Systematic Review and Meta-Analysis

Medicine®

OPEN

Identification of key genes in osteosarcoma – before and after CDK7 treatment

Yang An, MM[a,b], Yuanlin Wang, MM[c], Guoyong Xu, MM[a], Yinan Liao, MM[d], Ge Huang, MM[b], Xin Jin, MM[b], Chengxin Xie, MM[b], Qinglong Li, MM[b], Dong Yin, MD[a,b,*]

Abstract

Background: Osteosarcoma is one of the most common bone tumors, with a high degree of malignancy and a poor prognosis. Recent studies have shown that THZ2, a cyclin-dependent kinase 7 inhibitor, can exhibit strong antibone tumor effects in vivo and in vitro by inhibiting transcriptional activity. In this study, by screening the differentially expressed genes (DEGs) of osteosarcoma cells before and after THZ2 treatment, it provides new possible targets for the future targeted therapy of osteosarcoma.

Methods: Download the gene expression profile of GSE134603 from the Gene Expression Omnibus database, and use the R software package "limma Geoquery" to screen DEGs. DAVID database was used for gene ontology analysis of DEGs. Use search tool for the retrieval of interacting genes online database and Cytoscape software to construct protein–protein interaction network. Use the "MCODE" plugin in Cytoscape to analyze key molecular complexes (module) of DEGs, and use the "Cluego" plugin to perform Kyoto Encyclopedia of Genes and Genomes enrichment analysis on module genes. The Hub gene is selected from the genes in DEGs that coexist in the top 30 Degree and the Kyoto Encyclopedia of Genes and Genomes pathway.

Results: A total of 1033 DEGs were screened, including 800 up-regulated genes and 233 down-regulated genes. Gene ontology analysis showed that cell component is the main enrichment area of DEGs, mainly in the nucleus, cytoplasm, and nucleoplasm. In addition, in molecular function analysis, DEGs are mainly enriched in the process of protein binding. In biological process analysis, changes in DEGs can also be observed in transcription and regulation using DNA as a template. Twenty-nine module genes are enriched in the Ribosome biogenesis in eukaryotes pathway. Finally, 4 key genes are drawn: essential for mitotic growth 1, U3 SnoRNP protein 3 homolog, U3 small nucleolar RNA-associated protein 15 homolog, and WD repeat domain 3.

Conclusion: This study found that the 4 genes essential for mitotic growth 1, U3 SnoRNP protein 3 homolog, U3 small nucleolar RNA-associated protein 15 homolog, WD repeat domain 3, and the ribosome biogenesis in eukaryotes pathway play a very important role in the occurrence and development of osteosarcoma, and can become a new target for molecular targeted therapy of osteosarcoma in the future.

Abbreviations: CDK7 = cyclin-dependent kinase 7, DEGs = differentially expressed genes, EMG1 = essential for mitotic growth 1, GO = gene ontology, IMP3 = U3 SnoRNP protein 3 homolog, KEGG = Kyoto Encyclopedia of Genes and Genomes, PPI = protein–protein interaction, UTP15 = U3 small nucleolar RNA-associated protein 15 homolog, WDR3 = WD repeat domain 3.

Keywords: bioinformatics, differentially expressed genes, essential for mitotic growth 1, osteosarcoma, ribosome biogenesis in eukaryotes, U3 small nucleolar RNA-associated protein 15 homolog, U3 SnoRNP protein 3 homolog, WD repeat domain 3

1. Introduction

Osteosarcoma is the most common aggressive and malignant tumor in bones, mainly occurring in children and adolescents.[1] It usually occurs in the metaphysis of long bones, especially in the distal femur, proximal humerus and tibia.[2] Osteosarcoma is thought to originate from mesenchymal stem cells and consists of malignant osteoblasts that produce immature bone or bone-like tissue.[2–4] Current research indicates that the factors that affect the occurrence of osteosarcoma may be related to germline genetic variation, high birth weight and adolescent hormones.[5–7] The prognosis of its patients is poor, and the survival rate of osteosarcoma patients before the 1970s is less than 20%. After the introduction of chemotherapy, surgical resection with sufficient surgical margin and the use of new adjuvant chemotherapy with methotrexate, doxorubicin, cisplatin and ifosfamide can increase the survival rate to 60% to 70%,[8] which is also the current standard therapy for osteosarcoma. In addition, the occurrence of metastasis is also an important factor affecting the survival rate of osteosarcoma. The long-term survival rate of patients with localized osteosarcoma is about 65%, while the long-term survival rate of patients with metastatic

Editor: Jianxun Ding.

The authors have no conflicts of interest to disclose.

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

[a] Guangxi Collaborative Innovation Center for Biomedicine, Guangxi Medical University, Nanning, P. R. China, [b] Department of Orthopedics, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, P. R. China, [c] Graduate School, Tianjin Medical University, Tianjin, P. R. China, [d] Pharmaceutical College, Guangxi Medical University, Nanning, P. R. China.

[*] Correspondence: Dong Yin, Department of Orthopedics, The People's Hospital of Guangxi Zhuang Autonomous Region, No. 6 Taoyuan Road, Nanning 530021, P. R. China (e-mail: tangin2002@163.com).

