Identification of key genes in osteosarcoma by meta-analysis of gene expression microarray

JUNKUI SUN¹, HONGEN XU², MUGE Qi³, CHI ZHANG¹ and JIANXUANG SHI²,⁴

¹Department of Orthopedic Surgery, The First Affiliated Hospital of Zhengzhou University; ²Precision Medicine Center, Henan Institute of Medical and Pharmaceutical Sciences, Zhengzhou University, Zhengzhou, Henan 450052; ³Henan Center for Disease Control and Prevention, Zhengzhou, Henan 450016; ⁴Henan Key Laboratory of Tumor Epidemiology, College of Public Health, Zhengzhou University, Zhengzhou, Henan 450052, P.R. China

Received November 21, 2018; Accepted June 17, 2019

DOI: 10.3892/mmr.2019.10543

Abstract. Osteosarcoma (oS) is one of the most malignant tumors in children and young adults. To better understand the underlying mechanism, five related datasets deposited in the Gene Expression Omnibus were included in the present study. The Bioconductor ‘limma’ package was used to identify differentially expressed genes (deGs) and the ‘Weighted Gene Co-expression Network Analysis’ package was used to construct a weighted gene co-expression network to identify key modules and hub genes, associated with OS. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes overrepresentation analyses were used for functional annotation. The results indicated that 1,405 genes were dysregulated in oS, including 927 upregulated and 478 downregulated genes, when the cut off value was set at a ≥2 fold-change and an adjusted P-value of P<0.01 was used. Functional annotation of deGs indicated that these genes were involved in the extracellular matrix (ECM) and that they function in several processes, including biological adhesion, ECM organization, cell migration and leukocyte migration. These findings suggested that dysregulation of the ECM shaped the tumor microenvironment and modulated the OS hallmark. Genes assigned to the yellow module were positively associated with oS and could contribute to the development of oS. In conclusion, the present study has identified several key genes that are potentially druggable genes or therapeutics targets in OS. Functional annotations revealed that the dysregulation of the ECM may contribute to OS development and, therefore, provided new insights to improve our understanding of the mechanisms underlying OS.

Introduction

Osteosarcoma (oS) is a primary malignant bone tumor arising from primitive transformed cells of a mesenchymal origin (1) and it is the 8th most common form of childhood cancer (2). Although OS is a rare malignancy overall, it is the most common malignant tumor found in the bone tissue of children and usually requires chemotherapy and surgical treatment (3,4). The management of OS has improved over the past few decades, with the 5-year survival rate increasing from 20-30 to 60-70% (5). However, relapse and pulmonary metastasis remain big challenges in the management of OS (6). A greater understanding of the underlying mechanisms of OS will improve its management.

It has been reported that most cases of oS harbor chromosomal abnormalities and gene mutations (7,8). In total, ~70% of patients with OS showed loss-of-function mutations in the gene encoding the retinoblastoma-associated protein (9,10). Somatic mutations that lead to the loss of tumor suppressor functions are a pivotal step in oS pathogenesis, and there are a variety of genetic events that lead to the development of OS (11). Systematic research from the genetic perspective may help to improve our understanding of the mechanism underlying OS.

Gene microarrays are a powerful tool to obtain gene expression profiles. Comparisons made between normal and tumor samples can lead to the identification of dysregulated genes; most diseases have specific gene expression profiles and abnormal regulation patterns (12). A common practice for the identification of differentially expressed genes (DEGs) is to filter results using fold change, P-values and false discovery
rates (13,14). Weighted gene co-expression network analysis (WGCNA) can be used to identify groups of genes with similar functions, known as gene modules (15). Genes in the same module tend to have similar expression patterns and, therefore, may have similar functions. Genes with the most connectivity in a module are called hub genes; these genes are more relevant to the functionality of the module (14). WGCNA has been widely accepted as an investigation tool to identify hub genes in cancer studies.

