00[53.00,68.00]	.948	nonnorm

Exact = Fisher's exact test, nonnorm = non-normal distribution.

score of these 24 immune cells was analyzed by Spearman correlation.

2.4. Statistical analysis

In this study, all statistical analysis was carried out by using R (v3.6.2). Wilcoxon rank sum test was used to compare the expression of BCL7B in sarcoma samples with different clinicopathologic features. And we used univariate logistic regression to evaluate the relationships between clinical-pathologic features and the high and low expression of BCL7B. Kaplan–Meier method (Kaplan–Meier curve and log-rank test) with 95% confidence intervals (95% CI) was used to evaluate the follow-up duration. The survival analysis and prognostic factors such as residual tumor, age, race and gender were evaluated by using univariate Cox regression analysis and multivariate Cox regression analysis. The prognostic data is obtained from LIU.^[17] In all tests, hypothetical tests were 2-sided and P values <.05 were considered statistically significant.

2.5. Ethics statement

All analyses were based on the public GEO and TCGA databases, we did not need the informed consent of the patients, thus no ethical approval and patient consent are required.

3. Results

3.1. Association between BCL7B expression and clinicopathologic features

The characteristics of sarcoma patients in TCGA including tumor depth, tumor multifocal, residual tumor, metastasis, radiation therapy, gender, race, age and margin status were shown in Table 1. A total of 259 cases were analyzed, including 118 males and 141 females. The expression level of BCL7B was low in 130 cases (50.2%) and high in 129 cases (49.8%). The median age was 61 years (IQR 50.25–73) in low expression group of BCL7B

compared with 61 years (IQR 53–68) in high expression group. BCL7B was identified from various genes, because there were association between BCL7B expression and several clinicopathologic features of sarcoma, and BCL7B was a potential diagnosis biomarker for sarcomas. As showed in Figure 1, BCL7B was a potential biomarker for distinguishing normal and sarcoma tissues with the analysis of ROC curve (AUC=0.588). For further analysis, the BCL7B expression was significantly different (P < .05) in ACC, BLCA, CESC, CHOL, COAD, DLBC, ESCA, GBM, HNSC, KIRC, KIRP, LGG, LIHC, LUAD, LUSC, OV,

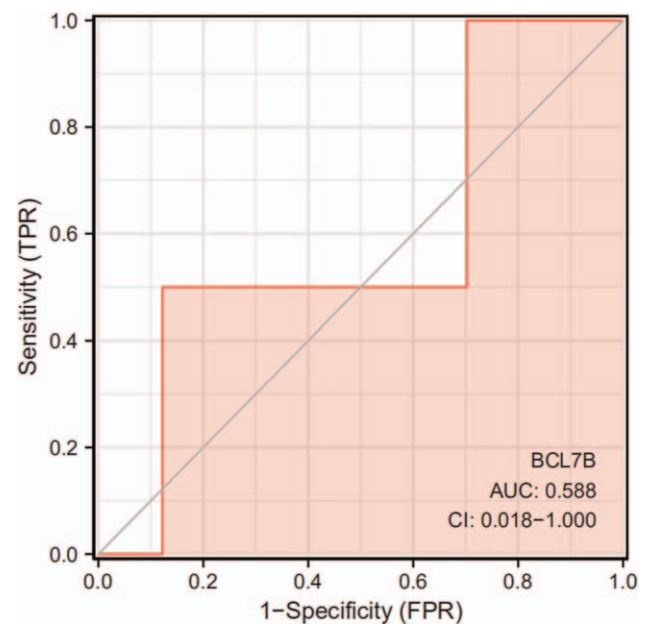


Figure 1. ROC curve showed the efficacy of BCL7B to distinguish sarcomas from normal tissues. The abscissa and ordinate represent the false positive rate and true positive rate respectively.

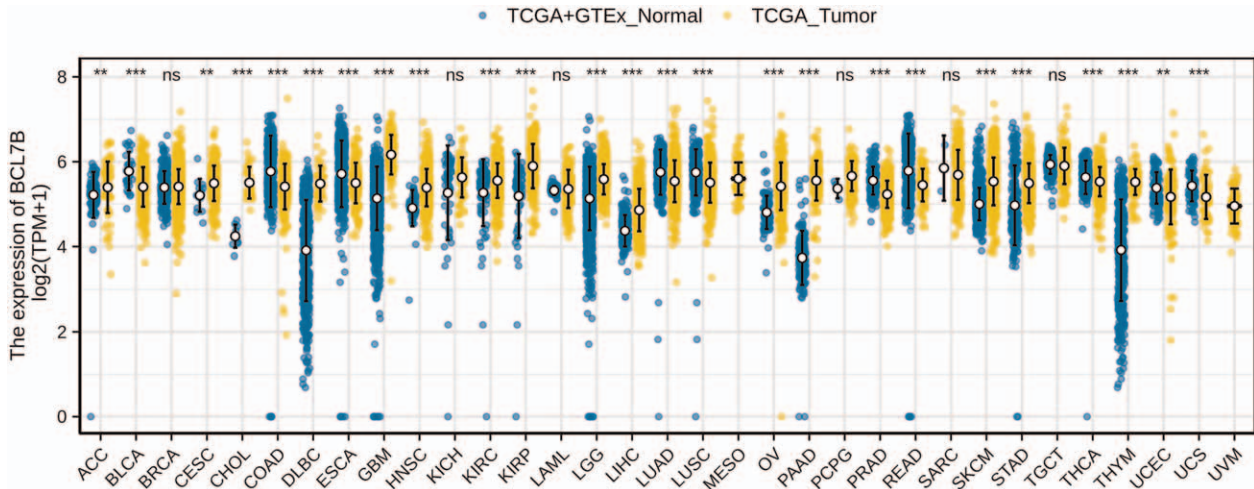


Figure 2. The BCL7B expression in 33 tumors compared with normal tissues. ns, $P \geq .05$; *, $P < .05$; **, $P < .01$; ***, $P < .001$.

