

# Identification of the Key Serum Biomarkers to Diagnose and Predict Metastasis of Osteosarcoma Based on the Analysis of Immune Cell Infiltration Characteristics

**Zhihao Chen**

Second Affiliated Hospital of Soochow University

**Liubing Li**

Second Affiliated Hospital of Soochow University

**Ziyuan Li**

Second Affiliated Hospital of Soochow University

**Xi Wang**

Second Affiliated Hospital of Soochow University

**Mingxiao Han**

Second Affiliated Hospital of Soochow University

**Zongshuai Gao**

Shanghai Jiao Tong University Affiliated Sixth People's Hospital

**Min Wang**

Second Affiliated Hospital of Soochow University

**Gangfeng Hu**

Second Affiliated Hospital of Soochow University

**Xiaolu Xie**

Second Affiliated Hospital of Soochow University

**Hong Du**

Second Affiliated Hospital of Soochow University

**Zonggang Xie**

Second Affiliated Hospital of Soochow University

**Haifang Zhang** (✉ [haifangzhang@suda.edu.cn](mailto:haifangzhang@suda.edu.cn))

Second Affiliated Hospital of Soochow University <https://orcid.org/0000-0001-6434-3561>

---

## Research Article

**Keywords:** circRNA, miRNA, immunotherapy, biomarkers, osteosarcoma

**Posted Date:** November 12th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-1047570/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The potential functions of circular RNAs (circRNAs) and micro RNAs (miRNAs) in osteosarcoma (OS) have not been fully elucidated. Especially, the behavior and mechanism of immune responses in OS development and progression have not been fully demonstrated. It was reported that circRNAs and miRNAs can serve as biomarkers for the diagnosis, prognosis, and therapy of many cancers. This study aimed to identify novel key serum biomarkers to diagnose and predict metastasis of OS based on the analysis of immune cell infiltration characteristics.

**Methods:** The differentially-expressed circRNAs (DEcircRNAs), differentially-expressed miRNAs (DEmiRNAs), and differentially-expressed mRNAs (DEmRNAs) of human OS were investigated based on the microarray data downloaded from Gene Expression Omnibus (GEO) datasets. Then, we analyzed immune characteristics pattern of tumor-infiltrating immune cells in OS. On this basis, we identified statistically-significant transcription factors and performed pathway enrichment analysis. Subsequently, we constructed protein-protein interaction (PPI) and competitive endogenous RNA (ceRNA) networks. Moreover, the biological characteristic of targets in ceRNA networks was proposed. Finally, the expression and diagnostic capability of these potential biomarkers from ceRNA network were confirmed by RT-qPCR in patients' serum.

**Results:** Seven differentially-expressed circRNAs (DEcircRNAs), 166 differentially-expressed miRNAs (DEmiRNAs) and 175 differentially-expressed mRNAs (DEmRNAs) were identified in total. The highest level of infiltration in OS patients were M0 macrophages, M2 macrophages and CD8+ T cells. Further, M0 macrophages and CD8+ T cells were showed the largest negative correlation coefficients. These significant immune characteristics pattern of tumor-infiltrating immune cells were revealed by the principal component analysis in OS. Moreover, we found 185 statistically-significant transcription factors in which the main significant molecules show the potential in immunotherapy of OS. Hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p, and FAM98A from ceRNA networks associated with immune cell infiltration were confirmed as the potential novel biomarkers for OS diagnosis, of which FAM98A could distinguish and predict metastasis. Most importantly, a novel diagnostic model consisting of the four promising biomarkers (hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p, and FAM98A) was highlighted with 0.928 AUC value.

**Conclusions:** In summary, the potential serum biomarkers to diagnose and predict metastasis of OS based on the analysis of immune cell infiltration characteristics were found, and a novel diagnostic model consisting of four promising serum biomarkers was proposed firstly. These results provided a new perspective for the immunotherapy of OS.

## Introduction

Osteosarcoma (OS) is the most common primary malignant tumor of bone in young people<sup>[1]</sup>. The combined application of surgery, chemotherapy, and radiotherapy has significantly improved the survival

rate of patients with OS<sup>[2]</sup>, but in spite of this the prognosis of patients with bone tumors remains very poor<sup>[3]</sup>. Most patients with OS will eventually die of metastases<sup>[4]</sup>. Now, Alkaline phosphatase (ALP) and lactate dehydrogenase (LDH) are most widely used serum targets to diagnose OS with unsatisfactory sensitivity and specificity<sup>[5-7]</sup>. Therefore, It is extremely important to find promising biomarkers for diagnose OS in early stage and predict metastasis progression of OS.

Circular RNA (circRNA) is a newly-discovered non-coding form of RNA with a covalent closed loop structure<sup>[8]</sup>, without a 5' cap structure or a 3' poly(A) tail, which makes circRNA resistant to exonucleases with the potential to become an ideal biomarker<sup>[9]</sup>. And the competitive endogenous RNA (ceRNA) network is considered to be the main mechanism for circRNA to perform biological functions<sup>[10]</sup>. Pan *et al.* found that circ\_0028171 acts as a sponge of micro RNA (miR)-218-5p to increase the expression of IKBKB which promotes the progression of OS and could be a potential novel marker for early diagnosis<sup>[11]</sup>. Li *et al.* reported that hsa\_circ\_0000073 can act as a sponge to inhibit miR-145-5p- and miR-151-3p-mediated down-regulation of NRAS, and promote the proliferation, migration, invasion and MTX resistance of OS cells<sup>[12]</sup>. Zhang *et al.* found that high expression of hsa\_circ\_0136666 predicts poor prognosis and promotes the development of OS through the miR-593-3p/ZEB2 pathway<sup>[13]</sup>. Although circRNA plays an important role in OS, the potential functions and complex mechanisms of most circRNAs in OS have not been fully elucidated.

The tumor microenvironment (TME) not only provides nutrients and growth factors for the proliferation and metastasis of tumor cells, but also limits the early detection of tumors and the efficacy of immunotherapy<sup>[14]</sup>. The bone micro-environmental cell composition of OS is complex, releasing a variety of growth factors and cytokines<sup>[15]</sup>. CIBERSORT is a new type of biological information tool. Through the deconvolution algorithm developed by Bindea *et al.*, CIBERSORT can estimate the cell composition of complex tissues based on standardized gene expression data<sup>[16, 17]</sup>. This method has been validated by flow cytometry in breast and lung cancer, and can be applied to analyze the gene expression profile data on a large scale<sup>[18-20]</sup>. Because the micro-environment plays an important role in tumor development, analyzing the immune micro-environment of OS can provide a better understanding of the pathogenesis of OS. Further, immunotherapy is thought to be a rising treatment option for treating OS<sup>[21-23]</sup>. Most importantly, circRNA is considered as a promising biomarker in cancer immunotherapy, especially involved in the TME<sup>[24, 25]</sup>. However, the report of ceRNA mechanism about TME in OS was quite few.

In this study, we collected circRNA, miRNA and mRNA expression profiles of OS *in vivo* obtained from OS patient and healthy individuals from the Gene Expression Omnibus (GEO) dataset. The significant immune characteristics pattern of tumor-infiltrating immune cells and functional enrichment of OS *in vivo* was revealed. Based on these results, we further constructed protein-protein interaction (PPI) and competitive endogenous RNA (ceRNA) networks.. We verified the targets of the ceRNA network by RT-qPCR in patients' serum and found that hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p, and FAM98A from ceRNA network associated with immune cell infiltration could be used as novel serum biomarkers for OS diagnosis in which FAM98A could distinguish and predict metastasis. Most importantly, a

promising diagnostic model consisting of these four targets (hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p, and FAM98A) was highlighted with 0.928 AUC value.

## Materials And Methods

### Raw data

The raw datasets GSE140256, GSE65071, GSE16088, GSE21257, GSE33382, and GSE124768 were obtained from the Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) which is an online public gene data repository for high-throughput sequencing research. GSE140256, a circRNA expression profile, included three cancer tissue samples and three paracancer samples, and miRNA sequencing data obtained from GSE65071 included 20 serum samples from patients with OS and 15 normal serum samples. In the mRNA expression profile data, GSE16088 included 17 cancer tissue samples and six normal tissue samples. The GSE21257, GSE33382, GSE124768 included 56 non-metastatic patients and 71 metastatic patients or patients metastatic within 5 years to explore the characteristics between non-metastatic patients and metastatic subgroup. The above raw data were extracted from the GEO database. Consequently, this research did not require any ethical review or informed consent because of the public availability of GEO data.

### Identification of DEcircRNAs, DE miRNAs and DE mRNAs

The DEcircRNAs, DE micro RNAs (miRNAs) and DE messenger RNAs (mRNAs) were identified by the difference in expression between the normal and OS samples in microarray data. The  $P$ -value and the absolute log value of fold-change ( $\log|FC|$ ) were analyzed in R language by the limma package. The  $\log|FC| > 1.0$  and  $p < 0.05$  were the selection criteria to define the differentially-expressed genes (DEGs).

### Tumor-infiltrating immune cell analysis

We used Cibersort software to analyze DE mRNAs differentially expressed between OS tissues and normal tissues. The distribution characteristics of 22 immune cells were calculated by the deconvolution method of Cibersort software. The 22 immune cell types included activated dendritic cells (DCs), resting DCs, activated mast cells, resting mast cells, activated natural killer cells (NKs), resting NKs, activated memory CD4+ T cells, resting CD4+ T cells, naïve CD4+ T cells, regulatory T cells (Tregs), T follicular helper cells (Tfhs), gamma delta T cells (Tgds), CD8+ T cells, eosinophils, neutrophils, monocytes, macrophages (M0s), type 1 macrophages (M1), type 2 macrophages (M2), memory B cells, naïve B cells, and plasma cells. Values of  $P < 0.05$  were considered as statistically significant.

### Analyses of differential expression transcription factors

Based on the DE miRNAs identified, FunRich (Version 3.1.3) was used to analyze and visualize the differentially-expressed transcription factors. FunRich is a stand-alone software tool used mainly for functional enrichment and interaction network analysis of genes and proteins.

