Check for updates

The whole transcriptome analysis using FFPE and fresh tissue samples identifies the molecular fingerprint of osteosarcoma

Bal Hari Poudel[1,2,3] and Sulev Koks[1,2]*

[1]Center for Molecular Medicine and Innovative Therapy, Murdoch University, Perth, WA, Australia, [2]Perron Institute of Neurological Diseases, Perth, WA, Australia, [3]Central Department of Biotechnology, Tribhuvan University, Kathmandu, Nepal

Abstract

Osteosarcoma is a form of bone cancer that predominantly impacts osteoblasts, the cells responsible for creating fresh bone tissue. Typical indications include bone pain, inflammation, sensitivity, mobility constraints, and fractures. Utilising imaging techniques such as X-rays, MRI scans, and CT scans can provide insights into the size and location of the tumour. Additionally, a biopsy is employed to confirm the diagnosis. Analysing genes with distinct expression patterns unique to osteosarcoma can be valuable for early detection and the development of effective treatment approaches. In this research, we comprehensively examined the entire transcriptome and pinpointed genes with altered expression profiles specific to osteosarcoma. The study mainly aimed to identify the molecular fingerprint of osteosarcoma. In this study, we processed 90 FFPE samples from PathWest with an almost equal number of osteosarcoma and healthy tissues. RNA was extracted from Paraffin-embedded tissue; RNA was sequenced, the sequencing data was analysed, and gene expression was compared to the healthy samples of the same patients. Differentially expressed genes in osteosarcoma-derived samples were identified, and the functions of those genes were explored. This result was combined with our previous studies based on FFPE and fresh samples to perform a meta-analysis. We identified 1,500 identical differentially expressed genes in PathWest osteosarcoma samples compared to normal tissue samples of the same patients. Meta-analysis with combined fresh tissue samples identified 530 differentially expressed genes. *IFITM5, MMP13, PANX3,* and *MAGEA6* were some of the most overexpressed genes in osteosarcoma samples, while *SLC4A1, HBA1, HBB, AQP7* genes were some of the top downregulated genes. Through the meta-analysis, 530 differentially expressed genes were identified to be identical among FFPE (105 FFPE samples) and 36 fresh bone samples. Deconvolution analysis with single-cell RNAseq data confirmed the presence of specific cell clusters in FFPE samples. We propose these 530 DEGs as a molecular fingerprint of osteosarcoma.

Impact statement

Although rare, osteosarcoma attracts global attention because of the unsatisfactory outcomes associated with current treatment approaches. In our investigation, we examined total RNA extracted from 90 FFPE paired samples, consisting of 50 tumoral bone specimens and their matched non-tumoral counterparts, sourced from osteosarcoma patients. By comparing our findings with previous studies, we uncovered differences in gene expression patterns between normal and affected bone tissues, particularly emphasizing changes in the regulation of collagen and extracellular matrix degradation and cell cycle regulation. These findings deepen our understanding of osteosarcoma and provide potential directions for future research endeavors.

Introduction

Bone cancer (osteosarcoma, OS) is the most common primary tumour of the bone in children and young adults [1, 2]. Osteosarcoma arises from the cells forming bone tissue and can be found in the bone's metaphysis, the region where the growth plate is located. The exact cause of osteosarcoma is not fully understood, but several risk factors have been identified [3]. Some individuals may have a genetic predisposition to developing osteosarcoma, although it is not typically inherited in a Mendelian pattern. Certain preexisting conditions, such as hereditary retinoblastoma (a rare eye cancer) and Paget's disease of bone, have been linked to an increased risk of osteosarcoma [4].

Symptoms of osteosarcoma may include pain and swelling in the affected bone or joint, limited range of motion in the nearby joint, a mass or lump in the affected area, fractures or bone weakening in the affected bone [5, 6]. Diagnosis typically involves a combination of imaging studies like X-rays, MRI, and CT scans, as well as a biopsy to confirm the presence of cancerous cells [5, 6]. Treatment for osteosarcoma usually involves a multimodal approach, which may include surgery, chemotherapy and radiation therapy [6], however, none of them are accurate and also have strong adverse effect with chemotherapy and radiation therapy. In addition, there is higher chances of getting a second cancer [7, 8].

The prognosis for osteosarcoma varies depending on factors such as the extent of the disease, response to treatment, and the presence of metastasis [9]. Early diagnosis and aggressive treatment can significantly improve the chances of successful outcomes. Long-term follow-up care is essential to monitor for any potential recurrence or late effects of treatment. Treatment plans are usually developed in collaboration with a team of oncologists, surgeons, and other healthcare professionals.

Whole transcriptome analysis is a powerful tool used in modern-day molecular biology to study the entries set of RNA molecules produced by a cell or a population of cells [10, 11]. Transcriptome analysis provides snapshots of the genes expressed

and can be compared with normal [11]. However, the differentially expressed genes (DEGs) may vary from sample to sample according to the quality of the samples. For rare diseases like osteosarcoma, there are not many options for getting fresh samples due to the death of patients and getting fresh bone samples is not easy; hence, formalin-fixed paraffin-embedded tissues (FFPEs) are the best way to analyse the sample. However, RNA gets degraded [12] due to formalin fixation in FFPE.

In this paper, we report a complex transcriptomic analysis that is based on four independent studies. We performed two original experiments and combined these results with previously published studies. This study provides the molecular signature of osteosarcoma to generate fundamental knowledge to develop new drugs against this disease. Through this study, we compared the DEGs in FFPE with fresh bone samples to identify whether or not the FFPE DEGs aligned with the fresh bone samples. We performed RNAseq analysis on a new set of 90 samples. We later compared it to the two other independent studies (FFPE and fresh) to conduct a meta-analysis of OS transcriptome. Identified molecular mechanisms and genes could be used as a potential target for osteosarcoma drug development and will be evaluated in future studies.

Materials and methods

The sample processing and whole transcriptome analysis

Our original sample analysis is based on two independent analyses of the formalin-fixed paraffin-embedded (FFPE) samples. The paraffin blocks were cut, and RNA was extracted. This way, any archived biological samples can be analysed for rare diseases like osteosarcoma; using FFPE samples is the only option to get meaningful and large samples for the complex transcriptomic analysis. The FFPE samples (N = 90: 50 OS + 40 healthy control) were obtained through PathWest Nedlands (QEII Medical Centre, Australia) and were processed in two batches, and total RNA was extracted using the Norgen FFPE RNA purification kit (#25300) using the manufacturer's standard protocol. The purified RNA samples were sent to the Australian Genome Research Facility (AGRF) Melbourne for sequencing also in two separate batches.

FFPE sample analysis

We combined two PathWest FFPE studies (90 samples altogether) into a single analysis to increase the formal statistical power of the analysis. Details of the samples are shown in Supplementary Table S1. Briefly, this table highlights the gender, sample group (tumour or normal), age at onset, deceased or alive and chemotherapy or non-chemotherapy. The raw FASTQ files obtained after sequencing were used for the data analysis. DEGs were detected by comparing the OS with normal samples using

**TABLE 1** List of top highly upregulated genes compared to normal tissue.

Gene name	Symbol	Base mean	log2FoldChange	Padj
Matrix metallopeptidase 9	MMP9	707.4	7.5	4.12E-25
Cellular communication network factor 4	CCN4	238.0	6.8	5.30E-24
Hyaluronan and proteoglycan link protein 1	HAPLN1	70.6	6.4	1.12E-24
Alkaline phosphatase, biomineralization associated	ALPL	158.1	6.4	8.19E-23
Matrix metallopeptidase 13	MMP13	276.9	6.3	8.61E-21
Pannexin 3	PANX3	117.7	6.1	9.43E-21
Rhomboid like 2	RHBDL2	51.6	6.0	8.19E-23
Collagen triple helix repeat containing 1	CTHRC1	123.4	6.0	8.92E-20
Glucosamine-phosphate N-acetyltransferase 1	GNPNAT1	137.6	5.9	1.41E-19
Transmembrane O-mannosyltransferase targeting cadherins 2	TMTC2	134.7	5.9	1.41E-19
Anillin, actin binding protein	ANLN	91.7	5.9	3.21E-21
Collagen type XXIV alpha 1 chain	COL24A1	61.4	5.8	3.90E-21
Collagen type XI alpha 1 chain	COL11A1	824.5	5.8	3.52E-14
Interferon induced transmembrane protein 5	IFITM5	32.5	5.7	1.04E-21
DLG associated protein 5	DLGAP5	41.9	5.7	9.00E-23
Integrin binding sialoprotein	IBSP	263.7	5.7	1.63E-16
Parathyroid hormone 1 receptor	PTH1R	100.8	5.7	2.45E-18
Lymphoid enhancer binding factor 1	LEF1	75.1	5.6	3.83E-20
Proline rich 11	PRR11	115.6	5.6	3.84E-18
SRY-box transcription factor 11	SOX11	53.2	5.5	8.81E-21
Cyclin B2	CCNB2	31.6	5.5	2.03E-22
DNA polymerase eta	POLH	99.4	5.5	1.12E-19
Fibrillin 2	FBN2	159.2	5.4	1.04E-15
Sp7 transcription factor	SP7	169.0	5.4	9.98E-16
Spindle apparatus coiled-coil protein 1	SPDL1	62.2	5.4	1.28E-19
Sterile alpha motif domain containing 5	SAMD5	182.1	5.3	1.16E-14
Collagen type X alpha 1 chain	COL10A1	103.3	5.3	6.69E-16
MyoD family inhibitor	MDFI	72.7	5.3	2.02E-17
Kinetochore scaffold 1	KNL1	149.9	5.3	4.68E-16
Neuron derived neurotrophic factor	NDNF	208.4	5.2	5.31E-14

Salmon [13]. Statistical analysis of data and differential gene expression was also performed by using the DESeq2 package of R [14, 15]. The magnitude of the differential gene expression between tumour and healthy samples was calculated by analysing the log-2-fold change of the genes (logFC, the cut-off value of 0.5). Benjamini-Hochberg method available in the DESeq2 package was used to adjust the nominal *p*-value (padj) in order to correct against false positive findings caused by multiple tests. The significance level was set at padj<0.05 [15].

The results from this study were used in the further meta-analysis and combined with two previously published data.