PAAD, PRAD, READ, SKCM, STAD, THCA, THYM, UCEC, and UCS compared with normal tissues (Fig. 2). The Wilcoxon rank sum test showed that the expression level of BCL7B was significantly correlated with histological type (leiomyosarcoma vs

dedifferentiated liposarcoma vs pleomorphic sarcoma, $P < .001$, Fig. 3A), gender (male vs female, $P = .018$, Fig. 3B), residual tumor (R0 vs R1 vs R2, $P = .012$, Fig. 3C) and tumor multifocal (no vs yes, $P = .026$, Fig. 3D), but not correlated with tumor depth

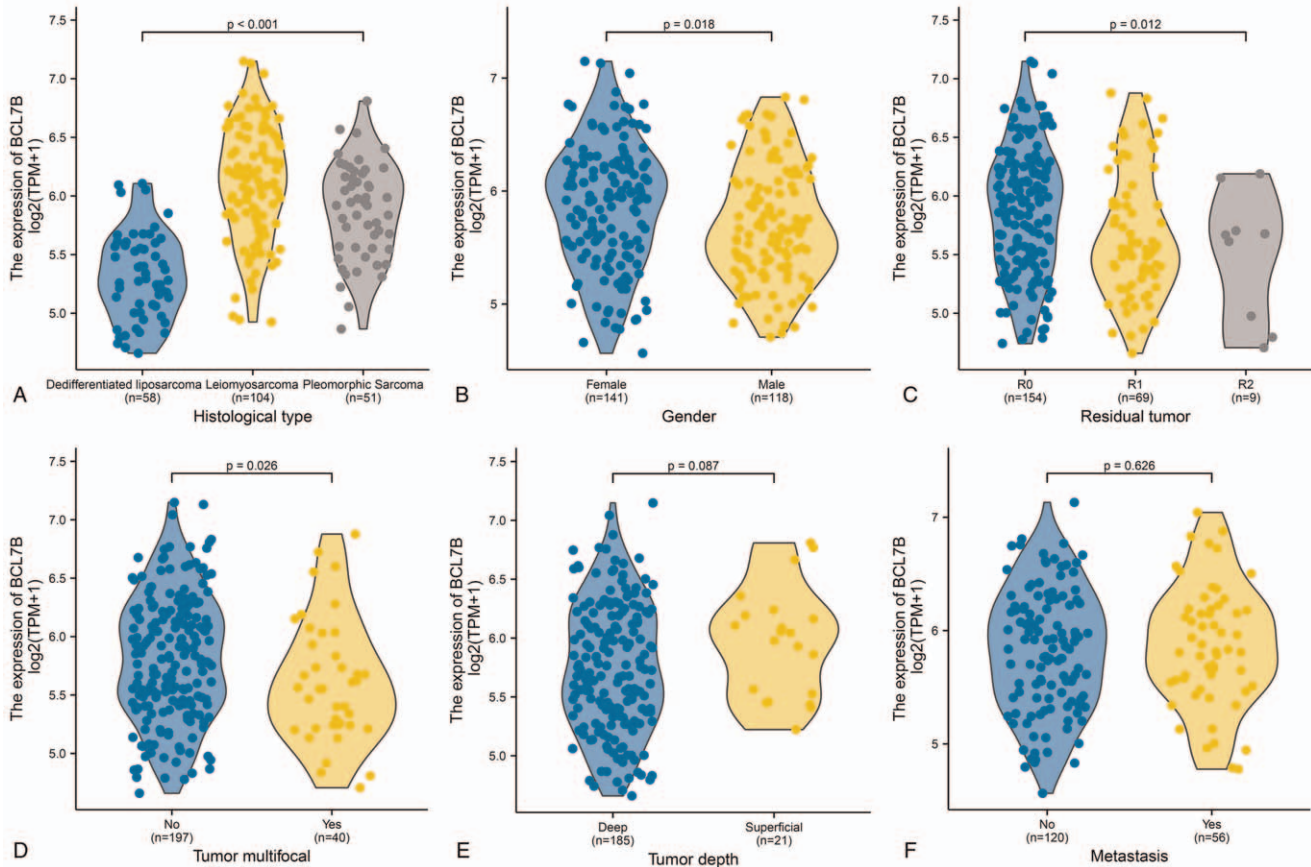


Figure 3. Association between BCL7B expression and clinicopathologic characteristics. (A) The higher expression level of BCL7B is associated with leiomyosarcoma, $P < .001$. (B) The expression level of BCL7B is higher in females than males, $P = .018$. (C) The higher expression level of BCL7B is associated with R0 residual tumor, $P = .012$. (D) The higher expression level of BCL7B is associated with no tumor multifocal, $P = .026$. (E) The expression level of BCL7B is not associated with tumor depth, $P = .087$. (F) The expression level of BCL7B is not associated with metastasis, $P = .626$.

Table 2
BCL7B expression* associated with clinicopathologic characteristics.