## Pathway enrichment analysis

We used GO annotation (<http://www.geneontology.org>) and KEGG pathway analysis to determine the potential functions of the DEmRNAs. The  $P$ -value  $< 0.05$ , as a screening condition, was considered highly credible by using R language of the clusterProfiler package.

## Construction of the ceRNA network

In the results of the differential analysis of the microarray data information, the DEcircRNAs were predicted by TSCD (tissue-specific circRNA database) (<http://gb.whu.edu.cn/TSCD>) and DEmiRNA target genes were predicted by the databases miRDB, miRTarBase and TargetScan. In this study, downstream molecules predicted in all three databases were identified as target genes of DEcircRNAs and DEmiRNAs.

We used the prediction of the DEcircRNAs and DEmiRNAs target genes to intersect with identified differentially-expressed downstream genes to further screen the prediction results. Finally, we used the prediction results to construct a ceRNA interaction network and visualize it by employing cytoscape (Version 3.8.0).

## PPI network and clustered sub-networks construction

The exploration of protein interactions helps to reveal the underlying pathological mechanism of OS. In this study, we used the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (<https://string-db.org/>) to construct a protein–protein interaction network. On this basis, the clustered subnetworks and hub genes were identified by employing Molecular Complex Detection (MCODE) and cytohuba in cytoscape (Version 3.8.0).

## Verification of diagnostic specificity using pan-cancer analysis

The function of ncRNAs lies in the regulation of target genes. Therefore, the exploration of pathological features about target genes is extremely important. To further determine the potential of gene targets in ceRNA network we constructed on the diagnosis of OS patients. We obtain the expression data of target genes in ceRNA network we constructed, as Fragments Per Kilobase of transcript per Million mapped reads (FPKM) from the Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx). We totally analyzed 32 different cancer type about 10,967 samples from TCGA and 17,382 samples from, GTEx ,including kidney renal clear cell carcinoma (KIRC); kidney renal papillary cell carcinoma (KIRP); kidney chromophobe (KICH); brain lower grade glioma (LGG); glioblastoma multiforme (GBM); breast cancer (BRCA); lung squamous cell carcinoma (LUSC); lung adenocarcinoma (LUAD); rectum adenocarcinoma (READ); colon adenocarcinoma (COAD); uterine carcinosarcoma (UCS); uterine corpus endometrial carcinoma (UCEC); ovarian serous cystadenocarcinoma (OV); head and neck squamous carcinoma (HNSC); thyroid carcinoma (THCA); prostate adenocarcinoma (PRAD); stomach adenocarcinoma (STAD); skin cutaneous melanoma (SKCM); bladder urothelial carcinoma (BLCA); liver hepatocellular carcinoma (LIHC); cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC); adrenocortical carcinoma (ACC); pheochromocytoma and paraganglioma (PCPG); sarcoma (SARC); pancreatic

adenocarcinoma (PAAD); esophageal carcinoma (ESCA); testicular germ cell tumors (TGCT); thymoma (THYM); uveal melanoma (UVM); lymphoid neoplasm diffuse large b-cell lymphoma (DLBC); cholangiocarcinoma (CHOL). The diagnostic specificity of gene targets we provided were better highlight between different types of cancer.

### **Verification of metastasis characteristics**

Metastasis has always been an important obstacle to treatment OS. So far, there is still a lack of effective biomarkers to diagnose OS in early stage and predict pathological progression of OS. The clinical data of patients from previous GEO datasets (GSE21257, GSE33382 and GSE124768) were employed to analyze the relationships between the identified potential targets in OS and metastasis using the limma package of R language. The potential of these gene targets were further explored in the comparison of non-metastatic patients with metastatic patients or patients with metastases within 5 years. Fold-change > 1.5 and  $P$ -value < 0.05 was used as the selection criteria

### **Verification of the diagnostic ability in patients' serum sample**

Specific validating primers (Table S1) for several differentially-expressed RNAs (DERNAs) were designed based on the sequence of linear transcripts. Nineteen OS serum samples and 19 healthy individuals serum samples were obtained. Ethical approval was obtained from the ethics committee of Second Affiliated Hospital of Soochow University. Total RNA was extracted from the tissue samples using TRIzol reagent TRIzol (Invitrogen USA), and treated with RNase-free DNase I (Vazyme Biotech Co., China) to eliminate traces of mixed DNA. Real-time PCR was performed on the Applied Biosystems StepOnePlus Real-Time PCR System (Thermo Fisher Scientific, USA) using qPCR SYBR Green master mix (Vazyme Biotech Co., China). The expression level of DEcircRNAs and DEmRNAs were normalized to the endogenous control of human glyceraldehyde-3-phosphate dehydrogenase (GAPDH) as the endogenous control. And the expression level of DEmiRNAs were normalized to the endogenous control of U6. The receiver operating characteristic (ROC) curve and logistic regression package in SPSS were used to constructed diagnostic model.

### **Statistical analyses**

Bioinformatic analysis was performed using R software (version 4.0.1), FunRich (version 3.1.3), and Cytoscape (version 3.8.0). GraphPad Prism (GraphPad Software, USA) and SPSS (IBM, USA) were employed for statistical analysis of qRT-PCR. Quantitative data were analyzed for statistical significance by t-test and expressed as mean  $\pm$  SD. A  $p$ -value < 0.05 was accepted as statistically significant.

## **Results**

### **Identification of DEcircRNAs, DEmiRNAs and DEmRNAs**

We performed a comprehensive bioinformatics analysis to elucidate the key circRNA-miRNA-mRNA axis in OS (flow chart was shown in supplementary Figure S1). Values of  $\log|FC| > 1.0$  and  $p < 0.05$  were the

selection criteria to define the DEGs. In the GSE140256 dataset, four up-regulated circRNAs and three down-regulated circRNAs were identified, which included three cancer tissue samples and three paracancer samples (Figure 1A, 1B). The GSE65071 dataset, including 20 serum samples from patients with OS and 15 normal serum samples, revealed a total of 78 upregulated miRNAs and 88 downregulated miRNAs (Figure 1C, 1D). In the mRNA expression profile data, which included 17 cancer tissue samples and six normal samples, the GSE16088 dataset, we identified 175 DEmRNAs, of which 149 mRNAs were up-regulated and 26 mRNAs were down-regulated (Figure 1E, 1F).

### **Tumor-infiltrating immune cell analyses**

We used Cibersort to evaluate the distribution of immune cells from the data of 17 OS patient tissues and six normal human tissues. Among them, we found that the immune cell types with the highest level of infiltration in OS patients were M0 macrophages, M2 macrophages and CD8+ T cells (Figure 2A). Compared with normal tissues, M0 macrophages ( $P = 0.010$ ), M2 macrophages ( $P = 0.010$ ), CD8+ T cells ( $P = 0.020$ ), memory B cells ( $P = 0.021$ ), plasma cells ( $P = 0.036$ ) and activated NK cells ( $P = 0.033$ ) showed statistically-significant differences in OS patient tissues ( $P < 0.05$ ) (Figure 2B). In the tissues of OS patients, immune cells with a larger positive correlation coefficient included CD8+ T cells and plasma cells (0.66), memory B cells and naïve B cells (0.59), plasma cells and naïve CD4+ T cells (0.58). In addition, immune cells with larger negative correlation coefficients included M0 macrophages and CD8+ T cells (0.83), resting mast cells and activated mast cells (0.62), plasma cells and M0 macrophages (0.61) (Figure 2C). Most importantly, results of principal component analysis (PCA) showed that immune cell infiltration patterns of 22 types of immune cell can effectively distinguish OS patients from healthy controls (Figure 2D).

### **Analyses of differentially-expressed transcription factors**

To assess potential of targets we noted, as tumor immune biomarkers, transcription factors analysis was performed. Transcription factors are important molecules that regulate immune in cancer. We analyzed differentially-expressed transcription factors associated with the 166 identified DEmiRNAs, and found 185 statistically-significant transcription factors. Among these results, EGR1, SP1, SP4, POU2F1 and NFIC are the most significant molecules in OS. EGR1 is the most significant target among the transcription factors we identified. And SP1 enjoys the largest percentage of these transcription targets (Figure 3). Interestingly, the top 5 transcription factors, highlighted in OS, all show the importance in immune. These results confirm that the statistically-significant transcription factors and the biomarkers performed transcription factors analysis enjoy promising potential in cancer immunotherapy.

### **Pathway enrichment analyses**

To better understand the potential biological functions of the identified DEmRNAs in OS, GO annotation and KEGG pathway analysis were performed. Through GO annotation analysis, we found that the DEmRNAs were significantly enriched in the terms of post-translational protein modification, RNA localization and extracellular structure organization in the biological process (BP) subgroup. Collagen-

containing extracellular matrix, endoplasmic reticulum lumen and blood microparticle were the most significant GO terms in the cellular component (CC) subgroup. The top three GO processes were single-stranded DNA binding, heat shock protein binding and extracellular matrix structural constituent in the molecular function (MF) subgroup for DEmRNAs (Figure 4A).

In KEGG pathway enrichment analysis, complement and coagulation cascades, RNA transport and drug metabolism–other enzymes were the most significant pathways enriched of DEmRNAs in OS (Figure 4B).

### **Construction of the ceRNA network**

In order to better understand the endogenous regulatory mechanism of DEGs in OS, we predicted the downstream target genes of DEcircRNAs by employing the TSCD databases. At the circRNA level, all of DEcircRNAs were predicted its downstream miRNAs. We used the predicted miRNAs as candidate miRNAs. Candidate miRNAs were cross-compared with DEmiRNAs identified in the microarray. In the analysis results, three intersecting miRNAs were obtained in data of the identified DEmiRNAs. At the miRNA level, a total of 153 miRNAs were predicted, and a total of 7,580 mRNAs were involved. In the results after comparison, four intersecting mRNAs were identified. Based on the predicted results of circRNA–miRNA and miRNA–mRNA pairs, we constructed a ceRNA network. In the ceRNA network, there were a total of one circRNA, three miRNAs and four mRNAs involved in the network construction (Figure 5).

