

RESEARCH ARTICLE

Open Access



# Bioinformatics integrated analysis to investigate candidate biomarkers and associated metabolites in osteosarcoma

Jun Wang, Mingzhi Gong, Zhenggang Xiong, Yangyang Zhao and Deguo Xing\*

## Abstract

**Background:** This study hoped to explore the potential biomarkers and associated metabolites during osteosarcoma (OS) progression based on bioinformatics integrated analysis.

**Methods:** Gene expression profiles of GSE28424, including 19 human OS cell lines (OS group) and 4 human normal long bone tissue samples (control group), were downloaded. The differentially expressed genes (DEGs) in OS vs. control were investigated. The enrichment investigation was performed based on DEGs, followed by protein–protein interaction network analysis. Then, the feature genes associated with OS were explored, followed by survival analysis to reveal prognostic genes. The qRT-PCR assay was performed to test the expression of these genes. Finally, the OS-associated metabolites and disease-metabolic network were further investigated.

**Results:** Totally, 357 DEGs were revealed between the OS vs. control groups. These DEGs, such as CXCL12, were mainly involved in functions like leukocyte migration. Then, totally, 38 feature genes were explored, of which 8 genes showed significant associations with the survival of patients. High expression of CXCL12, CEBPA, SPARCL1, CAT, TUBA1A, and ALDH1A1 was associated with longer survival time, while high expression of CFLAR and STC2 was associated with poor survival. Finally, a disease-metabolic network was constructed with 25 nodes including two disease-associated metabolites cyclophosphamide and bisphenol A (BPA). BPA showed interactions with multiple prognosis-related genes, such as CXCL12 and STC2.

**Conclusion:** We identified 8 prognosis-related genes in OS. CXCL12 might participate in OS progression via leukocyte migration function. BPA might be an important metabolite interacting with multiple prognosis-related genes.

**Keywords:** Osteosarcoma, Prognosis, Biomarkers, Metabolite

## Highlights

1. We identified 8 prognosis-related genes in OS.
2. CXCL12 might take part in OS via leukocyte migration function.
3. CXCL12 and STC2 might be used as novel biomarkers for OS.

4. Bisphenol A might be an important metabolite interacting with multiple genes.

## Background

Osteosarcoma (OS) is one of the most common malignant bone tumors in children and adolescents under 20 years old [1], accounting for about 5% of all children's tumors [2]. Current optimal treatment for OS consists of multi-agent chemotherapy and aggressive surgical resection of all sites of disease involvement [3]. Unfortunately, in patients with recurrent or metastatic OS, the

\* Correspondence: [xdgxdgtgy@qq.com](mailto:xdgxdgtgy@qq.com)

Department of Orthopedics and Trauma, The Second Hospital of Shandong University, No. 247 Beiyuan Street, Jinan 250033, China



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

long-term survival rate is only 20% [4]. Thus, it is important to further investigate the detailed pathological mechanism of OS.

The mutation of certain genes is responsible for syndromes that predispose to OS [5]. It has been proved that some genes are differentially expressed between OS samples and normal samples in humans, which can be further used for the prediction of chemotherapy response [6]. A previous study shows that the differentially expressed genes (DEGs) such as SRY-Box transcription factor 2 are taking part in the development of OS via influencing cell stemness and migration [7]. A recent study indicates that cyclin E1 is overexpressed in OS samples, which can be used as a prognostic biomarker and potential therapeutic target in OS [8]. Dong et al. proved that the lung adenocarcinoma transcript 1 that related with metastasis promoted the proliferation and metastasis of OS cells by stimulating the PI3K/Akt pathway [9]. Actually, the emerging biomarkers in OS have been revealed not only from genomics but also via metabolomics [10]. A previous study shows that catecholamines and their receptors can be potential molecular markers for OS [11]. It has been proved that metabolites such as parathyroid hormone peptides through the regulation of hyaluronan metabolism affect OS cell migration [12]. A previous global microarray analysis by Namløs et al. [13] revealed that several microRNA (miRNA)-gene interactions (such as miR-9/TGFBR2) were implicated in the development of OS. However, the potential biomarkers and their associated detail molecular mechanism during the OS progression is unknown.

Based on the microarray data provided by Namløs et al. [13], the current bioinformatics analysis was performed to reveal potential DEGs between human OS cell lines and human normal long bone tissue samples, followed by the function and pathway enrichment analysis based on these DEGs. The feature genes for OS were explored based on online databases, followed by the survival analysis to reveal prognostic genes. Finally, the OS-associated metabolites and networks were further investigated. Figure S1 shows the workflow of this study. We hoped to explore the potential biomarkers and associated molecular mechanisms during OS progression.

## Material and methods

### Data resource and preprocessing

A gene expression profile GSE28424 including 19 human OS cell lines (OS group) and 4 human normal long bone tissue samples (control group) was obtained from GEO database (platform: GPL570: Illumina HumanWG-6 v2.0 expression beadchip). As described in the study of Namløs et al. [13], the 19 human OS cell lines, HAL, HOS, 143B, IOR/MOS, IOR/OS9, IOR/OS10, IOR/OS14, IOR/OS15, IOR/OS18, SARG, KPD, MG-63, MHM, MNNG/HOS,

OHS, OSA, Saos-2, U-2 OS, and ZK-58 were derived from ATCC or different partner laboratories within EuroBioNeT. Cell line authentication was performed by STR DNA profiling using Powerplex 16. The four normal long bone tissue samples were obtained from amputations of cancer patients at the Norwegian Radium Hospital. The processed gene expression matrix file was obtained using Affy package in R software [14]. Probes with different gene symbols were excluded from this study, and the average value of genes matched to multiple probes was taken as the expression value of the probe.

### Differentially expression analysis

The limma package (version: 3.34.9) of R [15] was used to investigate DEGs between two groups based on the linear regression and empirical Bayesian methods. DEGs were screened by Benjamini & Hochberg (BH) adjusted  $P < 0.05$  and  $|\log_2 \text{fold change (FC)}| > 2$ . Then, the results were visualized by volcano plots and clustering heatmap.

### Functional enrichment analysis of DEGs

Gene ontology-biological process (GO-BP) function and The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of DEGs were performed using the clusterProfiler package (version: 3.2.11) in R software [16] with thresholds of  $P$  value  $< 0.05$  and count  $\geq 1$ .

### The difference analysis of KEGG between OS group and control group

Based on the enrichment background (c2.cp.kegg.v7.2.-symbols.gmt) in MSigDB v7.2 database [17], the enrichment scores of each KEGG in each sample of GSE28424 were calculated to obtain a scoring matrix using gene set variation analysis (GSVA) algorithm in R package. Then, the differential expression analysis between the OS group and control group was performed on each KEGG item using the limma package in R software. Finally, the adjusted  $P < 0.05$  was considered as the cut-off value for significantly different items.

### Protein-protein interaction (PPI) network

The Search Tool for the Retrieval of Interacting Genes (STRING) database (version: 11) [18] provides experimental and predicted interaction information. The hub-proteins associated with DEGs were selected according to the STRING database. Then PPI network was constructed by Cytoscape software (version: 3.6.1) [19] with medium confidence (score) = 0.7. The degree (number of the connections for the target protein) was used to evaluate the importance of the target gene.

### The feature gene prediction of OS

The relationship between disease (keyword: osteoporosis)-associated chemical molecular and target gene was revealed by Comparative Toxicogenomics Database (CTD) [20]. Then, based on VENN plot analysis, the screened genes in CTD and genes in GeneCards database [21] were intersected with nodes in the PPI network to explore the feature genes for OS.