Characteristics	Total (N)	Odds ratio (OR)	P value
Tumor depth (deep vs superficial)	206	0.43 (0.16–1.09)	.086
Tumor multifocal (yes vs no)	237	0.39 (0.18–0.80)	.012
Residual tumor (R1&R2 vs R0)	232	0.40 (0.22–0.69)	.001
Metastasis (yes vs no)	176	1.30 (0.69–2.48)	.425
Radiation therapy (yes vs no)	253	1.40 (0.82–2.41)	.223
Gender (male vs female)	259	0.48 (0.29–0.78)	.003
Race (White vs Asian&Black or African American)	250	0.29 (0.10–0.73)	.013
Age (≤ 60 vs > 60)	259	1.02 (0.62–1.65)	.951
Margin status (positive vs negative)	209	0.75 (0.42–1.32)	.321
Histological type (leiomyosarcoma vs dedifferentiated liposarcoma&pleomorphic sarcoma)	213	6.08 (3.39–11.21)	<.001

* Categorical dependent variable, greater or less than the median expression level.

(deep vs superficial, $P = .087$, Fig. 3E) and metastasis (no vs yes, $P = .626$, Fig. 3F). As showed in Table 1, the results of Fisher’s exact test and chi-square test indicated that BCL7B expression level was associated with tumor multifocal ($P = .017$), residual tumor ($P = .003$), gender ($P = .005$) and race ($P = .009$). In addition, logistic regression revealed that reduced BCL7B expression was significantly associated with tumor multifocal (OR=0.39 for yes vs no), larger residual tumor (OR=0.40 for R1,R2 vs R0), male gender (OR=0.48 for male vs female), and White race (OR=0.29 for White vs Asian, Black or African American), and increased BCL7B expression was associated with leiomyosarcoma histological type (OR=6.08 for leiomyosarcoma vs dedifferentiated liposarcoma, pleomorphic sarcoma) with all P values $< .05$ (Table 2). No correlation was found between BCL7B expression and other clinical features. These results indicated that lower BCL7B expression was associated with poor clinicopathologic characteristics.

3.2. Survival outcomes and analysis

Kaplan–Meier survival analysis was used to evaluate the correlation between BCL7B expression level and prognosis. As showed in Figure 4, lower BCL7B expression level was associated

with poor overall survival of sarcoma patients ($P = .004$). By performing univariate Cox regression analysis, the meaningful variables ($P < .1$) were incorporated into multivariate analysis, including tumor multifocal ($P < .001$), residual tumor ($P < .001$), metastasis ($P < .001$), margin status ($P = .013$) and BCL7B ($P = .004$). A multivariate Cox regression analysis indicated that BCL7B was independently associated with overall survival ($P = .008$), along with metastasis ($P < .001$) (Table 3).

3.3. BCL7B-related functions and signaling pathways based on GSEA

GSEA was performed to identify functions and signaling pathways associated with sarcomas between low and high BCL7B expression data sets. Significant differences (FDR $< .05$, adjust P value $< .05$) in enrichment of MSigDB Collection (c2.cp.v7.0.symbols.gmt) were observed. The GSEA results showed that 182 functions and signaling pathways were significantly enriched based on $|NES| > 1$, adjusted P value $< .05$, and FDR value < 0.25 (Table 4). In particular, 16 pathways, including GPCR ligand binding, secreted factors, class A1 rhodopsin like receptors, extracellular matrix organization, core matrisome, cytokine-cytokine receptor interaction, WNT signaling pathway, signaling

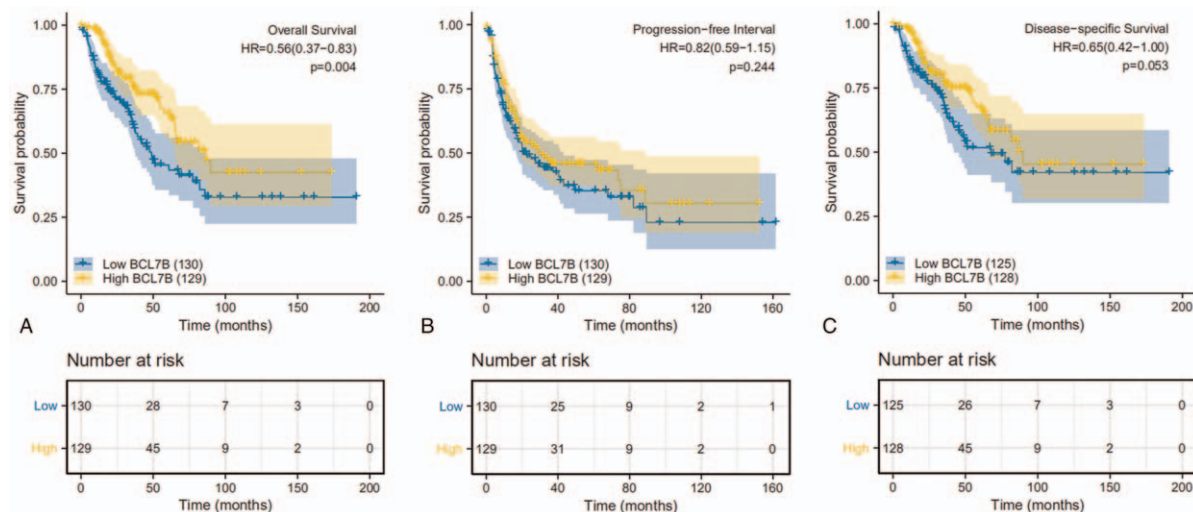


Figure 4. The Kaplan–Meier analysis of OS, PFI and DSS about BCL7B expression. (A) Lower BCL7B expression level was associated with poor OS of sarcoma patients, $P = .004$. (B) BCL7B expression level was not associated with PFI of sarcoma patients $P = .244$. (C) BCL7B expression level was not associated with DSS of sarcoma patients, $P = .053$. DSS=disease-specific survival, OS=overall survival, PFI=progression-free interval.

Table 3
Association between overall survival and clinicopathologic characteristics in sarcoma patients.