### **Construction of the protein–protein interaction (PPI) network**

We used the identified DEmRNAs to construct a PPI network. This network included 146 nodes and 687 edges, under the conditions that the comprehensive Gt score > 0.4 and unconnected points were removed. Among the 146 genes, 67 genes had a score > 100 when analyzed by the Maximal Clique Centrality (MCC) method in cytohubba. The top five hub genes were CDC20, MAD2L1, PCNA, KPNA2 and PRB1 (Figure 6A). In this study, we also defined the most closely-clustered subnetwork by employing the MCODE plug-in in Cytoscape. Using this, the most closely-clustered subnetwork was found to consist of 14 nodes and 90 edges (Figure 6B). In addition, we also identified two other clustered subnetworks, which contained 20 nodes and 97 edges, and 13 nodes and 35 edges, respectively (Figure 6C 6D). These groups were all composed of up-regulated mRNAs.

### **Pan-cancer analyses**

To further verify the diagnostic potential of the target gene, pan-cancer analysis was employed. In pan-cancer analysis, COL1A1 and FAM98A showed good specificity. There is no difference in the expression of COL1A1 in BLCA, KICH, KIRP, OV, SARC and UCS. It is worth mentioning that COL1A1 does not show differential expression in SARC, which indicates that COL1A1 has better specificity in the diagnosis of OS (Figure 7A). As for FAM98A, it had no differential expression in BLCA, OV, PCPG, SARC and UCS. In particular, the non-differential expression of FAM98A in SARC highlighted its specificity in OS diagnosis

(Figure 7B). RAN and HNRNPA2B1 showed differential expression in most cancers. And there were no significant expression differences in KICH and PCPG, LUCAD and PCPG, respectively (Figure 7C 7D).

### **Verification of metastasis characteristics**

To further verify the biological function of the target gene, public data from different regions, including Europe and the America, were employed to analyze the metastasis characteristics of the target gene we provided. We analyzed the metastatic patients and patients metastasized within 5 years, comparing with non-metastatic patients respectively.

Two different GEO datasets (GSE21257 and GSE124768) showed that differential expression of FAM98A was considered to be statistically significant between non-metastatic patients and metastatic patients in OS. Moreover, In the comparison between non-metastatic patients and patients metastasized within 5 years from GEO datasets (GSE21257 and GSE33382), the FAM98A was considered to be differentially expressed. As for another target with good specificity, COL1A1, its expression level showed no difference whether it is between non-metastatic and metastatic patients, or between non-metastatic patients and patients metastasized within 5 years (Table 1).

### **Verification of the expression of DERNA in serum sample**

Based on the previous analysis results, the expression of FAM98A, COL1A1 and their upstream targets, hsa-circ-0010220, hsa-miR-326, and hsa-miR-338-3p were validated by qPCR using 19 OS serum samples and 19 serum samples from healthy individuals. In the comparison between OS patients and healthy individuals, the hsa-circ-0010220 was considered to be differentially expressed (Figure 8A). The expression of hsa-miR-326 and hsa-miR-338-3p decreased significantly in OS serum samples compared to serum samples from healthy individuals (Figure 8B 8C). Further the expression of FAM98A greatly increased in serum sample from OS patients (Figure 8D). However, the results showed that there was no difference in the expression level of that COL1A1 in serum samples in the comparison between OS patients and healthy individuals (Figure 8E). So, the data indicated that hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p, FAM98A could be the candidate biomarkers to diagnose OS in early stage and predict pathological progression of OS.

In order to further explore the diagnostic role of validated targets in OS, we performed ROC curve analysis and proposed a good diagnostic model. The results demonstrated that hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p, and FAM98A, the validated targets, enjoyed high diagnostic sensitivity and specificity, and the AUC values of them were more than 0.7. In the four targets, we highlighted, hsa-circ-0010220 exhibited the highest specificity. As for sensitivity, hsa-miR-338-3p was considered as the best biomark. Moreover, we propose a diagnostic model consisting of these four targets. In the combined model, the AUC values reached 0.928. And its specificity and sensitivity were 89.47% and 94.74% respectively (Figure 8F).

## **Discussion**

CircRNA has long been considered to be non-functional and the product of abnormal gene splicing<sup>[26]</sup>. In recent years, more and more studies have found that circRNA is related to the occurrence and development of a variety of cancers<sup>[27–29]</sup>. OS is a common malignant tumor in orthopedics. Although the combined application of surgery, chemotherapy, and radiotherapy significantly improves the survival rate of patients with OS, their prognosis is still very poor because of cancer metastasis<sup>[3, 4]</sup>. Related studies have described the functional mechanisms of some circRNA in OS as potential biomarkers<sup>[30, 31]</sup>.

In this study, we revealed the key serum targets of diagnose and predict metastasis based on ceRNA network *in vivo* we new proposed and immune cell infiltration characteristics that immune cell infiltration patterns of 22 types of immune cell can effectively distinguish OS patients from healthy individual. The results showed that these serum targets and the combined model based on them from ceRNA network related immune cell infiltration are better diagnostic biological targets, compared with ,ALP and LDH, the commonly used serum biological targets in clinical practice<sup>[5–7]</sup>. Most importantly, FAM98A, considered as a note factor, indicates OS metastasis. However, there are few reports of serum targets as biomarkers of OS metastasis.

In our study, we obtained the raw datasets of OS from GEO public database to identify DEcircRNA, DEmiRNA and DEmRNA expression profiles of OS *in vivo*. On the basis, we conducted a biological analysis of transcription factors, functional characteristics, ceRNA and PPI network of OS *in vivo*. In immune infiltration characteristics analyses, PCA, as conclusion, showed that immune cell infiltration patterns of 22 types of immune cell can effectively distinguish OS patients from healthy controls. The related results further supported the key serum biomarkers, proposed from ceRNA network about immune infiltration. And pan-cancer analysis of large samples in TCGA and GTEx databases and clinical characteristics analyses from different regions were employed to highlight the functions of these targets and eliminate study errors. FAM98A and COL1A1 were considered promising targets with good specificity in in comparison of OS and other types of sarcoma. And the ability of FAM98A to diagnose metastasis and predict metastasis within 5 years was verified in multiple different dataset, GSE21257, GSE124768 and GSE33382, including 127 OS patients. So we chosen FAM98A, COL1A1, and their upstream targets and verify them in serum samples. The result shown that FAM98A, hsa-circ-0010220, hsa-miR-326, and hsa-miR-338-3p have the differential expression between OS serum samples and serum samples from healthy individuals by qRT-PCR. The ROC curve was employed to perform the specificity and sensitivity of verified targets. These serum targets, especially the combined model based on them, have important value as promising biomarkers in OS diagnosis and metastasis prediction.

The results of the present study contribute to our understanding of the targets in ceRNA network about immune infiltration we constructed, especially the targets in serum sample we verified. And the key serum biomarkers enjoyed interesting potential in immunity. In our study, the importance of M0 macrophages, M2 macrophages and CD8+ T cells in OS was promoted. In the study of Zhang *et al.*, M0 macrophages and M2 macrophages showed a significant correlation with prognosis<sup>[32]</sup>. The top 5 differentially-expressed transcription factors, EGR1, SP1, SP4, POU2F1 and NFIC, all showed the potencial in

immunotherapy<sup>[33-37]</sup>. It is worth mentioning that EGR1 is involved in the polarization of macrophages M2 induced by programmed death ligand 1 (PD-L1) to produce anti-PD-L1 treatment resistance and the promotion of metastasis by regulate angiogenic and osteoclastogenic factors<sup>[38, 39]</sup>. So on the basis of our study, EGR1 was an note immunotherapy biomarker to explore in OS. In addition, Li *et al.* reported that POU2F1 induces the immune escape by increasing the expression of PD-L1 in cancer<sup>[40]</sup>. So, potential relationships revealed in these part. The results further indicate the potential of key serum targets we verified in ceRNA network related immune cell infiltration as a biological target in cancer immunotherapy, especially hsa-miR-326 and hsa-miR-338-3p. They all showed the capability to regulate M2 macrophages to effect cancer characteristics<sup>[41, 42]</sup>. Hsa-miR-326 was invloved in regulation of PD-L1 and CD8+ T<sup>[43]</sup>. These intersting relationships firstly highlighted in this study. Bone has a highly specialized immune environment in which these results wouled promoted the importance of the key serum biomarkers we verified and the understanding of immunotherapy in OS without doubt. In the serum biomarker we proposed, the differential expression of hsa-circ-0010220 in serum showed the same trend in OS tissue<sup>[44]</sup>. As for miRNA, Wang *et al.* reported the long non-coding RNA SNHG1 acting as a sponge of hsa-miR-326 to regulate the expression of human NIN1 binding protein (NOB1) and affect the growth, migration and invasion of OS<sup>[45]</sup>. It is worth mentioning that our research has been verified to some extent by another independent research and promoted the understanding of hsa-miR-326 in OS. Cao L *et al.* reported that hsa-miR-326 was considered as diagnostic and prognostic target in OS which lower expression caused advanced clinical stage and more distant metastasis and regulates cell survival and apoptosis by targeting Bcl-2 in osteosarcoma<sup>[46]</sup>. We noticed that the lacking of the function research in the study for Bcl-2 *in vivo*. And the debate about the role of Bcl-2 in the metastasis and prognosis is difficult to elucidate the remarkable function of hsa-miR-326 in OS<sup>[47-49]</sup>. So, the proposition in our research may deserve more attention, that is, lower expression of hsa-miR-326 leads to worse clinical outcomes through FAM98A. It has been verified in our study based on the data from different races from regions that the high expression of the FAM98A indicates the metastasis and the metastasis within 5 years of OS. Most importantly, another new potential regulatory pathway was revealed about immune infiltration. Shao *et al.* reported that hsa-miR-326 regulates the expression of immune checkpoint molecules PD-L1 and B7-H3 to affect the cytokine profile of CD8+ T cells and tumor cell migration in lung adenocarcinoma<sup>[43]</sup>. The potential context focus on PD-L1 and B7-H3 has been identified for targeted immunotherapeutic intervention in OS<sup>[50]</sup>. The role of CD8+ T cells in OS was hightlighted in our study. Further, hsa-miR-326 has been confirmed to play an important role in the immune response of M2 macrophages through RohA in tumor<sup>[41]</sup>. And RohA has been reported to regulate the growth and metastasis of osteosarcoma through ROCK and Wnt5a signaling pathways<sup>[51-53]</sup>. So, the new strategy of using hsa-miR-326 in cancer immunotherapy of OS could be proposed.