### Survival analysis for feature genes

Survival analysis was used to identify biomarkers from significant feature genes. The  $\log_2(\text{fpkm}+1)$  expression data and clinical information of OS in the TARGET database were downloaded from the University of California Santa Cruz (UCSC) Genome Browser database (<https://xenabrowser.net/datapages/?cohort=GDC%20TARGET-OS&removeHub=https%3A%2F%2Fxenatreehouse.gi.ucsc.edu%3A443>) [22]. The results were visualized using Kaplan-Meier plots. The samples were divided into two groups (high and low) based on median expression level, followed by the overall survival computed between OS and normal by K-M survival curve.

### Real-time reverse transcription PCR (qRT-PCR)

Human OS cell line 143B was obtained from ATCC and was cultured in DMEM (Gibco) supplemented with 10% heat-inactivated fetal bovine serum (FBS). MC3T3, a murine pre-osteoblast cell line, was acquired from the Shandong Provincial Key Laboratory of Oral Tissue Regeneration. The cells were cultured in  $\alpha$ -minimal essential medium ( $\alpha$ -MEM) supplemented with 10% heat-inactivated FBS. The expression of all 8 prognostic genes in 143B cells and MC3T3 cells was detected by qRT-PCR. Briefly, total RNAs were extracted using TRizol reagent (cwbiotech, # CW0581) and reverse transcribed using HiFiScript cDNA Synthesis Kit (cwbiotech, # CW2569). The current assay was performed on ABI7900FAST (Thermo Fisher Scientific), and the primers were listed in Table 1. The PCR program was performed with thermocycling conditions: 50°C for 3 min, 95°C for 3 min, 40 cycles of 95°C for 10 s; 60°C for 30s, and melt curve 60 to 95°C (increment 0.5°C for 10s). The  $2^{-\Delta\Delta C_t}$  method was used for the investigation of gene expression.

### Predictive analysis of metabolites related to OS

The relationship between disease (keyword: osteoporosis)-associated chemical molecular and target gene was revealed by CDT, followed by the combination of prognosis-related genes obtained above to explore the key genes and compounds of OS. The chemicals associated with OS were mapped to metabolite ID with the application of the Compound ID Conversion tool in MetaboAnalyst database ([www.metaboanalyst.ca/](http://www.metaboanalyst.ca/)).

**Table 1** Amplified sequences of genes and their primers

Primer	Sequence
$\beta$ -actin	Forward: 5'-AGACCTGTACGCCAACACAG-3' Reverse: 5'-CGGACTCGTCATACTCTGC-3'
CXCL12	Forward: 5'-CTACAGATGCCCATGCCGAT-3' Reverse: 5'-CAGCCGGGCTACAATCTGAA-3'
CEBPA	Forward: 5'-AGAACAGCAACGAGTACCGG-3' Reverse: 5'-G GCGGTCATTGTCAGTGGTCA-3'
SPARCL1	Forward: 5'-ATGGCGATGATGATGGCGAT-3' Reverse: 5'-GATTGAGCTCTCTCGGCCTC-3'
CFLAR	Forward: 5'-TTGTGCCGGGATGTTGCTAT-3' Reverse: 5'-AGAGCAGTTCAGCCAAGTCC-3'
CAT	Forward: 5'-AGTGATCGGGGATTCCAGA-3' Reverse: 5'-AAGTCTCGCCGCATCTCAA-3'
STC2	Forward: 5'-CACTGTTTGGTCAACGCTGG-3' Reverse: 5'-AGCGTGGGCTTACATTTCA-3'
TUBA1A	Forward: 5'-CTATCCCGCATCCACTCC-3' Reverse: 5'-TTTACCATGGCGAGGGTCCAC-3'
ALDH1A1	Forward: 5'-ATCCTCTGACCCAGGAGTC-3' Reverse: 5'-AACACTGTGGGCTGGACAAA-3'

### Disease-metabolic network construction

KEGG pathway graph of DEGs was regarded as the analytic target to construct the differential pathway network that contained chemical compounds, reactions, enzymes, pathways, and KEGG modules using a heat diffusion model-based algorithm [23]. Specifically, the metabolic perturbation can be considered as the heat flow that traverses the KEGG graph. The null diffusive process highlighted that heat could flow out from nodes, which corresponded to the affected metabolites and across the whole differential pathway network. The p values of nodes were calculated according to the diffusion scores and ranked. The formula for the temperature diffusion score was as follows:  $T = -KI * G$ .

G indicator is input metabolite, it is 1 if the affected metabolite is entered; otherwise, 0. KI is the conductance matrix and equals to L plus B (L: the unnormalized graph Laplacian and B: the diagonal adjacency matrix). Notably,  $B(i, i) = 1$  if node i is a pathway; otherwise,  $B(i, i) = 0$ . Herein, the nodes with  $p < 0.05$  were retained and used to construct the disease-metabolic network.

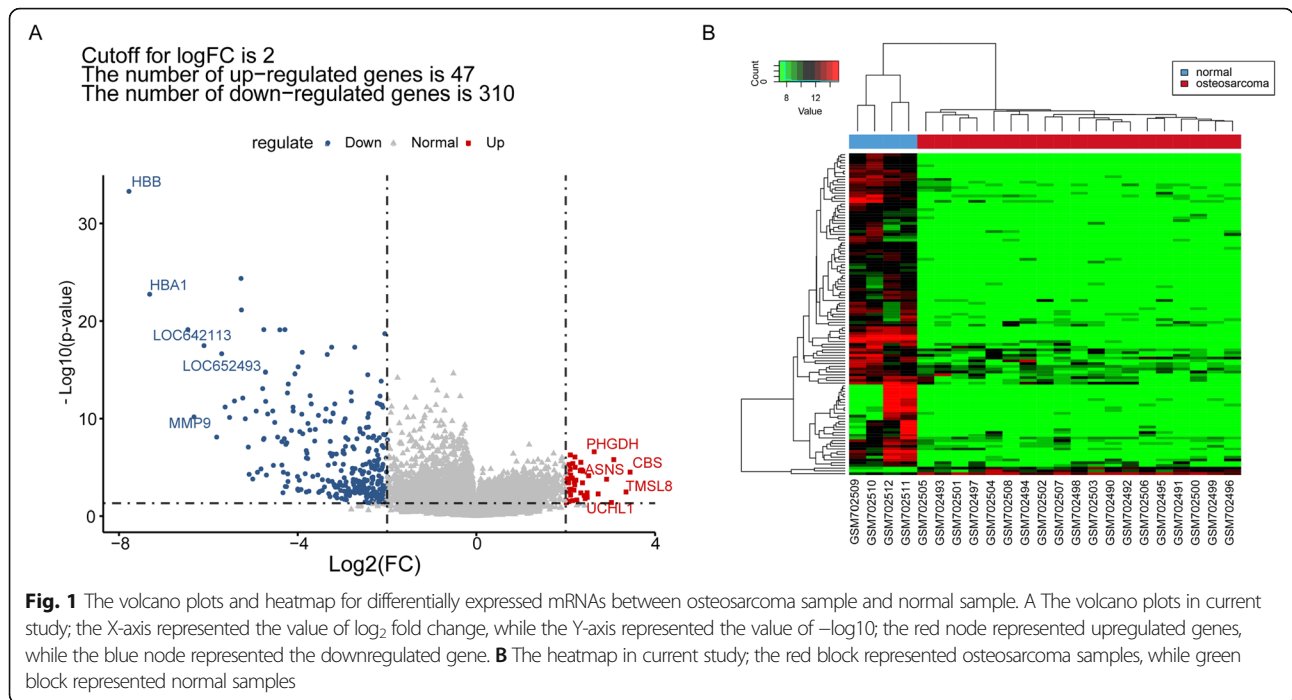
## Results

### DEGs between the OS group and control group

A total of 357 DEGs were revealed between the OS group and control group, containing 47 upregulated and 310 downregulated genes. The volcano plot showed that the upregulated genes and downregulated genes were significantly separated (Fig. 1A, B), suggesting a reliable result.