In the present study, microarray gene expression data derived from the same platform were extracted from the Gene Expression Omnibus (GEO) database to identify DEGs in OS. WGCNA was used to identify gene modules that were closely associated with OS. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were used for functional annotations. PharmGKB (16), oncoKB (17), Clinical Interpretations of Variants in Cancer (CIViC) (18) were used to check if potentially druggable targets could be found in closely related modules in OS. The results of the present study may increase the understanding of the molecular mechanisms underlying OS and contribute to the clinical management of OS.

Materials and methods

Search strategy. The GEO database (www.ncbi.nlm.nih.gov/geo/) was used to retrieve relevant studies (19-23) that used the Affymetrix Human Genome U133 Plus 2.0 platform (GPL570; Affymetrix; Thermo Fisher Scientific, Inc.) to explore the mRNA expression profiles in tumor tissues from patients with OS or bone marrow mesenchymal stromal cells (BM-MSCs) from healthy controls. Search terms including ‘osteosarcoma’, ‘cancer’ or ‘tumor’ or ‘neoplasm’ or ‘carcinoma’ or ‘sarcoma’, ‘mesenchymal stromal cells’ and ‘GPL570’ were used. The species was limited to Homo sapiens.

Study selection. Inclusion criteria: i) Studies that used OS tissues from patients or BM-MSCs from healthy controls to explore the RNA expression profiles; and ii) for studies that used mixed tissue types from patients with OS or healthy controls, only the data from OS tissues and normal BM-MSCs were included. Exclusion criteria: i) Studies that used cell lines derived from OS or human mesenchymal stromal cells were excluded; ii) studies that used BM-MSCs extracted from patients with osteoarthritis or osteoporosis were excluded; and iii) studies with only ‘chp’ type original files, other than ‘cel’ type files, were also excluded.

Data extraction and pre-processing procedures. Raw data from eligible studies were retrieved from GEO. GEO accession number, author, country, submission year, platform and detailed patient information, as well as available information from healthy controls, were obtained from the metadata. Data were extracted by two researchers independently and conflicts were resolved by consulting a third senior researcher. Raw data were normalized using the R ‘affy’ package (version 1.62.0; http://bioconductor.org/packages/affy) with robust-multi array average methods, as described previously (24-26). Mean expression values were calculated for genes measured by multiple detection probes. DEGs between patients with OS and control tissues were compared using the Bioconductor ‘limma’ package (version 3.40.2; http://bioconductor.org/packages/limma) (27). Genes with a fold change ≥2 and an adjusted P<0.01 were considered as DEGs.

Functional characterization of DEGs. Microarray probe IDs were converted to Ensemble IDs and gene symbols using ‘hgu133plus2.db’ R package (version 3.2.3; http://bioconductor.org/packages/hgu133plus2.db) (28). To interpret the biological significance of DEGs, GO enrichment of cellular component, biological process and molecular function, as well as KEGG pathway enrichment analysis were conducted using Bioconductor ‘clusterProfiler’ R package (version 3.10.0; http://bioconductor.org/packages/clusterProfiler) (29). The ‘Disease Ontology semantic and enrichment analysis’ (DOSE) package (version 3.10.0; http://bioconductor.org/packages/DOSE) (30) was used to find genes closely associated with OS.

Principal component analysis (PCA) of DEGs in patients with OS and controls. PCA analyses were conducted using the ClustVis online tool (https://biit.cs.ut.ee/clustvis/) developed by Metsalu et al (31). Due to limitations on the file size that can be uploaded, only gene expression values of DEGs were included in the PCA analysis. Groups (OS or control) and gender were two of the clinical traits that were used in the PCA analysis.

WGCNA. To identify key gene modules in OS, WGCNA was conducted with the R ‘WGCNA’ package (version 1.46) (32). Normalized gene expression data were used in WGCNA. Soft-connectivity was calculated using the default parameters. Topology networks and gene modules were constructed using one-step network construction.