As for hsa-miR-338-3p, involved in the novel regulatory axes we revealed in this study, was reported as tumor suppressor factor in OS by targeting RUNX2, CDK4, and AHSA1, as well as inhibition of the MAPK pathway<sup>[54, 55]</sup>. Furthermore, LncRNA CASC15 and circRNA CCDC66 was confirmed as sponge of miR-338-3p to regulate the progress of OS<sup>[56, 57]</sup>. Our results verified the differential expression of hsa-miR-

338-3p in serum sample which means simpler detection methods can be adopted to diagnose and predict change of OS. We noticed the immune mechanism involved in hsa-miR-338-3p. Coagulation factor X, secreted in the tumor microenvironment as the target of hsa-miR-338-3p, exhibited a potent chemotactic capacity to recruit and promoted macrophages toward M2 subtype polarization, regulating growth of tumor<sup>[42]</sup>. This indicates that hsa-miR-338-3p can be considered as a note factor in OS for tumor immunotherapy. So, hsa-miR-338-3p, verified as the junction of the different pathological processes of OS, its primer had been applied for a patent in our study. Hsa-miR-324-5p, which is in another novel regulatory axis we constructed, was considered to be involved in the regulation of ACE1, which enjoy the capability of DNA repair and redox regulation<sup>[33]</sup>, but its role in pathogenesis of OS was not elucidated<sup>[58]</sup>. And Molist *et al.* found that overexpression of hsa-miR-324-5p could make impairment of cell proliferation and reduce tumor growth in rhabdomyosarcoma<sup>[59]</sup>. Therefore, it is worth to study the potential functions of hsa-miR-324-5p in OS.

FAM98A, the novel important factors and promising markers, was verified in serum sample and highlighted about metastasis firstly in OS. Literatures were reported that FAM98A promotes cancer progression in endometrial carcinoma, non-small cell lung cancer, and breast cancer and show correlation of poor prognosis and metastasis<sup>[60-62]</sup>. And the most noteworthy is that FAM98A participates in the regulation of bone homeostasis and its inhibition leads to bone resorption defects. It provided a support for our research results that the high expression of FAM98A in serum sample indicates OS metastasis and metastasis may be within 5 years<sup>[63]</sup>. As for COL1A1, no statistically significant results were shown in serum samples. But, it was reported that in hFOB1.19 cells which were heterozygous for the c.3781A allele and the c.3781C allele, the expression of COL1A1 was regulated by hsa-miR-345-5p and affected alkaline phosphatase activity and substrate mineralization level<sup>[64]</sup>. In addition, Hawkins *et al.* found that activation of Wnt/beta-catenin signaling affected the expression level of COL1A1 on Ewing's sarcoma<sup>[65]</sup>. It is worth mentioning that the prognostic value of COL1A1 has been verified to a certain extent by other researchers. For example, it was reported that COL1A1 polymorphism is associated with risks of OS transfer and patient death in which COL1A1 polymorphism at rs1061970 has value for overall survival in Chinese OS patients<sup>[66]</sup>. So, we believe that COL1A1, as an OS biomarker, has good potential and is worthy of further research through more ways. And RAN, another mRNA in ceRNA network we constructed, it was firstly highlighted in OS although RAN is widely reported in various tumors, and its high expression is related to aggressiveness and metastasis<sup>[67-69]</sup>. Jain *et al.* reported that hsa-miR-197-3p regulates carcinogenicity in fibrosarcoma by targeting RAN<sup>[70]</sup>. It will be interesting to unveil the unknown molecular mechanism behind the role of the RAN in OS. Moreover, HNRNPA2B1 was reported that it is an independent risk factor for OS as an m6A-related regulators through humoral immune response<sup>[71]</sup>. This result is interesting. In our study, OS showed sufficiently significant immune characteristics, and the ceRNA network we established provides novel perspectives for HNRNPA2B1. HNRNPA2B1 may play a more important role through various regulation mechanism in OS.

Overall, this study is the first report of hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p, FAM98A, and the combined diagnosis model based on them from ceRNA network related immune cell infiltration could be

used promising biomarkers for the osteosarcoma diagnosis in which FAM98A could distinguish and predict metastasis. Most importantly, a novel diagnostic model consisting of the four targets was highlighted with 0.928 AUC value. And significant immune characteristics in OS we proposed provide novel perspective for further research on these targets. However, on the basis of a limited number of patients with individual characteristics, biased results may be included in our study. Validation of large samples would make our study more convincing and fulfill the potential of our results..

## Conclusions

Based on the analysis of immune cell infiltration characteristics in OS, we firstly identified hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p and FAM98A as promising biomarkers for the osteosarcoma diagnosis, which could combine a novel diagnosis model. More importantly, the clinical characteristics of FAM98A has been verified which could distinguish and predict metastasis in OS. Moreover, significant immune characteristics of OS we proposed provide a novel perspective for further research on these potential biomarkers.

## Declarations

We obtained the approval from the ethics committee of Second Affiliated Hospital of Soochow University.

### Author's contributions

Conception and design: HF Zhang, ZG Xie, ZH Chen

Development of methodology: ZH Chen, LB Li, ZY Li, ZS Gao

Acquisition of data: ZH Chen, LB Li, ZY Li, X Wang, MX Han, M Wang

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): HF Zhang, ZG Xie, ZH Chen, LB Li, X Wang, MX Han, M Wang, H Du

Writing, review and/or revision of the manuscript: ZH Chen, HF Zhang, ZG Xie

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): ZH Chen, LB Li, X Wang, MX Han, XL Xie, GF Hu

Study supervision: HF Zhang, ZG Xie

### Funding

This study is funded by "Basic Research Program of Jiangsu Province (Grant No. BK20181173)".

### Competing Interests

The authors have declared that no competing interest exists.

### **Consent for publication**

This study has not been published previously, and the authors agree the consent 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.

### **Acknowledgement**

This study was supported by the National Natural Science Foundation of China (82172332), the Natural Science Foundation of Jiangsu province, China (BK20181173), Gusu health youth talent of Suzhou (GSWS2019039, GSWS2020030), Jiangsu youth medical talents program (QNRC-866, 867, 877), Discipline Construction of The Second Affiliated Hospital of Soochow University (XKTJ-TD202001) and the Science and Technology Program of Suzhou (SYS2018101, SZS201715, SLT201934, SYS2020023, SS201764).

## **References**

- [1] Geller DS, Gorlick R. Osteosarcoma: a review of diagnosis, management, and treatment strategies. *Clin Adv Hematol Oncol*. 2010. 8(10): 705-18.
- [2] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *CA Cancer J Clin*. 2018. 68(1): 7-30.
- [3] Link MP, Goorin AM, Miser AW, et al. The effect of adjuvant chemotherapy on relapse-free survival in patients with osteosarcoma of the extremity. *N Engl J Med*. 1986. 314(25): 1600-6.
- [4] Luetke A, Meyers PA, Lewis I, Juergens H. Osteosarcoma treatment - where do we stand? A state of the art review. *Cancer Treat Rev*. 2014. 40(4): 523-32.
- [5] Shimose S, Kubo T, Fujimori J, Furuta T, Ochi M. A novel assessment method of serum alkaline phosphatase for the diagnosis of osteosarcoma in children and adolescents. *J Orthop Sci*. 2014. 19(6): 997-1003.
- [6] Marais LC, Bertie J, Rodseth R, Sartorius B, Ferreira N. Pre-treatment serum lactate dehydrogenase and alkaline phosphatase as predictors of metastases in extremity osteosarcoma. *J Bone Oncol*. 2015. 4(3): 80-4.
- [7] Fu Y, Lan T, Cai H, Lu A, Yu W. Meta-analysis of serum lactate dehydrogenase and prognosis for osteosarcoma. *Medicine (Baltimore)*. 2018. 97(19): e0741.

- [8] Memczak S, Jens M, Elefsinioti A, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature*. 2013. 495(7441): 333-8.
- [9] Hansen TB, Jensen TI, Clausen BH, et al. Natural RNA circles function as efficient microRNA sponges. *Nature*. 2013. 495(7441): 384-8.
- [10] Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature*. 2014. 505(7483): 344-52.
- [11] Pan F, Zhang J, Tang B, Jing L, Qiu B, Zha Z. The novel circ\_0028171/miR-218-5p/IKBKB axis promotes osteosarcoma cancer progression. *Cancer Cell Int*. 2020. 20: 484.
- [12] Li X, Liu Y, Zhang X, et al. Circular RNA hsa\_circ\_0000073 contributes to osteosarcoma cell proliferation, migration, invasion and methotrexate resistance by sponging miR-145-5p and miR-151-3p and upregulating NRAS. *Aging (Albany NY)*. 2020. 12(14): 14157-14173.
- [13] Zhang C, Zhou H, Yuan K, Xie R, Chen C. Overexpression of hsa\_circ\_0136666 predicts poor prognosis and initiates osteosarcoma tumorigenesis through miR-593-3p/ZEB2 pathway. *Aging (Albany NY)*. 2020. 12(11): 10488-10496.
- [14] Yang C, Tian Y, Zhao F, et al. Bone Microenvironment and Osteosarcoma Metastasis. *Int J Mol Sci*. 2020. 21(19).
- [15] Cersosimo F, Lonardi S, Bernardini G, et al. Tumor-Associated Macrophages in Osteosarcoma: From Mechanisms to Therapy. *Int J Mol Sci*. 2020. 21(15).
- [16] Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. 2013. 39(4): 782-95.
- [17] Charoentong P, Finotello F, Angelova M, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep*. 2017. 18(1): 248-262.
- [18] Gentles AJ, Newman AM, Liu CL, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med*. 2015. 21(8): 938-945.
- [19] Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015. 12(5): 453-7.
- [20] Ali HR, Chlon L, Pharoah PD, Markowitz F, Caldas C. Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. *PLoS Med*. 2016. 13(12): e1002194.