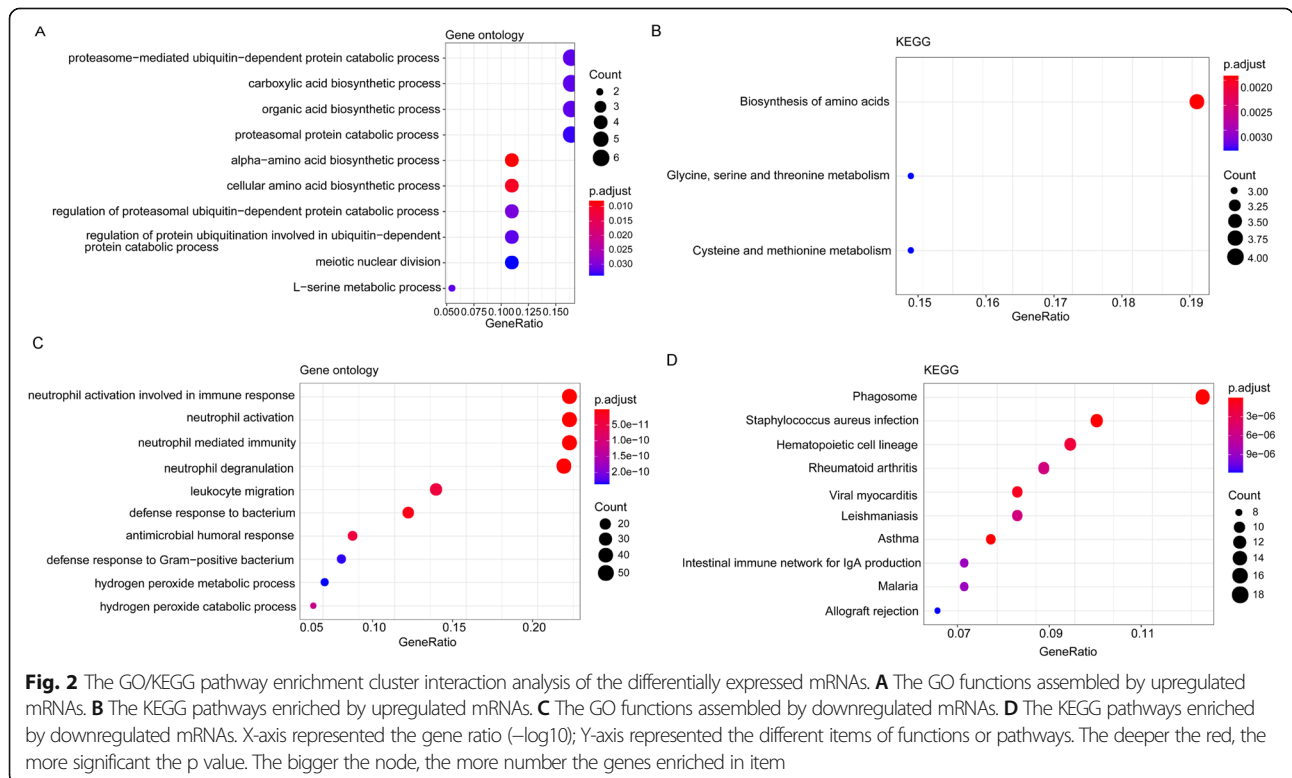
### Enrichment analysis and GSVA investigation

The upregulated DEGs mainly enriched in 14 GO-BP functions including leukocyte migration (GO:



0050900, Genes: C-X-C motif chemokine ligand 12 (CXCL12), etc.) (Fig. 2A) and 3 KEGG pathways such as biosynthesis of amino acids (hsa01230, genes: phosphoglycerate dehydrogenase (PHGDH), etc.) (Fig. 2B). Meanwhile, downregulated DEGs were

mainly involved in GO-BP functions including neutrophil activation (GO:0042119, genes: Fc fragment of IgG receptor IIIb (FCGR3B), etc.) and pathways like phagosome (hsa04145, genes: FCGR3B, etc.) (Fig. 2C, D).





Furthermore, the GSVA analysis on the KEGG pathway revealed that a totally 40 outstanding pathways showed the difference between OS and normal samples, such as vascular endothelial growth factor (VEGF) signaling pathway, cell adhesion molecules, and chemokine signaling pathway (Fig. 3).

#### Feature gene investigation and survival analysis

A PPI network was constructed in the current study based on 843 protein interactions and 23 genes. The detailed information for the current PPI network was showed in Fig. 4. Based on genes in the PPI network, CTD database (26355 genes and 3919 compounds) and GeneCards database (26355 genes), the feature genes for OS were explored using VENN plot analysis. The result showed that there were 38 intersected genes (feature genes) in the current study (Figure S2). Then, the survival analysis was performed on all feature genes (Table S1). Finally, a total of 8 feature genes was revealed as prognostic genes. Detailly, expression of genes including CXCL12, CCAAT enhancer-binding protein alpha (CEBPA), SPARC like 1 (SPARCL1), catalase (CAT), tubulin alpha 1a (TUBA1A), and aldehyde dehydrogenase 1 family member A1 (ALDH1A1) were positively correlated with overall survival of patients, while CASP8 and FADD-like apoptosis regulator (CFLAR) and

stanniocalcin 2 (STC2) were negatively correlated with overall survival of patients (Fig. 5).

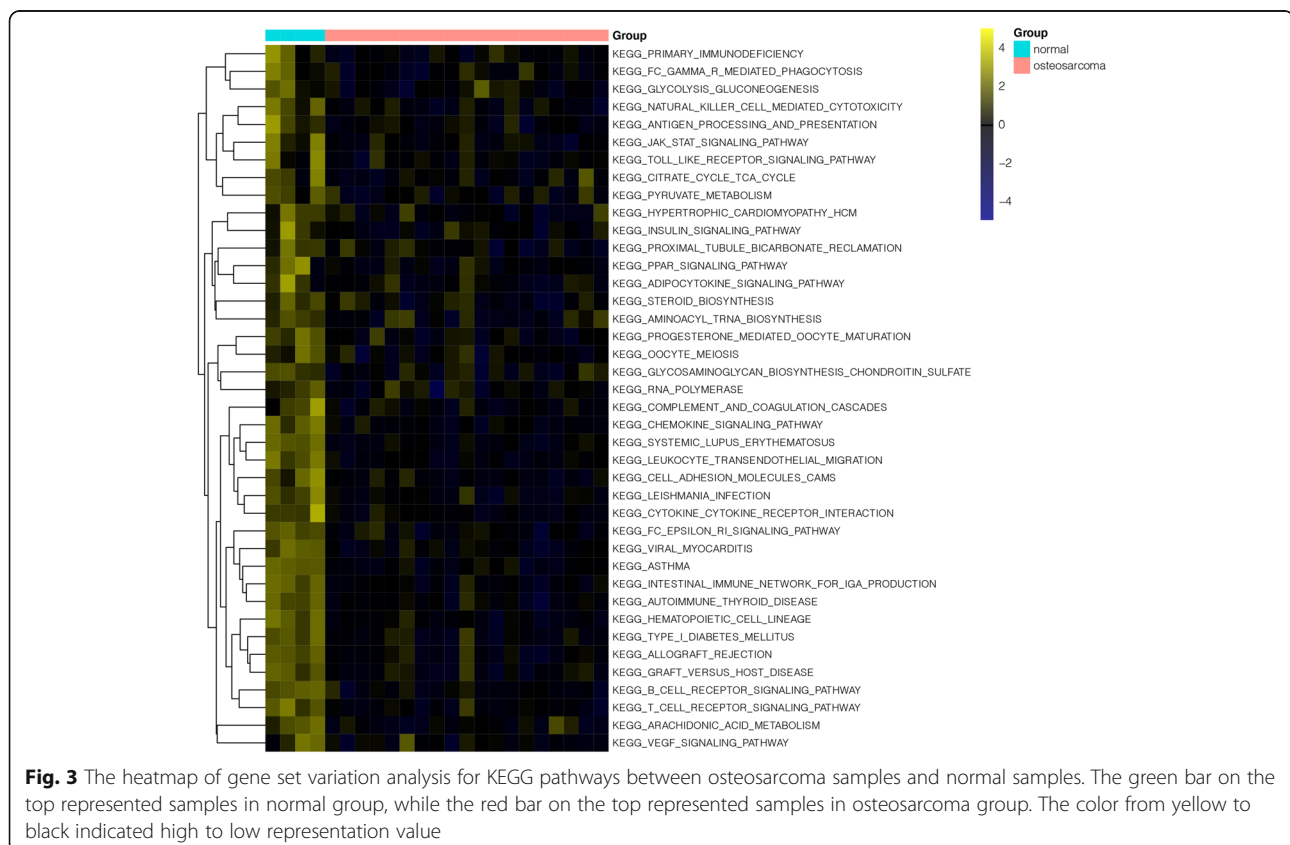
#### Verification for prognostic genes expression by qRT-PCR