Characteristics	Total (N)	HR (95% CI)		P value	
		Univariate analysis	Univariate analysis	Multivariate analysis	Multivariate analysis
Tumor depth (deep vs superficial)	206	2.920 (0.920–9.274)	.069	2.758 (0.362–21.028)	.328
Tumor multifocal (yes vs no)	237	2.404 (1.502–3.847)	<.001	1.518 (0.617–3.732)	.363
Residual tumor (R1&R2 vs R0)	232	2.554 (1.668–3.910)	<.001	1.604 (0.595–4.328)	.35
Metastasis (yes vs no)	176	3.014 (1.834–4.954)	<.001	3.071 (1.698–5.555)	<.001
Margin status (positive vs negative)	209	1.840 (1.138–2.974)	.013	1.456 (0.569–3.723)	.433
Radiation therapy (yes vs no)	253	0.850 (0.545–1.325)	.473		
Age (>60 vs ≤60)	259	1.354 (0.908–2.019)	.137		
Gender (male vs female)	259	0.855 (0.572–1.277)	.443		
Race (White vs Asian&Black or African American)	250	0.736 (0.356–1.525)	.41		
BCL7B (high vs low)	259	0.555 (0.371–0.831)	.004	0.424 (0.225–0.801)	.008

by ROBO receptors, rRNA processing, cell surface interactions at the vascular wall, peptide ligand binding receptors, ECM (extracellular matrix) glycoproteins, signaling by the B cell receptor BCR, influenza infection, Fc epsilon receptor I (FcεRI) mediated NF-κB activation, and TNFR2 non-canonical NF-κB pathway showed significantly differential enrichment in BCL7B low expression phenotype (Fig. 5).

3.4. The correlation between BCL7B expression and immune infiltration

The correlation between the expression level (TPM) of BCL7B and immune infiltration level (quantified by ssGSEA) was analyzed by Spearman correlation in sarcomas environment.

We found that Tfh were significantly negatively correlated with BCL7B expression (Spearman $R = -0.457$, $P < .001$). Besides, the expression level of BCL7B was negatively correlated with the infiltration of CD8T cells, eosinophils, Th1 cells, iDCs, Treg, macrophages, and T helper cells, and positively correlated with the infiltration of Tgd, Th2 cells and mast cells (Figs. 6 and 7).

4. Discussion

BCL7B, which is a protein encoding gene located on 7q11.23, belongs to a conserved gene family involved in early embryonic development.^[18] It was known that 7q11.23 was deleted with the reduction of BCL7B expression level in the Williams-Beuren contiguous gene deletion syndrome and pilocytic astrocyto-

Table 4
Gene set enrich analysis in high and low expression groups of BCL7B.

ID	setSize	enrichmentScore	NES	P value	P.adj	FDR	Rank	Leading_edge
REACTOME_GPCR_LIGAND_BINDING	437	-0.469	-1.739	.001	.027	0.022	8947	tags=39%, list=17%, signal=32%
REACTOME_CLASS_A_1_RHODOPSIN_LIKE_RECEPTORS	309	-0.448	-1.633	.001	.027	0.022	8929	tags=39%, list=17%, signal=33%
REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION	296	-0.48	-1.75	.001	.027	0.022	10194	tags=43%, list=20%, signal=34%
NABA_CORE_MATRISOME	265	-0.544	-1.966	.001	.027	0.022	6111	tags=42%, list=12%, signal=37%
KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	251	-0.432	-1.552	.001	.027	0.022	10502	tags=42%, list=21%, signal=33%
REACTOME_SIGNALING_BY_ROBO_RECEPTORS	210	-0.481	-1.71	.001	.027	0.022	12830	tags=44%, list=25%, signal=33%
REACTOME_RRNA_PROCESSING	191	-0.509	-1.793	.001	.027	0.022	14573	tags=46%, list=28%, signal=33%
REACTOME_CELL_SURFACE_INTERACTIONS_AT_THE_VASCULAR_WALL	190	-0.662	-2.333	.001	.027	0.022	4188	tags=36%, list=8%, signal=33%
REACTOME_PEPTIDE_LIGAND_BINDING_RECEPTORS	182	-0.525	-1.846	.001	.027	0.022	8828	tags=46%, list=17%, signal=38%
NABA_ECM_GLYCOPROTEINS	186	-0.502	-1.765	.001	.027	0.022	6481	tags=40%, list=13%, signal=35%
REACTOME_SIGNALING_BY_THE_B_CELL_RECEPTOR_BCR	161	-0.759	-2.649	.001	.027	0.022	2737	tags=34%, list=5%, signal=32%
KEGG_WNT_SIGNALING_PATHWAY	150	-0.485	-1.679	.001	.027	0.022	11617	tags=35%, list=23%, signal=27%
REACTOME_INFLUENZA_INFECTION	151	-0.536	-1.856	.001	.027	0.022	12618	tags=50%, list=25%, signal=38%
REACTOME_FCERI_MEDIATED_NF_KB_ACTIVATION	132	-0.801	-2.748	.001	.027	0.022	2737	tags=36%, list=5%, signal=35%
REACTOME_TNFR2_NON_CANONICAL_NF_KB_PATHWAY	96	-0.518	-1.717	.001	.027	0.022	10269	tags=27%, list=20%, signal=22%

Gene sets with |NES| > 1, adjusted P value < .05, and FDR value < 0.25 are considered as significant. FDR=false discovery rate, NES=normalized enrichment score, P.adj=adjusted P value.