- [21] Nabergoj S, Mlinarič-Raščan I, Jakopin Ž. Harnessing the untapped potential of nucleotide-binding oligomerization domain ligands for cancer immunotherapy. *Med Res Rev.* 2019. 39(5): 1447-1484.
- [22] Chen C, Xie L, Ren T, Huang Y, Xu J, Guo W. Immunotherapy for osteosarcoma: Fundamental mechanism, rationale, and recent breakthroughs. *Cancer Lett.* 2021. 500: 1-10.
- [23] Miwa S, Shirai T, Yamamoto N, et al. Current and Emerging Targets in Immunotherapy for Osteosarcoma. *J Oncol.* 2019. 2019: 7035045.
- [24] Li B, Zhu L, Lu C, et al. circNDUFB2 inhibits non-small cell lung cancer progression via destabilizing IGF2BPs and activating anti-tumor immunity. *Nat Commun.* 2021. 12(1): 295.
- [25] Huang XY, Zhang PF, Wei CY, et al. Circular RNA circMET drives immunosuppression and anti-PD1 therapy resistance in hepatocellular carcinoma via the miR-30-5p/snail/DPP4 axis. *Mol Cancer.* 2020. 19(1): 92.
- [26] Hsu MT, Coca-Prados M. Electron microscopic evidence for the circular form of RNA in the cytoplasm of eukaryotic cells. *Nature.* 1979. 280(5720): 339-40.
- [27] Kristensen LS, Ebbesen KK, Sokol M, et al. Spatial expression analyses of the putative oncogene ciRS-7 in cancer reshape the microRNA sponge theory. *Nat Commun.* 2020. 11(1): 4551.
- [28] Li J, Ma M, Yang X, et al. Circular HER2 RNA positive triple negative breast cancer is sensitive to Pertuzumab. *Mol Cancer.* 2020. 19(1): 142.
- [29] Wang H, Wei M, Kang Y, Xing J, Zhao Y. Circular RNA circ\_PVT1 induces epithelial-mesenchymal transition to promote metastasis of cervical cancer. *Aging (Albany NY).* 2020. 12.
- [30] Chen J, Liu G, Wu Y, et al. CircMYO10 promotes osteosarcoma progression by regulating miR-370-3p/RUVBL1 axis to enhance the transcriptional activity of  $\beta$ -catenin/LEF1 complex via effects on chromatin remodeling. *Mol Cancer.* 2019. 18(1): 150.
- [31] Shen S, Yao T, Xu Y, Zhang D, Fan S, Ma J. CircECE1 activates energy metabolism in osteosarcoma by stabilizing c-Myc. *Mol Cancer.* 2020. 19(1): 151.
- [32] Zhang C, Zheng JH, Lin ZH, et al. Profiles of immune cell infiltration and immune-related genes in the tumor microenvironment of osteosarcoma. *Aging (Albany NY).* 2020. 12(4): 3486-3501.
- [33] Lu D, Liu L, Sun Y, et al. The phosphatase PAC1 acts as a T cell suppressor and attenuates host antitumor immunity. *Nat Immunol.* 2020. 21(3): 287-297.
- [34] Dai Z, Liu P. High copy number variations, particular transcription factors, and low immunity contribute to the stemness of prostate cancer cells. *J Transl Med.* 2021. 19(1): 206.

- [35] Cai H, Miao M, Wang Z. miR-214-3p promotes the proliferation, migration and invasion of osteosarcoma cells by targeting CADM1. *Oncol Lett.* 2018. 16(2): 2620-2628.
- [36] Fujimoto H, Saito Y, Ohuchida K, et al. Deregulated Mucosal Immune Surveillance through Gut-Associated Regulatory T Cells and PD-1(+) T Cells in Human Colorectal Cancer. *J Immunol.* 2018. 200(9): 3291-3303.
- [37] Zhang J, Zhang Y, Lv H, et al. Human stem cells from the apical papilla response to bacterial lipopolysaccharide exposure and anti-inflammatory effects of nuclear factor I C. *J Endod.* 2013. 39(11): 1416-22.
- [38] Li L, Ameri AH, Wang S, et al. EGR1 regulates angiogenic and osteoclastogenic factors in prostate cancer and promotes metastasis. *Oncogene.* 2019. 38(35): 6241-6255.
- [39] Liu N, Zhang J, Yin M, et al. Inhibition of xCT suppresses the efficacy of anti-PD-1/L1 melanoma treatment through exosomal PD-L1-induced macrophage M2 polarization. *Mol Ther.* 2021. 29(7): 2321-2334.
- [40] Li F, Wang T, Huang Y. POU2F1 induces the immune escape in lung cancer by up-regulating PD-L1. *Am J Transl Res.* 2021. 13(2): 672-683.
- [41] Tu J, Wu F, Chen L, et al. Long Non-Coding RNA PCAT6 Induces M2 Polarization of Macrophages in Cholangiocarcinoma via Modulating miR-326 and RhoA-ROCK Signaling Pathway. *Front Oncol.* 2020. 10: 605877.
- [42] Zhang Y, Feng J, Fu H, et al. Coagulation Factor X Regulated by CASC2c Recruited Macrophages and Induced M2 Polarization in Glioblastoma Multiforme. *Front Immunol.* 2018. 9: 1557.
- [43] Shao L, He Q, Wang J, et al. MicroRNA-326 attenuates immune escape and prevents metastasis in lung adenocarcinoma by targeting PD-L1 and B7-H3. *Cell Death Discov.* 2021. 7(1): 145.
- [44] Ji X, Shan L, Shen P, He M. Circular RNA circ\_001621 promotes osteosarcoma cells proliferation and migration by sponging miR-578 and regulating VEGF expression. *Cell Death Dis.* 2020. 11(1): 18.
- [45] Wang J, Cao L, Wu J, Wang Q. Long non-coding RNA SNHG1 regulates NOB1 expression by sponging miR-326 and promotes tumorigenesis in osteosarcoma. *Int J Oncol.* 2018. 52(1): 77-88.
- [46] Cao L, Wang J, Wang PQ. MiR-326 is a diagnostic biomarker and regulates cell survival and apoptosis by targeting Bcl-2 in osteosarcoma. *Biomed Pharmacother.* 2016. 84: 828-835.
- [47] Trieb K, Sulzbacher I, Kubista B. Bcl-2 correlates with localization but not outcome in human osteosarcoma. *Oncol Lett.* 2013. 6(2): 559-561.

- [48] Fu T, Xia C, Li Z, Wu H. Lack of association between bcl-2 expression and prognosis of osteosarcoma: a meta-analysis. *Int J Clin Exp Med*. 2015. 8(6): 9093-9.
- [49] Ferrari S, Bertoni F, Zanella L, et al. Evaluation of P-glycoprotein, HER-2/ErbB-2, p53, and Bcl-2 in primary tumor and metachronous lung metastases in patients with high-grade osteosarcoma. *Cancer*. 2004. 100(9): 1936-42.
- [50] McEachron TA, Triche TJ, Sorenson L, Parham DM, Carpten JD. Profiling targetable immune checkpoints in osteosarcoma. *Oncoimmunology*. 2018. 7(12): e1475873.
- [51] Liu JL, Li J, Xu JJ, et al. MiR-144 Inhibits Tumor Growth and Metastasis in Osteosarcoma via Dual-suppressing RhoA/ROCK1 Signaling Pathway. *Mol Pharmacol*. 2019. 95(4): 451-461.
- [52] Zheng B, Zhou C, Qu G, et al. VEGFR2 Promotes Metastasis and PD-L2 Expression of Human Osteosarcoma Cells by Activating the STAT3 and RhoA-ROCK-LIMK2 Pathways. *Front Oncol*. 2020. 10: 543562.
- [53] Zhang A, Yan T, Wang K, Huang Z, Liu J. PI3Ka isoform-dependent activation of RhoA regulates Wnt5a-induced osteosarcoma cell migration. *Cancer Cell Int*. 2017. 17: 27.
- [54] Cao R, Shao J, Hu Y, et al. microRNA-338-3p inhibits proliferation, migration, invasion, and EMT in osteosarcoma cells by targeting activator of 90 kDa heat shock protein ATPase homolog 1. *Cancer Cell Int*. 2018. 18: 49.
- [55] Jia F, Zhang Z, Zhang X. MicroRNA-338-3p inhibits tumor growth and metastasis in osteosarcoma cells by targeting RUNX2/CDK4 and inhibition of MAPK pathway. *J Cell Biochem*. 2019. 120(4): 6420-6430.
- [56] Xiang D, Li Y, Lin Y. Circular RNA circCCDC66 Contributes to Malignant Phenotype of Osteosarcoma by Sponging miR-338-3p to Upregulate the Expression of PTP1B. *Biomed Res Int*. 2020. 2020: 4637109.
- [57] Zhang H, Wang J, Ren T, et al. LncRNA CASC15 is Upregulated in Osteosarcoma Plasma Exosomes and CASC15 Knockdown Inhibits Osteosarcoma Progression by Regulating miR-338-3p/RAB14 Axis. *Onco Targets Ther*. 2020. 13: 12055-12066.
- [58] Dai N, Zhong ZY, Cun YP, et al. Alteration of the microRNA expression profile in human osteosarcoma cells transfected with APE1 siRNA. *Neoplasma*. 2013. 60(4): 384-94.
- [59] Molist C, Navarro N, Giralt I, et al. miRNA-7 and miRNA-324-5p regulate alpha9-Integrin expression and exert anti-oncogenic effects in rhabdomyosarcoma. *Cancer Lett*. 2020. 477: 49-59.
- [60] Li Z, Li N, Sun X, Wang J. FAM98A promotes cancer progression in endometrial carcinoma. *Mol Cell Biochem*. 2019. 459(1-2): 131-139.