The qRT-PCR was performed to further investigate the expression of 8 prognostic genes in cells (Fig. 6). The results showed that the relative expression of CXCL12, CEBPA, and TUBA1A in tumor cells was significantly lower than that in normal cells (all  $P < 0.05$ ). Meanwhile, the relative expression of CFLAR and STC2 in tumor cells was significantly higher than that in normal cells (all  $P < 0.05$ ).

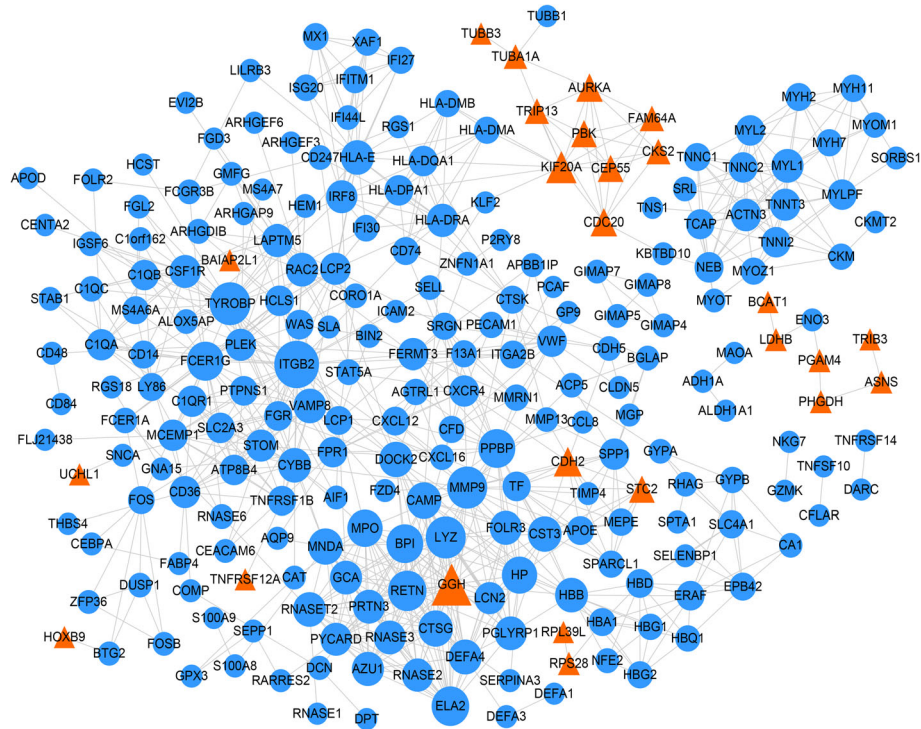
#### Disease-metabolic network investigation

As described in methods, a total of 45 compound-gene interactions involving 10 compounds and 8 prognostic genes were obtained from the CTD database (Table S2). Then, using the “Compound ID Conversion” tool provided by MetaboAnalyst database, 5 KEGG metabolite ID corresponding to these 10 compounds were obtained, including C06911, C01692, C13624, C07888, and C01661 (Table S3). These 5 metabolites were OS-associated metabolites.

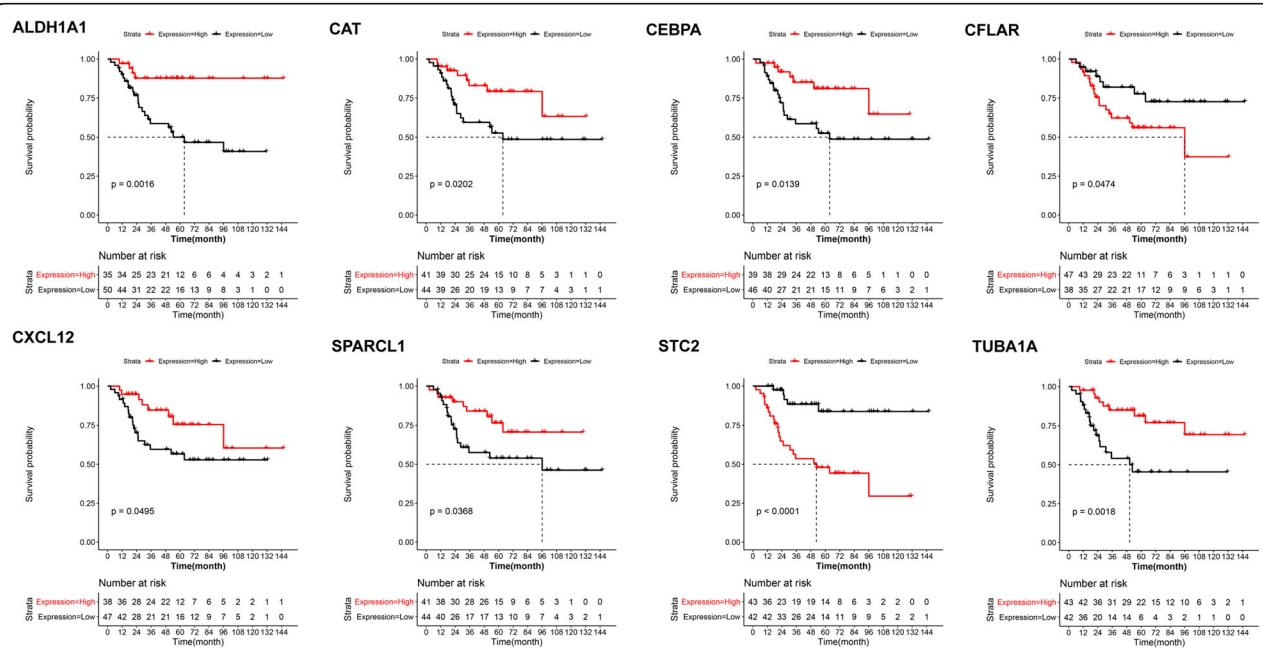
Then, a metabolic network was constructed with 10976 nodes and 32242 interactions. Among the 5 OS-associated metabolites, only C13624 and C07888