How to cite this article: An Y, Wang Y, Xu G, Liao Y, Huang G, Jin X, Xie C, Li Q, Yin D. Identification of key genes in osteosarcoma – before and after CDK7 treatment. Medicine 2021;100:39(e27304).

Received: 5 October 2020 / Received in final form: 13 July 2021 / Accepted: 2 September 2021

http://dx.doi.org/10.1097/MD.0000000000027304

osteosarcoma is less than 20%.[9,10] In the past 30 years, the survival rate of patients with osteosarcoma has not improved significantly, and patients with metastatic osteosarcoma have developed resistance to neoadjuvant chemotherapy drugs.[11] Therefore, searching for new biological targets or new molecular mechanisms is of great significance for the improvement of the therapeutic effect of osteosarcoma.

In recent years, due to the development of biology, genome and pathway analysis, and targeted drug research, many cytokines and pathways related to the occurrence and development of osteosarcoma have been discovered. Receptor activator of NF-kB ligand is considered to be dysregulated in bone tumors, and its homologous receptor TNF receptor superfamily member 11a signaling promotes the motility and anchorage-independent growth of osteosarcoma cells.[12] The targeted transmembrane glycoprotein neuromedin B exhibits in vitro cytotoxicity to osteosarcoma cells, and it is expressed at high levels in osteosarcoma cells.[13,14] Disialylganglioside is expressed by almost all osteosarcoma tumors, and this abnormal expression still exists after recurrence. Its targeted therapy is being studied for the treatment of relapsed and refractory osteosarcoma.[15,16] In addition, many targeted drugs such as PI3K/mTOR inhibitors, tyrosine kinase inhibitors, etc are also widely used in the targeted therapy of osteosarcoma.

Recent studies have shown that the malignant potential of osteosarcoma may be related to super enhancer-related genes.[17,18] Cyclin-dependent kinase (CDK) 7 is a type of CDK, which is related to the transcription initiation factor TFIIH. It is not only an effector CDK that phosphorylates CD II and other targets, but also at least one other CDK necessary for CDK involved in transcription activated kinase.[19] Knockout of CDK7 can reduce the phosphorylation of the C-terminal repeat domain of RNA polymerase II, which is rich in super enhancers and has inhibitory effects on a variety of cells.[20,21] Zhang et al[22] verified through experiments that THZ2, a specific small molecule inhibitor of CDK7, exhibits strong anti-osteosarcoma effects both in vivo and in vitro. This result suggests that using a specific CDK7 inhibitor THZ2 to target the super-enhancer-related oncogenic transcription program may be an effective treatment strategy for patients with osteosarcoma. It can be inferred that the differentially expressed genes (DEGs) in osteosarcoma cells before and after CDK7 treatment are the key genes for osteosarcoma targeted therapy.

In the present study, we analyzed the expression profile data (GSE134603) provided by Zhang et al,[22] and found a batch of DEGs after group comparison. Next, we performed gene ontology (GO) analysis, protein–protein interaction (PPI) network, module genes screening and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis on these DEGs. We screened out the key differential genes related to CDK7 therapy, revealed the abnormally expressed molecules during the occurrence of osteosarcoma, and provided new targets for the future targeted therapy of osteosarcoma.

2. Methods and materials

2.1. Microarray data

The GSE134603 data set is downloaded from the Gene Expression Omnibus (https://www.ncbi.nlm.nih.gov/geo/) database.[23] It is based on the chip gene expression data of the GPL15207 platform ([PrimeView] Affymetrix Human Gene Expression Array), which contains 24 samples in total. They are total RNA extracted from U2-OS and SJSA-1 cells treated with DMSO, 24 nM THZ2, 100 nM THZ2, and 400 nM THZ2 for 6 hours. Each treatment group has 3 replicate samples. The U2-OS cell line is an osteosarcoma cell derived from the tibia and has been used in basic research as a typical osteosarcoma cell line since 1964. Therefore, we selected U2-OS cell specimens (3 cases each) treated with DMSO and 400 nM THZ2 and divided them into 2 groups for study. Our data comes from an open database, so ethics and patient consent are not applicable.

2.2. Data processing

In the Rstudio software (https://www.rstudio.com/), use the R (v.4.0.2) software packages "limma Geoquery"[24] and "BioConductor"[25] for differential gene extraction and gene ID conversion to filter DEGs. Select P-value $< .01$ and $|\log \text{fold change}| > 1.5$ as the screening criteria, and the analysis results are saved in table form. Use the "ggplot2" software package (http://ggplot2-book.org/) to process DEGs, make a volcano map, and do a heat map analysis of the top 50 differential genes in the 2 groups.

2.3. GO enrichment analysis

Use the DAVID6.8 database[26] (https://david.ncifcrf.gov/) to perform GO analyses on the selected DEGs, and P-value $< .01$ is the admission standard. Use the "ggplot2" software package to visualize the GO enrichment analysis results and make bubble charts.