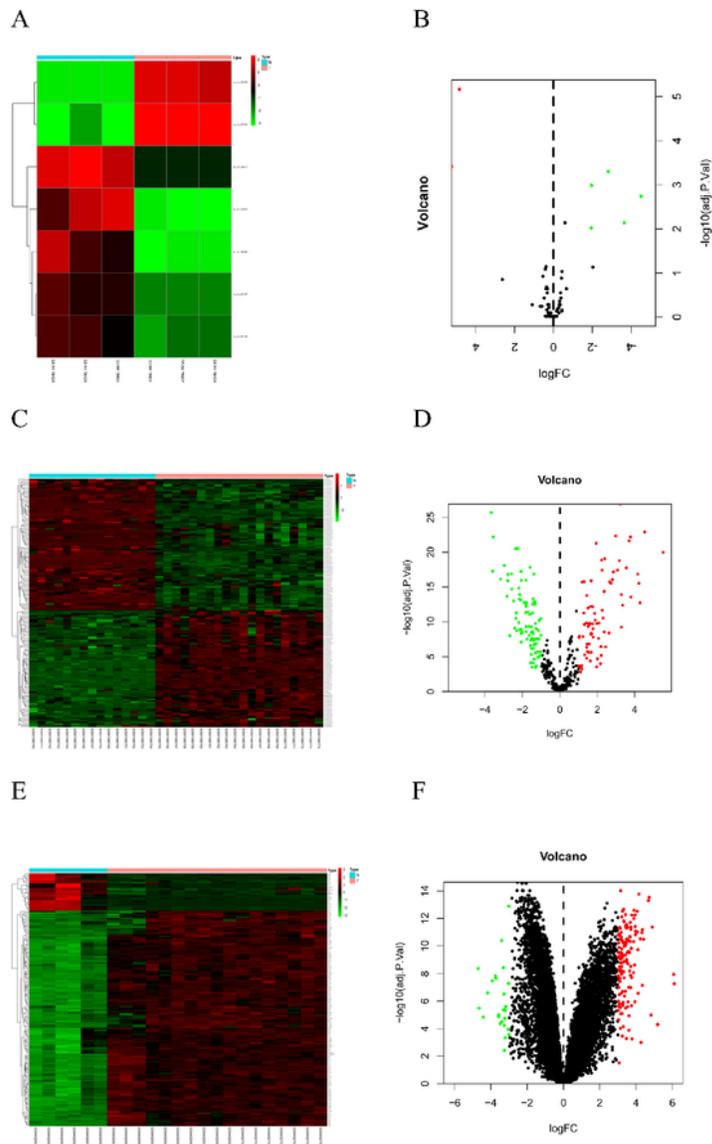
- [61] Zheng R, Liu Q, Wang T, Wang L, Zhang Y. FAM98A promotes proliferation of non-small cell lung cancer cells via the P38-ATF2 signaling pathway. *Cancer Manag Res.* 2018. 10: 2269-2278.
- [62] Liu T, Wang Z, Dong M, Wei J, Pan Y. MicroRNA-26a inhibits cell proliferation and invasion by targeting FAM98A in breast cancer. *Oncol Lett.* 2021. 21(5): 367.
- [63] Fujiwara T, Ye S, Castro-Gomes T, et al. PLEKHM1/DEF8/RAB7 complex regulates lysosome positioning and bone homeostasis. *JCI Insight.* 2016. 1(17): e86330.
- [64] Wang Y, Ye W, Liu Y, Mei B, Liu X, Huang Q. Osteoporosis genome-wide association study variant c.3781 C>A is regulated by a novel anti-osteogenic factor miR-345-5p. *Hum Mutat.* 2020. 41(3): 709-718.
- [65] Hawkins AG, Basrur V, da Veiga Leprevost F, et al. The Ewing Sarcoma Secretome and Its Response to Activation of Wnt/beta-catenin Signaling. *Mol Cell Proteomics.* 2018. 17(5): 901-912.
- [66] He M, Wang Z, Zhao J, Chen Y, Wu Y. COL1A1 polymorphism is associated with risks of osteosarcoma susceptibility and death. *Tumour Biol.* 2014. 35(2): 1297-305.
- [67] Yao L, Zhou Y, Sui Z, et al. HBV-encoded miR-2 functions as an oncogene by downregulating TRIM35 but upregulating RAN in liver cancer cells. *EBioMedicine.* 2019. 48: 117-129.
- [68] Zaoui K, Boudhraa Z, Khalifé P, Carmona E, Provencher D, Mes-Masson AM. Ran promotes membrane targeting and stabilization of RhoA to orchestrate ovarian cancer cell invasion. *Nat Commun.* 2019. 10(1): 2666.
- [69] Dell'Anno I, Barbarino M, et al. EIF4G1 and RAN as Possible Drivers for Malignant Pleural Mesothelioma. *Int J Mol Sci.* 2020. 21(14).
- [70] Jain N, Das B, Mallick B. Restoration of microRNA-197 expression suppresses oncogenicity in fibrosarcoma through negative regulation of RAN. *IUBMB Life.* 2020. 72(5): 1034-1044.
- [71] Li J, Rao B, Yang J, et al. Dysregulated m6A-Related Regulators Are Associated With Tumor Metastasis and Poor Prognosis in Osteosarcoma. *Front Oncol.* 2020. 10: 769.

## Tables

Table 1 Verification of metastasis characteristics

	GSE21257		GSE124768		GSE21257		GSE33382	
	non-metastasis vs metastasis		non-metastasis vs metastasis		non-metastasis vs metastasis within 5 years		non-metastasis vs metastasis within 5 years	
	logFC	P.Value	logFC	P.Value	logFC	P.Value	logFC	P.Value
COL1A1	0.176	0.354	0.398	0.186	0.233	0.174	0.261	0.144
FAM98A	1.330	0.003	0.936	0.009	0.817	0.037	1.020	0.005
RAN	-0.130	0.422	0.512	0.024	-0.145	0.253	-0.131	0.307
HNRNPA2B1	—	—	-0.413	0.005	—	—	—	—

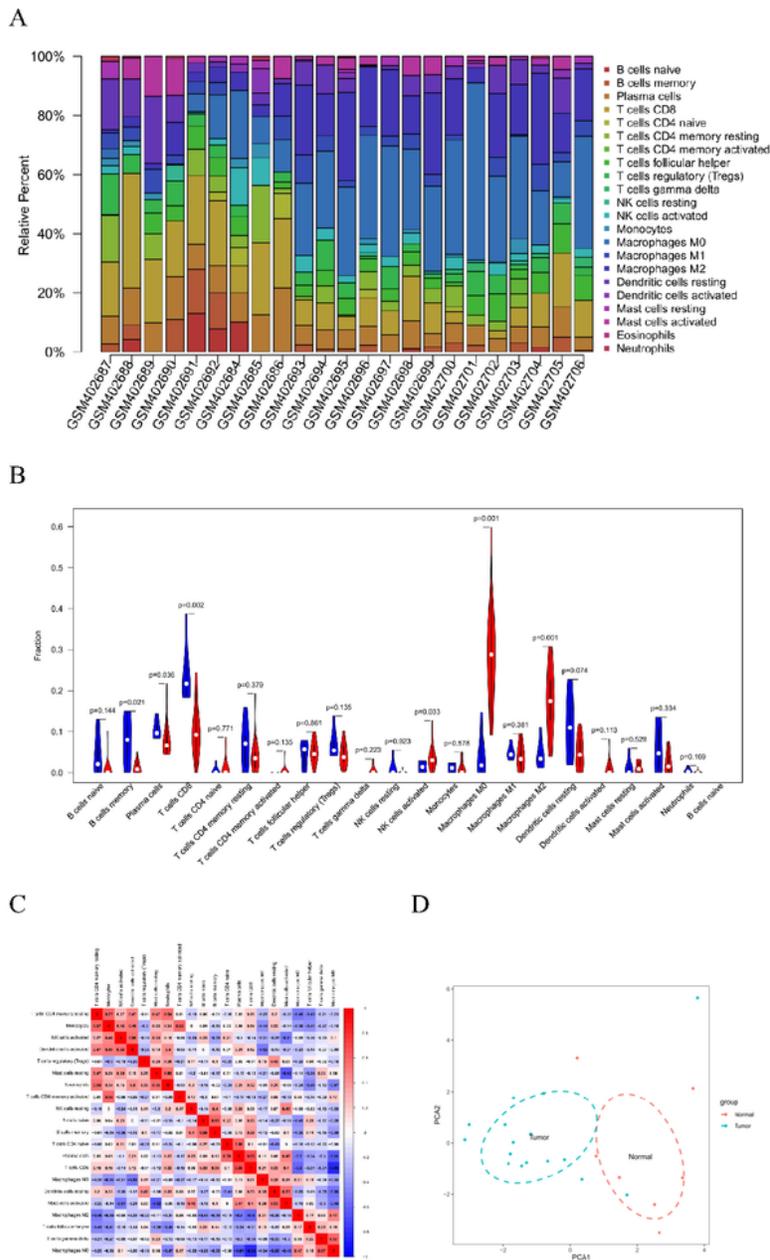
# Figures



**Figure 1**

Characteristics of differentially-expressed genes (DEGs) (A) Unsupervised clustering analysis of differentially-expressed (DE) circular RNAs (circRNAs). Red dots indicate significantly up-regulated circRNAs, green dots indicate significantly down-regulated circRNAs. (B) Volcano plots of circRNAs. Red