Hub genes are a group of genes that tend to have high connectivity with other genes and are expected to play pivotal biological roles. The connections between the top 30 hub genes were visualized using VisAnt software (version 5.51; http://visant.bu.edu). Functional annotations, including GO and KEGG enrichment analyses, were used to highlight the most overrepresented GO terms and KEGG pathways in modules that were closely associated with OS. To determine if any of the hub genes in the modules were abnormally expressed, log2 fold change (log2FC) was used to characterize the expression pattern and enrichment scores were used for characterizing the connectivity of genes in the yellow module.

Gene mutations may prevent the proper function of the corresponding protein by affecting protein structure or expression. PharmGKB (16), oncoKB (17), and CIViC (18) are three databases that provide information about the treatment implications of specific cancer gene alterations, and how these mutations affect response to treatment. In the present study, these databases were used to identify any potentially druggable targets in the modules that were found to be closely related with OS.

Tumor infiltrating immune cell profiling using CIBERSORT. Tumor microenvironments are critical to tumor cell survival and proliferation. Tumor infiltrating leukocytes are usually present in the microenvironment of solid tumors. The CIBERSORT
webtool (version 1.06; https://cibersort.stanford.edu) (33) was used to estimate the abundance of tumor-infiltrating immune cells in the OS microenvironment. LM22, which consisted of gene expression data from 22 distinct immune cell types, was used as reference in the present study (34).

Results

Characteristics of the included studies. In total, five eligible studies were included in the present study (19-23). RNA expression data were extracted from 48 patients with OS and 12 BM-MSCs from these previous studies (Table SI). Details of the included studies are shown in Table I. For GSE18043 and GSE36474, only data from three eligible individuals were included from each dataset (19,22). There were two biological replicates in GSE35331 (21), the results from the first set were included in the present study. More detailed information of the included individuals from each study can be found in Table SI. The original gene expression files from the included individuals were downloaded from the GEO website.

Identification and functional annotation of DEGs in OS. The Bioconductor ‘affy’ package was used to pre-process raw data for background correction and normalization. In total, expression values from 54,613 probes representing 20,188 known genes with symbols were analyzed in the present study. The mean expression values of multiple probes corresponding to each gene were calculated as the final expression value. The Bioconductor ‘limma’ package was used to identify DEGs. When the cutoff values were set as |log2Fc|>1 (adjusted P<0.01), 1,405 genes were found to be dysregulated (including 927 up- and 478 downregulated genes) in OS compared with controls (Table SII). When cutoff values were set as |log2Fc|>2 (adjusted P<0.01), there were 354 genes dysregulated (including 224 up- and 130 downregulated genes) in OS compared with controls. The top 10 most up- and downregulated genes are shown in Tables II and III. PCA analysis revealed that these DEGs could distinguish OS from normal controls and that there was no disparity between males and females (Fig. S1).


Table I. Characteristics of the included studies.

<table>
<thead>
<tr>
<th>Author, year</th>
<th>Country</th>
<th>GEO accession</th>
<th>Platform</th>
<th>OS cases</th>
<th>Controls</th>
<th>Samples type</th>
<th>(Refs.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vella et al, 2016</td>
<td>The Netherlands</td>
<td>GSE87437</td>
<td>GPL570</td>
<td>21</td>
<td>NA</td>
<td>High-grade osteosarcoma</td>
<td>(23)</td>
</tr>
<tr>
<td>Kobayashi et al, 2009</td>
<td>Japan</td>
<td>GSE14827</td>
<td>GPL570</td>
<td>27</td>
<td>NA</td>
<td>Fresh frozen tumor specimens</td>
<td>(20)</td>
</tr>
<tr>
<td>Hamidouche et al, 2009</td>
<td>Germany</td>
<td>GSE18043</td>
<td>GPL570</td>
<td>NA</td>
<td>3</td>
<td>BM-MSCs</td>
<td>(19)</td>
</tr>
<tr>
<td>André et al, 2013</td>
<td>Belgium</td>
<td>GSE36474</td>
<td>GPL570</td>
<td>NA</td>
<td>3</td>
<td>BM-MSCs</td>
<td>(22)</td>
</tr>
<tr>
<td>Guilloton et al, 2012</td>
<td>France</td>
<td>GSE35331</td>
<td>GPL570</td>
<td>NA</td>
<td>6</td>
<td>BM-MSCs</td>
<td>(21)</td>
</tr>
</tbody>
</table>