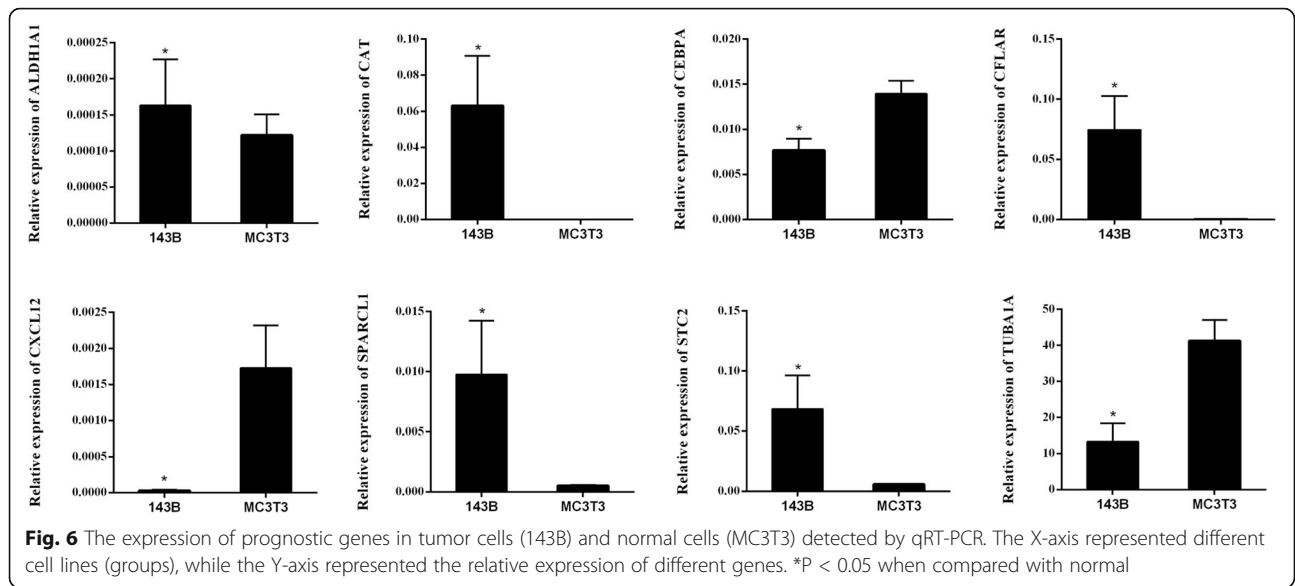
**Fig. 3** The heatmap of gene set variation analysis for KEGG pathways between osteosarcoma samples and normal samples. The green bar on the top represented samples in normal group, while the red bar on the top represented samples in osteosarcoma group. The color from yellow to black indicated high to low representation value



**Fig. 4** Protein–protein interaction network in current study. The blue circle represented downregulated gene, while the orange triangle represented upregulated gene. The larger the node, the bigger the degree



**Fig. 5** The survival analysis for 8 feature genes. The X-axis represented the survival time (month), while Y-axis represented survival probability

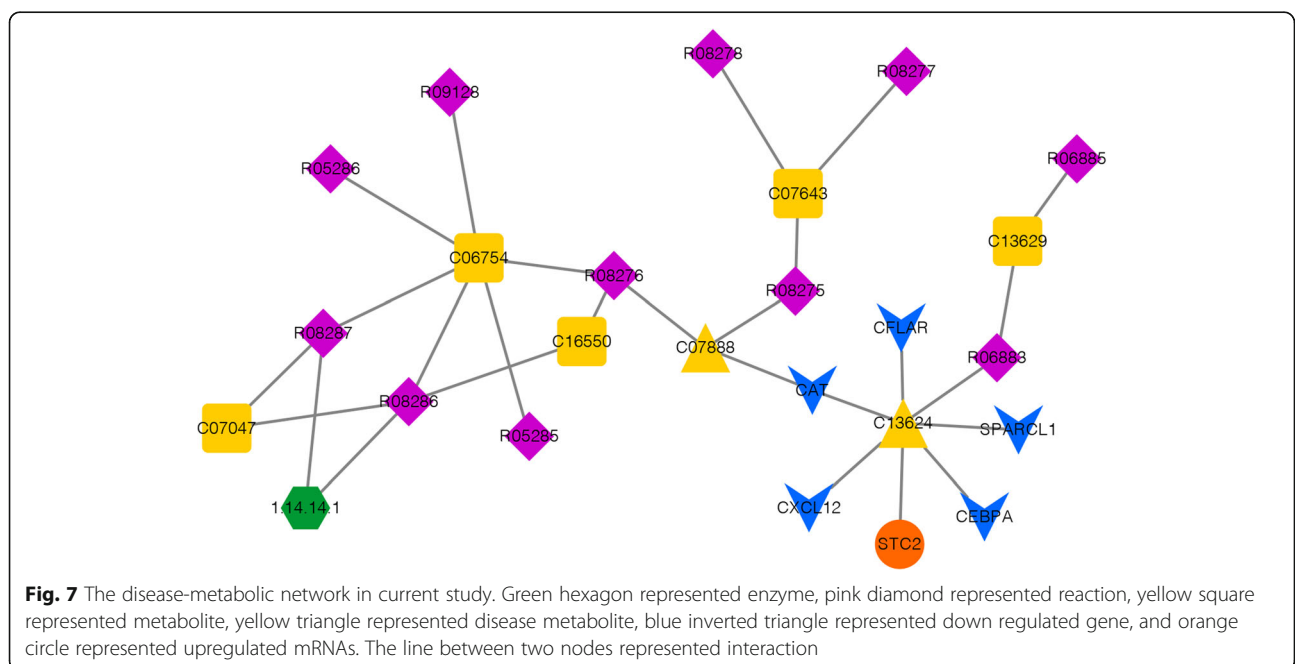


were matched to the network. With  $P < 0.05$ , totally, 19 nodes in the metabolic network were retained, then the DEG-compound interactions were added to the network to generate a disease-metabolic network. The result showed there were 11 reactions, 5 metabolites, 2 disease-metabolites [including C13624 (Bisphenol A, BPA) and C07888 (Cyclophosphamide)], 1 enzyme and 6 DEGs (5 downregulated and 1 upregulated), and 27 interactions in the current disease-metabolic network (Fig. 7, Table 2). Among the disease-metabolite network, C13624 (BPA) showed

interactions with all the 6 DEGs was considered as the key OS-associated metabolite.

### Discussion

In this study, we screened 357 genes that differentially expressed in human OS cell lines and human normal bone tissue samples. These genes were considered as the key genes involved in the development of OS. Functional enrichment analysis showed that the downregulated genes were mainly enriched in various immune-related functions (such as leukocyte migration and neutrophil



**Table 2** Detail information for nodes in current disease-metabolic network

KEGG ID	Entry type	KEGG name	P score
1.14.14.1	Enzyme	Unspecific monooxygenase	4.17E-03
R05285	Reaction	2-Chloroethanol:cytochrome c oxidoreductase	1.00E-06
R05286	Reaction	Chloroacetaldehyde:NAD+ oxidoreductase	1.00E-06
R06883	Reaction	Bisphenol A + NADH + H+ + Oxygen <=> 1,2-Bis(4-hydroxyphenyl)-2-propanol + NAD+ + H <sub>2</sub> O	1.00E-06
R06885	Reaction	1,2-Bis(4-hydroxyphenyl)-2-propanol <=> 4,4'-Dihydroxy-alpha-methylstilbene + H <sub>2</sub> O	1.00E-06
R08275	Reaction	Cyclophosphamide + NADPH + H+ + Oxygen <=> 4-Hydroxycyclophosphamide + NADP+ + H <sub>2</sub> O	1.00E-06
R08276	Reaction	Cyclophosphamide <=> Dechloroethylcyclophosphamide + Chloroacetaldehyde	1.00E-06
R08277	Reaction	4-Hydroxycyclophosphamide <=> 4-Ketocyclophosphamide	3.23E-02
R08278	Reaction	4-Hydroxycyclophosphamide <=> Aldophosphamide	2.08E-03
R08286	Reaction	Ifosfamide <=> Dechloroethylcyclophosphamide + Chloroacetaldehyde	1.00E-06
R08287	Reaction	Ifosfamide <=> 2-Dechloroethylifosfamide + Chloroacetaldehyde	1.54E-05
R09128	Reaction	2-Chloroethanol:cytochrome c oxidoreductase	1.00E-06
C06754	Compound	Chloroacetaldehyde	1.00E-06
C07047	Compound	Ifosfamide	5.24E-05
C07643	Compound	4-Hydroxycyclophosphamide	6.18E-04
C07888	Compound	Cyclophosphamide	1.00E-06
C13624	Compound	Bisphenol A	1.00E-06
C13629	Compound	1,2-Bis(4-hydroxyphenyl)-2-propanol	1.00E-06
C16550	Compound	Dechloroethylcyclophosphamide	1.00E-06