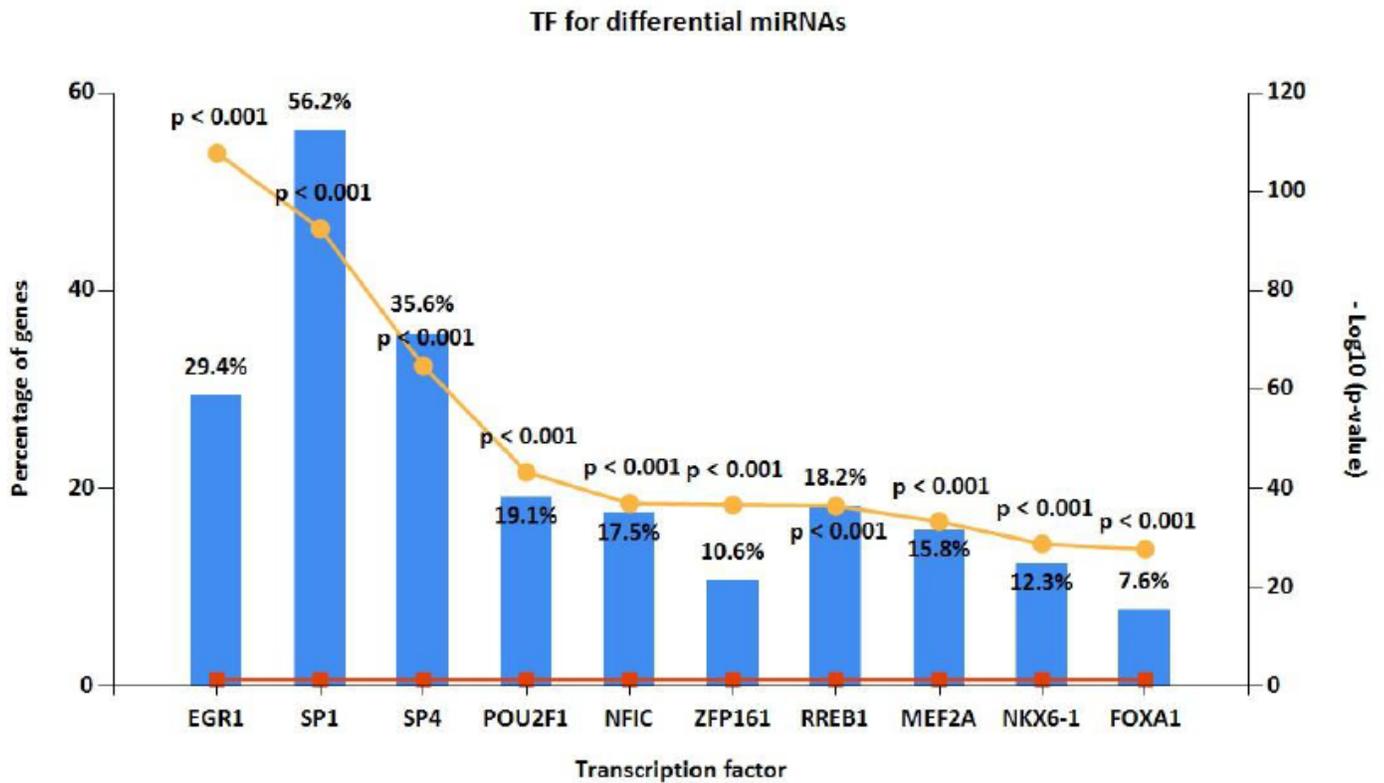
dots indicate up-regulated DEcircRNAs, green dots indicate down-regulated DEcircRNAs, black dots indicate non-differentially expressed circRNAs. (C) Unsupervised clustering analysis of DE micro RNAs (miRNAs). Red dots indicate significantly up-regulated miRNAs, green dots indicate significantly down-regulated miRNAs. (D) Volcano plots of miRNAs. Red dots indicate up-regulated DEmiRNAs, green dots indicate down-regulated DEmiRNAs, black dots indicate non-differentially expressed miRNAs. (E) Unsupervised clustering analysis of the DE messenger RNAs (mRNAs). Red dots indicate significantly up-regulated mRNAs, green dots indicate significantly down-regulated mRNAs. (F) Volcano plots of mRNAs. Red dots indicate up-regulated DEmRNAs, green dots indicate down-regulated DEmRNAs, black dots indicate non-differentially expressed mRNAs.



**Figure 2**

Tumor-infiltrating immune cells analyzed by the identified DEmRNAs (A) The mean proportion of 22 immune cells in the 17 OS patient tissues and six normal human tissues. (B) Violin plot of OS patient tissues (red) and normal human tissues (blue). The P values showed different infiltrating types of immune cells. (C) Correlation matrix of 21 immune cell proportions and immune/stromal score. Variables have been ordered by average linkage clustering. (D) Principal component analysis of immune cell

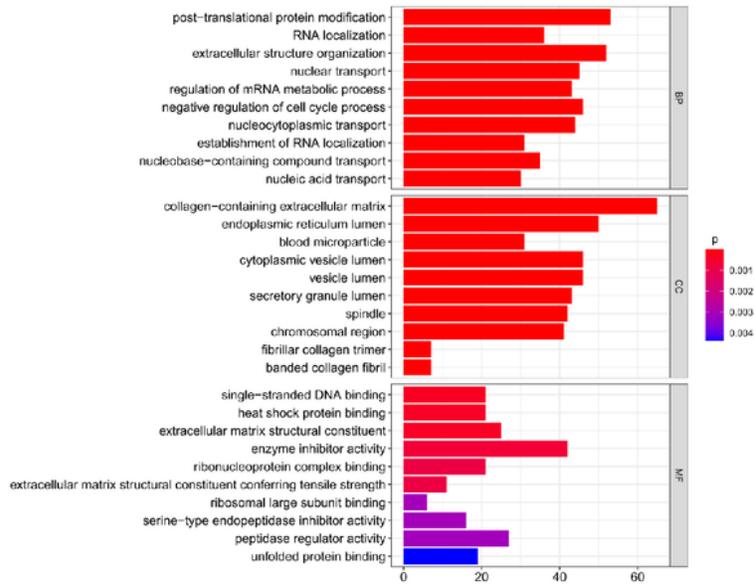
infiltration patterns of 22 types of immune cell between the OS patient tissues and normal human tissues. Red circles indicate normal human tissues, blue circles indicate OS patient tissues.



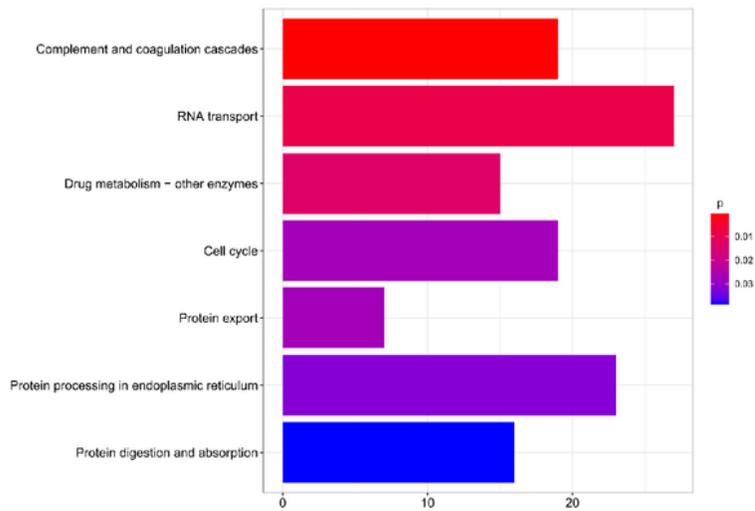
**Figure 3**

Analysis of differentially-expressed transcription factors associated with the identified DE miRNAs. EGR1, SP1, SP4, POU2F1 and NFIC are the most significant molecules in OS.