OS, osteosarcoma; NA, not available; BM-MSCs, bone marrow mesenchymal stromal cells.

Table II. Top 10 upregulated genes in osteosarcoma.

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Log.FC</th>
<th>AveExp</th>
<th>t-score</th>
<th>P-value</th>
<th>Padj</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPE</td>
<td>6.01</td>
<td>9.77</td>
<td>13.95</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MMP9</td>
<td>5.96</td>
<td>10.45</td>
<td>9.99</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>SPARC</td>
<td>5.93</td>
<td>8.60</td>
<td>13.18</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>S100A4</td>
<td>5.29</td>
<td>8.75</td>
<td>12.64</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CA2</td>
<td>5.60</td>
<td>8.08</td>
<td>9.62</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>COL15A1</td>
<td>5.29</td>
<td>8.75</td>
<td>12.64</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ACP5</td>
<td>5.19</td>
<td>9.01</td>
<td>7.89</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>CIQC</td>
<td>5.18</td>
<td>8.93</td>
<td>16.32</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MMP13</td>
<td>5.17</td>
<td>8.91</td>
<td>6.80</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>MRC1</td>
<td>5.00</td>
<td>7.55</td>
<td>15.24</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

AveExp, average expression across all samples; Log.FC, Log2(fold change); Padj, adjusted P-value; t-score, statistic value for t-test.

Table III. Top 10 downregulated genes in osteosarcoma.

<table>
<thead>
<tr>
<th>Gene symbol</th>
<th>Log.FC</th>
<th>AveExp</th>
<th>t-score</th>
<th>P-value</th>
<th>Padj</th>
</tr>
</thead>
<tbody>
<tr>
<td>KRTAP1-5</td>
<td>-5.36</td>
<td>5.76</td>
<td>-17.48</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DKK1</td>
<td>-4.68</td>
<td>7.14</td>
<td>-7.45</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>STC2</td>
<td>-4.25</td>
<td>6.78</td>
<td>-16.57</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>TFP12</td>
<td>-4.23</td>
<td>5.06</td>
<td>-11.62</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>PTX3</td>
<td>-4.19</td>
<td>8.95</td>
<td>-7.46</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>RGS4</td>
<td>-4.08</td>
<td>6.30</td>
<td>-12.49</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>NPR3</td>
<td>-4.07</td>
<td>5.77</td>
<td>-16.82</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>DSP</td>
<td>-4.00</td>
<td>6.85</td>
<td>-8.51</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>LTBP2</td>
<td>-3.99</td>
<td>7.91</td>
<td>-12.64</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>VGLL3</td>
<td>-3.91</td>
<td>5.93</td>
<td>-8.69</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

AveExp, average expression across all samples; Log.FC, Log2(fold change); Padj, adjusted P-value; t-score, statistic value for t-test.
 binding’. These DEGs participated in biological processes including ‘angiogenesis’, ‘extracellular structure organization’, ‘leukocyte migration’, ‘extracellular matrix organization’, ‘cell chemotaxis’, ‘positive regulation of cell migration’, ‘positive regulation of cell motility’, ‘ossification’ and ‘positive regulation of locomotion’. More detailed information can be found in Fig. 1. Dysregulation of the extracellular matrix (ECM) could shape the tumor microenvironment and further modulate cancer hallmarks (35). These results also suggest that dysregulation of the ECM may contribute to OS development and metastasis (Fig. 2).