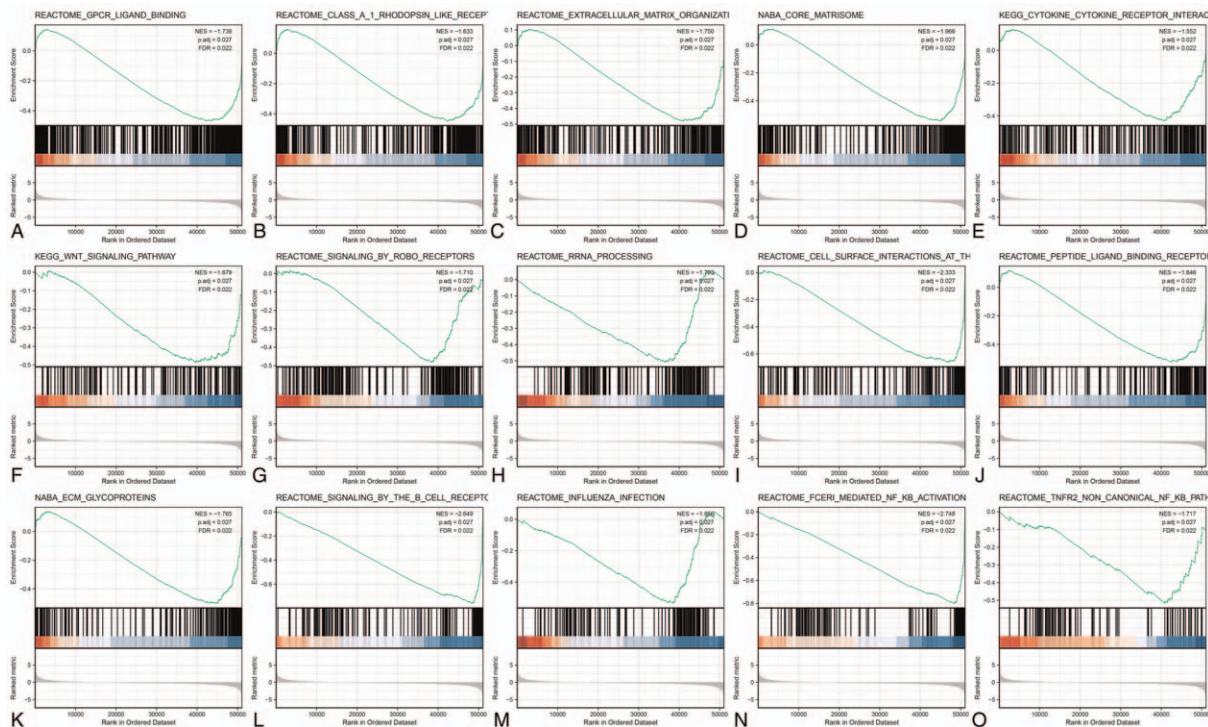


Figure 5. Enrichment plots from the gene set enrichment analysis (GSEA). (A–O) Several BCL7B-related pathways and functions were observed in sarcomas, including GPCR ligand binding, secreted factors, class A1 rhodopsin like receptors, extracellular matrix organization, core matrisome, cytokine-cytokine receptor interaction, WNT signaling pathway, signaling by ROBO receptors, rRNA processing, cell surface interactions at the vascular wall, peptide ligand binding receptors, ECM glycoproteins, signaling by the B cell receptor BCR, influenza infection, and TNFR2 non-canonical NF-κB pathway. FDR=false discovery rate, NES= normalized ES, *P*.adj=adjusted *P* value.