Meta-analysis

The results from four different whole transcriptome studies of osteosarcoma were combined to identify the differentially expressed genes. Two studies are from the present analysis,

FIGURE 1
Heat map of top differentially expressed genes with largest fold change differences between tumor and normal samples.
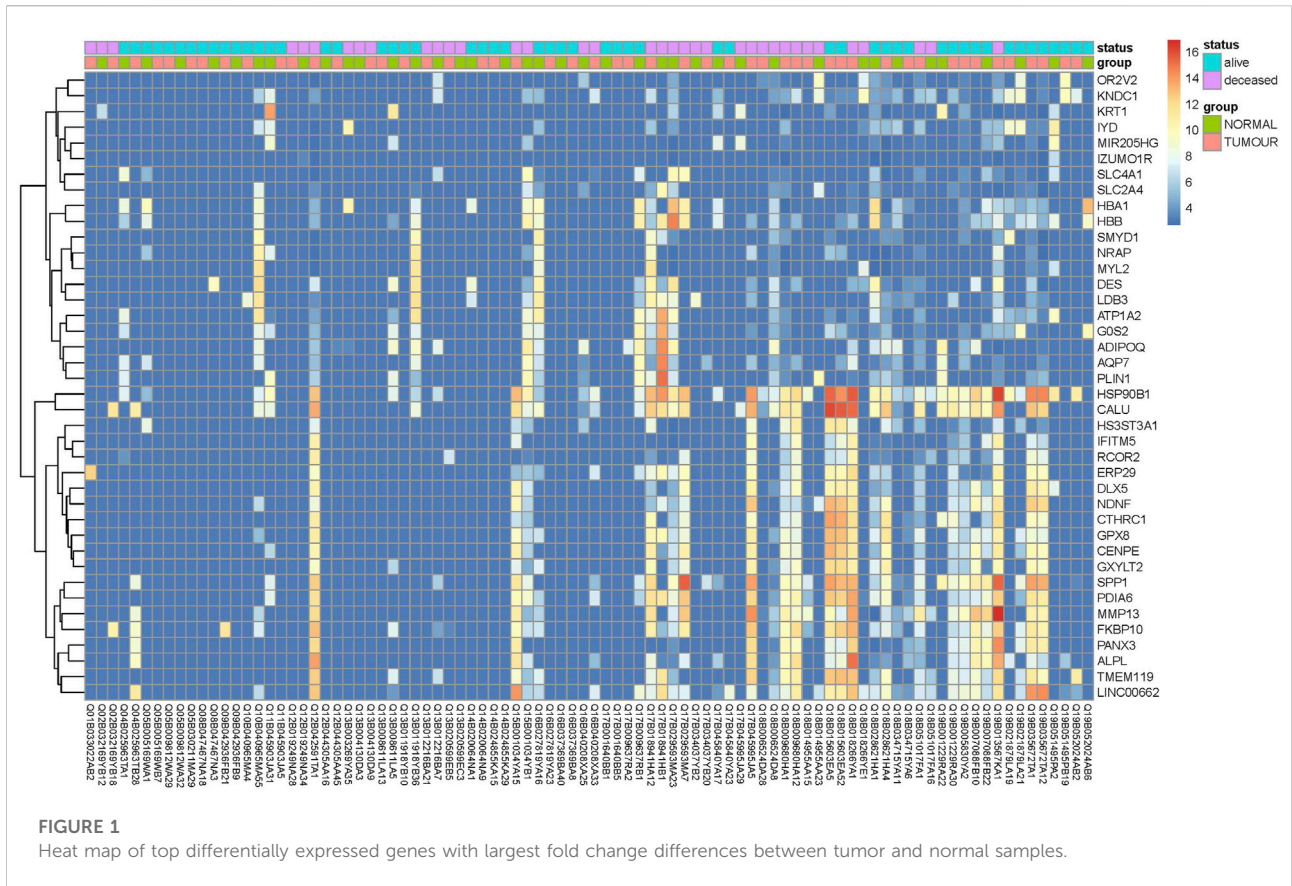
two others are from our previously published papers. Three studies were based on FFPE samples and one previous study from Estonia and Vietnam was based on fresh OS samples [16]. (Study 1: Estonia + Vietnam FFPE samples N = 15, Study 2: PathWest FFPE 1: N = 24 OS + 16 healthy controls, Study 3: PathWest FFPE 2N = 26 OS + 24 healthy controls and Study 4: Estonia and Vietnam fresh bones N = 18 OS + 18 healthy controls). The DEGs found in individual studies were compared to each other to find common DEGs.

Deconvolution with single-cell transcriptomics

Deconvolution was based on a previously published single-cell transcriptomic study in OS [17]. Cell profiles in scRNAseq study were from six OS patient samples (GSE162454). 'Seurat' package was used to find conserved markers that define cell clusters. To identify canonical cell type marker genes that are conserved across all conditions (tumour and treatment), we used the "FindConservedMarkers ()" function. This function tests differential gene expression for each dataset/group and combines the *p*-values using meta-analysis methods from the "MetaDE" R package [18]. "Granulator" package was used to

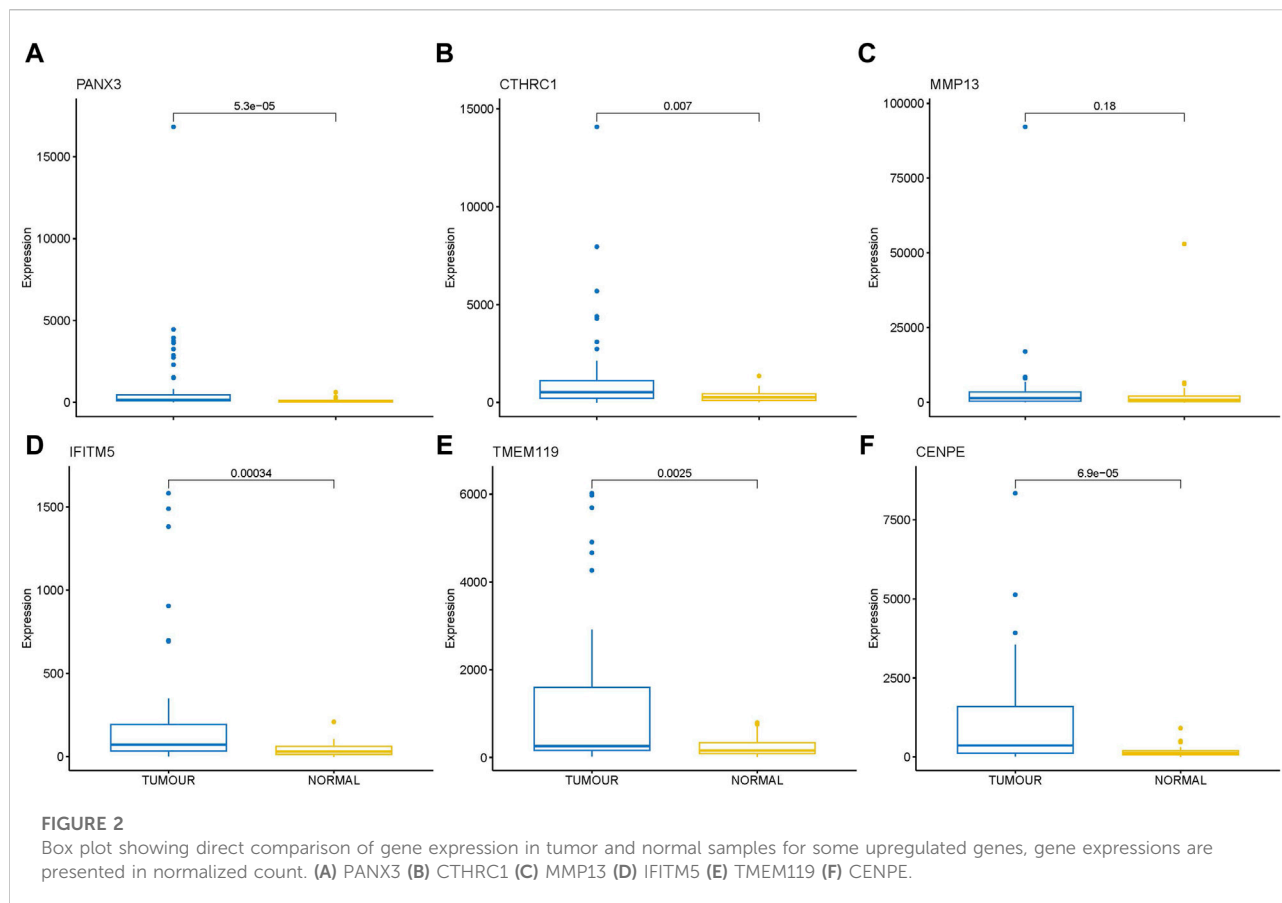identify the cluster-specific profiles in bulk FFPE RNAseq data.

Functional analysis of the differentially expressed genes

To understand the variation of different groups of gene functions in osteosarcoma, gene ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) were used to interpret the change in gene expression along with their cellular locations, biological processes, and involvement in the molecular pathway [19, 20].

Results

FFPE RNA-seq and meta-analysis

Transcriptome analysis of the PathWest FFPE samples showed a large number of differentially expressed genes in tumours compared to healthy bone, as shown in Table 1 and in Figure 1 (heatmap of the top genes form the FFPE 90 sample). Some of the overexpressed genes were *MMP9, CCN4, HAPLN1,*
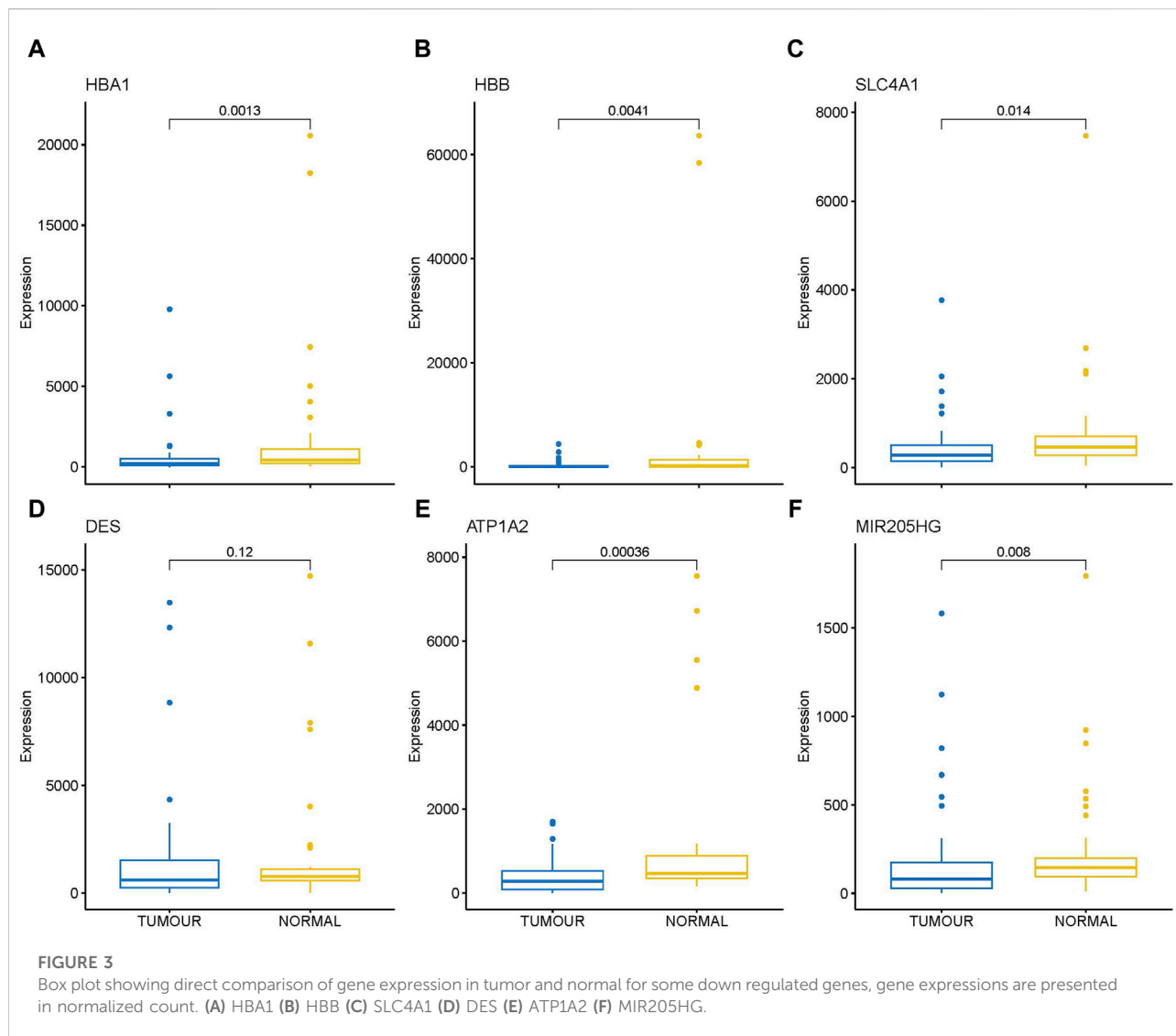
FIGURE 2
Box plot showing direct comparison of gene expression in tumor and normal samples for some upregulated genes, gene expressions are presented in normalized count. **(A)** PANX3 **(B)** CTHRC1 **(C)** MMP13 **(D)** IFITM5 **(E)** TMEM119 **(F)** CENPE.