Identification of key modules and genes closely associated with OS. Gene expression values from all genes and samples were included in WGCNA. Soft-thresholding was selected with a power of 12, a minimum module size of 30 and a medium sensitivity to cluster splitting (Fig. S2). A module-trait association heatmap was plotted to identify modules that were significantly associated with clinical traits (Fig. 3). As shown in Fig. 3, the yellow, red and pink modules are positively related with OS status, osteoblastic tumor type and chemotherapy response; these three modules are negatively associated with age. The blue module is negatively associated with OS and positively associated with age. Modules with a height of <0.25 were merged. In total, 15 gene modules were identified and the dendrogram displayed together with the color assignment is shown in Fig. 4.

Genes in these modules may play a pivotal role in OS. The results showed that 383 out of 749 genes in the yellow module were dysregulated, including 358 upregulated genes and 25 downregulated genes when the cutoff value was set as |log2 FC|>1 and adjusted P<0.01. The top 30 hub genes from the yellow module were extracted to visualize their connections using VisAnt software (Fig. 5). As shown in Fig. 5, C1QC and MRC1 are two hub genes that were upregulated in OS, which indicated that these two genes may play important roles in OS.

DEGs from the yellow module were investigated using the PharmGKB, oncoKB and CIViC databases to identify potentially druggable targets (16-18). The results showed that MERTK and SYK were druggable according to the CIViC database (18,36,37). Gene variations in CXCR4, FCGR2A, MGAT4A, NCOAI, PIK3R1, RG55, RRA52 and SOD2 may be predictive markers or have targetable variations according to the PharmGKB database (16,38-47). A total of 10 oncogenes (CXCR4, ERG, FLT1, IGF1, KDR, LYN, MITF, PIK3CG, REL and SYK), six tumor suppressor genes (MITF, MAP3K1, MOB3B, NFkBIA, PRD1 and SAMHD1) and one gene (CSF1R), belonging to neither oncogene or tumor suppressor gene categories, were identified, according to the oncoKB database. Some mutations in MITF are oncogenic while others can repress the development of cancer (17). Therefore, MITF was classified as both a tumor suppressor and an oncogenic gene. However, no further evidence could support these findings in OS as the present study only included gene expression results from microarray analysis where no gene variation data were available. These findings were predominantly reported in breast cancer, colorectal cancer and prostatic neoplasms, with only a few studies in OS.

Tumor infiltrating immune cell profiles in OS. The results of the present study showed that M0 and M2 macrophages were two major types of immune cells found in OS tissues. Some memory resting CD4+ T cells were present in OS tissues. In BM-MSCs, memory resting CD4+ T cells and naïve B cells were the two major types of immune cells identified (Fig. S3).
Figure 2. Top 10 terms of KEGG enrichment analysis (adjusted P<0.01). (A) Overrepresented and (B) enriched KEGG pathways. (C) Enriched KEGG modules. The x-axis shows the ratio of genes enriched in a KEGG pathway. KEGG, Kyoto Encyclopedia of Genes and Genomes.