A

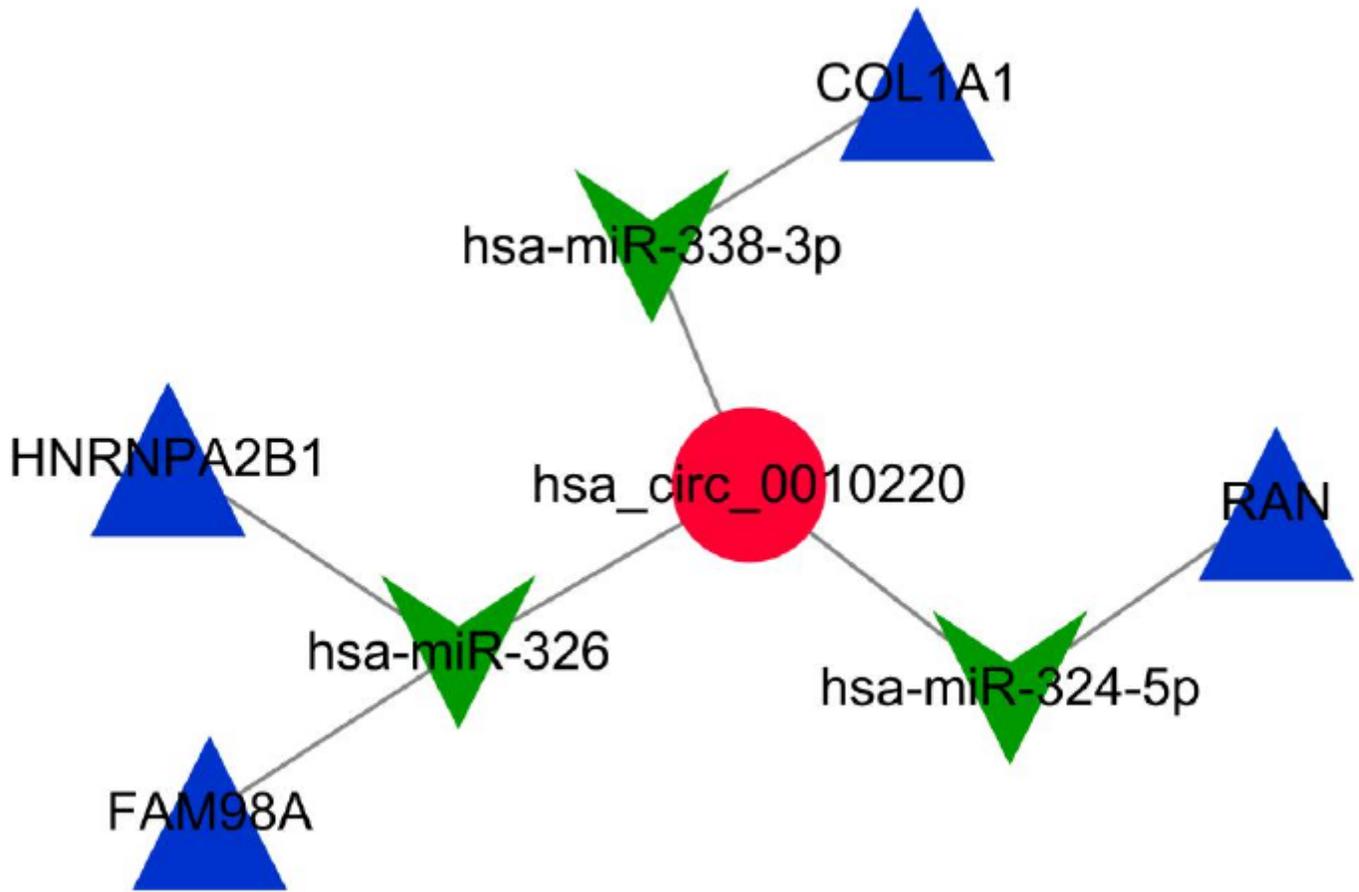


B



**Figure 4**

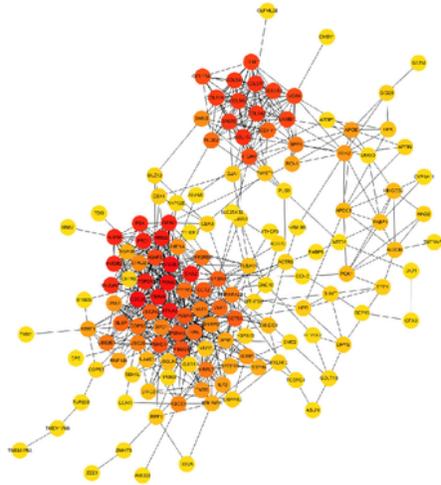
Pathway enrichment analysis (A) GO enrichment analysis of the DEMRNAs in biological process (BP), cellular component (CC) and molecular function (MF) subgroups. (B) KEGG enrichment analysis of DEMRNAs.



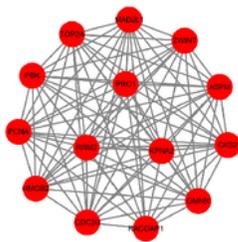
**Figure 5**

The interaction network of DEGs in OS Based on the predicted results of circRNA–miRNA and miRNA–mRNA pairs, we constructed a ceRNA network including one circRNAs, three miRNAs and four mRNAs.

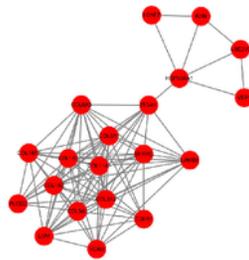
A



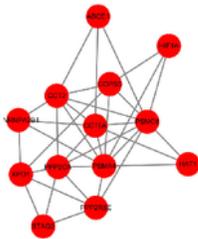
B



C



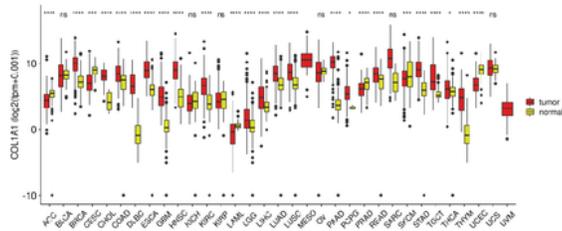
D



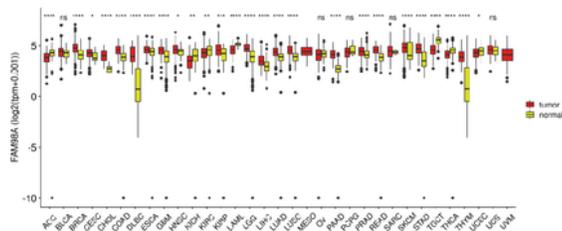
## Figure 6

The PPI network of the target genes of the identified DEmRNAs (A) The target genes of the identified DEmRNAs were ranked in the PPI network. The depth of red indicates the importance of genes in the network. (B) The most closely-clustered subnetwork was composed of 14 nodes and 90 edges. (C) The clustered subnetwork identified by the MCODE plug-in in Cytoscape had 20 nodes and 97 edges. (D) The clustered subnetwork identified by the MCODE plug-in in Cytoscape which had 13 nodes and 35 edges.

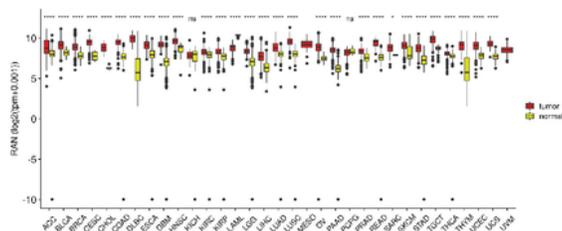
A



B



C



D

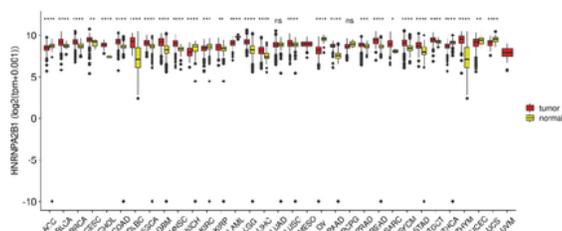
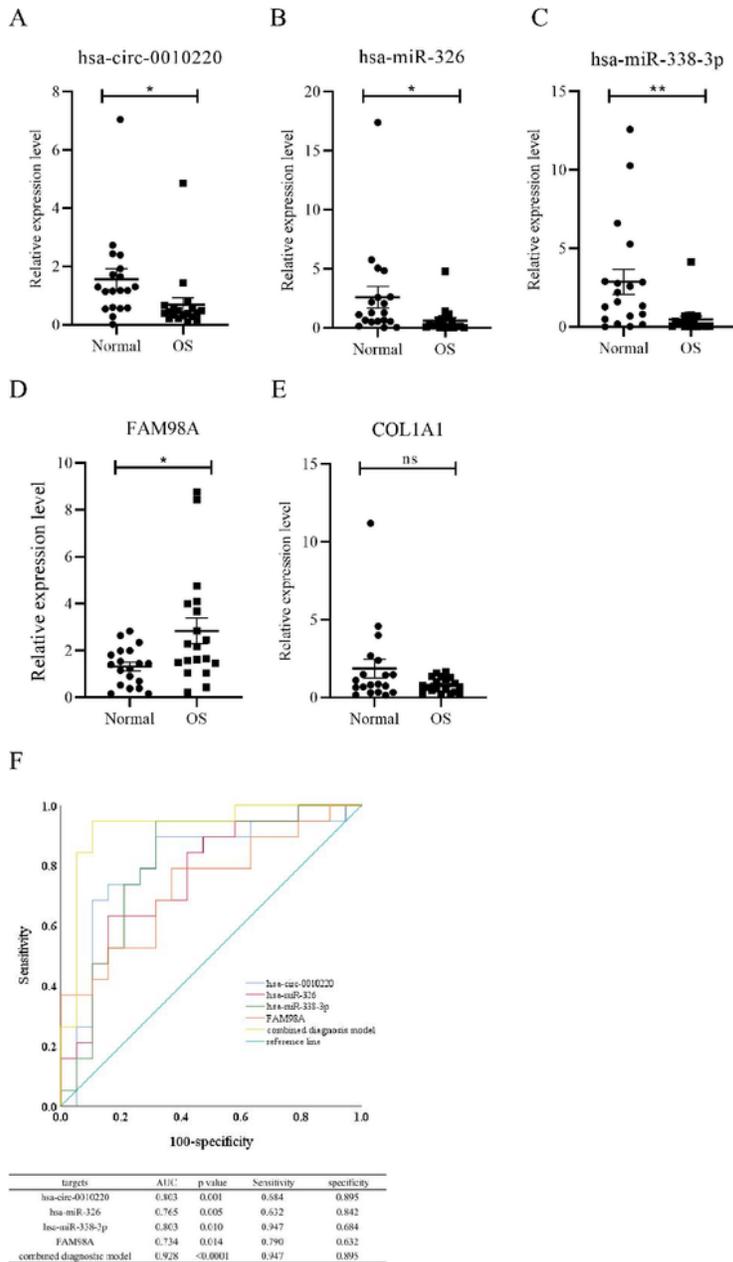


Figure 7

Pan-cancer analysis (A) The expression level of COL1A1 in 32 different cancer types.. (B) The expression level of FAM98A in 32 different cancer types..(C) The expression level of RAN in 32 different cancer types. (D) The expression level of HNRPHA2B1 in 32 different cancer types.



**Figure 8**

Validation the expression of DERNAs in serum sample by RT-qPCR (A) Hsa-circ-0010220 was down-regulated significantly in OS serum sample compared with serum sample from healthy individuals ( $p < 0.05$ ). (B) Hsa-miR-326 was down-regulated significantly in OS serum sample compared with serum sample from healthy individuals ( $p < 0.05$ ). (C) Hsa-miR-338-3p was down-regulated significantly in OS serum sample compared with serum sample from healthy individuals ( $p < 0.01$ ). (D) FAM98A was up-

regulated significantly in OS serum sample compared with serum sample from healthy individuals. (p<0.05) (E) COL1A1 was not significantly differential expression in OS serum sample compared with serum sample from healthy individuals. (p=0.09) (F) The ROC curve of hsa-circ-0010220, hsa-miR-326, hsa-miR-338-3p and FAM98A and the combined diagnostic model.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1Awholeflowchartofthisstudy.jpg](#)
- [TableS1RTqPCRprimers1.pdf](#)