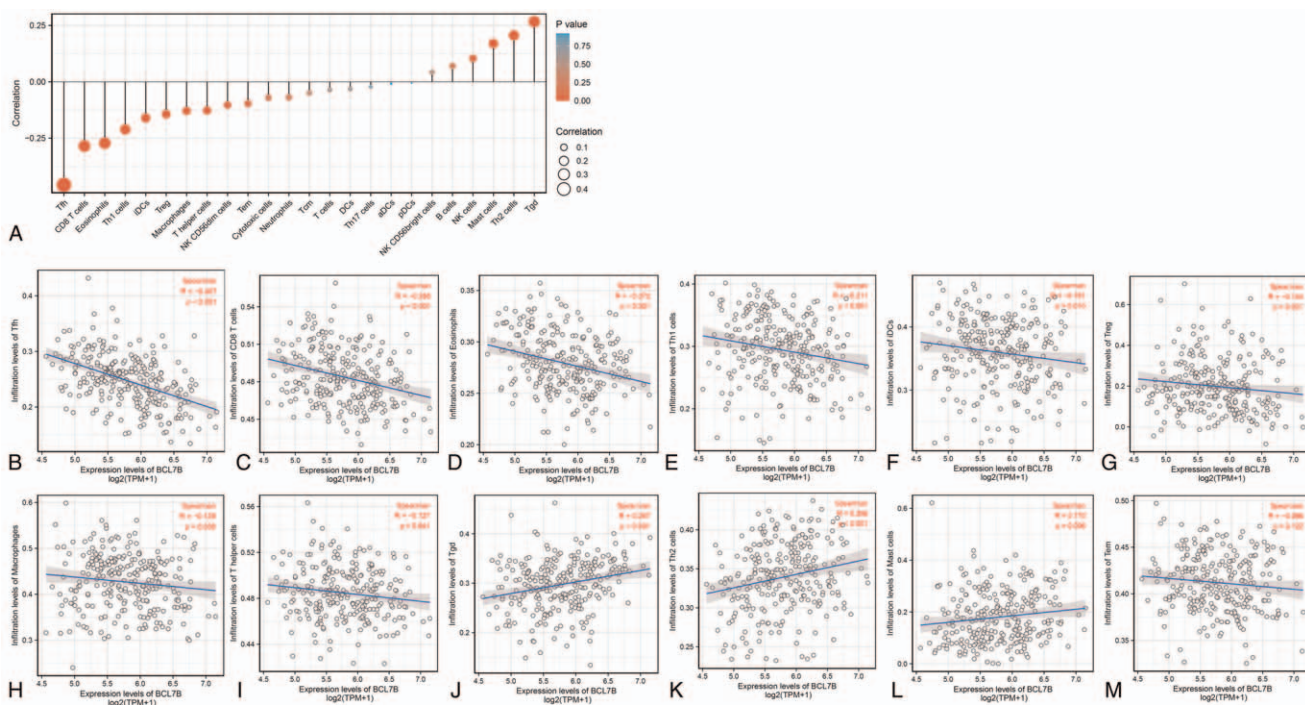


Figure 6. The expression level of BCL7B was associated with the immune infiltration in the tumor microenvironment. (A) Correlation between BCL7B expression level and the relative abundances of 24 immune cells. (B–M) Correlation between BCL7B expression level and the relative enrichment score of Tfh, CD8T cells, eosinophils, Th1 cells, iDCs, Treg, macrophages, T helper cells, Tgd, Th2 cells, mast cells, and Tem. iDCs=immature DCs, Tem=T effector memory, Tfh=T follicular helper, Tgd=T gamma delta, Th=helper T cells, Treg=regulatory T cells.

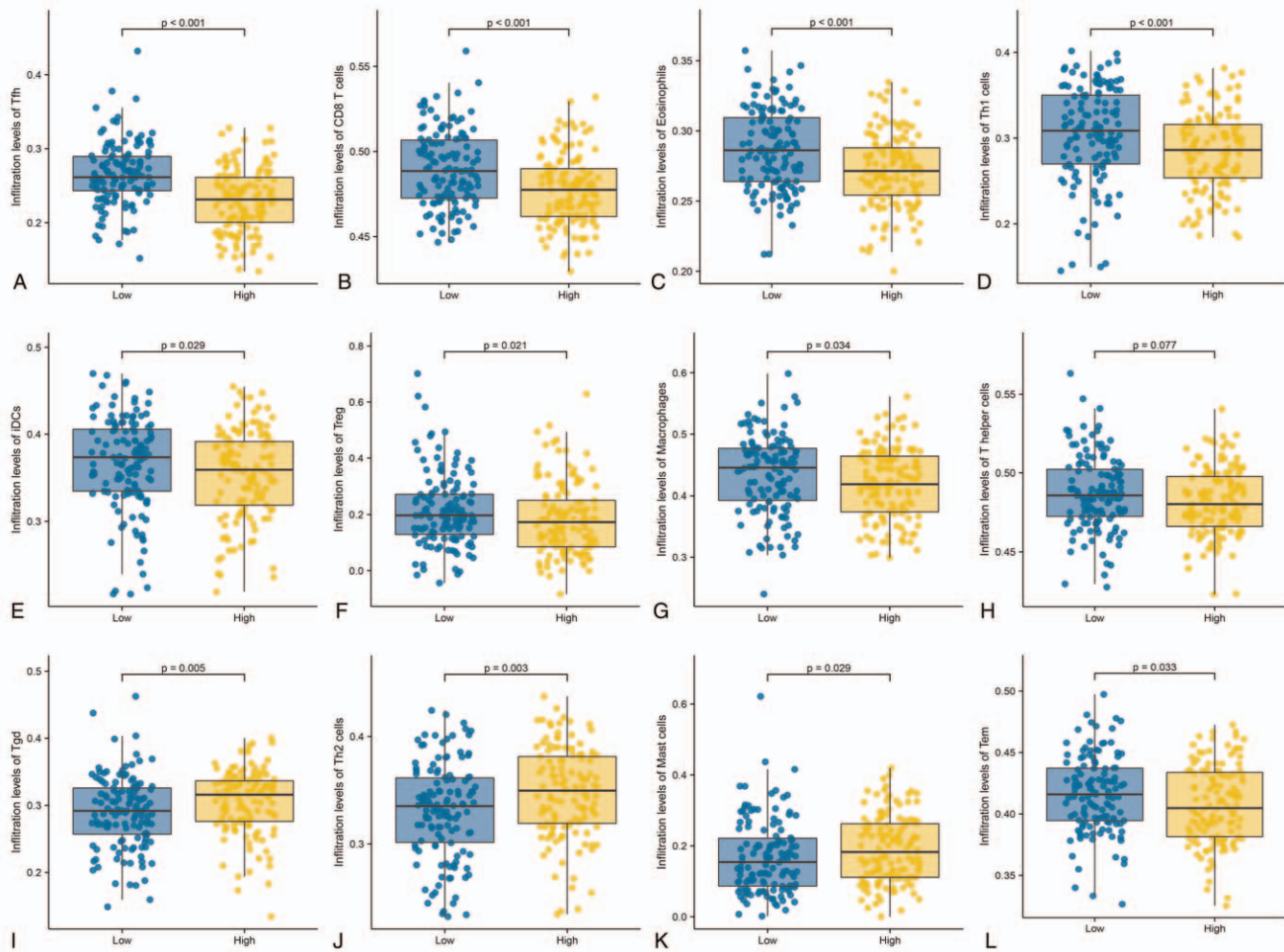


Figure 7. The expression level of BCL7B was associated with the immune infiltration in the tumor microenvironment. (A–L) The infiltration levels of Tfh, CD8T cells, eosinophils, Th1 cells, iDCs, Treg, macrophages, T helper cells, Tgd, Th2 cells, mast cells, and Tem in high and low expression groups of BCL7B. iDCs=immature DCs, Tem=T effector memory, Tfh=T follicular helper, Tgd=T gamma delta, Th=helper T cells, Treg=regulatory T cells.

mas.^[7,10,19] But its expression level in sarcomas is still unclear. Previous study showed that BCL7B/BCL-7 played an important role in maintaining the structure of nuclei and the absence of BCL7B may lead to the risk of malignant tumors.^[11] However, the high BCL7B expression was associated with poor overall survival of pancreatic cancer patients and BCL7B promoted the motility and invasion of pancreatic cancer cells through the dephosphorylation of cAMP response element-binding protein (CREB).^[18] This indicates that BCL7B may have opposite effects on the progression of different cancers. WNT and apoptosis signaling pathways are found to be closely related to BCL7B in tumors, which was consistent with our study.^[11] Despite BCL7B has been studied, its diagnostic and prognostic value is still unclear in sarcomas. In this study, a bioinformatics analysis in sarcomas was performed to evaluate the expression level, diagnostic and prognostic value by using RNA-seq data obtained from TCGA.