Figure 3. Module-trait relationship heatmap for different traits and gene modules. The yellow, red and pink gene modules are positively related to OS, osteoblastic status and CT response; these modules are negatively associated with age. The brown and greenyellow modules are positively related to age and negatively related to OS, osteoblastic status and CT response. Values in the figure indicate the correlation coefficient between modules and clinical traits. Values in brackets are the P-values for the association test. OS, osteosarcoma; CT, chemotherapy; ME, module.
Bioinformatic approaches are widely used for the clinical prediction of cancer diagnosis and gene research. By using DEG analysis in combination with WGCNA, biologically meaningful genes and gene modules can be identified as candidate biomarkers (32,48). The present study identified five previous studies related to OS from GEO; 1,405 genes were found dysregulated in OS compared with BM-MSCs. These genes were found to be involved in the ECM, according to the results from functional annotations. WGCNA analysis showed that the yellow module was positively associated with OS. A total of 30 hub genes were selected to visualize their connections. Several DEGs in the yellow module were found to be potentially druggable genes, according to the Clin, PharmGKB and oncoKB databases. CXCR4 belongs to the chemokine receptor family and is an oncogene that can mediate metastasis in cancers (49). It is overexpressed in breast cancer (50), ovarian cancer (51), melanoma (52), and prostate cancer (53,54). A previous study investigating gastric cancer showed that CXCR4 mRNA expression was positively correlated with docetaxel sensitivity (55), indicating that docetaxel may be effective in patients with OS who have a high level of CXCR4 mRNA expression. These findings may be helpful for guiding the clinical management of OS.

Yang et al (56) conducted a meta-analysis of OS microarray data in 2014 to better understand the underlying mechanism of OS. In this previous study, data was included from different microarray platforms, and results from OS tissue samples and cell lines were also included. The study revealed that ‘ECM-receptor interaction’ and the ‘cell cycle’ were highly enriched KEGG pathways, and several hub genes were identified, including PTBP2, RGS4 and FXYD6 (56). The present study provided some improvements in the inclusion criteria and the analytic methods used. Evidence
suggests that cell lines from different laboratories are heterogeneous in various ways (57), therefore, gene expression profiles generated from OS cell lines were excluded from the present study. BM-MSCs from individuals with no signs of malignancies were selected as controls as OS has been reported to originate from BM-MSCs (1). Furthermore, as different microarray platforms may have different probes to represent the same gene, results from different platforms are not usually directly comparable. In the present study, only RNA expression data generated from Human Genome U133 Plus 2.0 array (GPL570) using tissue samples were included to reduce bias. Public databases were used to investigate key genes in order to identify potentially druggable genes or therapeutic targets.

Results from GO and KEGG enrichment analysis suggested that the DEGs identified participated in the ECM; this may contribute to OS development, which was consistent with a previous study by Yang et al (56). Many ECM proteins are significantly dysregulated during the progression of cancer, causing both biochemical and biomechanical changes (58). It has been reported that cancer cells can degrade the ECM, and promote metastasis by facilitating tumor associated angiogenesis and inflammation (59,60). Accumulation of ECM proteins can provide a suitable microenvironment to promote cancer cell proliferation and metastasis (61-63). Miyata et al (64) reported that MMP2 and MMP9 could promote the mobility of vascular epithelial cells by remodeling the ECM. In the present study, both MMP9 and MMP13 were found to be overexpressed in OS.