activation involved in immune response), while the up-regulated genes were mainly enriched in various biosynthetic processes. Similarly, GSVA revealed that various pathways, such as B/T cell receptor signaling pathway, VEGF signaling pathway, cell adhesion molecules pathway, and chemokine signaling pathways showed significant differences between these two groups. The bone shows a highly specialized immune environment, and various immune-related processes and pathways are involved in bone homeostasis. The success of mifamotide, an innate immune stimulator, in adjuvant therapy for non-metastatic OS demonstrates the potential for immune-based therapy to improve the prognosis of patients with OS [24, 25]. Miao et al. suggested that leukocyte recruitment-associated myokines were down-regulated in OS, indicating escaping from the host immune system would contribute to the development of OS [26]. Chemokines and their receptors play important roles in the regulation of tumor-mediated immune response in tumors. For example, CXCL1 together with its receptor CXCR2 were implicated in assisting with the homing of neutrophils into the tumor microenvironment in OS [27]. Chemokine receptor CXCR3 expression was found to positively correlated with the abundance of tumor-infiltrating immune cells, such as macrophages M1, CD8 T cells, and activated NK cells in OS [28]. VEGF is an important angiogenesis-promoting factor in various tumors, that has been reported to contribute to

the growth and aggressive behavior in OS [29, 30], promote angiogenesis, and inhibit cell apoptosis [31]. Considering these studies, we speculated that the DEGs were involved in the development of OS by regulating pathways associated with immune, chemokines, and VEGF signalings.

In order to screen most valuable DEGs, Venn analysis was performed on genes in the PPI network, and OS-related genes from CTD and Genecards databases, and 38 overlapped genes were identified, among which, 8 genes associated with survival of patients were considered as most valuable genes, including STC2, SPARC1, ALDH1A1, CFLAR, CEBPA, CAT, TUBA1A, and CXCL12. STC2 is a secretory glycoprotein hormone that can regulate cell proliferation and cancer cell lesions [32]. A previous study shows that STC2 is an outstanding immune-related gene during the progression of OS [33]. A recent study indicates that 20 genes including STC2 signatures are identified related to OS, which can be helpful for predicting prognosis of patients with OS [34]. SPARC1 protein can affect osteoblast differentiation, tumorigenesis and tumor metastasis [35]. A previous study indicates that SPARC1 can block the metastasis of human OS by the upregulation of canonical signaling [36]. As an aldehyde dehydrogenase, ALDH1A1 has been found to be differentially expressed between low and highly metastatic OS [37]. Qi et al. proved that ALDH1A1 was upregulated in the



development of OS [38]. In addition, CFLAR is a common therapeutic target in various human cancers including OS [39]. A previous study shows that miR-20a can be used to suppress OS cell proliferation and invasion through CFLAR [40]. CAT encodes catalase, which was found to be involved in regulating the generation of reactive oxygen species, thereby affecting the cytotoxic effects of pimozone on OS cells [41] and regulating apoptosis of p53 null OS MG63 cells [42]. Furthermore, as a potential target in OS treatment, CXCL12 has been proved to participate in the progression and metastasis of bone sarcomas [43]. It is believed that the upregulation of CXCL12 contributes to the positive OS outcome [44]. Actually, the expression of CXCL12 is commonly realized via certain biological functions [45]. Gulino et al. showed that CXCL12 takes part in the variation of mature polymorphonuclear via altered leukocyte response [46]. A previous study indicates that CXCL12 takes part in the trauma and sterile inflammation via leukocyte migration [47]. These studies emphasized the important roles of these genes in the development and progression of OS.

Finally, a disease-metabolic network was constructed including two disease-associated metabolites BPA and cyclophosphamide. Among the disease-metabolites network, C13624 (BPA) showed interactions with all the 6 DEGs were considered as the key OS associated-metabolite. BPA is a widely studied typical endocrine-disrupting chemical [48]. It is closely associated with the clinical treatment of OS [49]. A previous study shows that BPA contributes to the decreasing activity of OS cells and the inhabitation of cell proliferation [50]. Actually, BPA is considered as a prioritized effect biomarker for human biomonitoring [51]. The relationship between BPA and disease risk prediction has already been investigated in a previous study [52]. Our results revealed that BPA interacted with multiple prognosis-related genes, such as CXCL12, STC2, and CFLAR. Thus, we speculated that BPA might be an important metabolite involved in the development of OS by interacting with different genes.

There were some limitations in the current study. (1) The selected GEO dataset was generated from 19 human OS cell lines and 4 human normal bone tissue samples. There might be differences between human tissue samples and cell lines, and differences among different cell lines. In addition, the validation qPCR was performed on the human OS 143B cell line and murine pre-osteoblast MC3T3 cell line. Therefore, a study based on human OS tissue samples and matched adjacent normal samples should be carried out to eliminate the confounding factors. (2) The disease-associated metabolites and metabolite-gene interactions were predicted using online databases. To

obtain more reliable results, an integrated analysis based on metabolomics data and transcriptomics should be performed to investigate the differential metabolites and their roles in the development of OS. (3) Eight prognosis-related genes were screened. However, their prognostic value had not been confirmed, and their clinical value should be further evaluated by clinical data.

## Conclusion

In conclusion, eight prognosis-related genes were identified in OS. The downregulated CXCL12 might take part in the progression of OS via participating in the leukocyte migration function. Moreover, mRNAs including CXCL12 and STC2 might be two novel biomarkers for OS. Furthermore, BPA might be an important metabolite interacting with multiple prognosis-related genes in OS.

## Abbreviations

OS: Osteosarcoma; DEGs: Differentially expressed genes; miRNAs: MicroRNAs; GO-BP: Gene ontology-biological process; KEGG: The Kyoto Encyclopedia of Genes and Genomes; STRING: Search Tool for the Retrieval of Interacting Genes; GSVA: Gene set variation analysis; CTD: Comparative toxicogenomics database; UCSC: University of California Santa Cruz; FBS: Fetal bovine serum;  $\alpha$ -MEM:  $\alpha$ -Minimal essential medium; CXCL12: C-X-C motif chemokine ligand 12; PHGDH: Phosphoglycerate dehydrogenase; FCGR3B: Fc fragment of IgG receptor IIIb; CEBPA: CCAAT enhancer-binding protein alpha; SPARCL1: SPARC like 1; CAT: Catalase; TUBA1A: Tubulin alpha 1a; ALDH1A1: Aldehyde dehydrogenase 1 family member A1; CFLAR: CASP8 and FADD-like apoptosis regulator; STC2: Stanniocalcin 2; VEGF: Vascular endothelial growth factor; BPA: Bisphenol A

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13018-021-02578-0>.