The BCL7B expression in sarcomas was associated with clinicopathologic variables (histological type, gender, residual tumor and tumor multifocal), patient characteristics (tumor multifocal, residual tumor, gender and race) and poor overall survival. These results indicated that BCL7B was a potential biomarker for early diagnosis and prognosis of sarcomas.

GSEA was performed to further investigate the function and associated mechanism of BCL7B in sarcomas. Some functions of BCL7B in sarcomas have been reported previously. In this study, we found that FcεRI mediated NF-κB activation in sarcomas was differentially enriched in BCL7B low-expression phenotype. NF-κB is a family of eukaryotic transcription factors.^[20] Some downstream genes of NF-κB are closely related with tumors, such as CyclinD1 and c-Myc. Continual activation of NF-κB leads to uncontrolled cell growth, inhibition of apoptosis, tumor metastasis and angiogenesis.^[21] So the knockdown of BCL7B may result in suppressed cell death. A study has found that glucagon-like peptide 2 inhibits the growth of osteosarcoma cells by inhibiting the activity of NF-κB, triggering the decrease of pyruvate kinase M2, Myc, and CyclinD1.^[22] These findings suggested that BCL7B may participate in the regulation of cancer cell growth during the development of sarcoma, but the underlying regulatory mechanism needed further clarification. GPCR ligand binding, extracellular matrix organization, cytokine-cytokine receptor interaction, WNT signaling pathway, cell surface interactions at the vascular wall, ECM glycoproteins, signaling by the B cell receptor BCR and TNFR2 non-canonical NF-κB pathway also showed significantly differential enrichment

in BCL7B low expression phenotype. They may be the potential mechanism of sarcoma progression and metastases.

Another important aspect of this study was that BCL7B expression was significant positively correlated with the infiltration of Tgd and negatively correlated with the infiltration of Tfh. Tgd cells are important immunocytes which are activated in varieties of tumors *in vivo*.^[23] Because of their ability to produce inflammatory cytokines that enhance the anti-tumor properties of other immunocytes and directly kill tumor cells, Tgd cells contribute to protection against cancers.^[24] *In vitro*, Tgd cells lysed hepatoma cells and significantly reduced the activity of hepatoma cells.^[25] According to previous studies, activated (CD44high) V γ 4+ γ δ T cells participated in immune surveillance by secreting IFN- γ in TCR δ -/- mice.^[26] Our study indicated that BCL7B may play a potential role in tumor resistance owing to the infiltration of Tgd cells in sarcoma microenvironment. Tfh, which is a subgroup of helper T cells, plays an important role in human immunity and tumor development. For example, IL-12 secreted by Tfh improved the killing effects of NK cells on tumors^[27] and inhibited the proliferation and sprouting of activated vascular endothelial cells *in vitro*,^[28] through which Tfh can resist the development and metastasis of tumors. The knockout of BCL7B may improve the infiltration and anti-tumor efficient of Tfh owing to the potential negative correlation between BCL7B and Tfh. These findings indicate that BCL7B may play a potential role in immune cell infiltration and the treatment of sarcomas. But we can't conclude if BCL7B is good or bad for sarcoma based on available evidence, there are many factors involved in the development of sarcoma except for immune cells.

Although this study improves our knowledge of the relationship between BCL7B and sarcomas, there are still some limitations. First, our study is just bioinformatics analysis, the results should be verified by using clinical samples, and cell and animal experiments are also needed for exploring mechanism. Based on these, we have planned to do some experiments in the sooner future. Secondly, our understanding of BCL7B is still not comprehensive, further multi-omics study is needed especially BCL7B protein level and its functional mechanism. Besides, retrospective studies have limits such as lacking of some information and non-uniform intervening measures. Therefore, prospective studies are required in the future to avoid bias analysis led by the retrospective studies.

5. Conclusion

In summary, our study revealed that BCL7B was significantly associated with poor survival, cancer progression, signaling pathways and immune infiltrations in sarcomas. BCL7B may be a potential biomarker for early diagnose and prognosis prediction of sarcomas. Further experiments should be performed to explore the mechanism and clinical value for sarcoma patients.

Author contributions

Conceptualization: Dinglong Yang, Zhiqiang Zhang.

Formal analysis: Dinglong Yang, Junjun Shi, Yujing Chen.

Funding acquisition: Chunjiang Li.

Investigation: Dinglong Yang, Xiaodong Gu.

Methodology: Dinglong Yang, Xiaodong Gu, Chunjiang Li, Junjun Shi, Zhiqiang Zhang.

Project administration: Zhiqiang Zhang.

Software: Dinglong Yang.

Supervision: Zhiqiang Zhang.

Validation: Zhiqiang Zhang.

Visualization: Dinglong Yang, Xiaodong Gu, Chunjiang Li.

Writing – original draft: Dinglong Yang.

Writing – review & editing: Dinglong Yang, Mingjie Dong, Zhiqiang Zhang.