The yellow gene module was positively related to OS. The top 30 hub genes were selected to visualize the gene-gene interactions. Each of these genes were upregulated >2-fold. In total, 17 out of the 30 hub genes identified were from the innate immune system, including C1QA, C1QB, C1QC, CD14, CTSS, HCK, HLA-DMB, IL10RA, LCP2, MRC1, PECAM1, PIK3AP1, PTTPR, RNASE6, TNFSF13B, TYROBP and VAMP8. C1QA, C1QB and C1QC encode the A-, B- and C-chain of the serum complement subcomponent C1q, respectively. The protein encoded by CD14 is a component of the innate immune system and CD163 functions as an innate immune sensor in bacteria (65,66). CTSS acts as a pivotal role in antigen presentation and can promote cell growth during tumorigenesis (67). GGTA1P can produce an immunogenic carbohydrate structure in Homo sapiens and the aberrant expression of this gene is associated with autoimmune disorders. GIMAP6 and GIMAP8 are members of the GTPase of immunity-associated protein family, regulating lymphocyte survival and homeostasis (68). The protein encoded by HCK may play a role in neutrophil migration and neutrophil degranulation (66). Expression of HCLS1 is not restricted to hematopoietic cell lineages (69). HLA-DM and HLA-DMB are both required for the normal assembly of peptides onto major histocompatibility complex class II molecules (70,71). HLA-DMB is upregulated in the tumor tissues of patients of a Caucasian decent, but not of African-American descent, and is positively correlated with an increase in T cell infiltration and an improved prognosis (72). In mouse models, PECAM1 is associated with dysregulated osteoclastogenesis and hematopoiesis (73). PTTPRC is also known as the CD45 antigen, which belongs to the protein tyrosine phosphatase (PTP) family (74). PTPs can regulate a variety of cellular processes, including cell growth, differentiation, the mitotic cell cycle and oncogenic transformation (74). CENTA2, encoded by ADAP2, can bind β-tubulin and increase its stability (75). AIF1 expression is induced by cytokines and interferon, and may promote the activation of macrophages and the growth of vascular smooth muscle cells (76). It has been reported that tumor-associated macrophages can suppress the T cell-mediated anti-tumor immune response (77). OS is a type of cold tumor, largely due to the anti-inflammatory M2 macrophages enriched in the tumor microenvironment, which can repress tumor-infiltrating T cells (78,79). Results from CIBERSORT revealed that M0 and M2 macrophages were two major types of immune cell found in OS compared with BM-BMCs. This is consistent with the results of previous studies that have been reviewed by Kelleher et al (78), indicating that AIF1 may play an important role in the OS microenvironment. Programmed cell death protein 1 (PD-1) and cytotoxic T-lymphocyte antigen 4 can downregulate the immune system by suppressing T cell-mediated inflammatory activity in order to prevent the immune system from killing cancer cell (80). However, a previous study showed that a PD-1 inhibitor is only effective in metastatic OS, as only metastatic OS expressed PD-1 (81). In the present study, PD-1 and PD-L1 were downregulated in OS compared with BM-MSCs.

The protein encoded by MGAT4A can regulate the availability of serum glycoproteins, and may participate in oncogenesis and differentiation (82). As OS is frequently infiltrated by immune cells, including M2 macrophages and T cells (83,84), these findings may be instrumental in developing a better understanding of the mechanisms underlying OS. Further studies are warranted to explore whether and how personalized chemotheraphy along with targeted therapy, including PDI inhibitors, can benefit patients with primary OS.

Several advantages of the present study should be mentioned. Firstly, the inclusion criteria for relevant studies has been improved, only RNA expression data from the same platform and from tissues were included in the present study. Additionally, DEG analysis was combined with WGCNA analysis, which reduced the number of genes closely related to OS. Key genes from the yellow module were further compared using the CiviC, PharmGKB and oncoKB databases, and several promising druggable targets were identified. However, there were also several limitations to the present study. The five studies included were from The Netherlands, Japan, Germany, Belgium and France; stratified analysis was not performed on these data due to the relatively small samples in each study, and the heterogeneity of the tissues used in the different studies were unmodifiable, which should be improved in further studies.

In conclusion, the present study identified a group of DEGs in OS using meta-analysis and bioinformatics analysis, and several key genes that may contribute to OS were identified. Functional annotations of these hub genes indicated that the ECM is involved in the development of OS. The present study improved our understanding of the mechanisms underlying the development of OS.
Acknowledgements

Not applicable.

Funding

The present study was supported by the Key Research Program for Higher Educational Institutes of Henan Province (grant no. 19B310007) and the Key R&D and Transfer Program for the Institute of Science and Technology of Henan Province (grant nos. 192102310402 and 192102310034). The funding body had no role in the study design, collection, analysis or interpretation of the data, writing the manuscript or the decision to submit the paper for publication.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

JSh and CZ conceived the idea. JSu, HX and MQ retrieved and screened all relevant studies, and conducted the bioinformatic analyses. JSh prepared the manuscript with all other authors contributing to its completion.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References