**Additional file 1: Figure S1.** The workflow of the current study.

**Additional file 2: Figure S2.** The VENN plot analysis for feature genes of osteosarcoma.

**Additional file 3: Table S1.** Survival data and results for survival analysis of all feature genes.

**Additional file 4: Table S2.** The 45 compounds-genes interactions obtained from CTD database.

**Additional file 5: Table S3.** The five KEGG metabolites ID corresponding to these 10 compounds.

## Acknowledgements

Not applicable.

## Authors' contributions

Jun Wang carried out the conception and design of the research, and Zhenggang Xiong participated in the acquisition of data. Yangyang Zhao carried out the analysis and interpretation of data. Deguo Xing participated in the design of the study and performed the statistical analysis. Jun Wang and Mingzhi Gong conceived of the study, and participated in its design and coordination and helped to draft the manuscript and revision of the manuscript for important intellectual content. The authors read and approved the final manuscript.

## Funding

Not applicable.

**Availability of data and materials**

Not applicable. This study was only the primary research, and further study has been in progress.

**Declarations****Ethics approval and consent to participate**

Our study did not require an ethical board approval because it did not contain human or animal trials.

**Consent for publication**

Not applicable.

**Competing interests**

The authors declare that they have no competing interests.

Received: 17 March 2021 Accepted: 24 June 2021

Published online: 05 July 2021

**References**

- de Azevedo JWW, Fernandes TAAM, Fernandes JV, de Azevedo JCV, Lanza DCF, Bezerra CM, et al. Biology and pathogenesis of human osteosarcoma. *Oncol Lett.* 2020;19(2):1099–116. <https://doi.org/10.3892/ol.2019.11229>.
- Lindsey BA, Markel JE, Kleinerman ES. Osteosarcoma overview. *Rheumatology and therapy.* 2017;4(1):25–43. <https://doi.org/10.1007/s40744-016-0050-2>.
- Harrison DJ, Geller DS, Gill JD, Lewis VO, Gorlick R. Current and future therapeutic approaches for osteosarcoma. *Expert Rev Anticancer Ther.* 2018;18(1):39–50. <https://doi.org/10.1080/14737140.2018.1413939>.
- Chou AJ, Geller DS, Gorlick R. Therapy for osteosarcoma. *Pediatr Drugs.* 2008;10(5):315–27. <https://doi.org/10.2165/00148581-200810050-00005>.
- Czarnecka AM, Synoradzki K, Firlej W, Bartnik E, Sobczuk P, Fiedorowicz M, et al. Molecular biology of osteosarcoma. *Cancers.* 2020;12(8):2130. <https://doi.org/10.3390/cancers12082130>.
- Xu J-F, Wang Y-P, Zhang S-J, Chen Y, Gu H-F, Dou X-F, et al. Exosomes containing differential expression of microRNA and mRNA in osteosarcoma that can predict response to chemotherapy. *Oncotarget.* 2017;8(44):75968–78. <https://doi.org/10.18632/oncotarget.18373>.
- Wang Z, Tan M, Chen G, Li Z, Lu X. LncRNA SOX2-OT is a novel prognostic biomarker for osteosarcoma patients and regulates osteosarcoma cells proliferation and motility through modulating SOX2. *IUBMB Life.* 2017;69(11):867–76. <https://doi.org/10.1002/iub.1681>.
- Wei R, Thanindratarn P, Dean DC, Hornicek FJ, Guo W, Duan Z. Cyclin E1 is a prognostic biomarker and potential therapeutic target in osteosarcoma. *J Orthop Res.* 2020;38(9):1952–64.
- Dong Y, Liang G, Yuan B, Yang C, Gao R, Zhou X. MALAT1 promotes the proliferation and metastasis of osteosarcoma cells by activating the PI3K/Akt pathway. *Tumor Biol.* 2015;36(3):1477–86. <https://doi.org/10.1007/s13277-014-2631-4>.
- Dean DC, Shen S, Hornicek FJ, Duan Z. From genomics to metabolomics: emerging metastatic biomarkers in osteosarcoma. *Cancer Metastasis Rev.* 2018;37(4):719–31. <https://doi.org/10.1007/s10555-018-9763-8>.
- Bandala C, Ávila-Luna A, Gómez-López M, Estrada-Villaseñor E, Montes S, Alfaro-Rodríguez A, et al. Catecholamine levels and gene expression of their receptors in tissues of adults with osteosarcoma. *Arch Physiol Biochem.* 2019;1–7. <https://doi.org/10.1080/13813455.2019.1638942>.
- Berdiaki A, Datsis GA, Nikitovic D, Tsatsakis A, Katonis P, Karamanos NK, et al. Parathyroid hormone (PTH) peptides through the regulation of hyaluronan metabolism affect osteosarcoma cell migration. *IUBMB Life.* 2010;62(5):377–86. <https://doi.org/10.1002/iub.320>.
- Namlós HM, Meza-Zepeda LA, Barøy T, Østensen IH, Kresse SH, Kuijjer ML, et al. Modulation of the osteosarcoma expression phenotype by microRNAs. *PLoS One.* 2012;7(10):e48086. <https://doi.org/10.1371/journal.pone.0048086>.
- Gautier L, Cope L, Bolstad B, Irizarry R. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics.* 2004;20(3):307–15. <https://doi.org/10.1093/bioinformatics/btg405>.
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7):e47–e.
- Yu G, Wang L, Han Y, He Q. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16(5):284–7. <https://doi.org/10.1089/omi.2011.0118>.
- Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics.* 2013;14(1):7.
- Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 2015 Jan;43(Database issue):D447–52. <https://doi.org/10.1093/nar/gku1003>.
- Shannon P, Markiel A, Ozier O, Baliga N, Wang J, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498–504. <https://doi.org/10.1101/gr.1239303>.
- Davis AP, Grondin CJ, Lennon-Hopkins K, Saraceni-Richards C, Sciaky D, King BL, et al. The Comparative Toxicogenomics Database's 10th year anniversary: update 2015. *Nucleic Acids Res.* 2015;43(D1):D914–D20. <https://doi.org/10.1093/nar/gku935>.
- Safran M, Dalah I, Alexander J, Rosen N, Iny Stein T, Shmoish M, et al. GeneCards Version 3: the human gene integrator. *Database.* 2010;2010(0). <https://doi.org/10.1093/database/baq020>.
- Tyner C, Barber GP, Casper J, Clawson H, Diekhans M, Eisenhart C, et al. The UCSC Genome Browser database: 2017 update. *Nucleic Acids Res.* 2017;45(Database issue):D626–D34.
- Picart-Armada S, Fernandez-Albert F, Vinaixa M, Rodriguez MA, Aivio S, Stracker TH, et al. Null diffusion-based enrichment for metabolomics data. *PLoS One.* 2017;12(12):e0189012. <https://doi.org/10.1371/journal.pone.0189012>.
- Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. *Nat Rev Cancer.* 2014 Nov;14(11):722–35. <https://doi.org/10.1038/nrc3838>.
- Wang Z, Wang Z, Li B, Wang S, Chen T, Ye Z. Innate immune cells: a potential and promising cell population for treating osteosarcoma. *Front Immunol.* 2019;10:1114. <https://doi.org/10.3389/fimmu.2019.01114>.
- Miao Y, Hu B, Wang Q, Yang Q, Zhou S. Myokines related to leukocyte recruitment are down-regulated in osteosarcoma. *Int J Med Sci.* 2018;15(9):859–66. <https://doi.org/10.7150/ijms.24928>.
- Chao CC, Lee CW, Chang TM, Chen PC, Liu JF. CXCL1/CXCR2 paracrine axis contributes to lung metastasis in osteosarcoma. *Cancers (Basel).* 2020 Feb 17;12(2):459. <https://doi.org/10.3390/cancers12020459>.
- Tang Y, Gu Z, Fu Y, Wang J. CXCR3 from chemokine receptor family correlates with immune infiltration and predicts poor survival in osteosarcoma. *Biosci Rep.* 2019;39:11.
- Daft PG, Yang Y, Napierala D, Zayzafoon M. The growth and aggressive behavior of human osteosarcoma is regulated by a CaMKII-controlled autocrine VEGF signaling mechanism. *PLoS One.* 2015;10(4):e0121568. <https://doi.org/10.1371/journal.pone.0121568>.
- Ohba T, Cates JM, Cole HA, Slosky DA, Haro H, Ando T, et al. Autocrine VEGF/VEGFR1 signaling in a subpopulation of cells associates with aggressive osteosarcoma. *Mol Cancer Res.* 2014 Aug;12(8):1100–11. <https://doi.org/10.1158/1541-7786.MCR-14-0037>.
- Peng N, Gao S, Guo X, Wang G, Cheng C, Li M, et al. Silencing of VEGF inhibits human osteosarcoma angiogenesis and promotes cell apoptosis via VEGF/PI3K/AKT signaling pathway. *Am J Transl Res.* 2016;8(2):1005–15.
- Li Q, Zhou X, Fang Z, Pan Z. Effect of STC2 gene silencing on colorectal cancer cells. *Mol Med Rep.* 2019;20(2):977–84.
- Li L-Q, Zhang L-H, Zhang Y, Lu X-C, Zhang Y, Liu Y-K, et al. Construction of immune-related gene pairs signature to predict the overall survival of osteosarcoma patients. *Aging (Albany NY).* 2020;12(22):22906.
- Qiu Z, Du X, Chen K, Dai Y, Wang S, Xiao J, et al. Gene signatures with predictive and prognostic survival values in human osteosarcoma. *PeerJ.* 2021;9:e10633. <https://doi.org/10.7717/peerj.10633>.
- Bradshaw AD. Diverse biological functions of the SPARC family of proteins. *Int J Biochem Cell Biol.* 2012;44(3):480–8. <https://doi.org/10.1016/j.biocel.2011.12.021>.
- Zhao S, Jiang Y, Xu N, Li Q, Zhang Q, Wang S, et al. SPARCL1 suppresses osteosarcoma metastasis and recruits macrophages by activation of canonical WNT/β-catenin signaling through stabilization of the WNT-receptor complex. *Oncogene.* 2018;37(8):1049–61. <https://doi.org/10.1038/onc.2017.403>.
- Mandell JB, Douglas N, Ukani V, Anderson C, Beumer J, Watters R, et al. Altered ALDH1A1 expression and cellular copper levels between low and