References

- [1] Blay JY, Coindre JM, Ducimetiere F, Ray-Coquard I. The value of research collaborations and consortia in rare cancers. *Lancet Oncol* 2016;17:e62–9.
- [2] Soomers V, Husson O, Young R, Desar I, Van der Graaf W. The sarcoma diagnostic interval: a systematic review on length, contributing factors and patient outcomes. *ESMO Open* 2020;5:e000592.
- [3] Aung T, Asam C, Haerteis S. Ion channels in sarcoma: pathophysiology and treatment options. *Pflugers Arch* 2019;471:1163–71.
- [4] El Beaino M, Jupiter DC, Assi T, et al. Diagnostic value of TLE1 in synovial sarcoma: a systematic review and meta-analysis. *Sarcoma* 2020;2020:7192347.
- [5] Zani VJ, Asou N, Jadayel D, et al. Molecular cloning of complex chromosomal translocation t(8;14;12)(q24.1;q32.3;q24.1) in a Burkitt lymphoma cell line defines a new gene (BCL7A) with homology to caldesmon. *Blood* 1996;87:3124–34.
- [6] Carbone A, Bernardini L, Valenzano F, et al. Array-based comparative genomic hybridization in early-stage mycosis fungoides: recurrent deletion of tumor suppressor genes BCL7A, SMAC/DIABLO, and RHOA. *Genes Chromosomes Cancer* 2008;47:1067–75.
- [7] Potter N, Karakoula A, Phipps KP, et al. Genomic deletions correlate with underexpression of novel candidate genes at six loci in pediatric pilocytic astrocytoma. *Neoplasia* 2008;10:757–72.
- [8] Morton LM, Purdue MP, Zheng T, et al. Risk of non-Hodgkin lymphoma associated with germline variation in genes that regulate the cell cycle, apoptosis, and lymphocyte development. *Cancer Epidemiol Biomarkers Prev* 2009;18:1259–70.
- [9] van Doorn R, Zoutman WH, Dijkman R, et al. Epigenetic profiling of cutaneous T-cell lymphoma: promoter hypermethylation of multiple tumor suppressor genes including BCL7A, PTPRG, and p73. *J Clin Oncol* 2005;23:3886–96.
- [10] Perez Jurado LA, Peoples R, Kaplan P, Hamel BC, Francke U. Molecular definition of the chromosome 7 deletion in Williams syndrome and parent-of-origin effects on growth. *Am J Hum Genet* 1996;59:781–92.
- [11] Uehara T, Kage-Nakadai E, Yoshina S, Imae R, Mitani S. The tumor suppressor BCL7B functions in the Wnt signaling pathway. *PLoS Genet* 2015;11:e1004921.
- [12] Bonner AE, Lemon WJ, Devereux TR, Lubet RA, You M. Molecular profiling of mouse lung tumors: association with tumor progression, lung development, and human lung adenocarcinomas. *Oncogene* 2004;23:1166–76.
- [13] Mathies LD, Aliev F, Investigators C, Davies AG, Dick DM, Bettinger JC. Variation in SWI/SNF chromatin remodeling complex proteins is associated with alcohol dependence and antisocial behavior in human populations. *Alcohol Clin Exp Res* 2017;41:2033–40.
- [14] Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
- [15] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 2012;16:284–7.
- [16] Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity* 2013;39:782–95.
- [17] Liu J, Lichtenberg T, Hoadley KA, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* 2018;173:400–16. e411.
- [18] Taniuchi K, Furihata M, Naganuma S, Dabanaka K, Hanazaki K, Saibara T. BCL7B, a predictor of poor prognosis of pancreatic cancers, promotes cell motility and invasion by influencing CREB signaling. *Am J Cancer Res* 2018;8:387–404.
- [19] Decimi V, Fazio G, Dell'Acqua F, et al. Williams syndrome and mature B-Leukemia: a random association? *Eur J Med Genet* 2016;59:634–40.
- [20] Betzler AC, Theodoraki MN, Schuler PJ, et al. NF-kappaB and its role in checkpoint control. *Int J Mol Sci* 2020;21:3949.

- [21] Tang X, Cao T, Zhu Y, et al. PIM2 promotes hepatocellular carcinoma tumorigenesis and progression through activating NF-kappaB signaling pathway. *Cell Death Dis* 2020;11:510.
- [22] Lu Y, Lu D, Hu Y. GLP2 promotes directed differentiation from osteosarcoma cells to osteoblasts and inhibits growth of osteosarcoma cells. *Mol Ther Nucleic Acids* 2018;10:292–303.
- [23] Seki S, Abo T, Masuda T, et al. Identification of activated T cell receptor gamma delta lymphocytes in the liver of tumor-bearing hosts. *J Clin Invest* 1990;86:409–15.
- [24] Catros V, Toutirais O, Bouet F, Cabillic F, Desille M, Fournie JJ. Tgammadelta lymphocytes in oncology: unconventional killer lymphocytes. *Med Sci (Paris)* 2010;26:185–91.
- [25] Hoh A, Dewerth A, Vogt F, et al. The activity of gammadelta T cells against paediatric liver tumour cells and spheroids in cell culture. *Liver Int* 2013;33:127–36.
- [26] He W, Hao J, Dong S, et al. Naturally activated V gamma 4 gamma delta T cells play a protective role in tumor immunity through expression of eomesodermin. *J Immunol* 2010;185:126–33.
- [27] Denman CJ, Senyukov VV, Somanchi SS, et al. Membrane-bound IL-21 promotes sustained ex vivo proliferation of human natural killer cells. *PLoS One* 2012;7:e30264.
- [28] Castermans K, Tabruyn SP, Zeng R, et al. Angiostatic activity of the antitumor cytokine interleukin-21. *Blood* 2008;112:4940–7.