ALPL, MMP13, PANX3, IFITM5, , CTHRC1, , TMEM119,. In addition, *COLLA11A1, FBN2, IBSP, COLA10A1, MDFI, GPX8* were also found to be overexpressed. A comparison of gene expressions in normal and tumours for some of the top differentially (upregulated and downregulated) expressed genes is presented in Figures 2, 3.

For the meta-analysis, two PathWest studies were treated as independent studies because the samples were collected, extracted and sequenced independently from each other. During the meta-analysis we observed 6,536 DEGs similar among our PathWest FFPE 2 study and EE-VN study. However, the number of overlapping DEGs was smaller (1,505) in PathWest FFPE 2 and PathWest FFPE 1 (Figures 4A, B). In addition, 1,211 DEGs were similar among PathWest FFPE1 and EE-VN FFPE study. Among those 105 FFPE samples from three studies, we observed 1,072 identical DEGs. The DEGs observed with fresh bone OS samples were then compared with combined FFPE samples, and we observed 530 similar genes to be differentially expressed as indicated in Figure 4, and the number of similar and unique genes across different samples are presented in the Venn diagram (Figure 4B). Some of the top differentially expressed genes observed in OS compared to normal samples are presented in Figure 2. The clustering shows that OS has a remarkable transcriptome heterogeneity

and indicates the molecular heterogeneity in the pathogenetic mechanisms causing the osteosarcoma. Our previous study found similar heterogeneity with only fresh tissue OS samples. Malignancies are known to have multiple mutations and complex molecular mechanisms that drive pathogenesis.

Deconvolution analysis was performed to identify the cellular populations from the FFPE-derived bulk RNAseq. Using the Seurat package and publicly available scRNA seq data from the OS sample, we could identify 11 clusters characteristic of the OS (Figures 5, 6). The clusters had specific patterns that could be designated to different cell populations, and cluster 2 was confirmed to be osteoblast-specific as it had typical osteoblast markers expressed. Our 90 samples from the Pathwest FFPE collection all had a stable proportion (20%) of osteoblastic cells Figure 5. Cluster 9 varied quite a lot between samples. The exact identity of these cellular clusters needs further analysis, but remarkably, deconvolution is possible from the FFPE-based bulk RNAseq data.

In addition, we compared the differential gene expression among the PathWest FFPE samples for chemo and non-chemo and observed 615 common DEGs (Figure 7). These genes are not affected by chemotherapy and are specific for the pathogenesis of the OS. Chemotherapy itself affected a large number of genes and could be a significant confounding factor. This is why

FIGURE 3
Box plot showing direct comparison of gene expression in tumor and normal for some down regulated genes, gene expressions are presented in normalized count. **(A)** HBA1 **(B)** HBB **(C)** SLC4A1 **(D)** DES **(E)** ATP1A2 **(F)** MIR205HG.
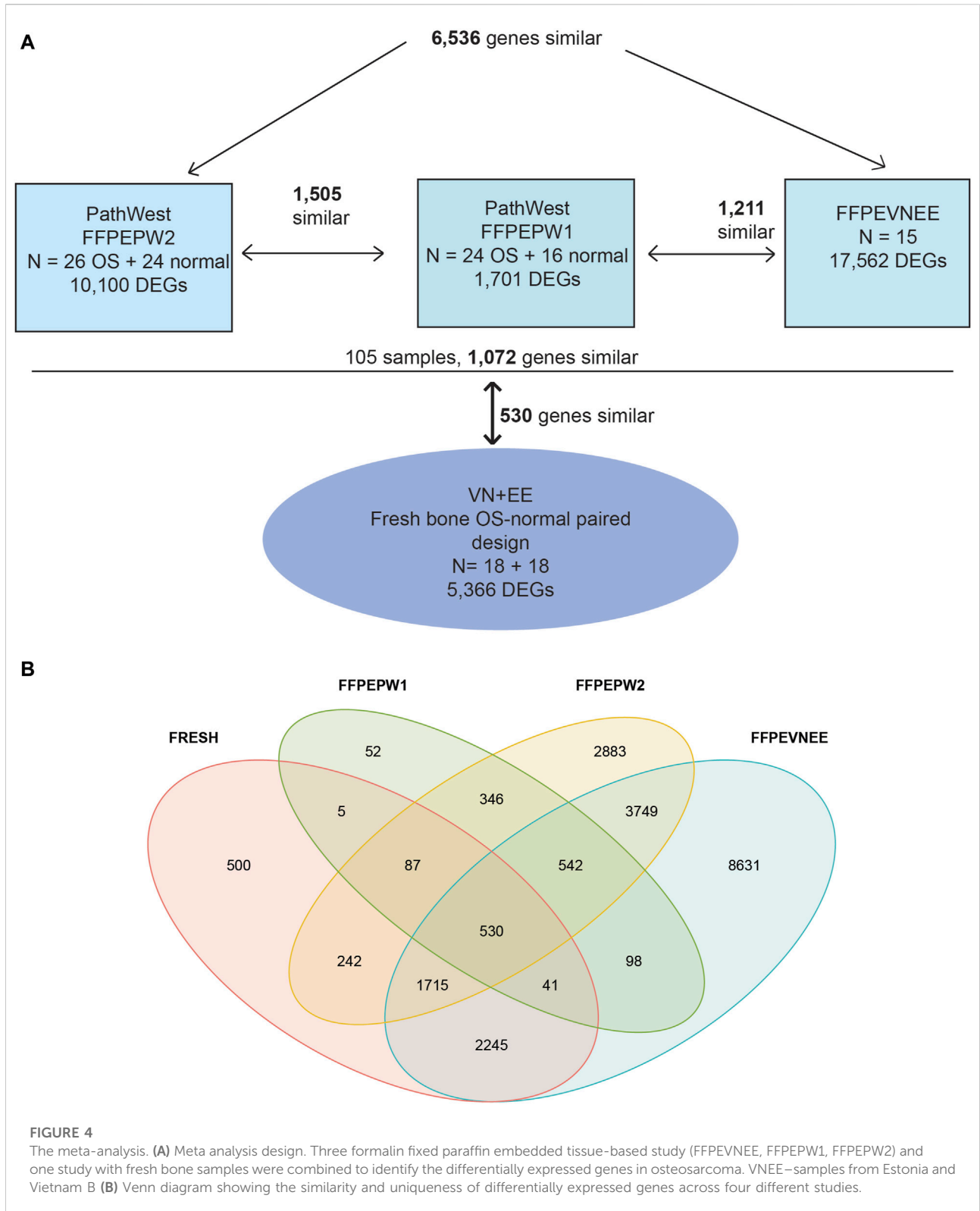
meta-analysis is needed to combine the statistical outcomes from independent studies. It is almost impossible to get clinical samples without treatment effects, and this is a common challenge for all real-life data-based studies, including genomics. We also observed many genes to be downregulated in OS compared to normal, as shown in Table 2. Some of the downregulated genes were SLC4A1, HBA1, HBA2, HBB, DES, ATP1A2, KRT1, LDB3, AQP7, FosB, G0S2, NRAP, and PLIN4, as shown in Table 2 with a log2FC value.

The volcano plot (Figure 8) shows the top highly expressed and top downregulated genes in osteosarcoma compared to normal samples, with some of the genes labelled and indicated on the plot. This gives a good overview of the extent and statistical significance of the differentially regulated genes in all studies combined and illustrates the transcriptomic fingerprint of osteosarcoma.

Functional analysis

GO and KEGG analyses were carried out to investigate the activation of the cellular pathways and common functional themes activated by the differentially expressed genes. This analysis is based on the differences among gene groups with various expressions and provides a preliminary interpretation of the biological activities of the differentially expressed genes. Similarly, disease ontology was performed to identify the disease patterns in the differentially expressed genes. We identified that the gene expression pattern we found matches with the osteosarcoma verifying our results again. A disease ontology study showed (Figure 9) that the DEGs are highly associated with osteosarcoma, bone cancer, connective tissue cancer, bone development disease and many other cancers (Figure 9) and also observed to be correlated with movement disease.

FIGURE 4
The meta-analysis. **(A)** Meta analysis design. Three formalin fixed paraffin embedded tissue-based study (FFPEVNEE, FFPEPW1, FFPEPW2) and one study with fresh bone samples were combined to identify the differentially expressed genes in osteosarcoma. VNEE−samples from Estonia and Vietnam B **(B)** Venn diagram showing the similarity and uniqueness of differentially expressed genes across four different studies.

Functional analysis was done to identify the cellular and molecular pathways that had changed in the OS samples. This lets us identify the common themes that connect these genes in this list.