- highly metastatic osteosarcoma provides a case for novel repurposing of disulfiram: AACR; 2020.
38. X-t Q, Y-l L, Zhang Y-q XT, Lu B, Fang L, et al. KLF4 functions as an oncogene in promoting cancer stem cell-like characteristics in osteosarcoma cells. *Acta Pharmacol Sin.* 2019;40(4):546–55.
  39. Fulda S. Targeting c-FLICE-like inhibitory protein (CFLAR) in cancer. *Expert Opin Ther Targets.* 2013;17(2):195–201. <https://doi.org/10.1517/14728222.2013.736499>.
  40. Lu S, Liao Q, Tang L. MiR-155 affects osteosarcoma cell proliferation and invasion through regulating NF- $\kappa$ B signaling pathway. *Eur Rev Med Pharmacol Sci.* 2018;22(22):7633–9. [https://doi.org/10.26355/eurev\\_201811\\_16380](https://doi.org/10.26355/eurev_201811_16380).
  41. Cai N, Zhou W, Ye LL, Chen J, Liang QN, Chang G, et al. The STAT3 inhibitor pimozide impedes cell proliferation and induces ROS generation in human osteosarcoma by suppressing catalase expression. *Am J Transl Res.* 2017; 9(8):3853–66.
  42. Wang Y, Wei Y, Zhang H, Shi Y, Li Y, Li R. Arsenic trioxide induces apoptosis of p53 null osteosarcoma MG63 cells through the inhibition of catalase. *Med Oncol.* 2012 Jun;29(2):1328–34. <https://doi.org/10.1007/s12032-011-9848-5>.
  43. Brennecke P, Arlt MJ, Campanile C, Husmann K, Gvozdenovic A, Apuzzo T, et al. CXCR4 antibody treatment suppresses metastatic spread to the lung of intratibial human osteosarcoma xenografts in mice. *Clin Exp Metastasis.* 2014;31(3):339–49. <https://doi.org/10.1007/s10585-013-9632-3>.
  44. Baumhoer D, Smida J, Zillmer S, Rosemann M, Atkinson MJ, Nelson PJ, et al. Strong expression of CXCL12 is associated with a favorable outcome in osteosarcoma. *Mod Pathol.* 2012;25(4):522–8. <https://doi.org/10.1038/modpathol.2011.193>.
  45. Liang Z, Brooks J, Willard M, Liang K, Yoon Y, Kang S, et al. CXCR4/CXCL12 axis promotes VEGF-mediated tumor angiogenesis through Akt signaling pathway. *Biochem Biophys Res Commun.* 2007;359(3):716–22. <https://doi.org/10.1016/j.bbrc.2007.05.182>.
  46. Gulino AV, Moratto D, Sozzani S, Cavadini P, Otero K, Tassone L, et al. Altered leukocyte response to CXCL12 in patients with warts hypogammaglobulinemia, infections, myelokathexis (WHIM) syndrome. *Blood.* 2004;104(2):444–52. <https://doi.org/10.1182/blood-2003-10-3532>.
  47. Venereau E, Schiraldi M, Ugucioni M, Bianchi ME. HMGB1 and leukocyte migration during trauma and sterile inflammation. *Mol Immunol.* 2013;55(1): 76–82. <https://doi.org/10.1016/j.molimm.2012.10.037>.
  48. Eladak S, Grisin T, Moison D, Guerquin M-J, N'Tumba-Byn T, Pozzi-Gaudin S, et al. A new chapter in the bisphenol A story: bisphenol S and bisphenol F are not safe alternatives to this compound. *Fertil Steril.* 2015;103(1):11–21. <https://doi.org/10.1016/j.fertnstert.2014.11.005>.
  49. Fic A, Mlakar SJ, Juvan P, Mlakar V, Marc J, Dolenc MS, et al. Genome-wide gene expression profiling of low-dose, long-term exposure of human osteosarcoma cells to bisphenol A and its analogs bisphenols AF and S. *Toxicol in Vitro.* 2015;29(5):1060–9. <https://doi.org/10.1016/j.tiv.2015.03.014>.
  50. Kidani T, Yasuda R, Miyawaki J, Oshima Y, Miura H, Masuno H. Bisphenol A inhibits cell proliferation and reduces the motile potential of murine LM8 osteosarcoma cells. *Anticancer Res.* 2017;37(4):1711–22. <https://doi.org/10.21873/anticancer.11503>.
  51. Mustieles V, d'Cruz SC, Couderq S, Rodríguez-Carrillo A, Fini J-B, Hofer T, et al. Bisphenol A and its analogues: a comprehensive review to identify and prioritize effect biomarkers for human biomonitoring. *Environ Int.* 2020; 144:105811. <https://doi.org/10.1016/j.envint.2020.105811>.
  52. Eng DS, Lee JM, Gebremariam A, Meeker JD, Peterson K, Padmanabhan V. Bisphenol A and chronic disease risk factors in US children. *Pediatrics.* 2013; 132(3):e637–e45. <https://doi.org/10.1542/peds.2013-0106>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