KEGG pathway analysis is shown in Figure 10, and it describes the activation of the ribosome and protein processing in the endoplasmic reticulum. The KEGG pathway

FIGURE 5
Proportions of cell clusters in different normal and tumour samples.
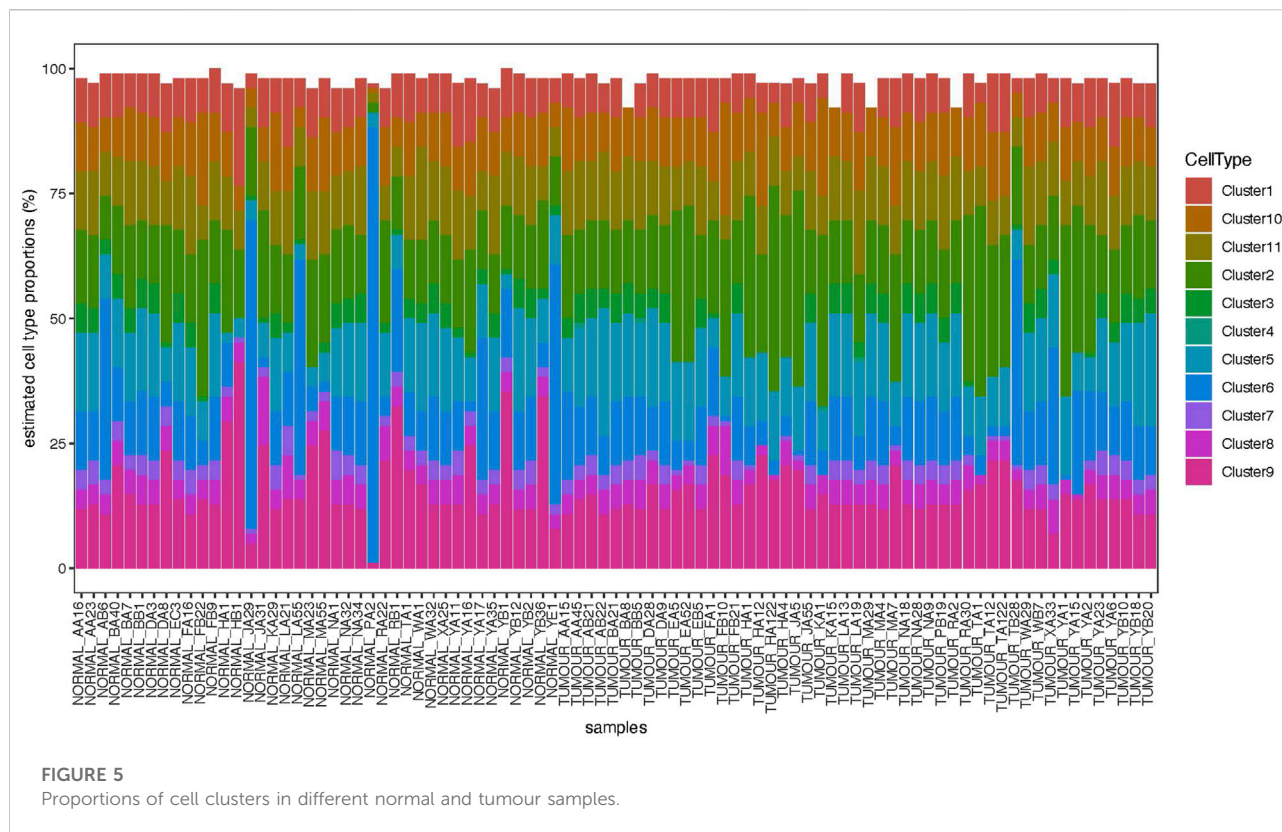
system is a unified canonical system to develop pathway maps for different cellular and biological functions. These maps comprehensively describe variable biological functions and provide maps for visualisation. The maps help us to define these molecular functions that are affected by osteosarcoma. KEGG pathway analysis also identified viral carcinogenesis pathways with some immune-regulated pathways.

Interestingly, a Parkinson's disease-related pathway was also identified (Figure 10). This finding possibly reflects the overlap between protein processing and Parkinson's pathogenesis pathways. It is important to mention here that the *PARK7* gene was identified as differentially expressed in all OS studies and samples we have analysed and the upregulation of P*ARK7* is a constant finding in OS cases. Therefore, activation of this pathway seems to be feasible and functionally plausible. *PARK7* is a gene that supports cell survival, and its upregulation has been shown to be related to aggressive brain cancer progression [21]. The functional deficiency of *PARK7* and mutations in this gene cause autosomal recessive early-onset Parkinson's disease.

Gene set enrichment analysis (GSEA) pathway enrichment

GSEA pathway enrichment was carried out to describe the common themes in the activated gene lists. This is based on the

gene set enrichment comparison by using the ranked gene list of the differential expression output. This way only statistically significant differences will be analysed against the existing genes sets [22]. In the OS samples, with the GSEA analysis, it was observed that many gene sets related to the cell cycle and mitosis regulation indicating presence of the malignant disease in OS in comparison to normal. To visualise the GSEA analysis results, we developed dot plots to illustrate the gene number and the statistical significance level for different gene sets and pathway that were activated in OS samples. Figure 11 shows the reactome pathways that are enriched in OS samples. Top 15 highly significant reactome pathways are shown with corresponding adjusted *p*-value and normalized count (Figure 11). Some of the major identified pathways in the plot are protein translation, cell cycle check points, gene silencing and keratinization.

Figures 11, 12 show statistical significance and gene numbers in different GSEA pathways. These illustrations combine the magnitude of the gene activation with the statistical significance. In both the figures we see the activation of cell cycle related pathways and mitosis signals that all reflect the activated malignant processes (Figure 12). Therefore, the genes we identified as differentially expressed in osteosarcoma samples with some indicated in Table 2 is reflective of the malignant process. Functional analysis of the gene activation networks can be even more specific and can go
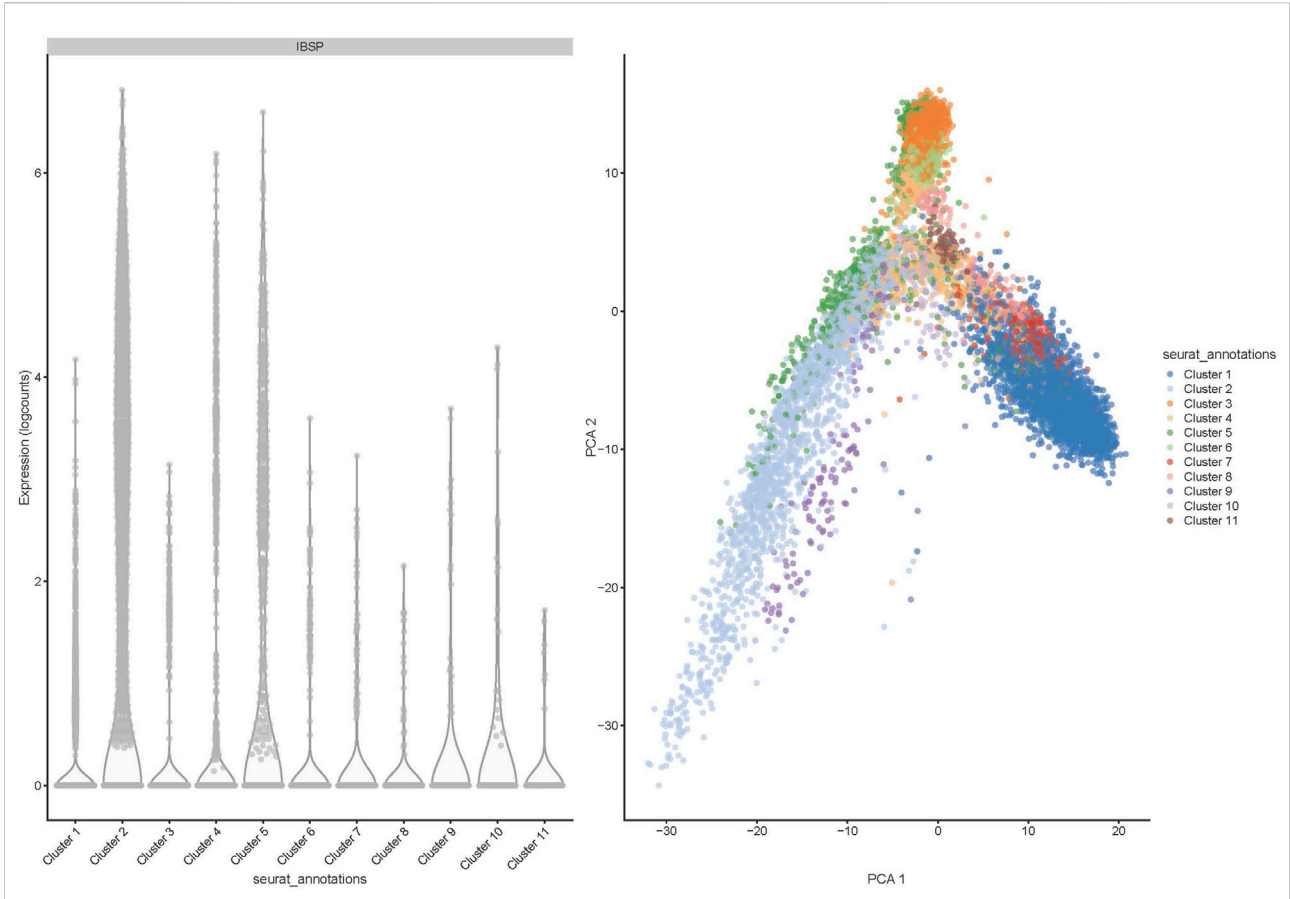
FIGURE 6
Distribution of 11 clusters of cells (shown in PCA1) and bone sialoprotein or integrin-binding sialoprotein (IBSP) expression in 11 different cell clusters shown as logcounts.
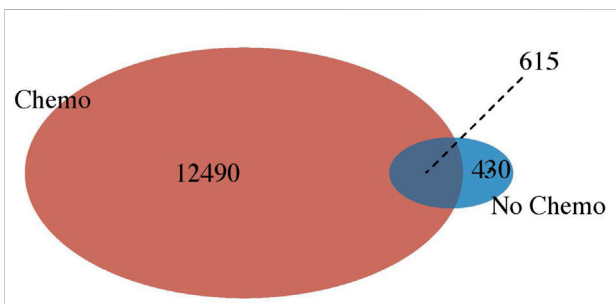

FIGURE 7
Venn diagram showing relationship between chemo-therapy verses non chemo-therapy.

from canonical pathways to the practical analysis of the disease pathologies. Gene set enrichment data were further analysed in the context of known pathologies and diseases by using disease ontology annotation. Interestingly, our dataset only indicated malignant diseases including different cancers, bone development and movement disease (Figure 9). Most

remarkably the top disease ontology matches were osteosarcoma and bone cancer.

Discussion

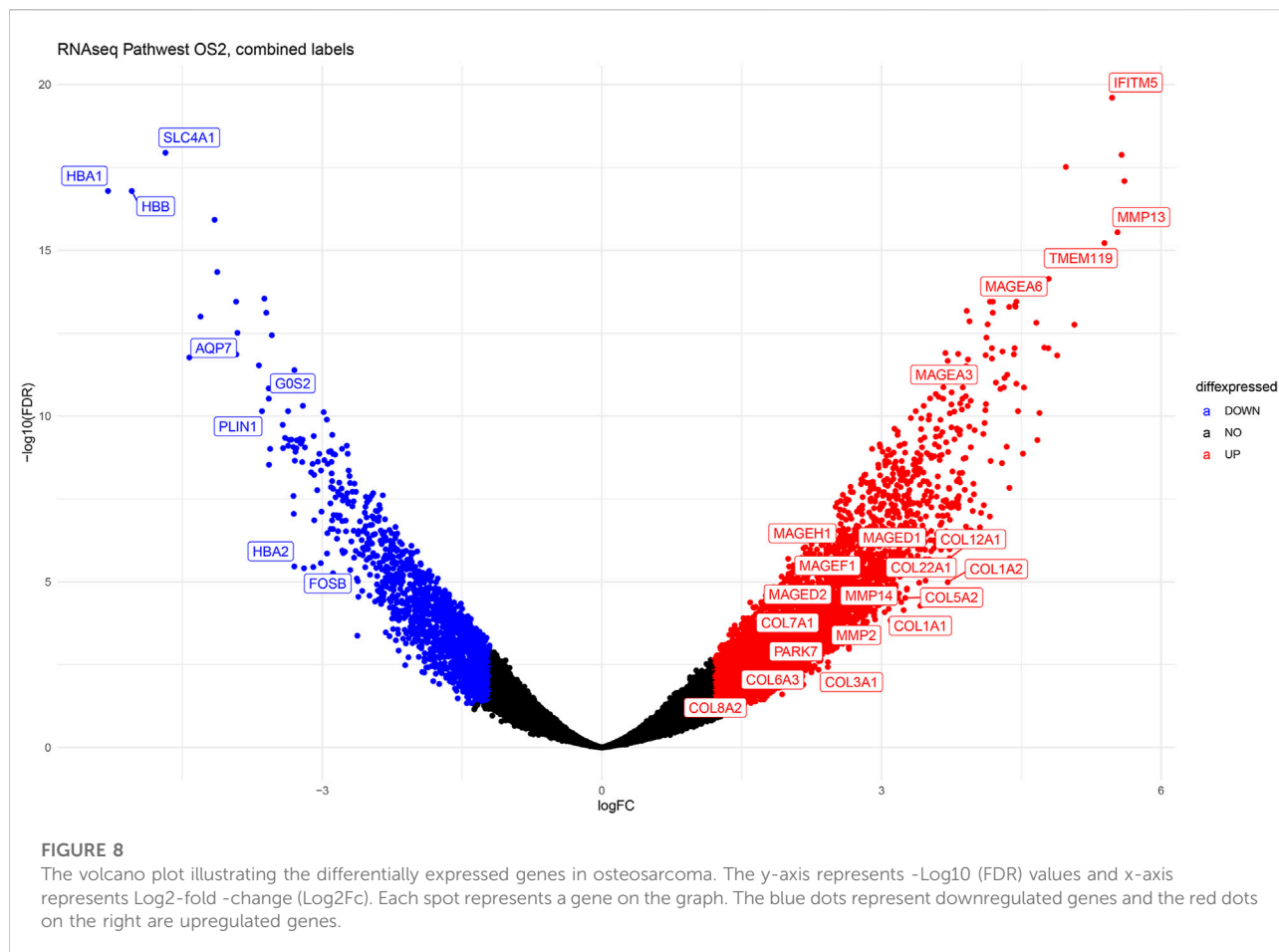### Highly upregulated genes in OS

IFITM5 (Interferon-induced transmembrane protein 5) gene was overexpressed in osteosarcoma samples compared to normal samples. It has been shown that *IFITM5* is expressed predominantly in bone tissue and plays a key role in the formation and mineralisation of bone [23, 24]. In particular, several studies have shown that *IFITM5* is overexpressed in osteosarcoma and also it was reported that the overexpression of IFITM5 has been shown to promote the growth and invasion of osteosarcoma cells *in vitro* and *in vivo* [24, 25]. The exact mechanism by which *IFITM5* promotes the development and progression of osteosarcoma is still not fully understood. However, it is thought that *IFITM5* may regulate the activity

TABLE 2 Top downregulated genes in osteosarcoma based on FFPE PathWest 2 study.

Gene name	Symbol	Base mean	Log2FoldChange	Padj
Hemoglobin subunit alpha 1	HBA1	133.6	−6.3	7.21E-25
Regulator of G protein signaling 6	RGS6	13.9	−4.2	2.36E-17
ATPase Na+/K+ transporting subunit alpha 2	ATP1A2	44.0	−4.1	2.62E-11
Keratin 2	KRT2	10.7	−4.0	1.51E-13
Hemoglobin subunit beta	HBB	61.0	−4.0	3.33E-11
LIM domain binding 3	LDB3	40.4	−3.9	8.24E-11
Desmocollin 1	DSC1	9.8	−3.9	5.88E-14
Polymeric immunoglobulin receptor	PIGR	9.6	−3.8	2.33E-13
Myosin binding protein C1	MYBPC1	29.4	−3.7	8.11E-09
Keratin 1	KRT1	7.7	−3.6	6.32E-13
WNK lysine deficient protein kinase 2	WNK2	23.3	−3.3	5.73E-09
Perilipin 4	PLIN4	19.8	−3.3	1.46E-09
Adenylate cyclase 5	ADCY5	14.2	−3.3	1.06E-09
GPR176 divergent transcript	GPR176-DT	10.1	−3.3	5.11E-11
Hemoglobin subunit alpha 2	HBA2	40.6	−3.2	1.08E-07
XK related 4	XKR4	20.8	−3.2	1.36E-08
Complement C7	C7	20.2	−3.1	1.22E-07
Cardiomyopathy associated 5	CMYA5	65.3	−3.1	1.12E-06
Fatty acid binding protein 4	FABP4	10.0	−3.1	4.29E-10
FosB proto-oncogene, AP-1 transcription factor subunit	FOSB	227.3	−3.1	1.55E-05
Tubulointerstitial nephritis antigen like 1	TINAGL1	27.2	−3.1	1.27E-07
G0/G1 switch 2	G0S2	19.4	−3.0	1.86E-08
Adiponectin, C1Q and collagen domain containing	ADIPOQ	36.3	−3.0	1.31E-06
Synaptophysin like 2	SYPL2	12.0	−3.0	2.06E-08
Potassium voltage-gated channel subfamily A regulatory beta subunit 1	KCNAB1	14.8	−3.0	3.75E-08
Hemicentin 2	HMCN2	17.9	−2.9	1.87E-07
G protein-coupled receptor class C group 5 member A	GPRC5A	19.4	−2.8	3.85E-07
Ceruloplasmin	CP	12.2	−2.8	1.30E-07
Enamelin	ENAM	8.7	−2.8	3.91E-09
Sorbin and SH3 domain containing 1	SORBS1	60.6	−2.8	6.77E-06

of osteoblasts and osteoclasts, which are cells that form and break down bone tissue, respectively. In addition, *IFITM5* may modulate the expression of genes that are important for cell proliferation, invasion, and survival. Overall, while more research is needed to fully understand the role of *IFITM5* in bone cancer, these findings suggest that *IFITM5* may be a potential target for the development of new treatments for osteosarcoma and other bone cancers.

Collagen triple helix repeat containing 1 (*CTHRC1*) is a gene that has been repeatedly shown to be overexpressed in osteosarcoma and is not only involved in carcinogenesis but is also a prognostic marker for malignancy, progression, and OS survival [26]. The high expression of *CTHRC1* has been shown to be correlated with differentiation, recurrency, chemotherapy response, and metastasis in patients with OS [27]. In addition, the survival analysis suggested that high expression of *CTHRC1*

FIGURE 8
The volcano plot illustrating the differentially expressed genes in osteosarcoma. The y-axis represents -Log10 (FDR) values and x-axis represents Log2-fold -change (Log2Fc). Each spot represents a gene on the graph. The blue dots represent downregulated genes and the red dots on the right are upregulated genes.
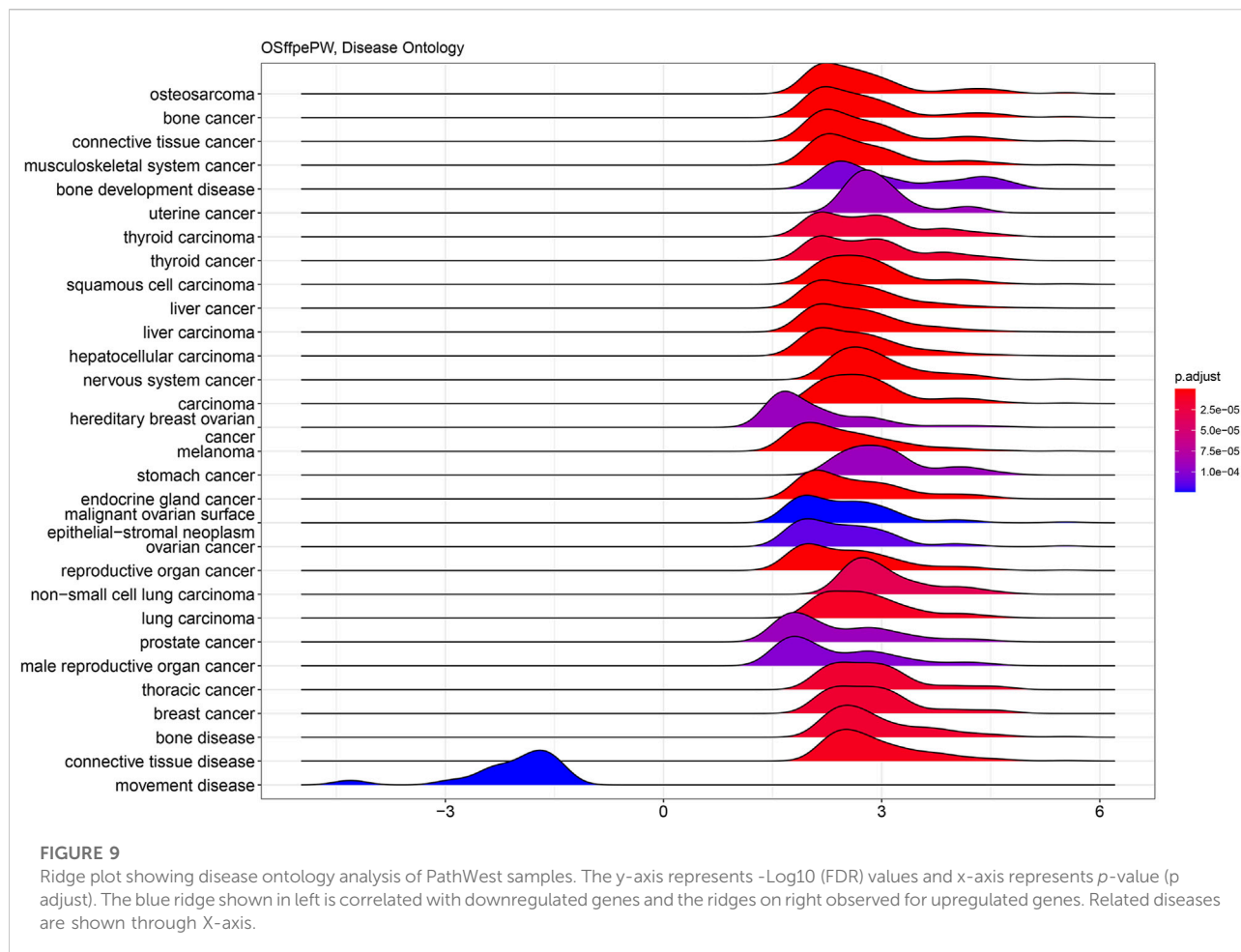
in OS patients correlates with a significantly shorter survival time. *CTHRC1* is related to metastasis development and osteosarcoma invasion. Recent research has shown that *CTHRC1* plays an important role in osteosarcoma progression. Lentivirus-mediated short hairpin RNA (shRNA) against *CTHRC1* significantly inhibited cell proliferation and colony formation in U2OS and SW1353 cells [26]. Flow cytometry assay showed that knockdown of *CTHRC1* increased the cell percentage of G0/G1 phase, resulting in cell cycle arrest in U2OS cells. Moreover, *CTHRC1* silencing induced the cell cycle arrest by a decrease in the cell percentage in G0/G1 phase and increased in G2/M phase in SW1353 cells. In addition, crystal violet staining suggested *CTHRC1* silencing inhibited migration of U2OS and SW1353 cells. Downregulation of *CTHRC1* would be an excellent target to stop the metastases of osteosarcoma.

Pannexin 3 is a member of the pannexin family of proteins and observed to be highly expressed in osteosarcoma compared to normal sample. *PANX3* regulates bone growth and differentiation [28] and is expressed in cartilage and regulates chondrocyte proliferation and differentiation [28]. A phenotypic analysis of Panx3-KO mice has indicated that *PANX3* regulates

the terminal differentiation of chondrocytes by promoting vascular endothelial growth factor (VEGF) and matrix metalloproteinase (MMP13) [29]. Therefore, *PANX3* is directly involved in the proliferation of the bone progenitor cells. In our studies, we observed *PANX3* was overexpressed in OS samples every time.

The next genes in our list with significant upregulation was the matrix metallopeptidase 13 (*MMP13*). This gene is very well known to be involved in the development of osteosarcoma [30, 31]. *MMP13* is involved in aggressive invasion and migration of the OS. It has been repeatedly shown that inhibition of the MMP13 expression will stop the invasion and growth of the OS [32]. Moreover, MMP13 interacts with other MMPs to form a network for osteosarcoma genes. *MMP13, MMP2, and MMP14* have been identified to interact with each other and to promote the progression and invasion of osteosarcoma [33, 34]. This network also interacts with the COL gene family. Furthermore, the upregulation of matrix metallopeptidase genes, including *MMP1, MMP2, MMP9, MMP11*, and *MMP16*, has been observed in osteosarcoma (OS) [35]. Previous studies have demonstrated the overexpression of
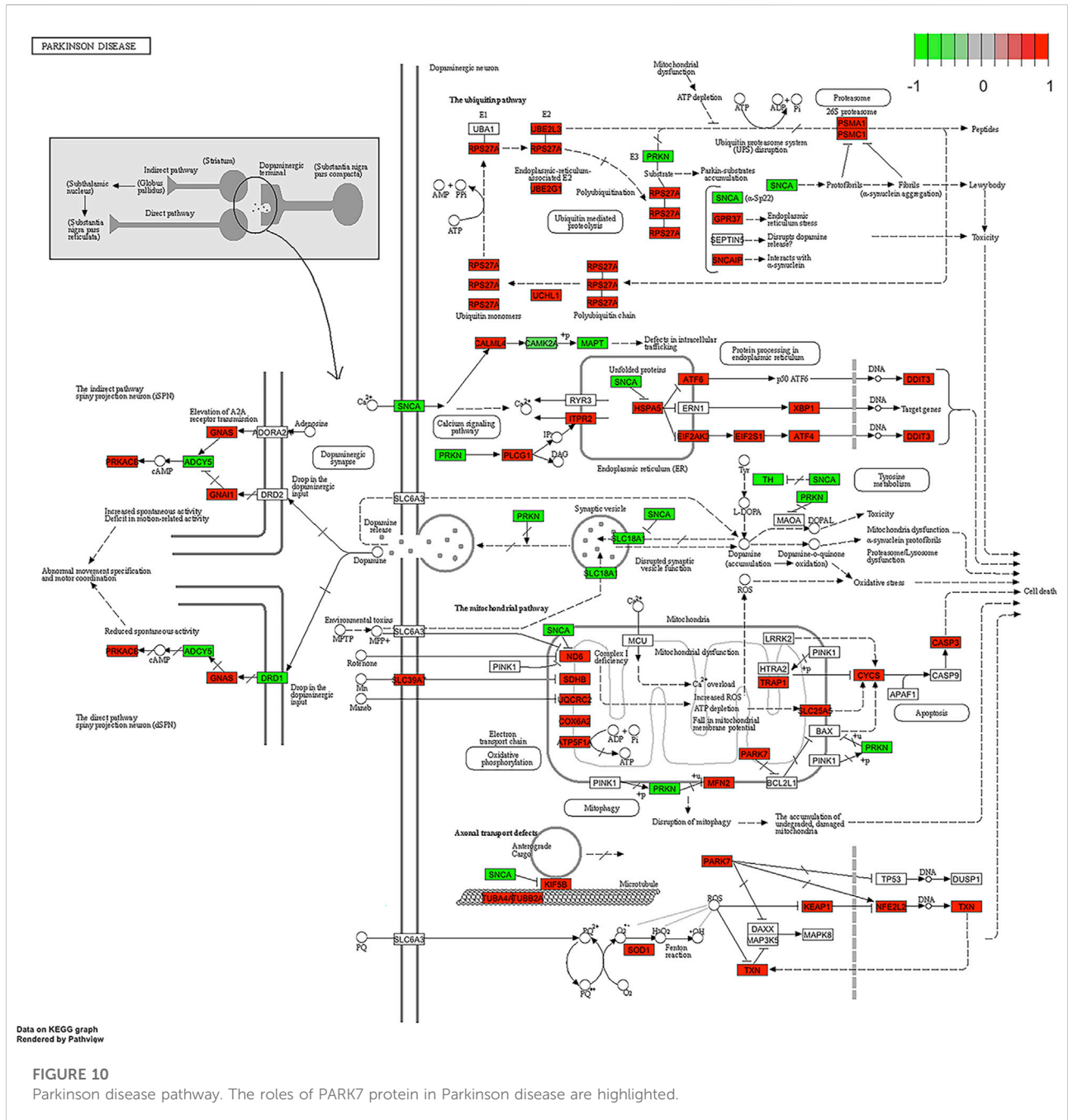
FIGURE 9

Ridge plot showing disease ontology analysis of PathWest samples. The y-axis represents -Log10 (FDR) values and x-axis represents *p*-value (p adjust). The blue ridge shown in left is correlated with downregulated genes and the ridges on right observed for upregulated genes. Related diseases are shown through X-axis.

these genes not only in OS but also in various types of cancers, where they play a crucial role in cancer survival [36, 37]. Consequently, MMPs could serve as promising diagnostic markers and potential drug targets for osteosarcoma.

TMEM119 (transmembrane protein 119) was observed to be highly upregulated in osteosarcoma derived samples as compared to healthy samples. The level of *TMEM119* protein expression was shown to be strongly associated with tumour size, clinical stage, distant metastases, and overall survival time [38]. Moreover, the gene set enrichment analysis revealed that *TMEM119* expression in osteosarcoma tissues is positively correlated with cell cycle, apoptosis, metastasis and TGF-β signaling. The reduction of the *TMEM119* expression in U2OS and MG63 cells using small interfering RNA revealed that downregulation of *TMEM119* could inhibit the proliferation of osteosarcoma cells by inducing cell cycle arrest in G0/G1 phase and apoptosis [38]. It was also found that *TMEM119* knockdown significantly inhibited cell migration and invasion and decreased the expression of TGF-β pathway-related factors. Several genes of the TMEM family have been shown to predict the survival of osteosarcoma patients.

Another important gene centromere protein F (*CENPF*) was found to be overexpressed in OS samples in compared to normal samples. *CENPF* has shown to play a key role in regulating the cell cycle and it was also shown that the increased level contributed to accelerate the cell proliferation by mediating apoptosis and cell cycle in OS [39].
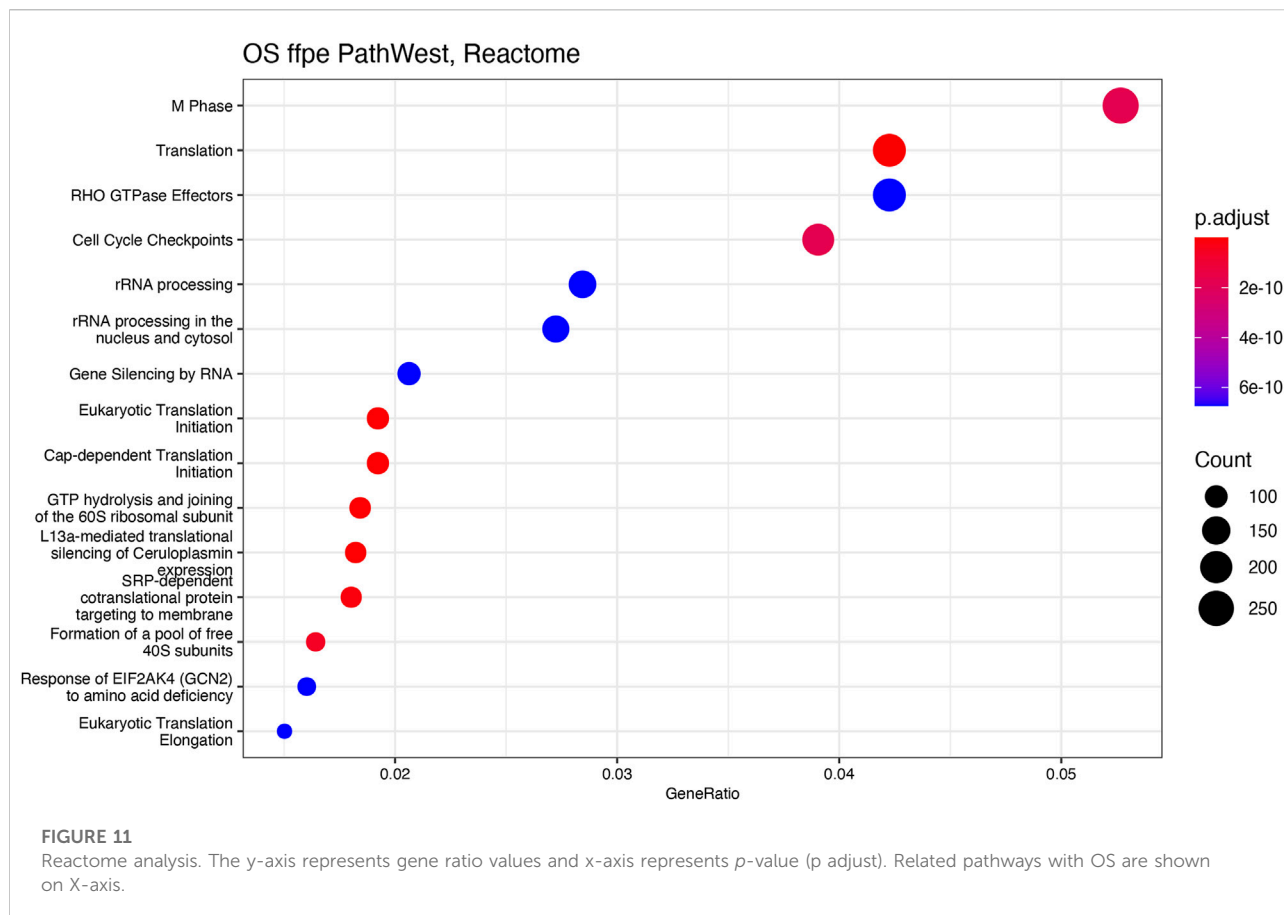
Haemoglobin related genes are downregulated in OS

Similarly, *HBA1, HBA2,* and *HBB* were downregulated in osteosarcoma samples. These are genes that encode haemoglobin subunits and are additionally involved in the malignancies. It was shown that overexpression of *HBA1* and *HBB* inhibits the cell proliferation, induces cellular apoptosis and block the cell cycle at the G2/M phase [40]. *HBB* and *HBA1* are now anti-metastatic factors in other cancers [40, 41] and the downregulation of these genes may be indicative for enhanced formation of metastases. *HBA1* and *HBB* mediate apoptosis and growth arrest on malignant cells, therefore the upregulation of these genes might be beneficial for the OS patients.

FIGURE 10
Parkinson disease pathway. The roles of PARK7 protein in Parkinson disease are highlighted.

FosB and cell cycle check point protein G0S2 are reduced in OS

FosB is a gene named as FBJ murine osteosarcoma viral oncogene homolog B and this gene was downregulated in OS samples. *FosB* is thought to play a role in the development and progression of osteosarcoma by promoting the proliferation and survival of cancer cells. In addition, *FosB* has been found to be involved in the regulation of genes that are important for bone formation, such as osteocalcin and collagen. This suggests that

FosB may also contribute to the development of osteosarcoma by altering the normal processes of bone formation and remodeling [42, 43]. Overall, while the exact role of *FosB* in osteosarcoma is still being studied, there is evidence to suggest that it may be a potential target for the development of new treatments for this aggressive form of bone cancer. The overexpression of *FosB* gene attenuated lung cancer growth and induced the death of the cancer cells. The downregulation of *FosB* has been shown to be negatively correlated with the cancer grade. Interestingly, the studies from other groups have found that *FosB* is a tumour

FIGURE 11
Reactome analysis. The y-axis represents gene ratio values and x-axis represents *p*-value (p adjust). Related pathways with OS are shown on X-axis.

suppressor gene [44]. Research has shown that the expression of *FosB* is reduced in osteosarcoma cells compared to normal bone cells.
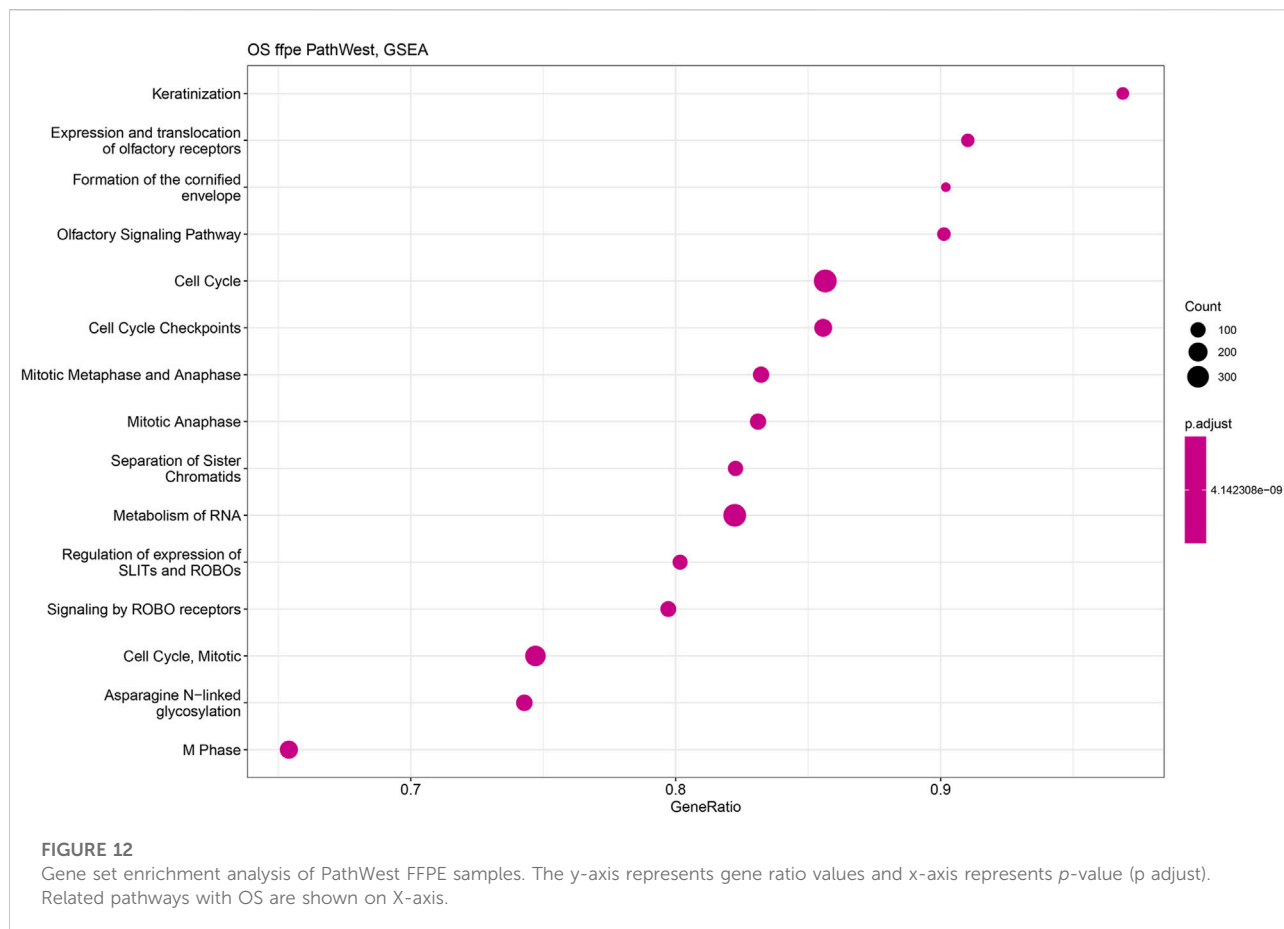
G0S2 is a G0/G1 switch gene 2 which was observed to be downregulated in osteosarcoma. It is a basic protein that inhibits the proliferation of stem cells. *G0S2* gene is a switch that has been reported to be involved in migration and invasiveness of the malignant cells [45, 46]. *G0S2* gene controls the cell cycle and the down regulation of this gene in the OS samples possibly indicates the mechanisms how OS evades the cell cycle control [47].

This study indicates that even using the partially degraded FFPE samples and with variable quality, the signature of the molecular genetic changes that are indicative for osteosarcoma can be identified. This finding validates our approach and makes the genes listed in Table 2 highly accurate for the OS analysis. The number of similar DEGs that exactly matched across our different independent studies were 530, this number seems to be low in comparison to the observed DEGs in some populations. This could be due to the heterogenic nature of our samples. However, the listed genes do not match with the top DEGs published by Yang et al., in 2014 [48]. It might be due to the heterogeneity of samples and the type of control samples used for the comparison. The enrichment of the GSEA sets and its comparison to the disease ontology showed very clearly that

the transcriptional profile we have discovered is related to the osteosarcoma and the genes we have identified are relevant for the OS pathology. GSEA enrichment indicated highly significant enrichment of the genetic networks related to osteosarcoma, bone cancer and connective tissue cancer. The first part of our analysis can be concluded that the FFPE sample analysis resulted in functionally meaningful and feasible results as we were able to identify the gene expression profile that is recognised by the complex computational enrichment analysis as osteosarcoma molecular network. Moreover, genes in Table 2 that is our formal list of statistically significant genes, contains many genes that are identified malignancy genes in literature. These malignancy genes are involved in osteosarcoma, other sarcomas, and in brain cancer.

Conclusion

Although the prevalence of osteosarcoma is relatively modest compared to other human cancers, the degree of malignancy is extremely high and poses a serious hazard to young patients [2, 3, 6]. The fact that neither the newly developed gene targeting therapy nor immunotherapy, which have been showing inspiring clinical effects in many other tumours, have received a positive

FIGURE 12
Gene set enrichment analysis of PathWest FFPE samples. The y-axis represents gene ratio values and x-axis represents *p*-value (p adjust). Related pathways with OS are shown on X-axis.

response in osteosarcoma, has contributed no notable improvement in patient survival over the past 30 years [6, 49]. It is important to explore the genetics of osteosarcoma in order to screen for prospective genes involved in cancer development and find promising targets. Analysis of differentially expressed gene expression could be a first step to identifying and establishing the marker and may potentially lead to the development of a potential drug target specific to osteosarcoma. From this study we were able to report the top upregulated and downregulated genes specific to OS in comparison to normal tissue which could assist to develop an efficient prognostic marker and therapeutic intervention specific to osteosarcoma.

We observed 530 differentially expressed genes in OS comparison to normal across our four different studies. The top highly expressed genes in OS comparison to normal controls are mostly found to be associated with extra cellular matrix related genes and have shown promise as clinical biomarker and therapeutics target in OS which needs further evaluations in preclinical models. Nevertheless, additional data analysis is required using fresh bone samples and validation of differential gene expression with qPCR and immunocytochemistry. The challenge lies in obtaining fresh bone samples, as not all patients undergo surgery. Another limitation in our study was the heterogeneous nature of the samples in terms of treatment; some underwent drug therapy, and such treatments may impact the gene expression pattern.

Author contributions

BP analyzed the data and wrote the manuscript. SK analyzed the data and reviewed and edited the manuscript.

Data availability

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, GSE253548.

Ethics statement

Funding

Acknowledgments

Conflict of interest

Supplementary material

The Supplementary Material for this article can be found online at: https://www.ebm-journal.org/articles/10.3389/ebm.2024.10161/full#supplementary-material

References

1. Jafari F, Javdansirat S, Sanaie S, Naseri A, Shamekh A, Rostamzadeh D, et al. Osteosarcoma: a comprehensive review of management and treatment strategies. *Ann Diagn Pathol* (2020) **49**:151654. doi:10.1016/j.anndiagpath.2020.151654

2. Taran R, Taran S, Malipatil NB. Pediatric osteosarcoma: an updated review. *Indian J Med Paediatr Oncol* (2017) **38**(1):33–43. doi:10.4103/0971-5851.203513

3. Ritter J, Bielack SS. Osteosarcoma. *Ann Oncol* (2010) **21**(Suppl. 7):vii320–5. doi:10.1093/annonc/mdq276

4. Fuchs B, Pritchard DJ. Etiology of osteosarcoma. *Clin Orthopaedics Relat Res* (2002) **397**:40–52. doi:10.1097/00003086-200204000-00007

5. Isakoff MS, Bielack SS, Meltzer P, Gorlick R. Osteosarcoma: current treatment and a collaborative pathway to success. *J Clin Oncol* (2015) **33**(27):3029–35. doi:10.1200/jco.2014.59.4895

6. Rathore R, Van Tine BA. Pathogenesis and current treatment of osteosarcoma: perspectives for future therapies. *J Clin Med* (2021) **10**(6):1182. doi:10.3390/jcm10061182

7. Dracham CB, Shankar A, Madan R. Radiation induced secondary malignancies: a review article. *Radiat Oncol J* (2018) **36**(2):85–94. doi:10.3857/roj.2018.00290

8. Morton LM, Onel K, Curtis RE, Hungate EA, Armstrong GT. The rising incidence of second cancers: patterns of occurrence and identification of risk factors for children and adults. *Am Soc Clin Oncol Educ Book* (2014) **34**:e57–67. doi:10.14694/edbook_am.2014.34.e57

9. Xin S, Wei G. Prognostic factors in osteosarcoma: a study level meta-analysis and systematic review of current practice. *J Bone Oncol* (2020) **21**:100281. doi:10.1016/j.jbo.2020.100281

10. Jiang Z, Zhou X, Li R, Michal JJ, Zhang S, Dodson MV, et al. Whole transcriptome analysis with sequencing: methods, challenges and potential solutions. *Cell Mol Life Sci* (2015) **72**(18):3425–39. doi:10.1007/s00018-015-1934-y

11. Yang IS, Kim S. Analysis of whole transcriptome sequencing data: workflow and software. *Genomics Inform* (2015) **13**(4):119–25. doi:10.5808/gi.2015.13.4.119

12. von Ahlfen S, Missel A, Bendrat K, Schlumpberger M. Determinants of RNA quality from FFPE samples. *PLoS One* (2007) **2**(12):e1261. doi:10.1371/journal.pone.0001261

13. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* (2017) **14**(4):417–9. doi:10.1038/nmeth.4197

14. Varet H, Brillet-Guéguen L, Coppée JY, Dillies MA. SARTools: a DESeq2-and EdgeR-based R pipeline for comprehensive differential analysis of RNA-seq data. *PLoS One* (2016) **11**(6):e0157022. doi:10.1371/journal.pone.0157022

15. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* (2014) **15**(12):550. doi:10.1186/s13059-014-0550-8

16. Ho XD, Phung P, Q Le V, H Nguyen V, Reimann E, Prans E, et al. Whole transcriptome analysis identifies differentially regulated networks between osteosarcoma and normal bone samples. *Exp Biol Med (Maywood)* (2017) **242**(18):1802–11. doi:10.1177/1535370217736512

17. Liu Y, Feng W, Dai Y, Bao M, Yuan Z, He M, et al. Single-cell transcriptomics reveals the complexity of the tumor microenvironment of treatment-naive osteosarcoma. *Front Oncol* (2021) **11**:709210. doi:10.3389/fonc.2021.709210

18. Polanin JR, Hennessy EA, Tanner-Smith EE. A review of meta-analysis packages in R. *J Educ Behav Stat* (2016) **42**:206–42. doi:10.3102/1076998616674315

19. Du J, Yuan Z, Ma Z, Song J, Xie X, Chen Y. KEGG-PATH: Kyoto encyclopedia of genes and genomes-based pathway analysis using a path analysis model. *Mol Biosyst* (2014) **10**(9):2441–7. doi:10.1039/c4mb00287c

20. Chen L, Zhang YH, Wang S, Zhang Y, Huang T, Cai YD. Prediction and analysis of essential genes using the enrichments of gene ontology and KEGG pathways. *PLoS One* (2017) **12**(9):e0184129. doi:10.1371/journal.pone.0184129

21. Jin W. Novel insights into PARK7 (DJ-1), a potential anti-cancer therapeutic target, and implications for cancer progression. *J Clin Med* (2020) **9**(5):1256. doi:10.3390/jcm9051256

22. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, 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**(43):15545–50. doi:10.1073/pnas.0506580102

23. Sun Y, Zhang C, Fang Q, Zhang W, Liu W. Abnormal signal pathways and tumor heterogeneity in osteosarcoma. *J Transl Med* (2023) **21**(1):99. doi:10.1186/s12967-023-03961-7

24. Patoine A, Husseini A, Kasaai B, Gaumond MH, Moffatt P. The osteogenic cell surface marker BRIL/IFITM5 is dispensable for bone development and homeostasis in mice. *PLoS One* (2017) **12**(9):e0184568. doi:10.1371/journal.pone.0184568

25. Jiang H, Du H, Liu Y, Tian X, Xia J, Yang S. Identification of novel prognostic biomarkers for osteosarcoma: a bioinformatics analysis of differentially expressed genes in the mesenchymal stem cells from single-cell sequencing data set. *Translational Cancer Res* (2022) **11**(10):3841–52. doi:10.21037/tcr-22-2370

26. Sang W, Zhu L, Ma J, Lu H, Wang C. Lentivirus-mediated knockdown of CTHRC1 inhibits osteosarcoma cell proliferation and migration. *Cancer Biother Radiopharm* (2016) **31**(3):91–8. doi:10.1089/cbr.2014.1758

27. Mei D, Zhu Y, Zhang L, Wei W. The role of CTHRC1 in regulation of multiple signaling and tumor progression and metastasis. *Mediators Inflamm* (2020) **2020**:9578701–13. doi:10.1155/2020/9578701

28. Iwamoto T, Nakamura T, Doyle A, Ishikawa M, de Vega S, Fukumoto S, et al. Pannexin 3 regulates intracellular ATP/cAMP levels and promotes chondrocyte differentiation. *J Biol Chem* (2010) **285**(24):18948–58. doi:10.1074/jbc.m110. 127027

29. Ishikawa M, Yamada Y. The role of pannexin 3 in bone biology. *J Dent Res* (2017) **96**(4):372–9. doi:10.1177/0022034516678203

30. Cui J, Dean D, Hornicek FJ, Chen Z, Duan Z. The role of extracelluar matrix in osteosarcoma progression and metastasis. *J Exp Clin Cancer Res* (2020) **39**(1):178. doi:10.1186/s13046-020-01685-w

31. Li S, Pritchard DM, Yu LG. Regulation and function of matrix metalloproteinase-13 in cancer progression and metastasis. *Cancers (Basel)* (2022) **14**(13):3263. doi:10.3390/cancers14133263

32. Ma O, Cai WW, Zender L, Dayaram T, Shen J, Herron AJ, et al. MMP13, Birc2 (cIAP1), and Birc3 (cIAP2), amplified on chromosome 9, collaborate with p53 deficiency in mouse osteosarcoma progression. *Cancer Res* (2009) **69**(6): 2559–67. doi:10.1158/0008-5472.can-08-2929

33. Tang H, Tang Z, Jiang Y, Wei W, Lu J. Pathological and therapeutic aspects of matrix metalloproteinases: implications in osteosarcoma. *Asia-Pacific J Clin Oncol* (2019) **15**(4):218–24. doi:10.1111/ajco.13165

34. Wang Z, Jian M, Li X. Profiling of multiple matrix metalloproteinases activities in the progression of osteosarcoma by peptide microarray-based fluorescence assay on polymer brush-coated zinc oxide nanorod substrate. *Methods Mol Biol* (2023) **2578**:161–75. doi:10.1007/978-1-0716-2732-7_11

35. Zhou J, Liu T, Wang W. Prognostic significance of matrix metalloproteinase 9 expression in osteosarcoma: a meta-analysis of 16 studies. *Medicine (Baltimore)* (2018) **97**(44):e13051. doi:10.1097/md.0000000000013051

36. Amakye D, Gyan PO, Santa S, Aryee NA, Adu-Bonsaffoh K, Quaye O, et al. Extracellular matrix metalloproteinases inducer gene polymorphism and reduced serum matrix metalloprotease-2 activity in preeclampsia patients. *Exp Biol Med (Maywood)* (2023) **248**(18):1550–5. doi:10.1177/15353702231199464

37. Quintero-Fabian S, Arreola R, Becerril-Villanueva E, Torres-Romero JC, Arana-Argáez V, Lara-Riegos J, et al. Role of matrix metalloproteinases in angiogenesis and cancer. *Front Oncol* (2019) **9**:1370. doi:10.3389/fonc.2019.01370

38. Jiang ZH, Peng J, Yang HL, Fu XL, Wang JZ, Liu L, et al. Upregulation and biological function of transmembrane protein 119 in osteosarcoma. *Exp Mol Med* (2017) **49**(5):e329. doi:10.1038/emm.2017.41

39. Zou PA, Yang ZX, Wang X, Tao ZW. Upregulation of CENPF is linked to aggressive features of osteosarcoma. *Oncol Lett* (2021) **22**(3):648. doi:10.3892/ol. 2021.12909

40. Luo P, Liu X, Tang Z, Xiong B. Decreased expression of HBA1 and HBB genes in acute myeloid leukemia patients and their inhibitory effects on growth of K562 cells. *Hematology* (2022) **27**(1):1003–9. doi:10.1080/16078454.2022. 2117186

41. Liu X, Gao Y, Zhao B, Li X, Lu Y, Zhang J, et al. Discovery of microarray-identified genes associated with ovarian cancer progression. *Int J Oncol* (2015) **46**(6):2467–78. doi:10.3892/ijo.2015.2971

42. Inoue D, Kido S, Matsumoto T. Transcriptional induction of FosB/ΔFosB gene by mechanical stress in osteoblasts. *J Biol Chem* (2004) **279**(48):49795–803. doi:10.1074/jbc.m404096200

43. Lam SW, Cleven AHG, Kroon HM, Briaire-de Bruijn IH, Szuhai K, Bovée JVMG. Utility of FOS as diagnostic marker for osteoid osteoma and osteoblastoma. *Virchows Arch* (2020) **476**(3):455–63. doi:10.1007/s00428-019-02684-9

44. Tang C, Jiang Y, Shao W, Shi W, Gao X, Qin W, et al. Abnormal expression of FOSB correlates with tumor progression and poor survival in patients with gastric cancer. *Int J Oncol* (2016) **49**(4):1489–96. doi:10.3892/ijo. 2016.3661

45. Fukunaga T, Fujita Y, Kishima H, Yamashita T. Methylation dependent down-regulation of G0S2 leads to suppression of invasion and improved prognosis of IDH1-mutant glioma. *PLoS One* (2018) **13**(11):e0206552. doi:10.1371/journal. pone.0206552

46. Yim CY, Sekula DJ, Hever-Jardine MP, Liu X, Warzecha JM, Tam J, et al. G0S2 suppresses oncogenic transformation by repressing a MYC-regulated transcriptional program. *Cancer Res* (2016) **76**(5):1204–13. doi:10.1158/0008-5472.can-15-2265

47. Heckmann BL, Zhang X, Xie X, Liu J. The G0/G1 switch gene 2 (G0S2): regulating metabolism and beyond. *Biochim Biophys Acta (Bba) - Mol Cell Biol Lipids* (2013) **1831**(2):276–81. doi:10.1016/j.bbalip.2012.09.016

48. Yang Z, Chen Y, Fu Y, Yang Y, Zhang Y, Chen Y, et al. Meta-analysis of differentially expressed genes in osteosarcoma based on gene expression data. *BMC Med Genet* (2014) **15**:80. doi:10.1186/1471-2350-15-80

49. Prudowsky ZD, Yustein JT. Recent insights into therapy resistance in osteosarcoma. *Cancers (Basel)* (2020) **13**(1):83. doi:10.3390/cancers13010083