2.4. PPI network construction

To analyze the relationship among proteins, functional protein association networks database STRING (version 11.0) was used to evaluate the interaction between DEGs; a composite score $>$ 0.4 was considered a statistically significant interaction. After exporting the results, Cytoscape software (v.3.7.2; https://cytoscape.org/)[27] was used to make visual adjustments.

2.5. Screening of module genes and KEGG enrichment analysis

Due to the large number of DEGs, we screened the "MCODE" plugin through the molecular complex detection in Cytoscape, Degree cutoff > 2, node score cutoff > 0.2, K score > 2 to screen out key molecular complexes (module). For the module results, use the "Gluego" plugin[28] to perform the KEGG pathway enrichment analysis, to screen out the pathways in the module concentration, and P-value $< .05$ is the admission standard.

2.6. Screening of Hub gene

We sorted the genes in the entire PPI map according to the number of degrees. Among the top 30 genes, compared with the KEGG result pathway obtained in 2.5., the overlapping genes are the selected Hub genes.

3. Results

3.1. Changes in the expression level of DEGs before and after THZ2 treatment

After conducting differential gene screening and gene ID conversion according to the R language software packages

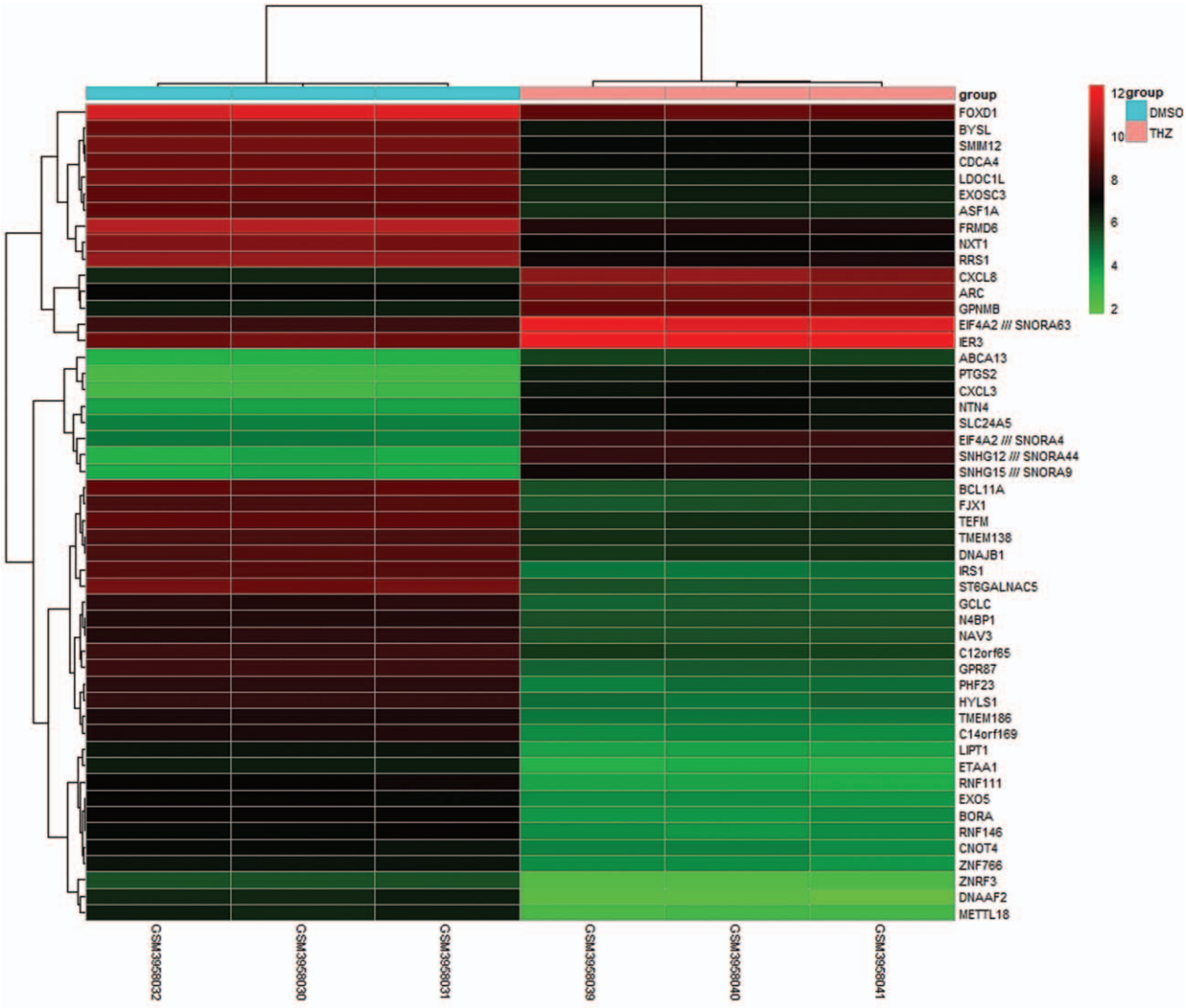

Figure 1. Heat map of the top 50 differentially expressed genes.

"limma" and "BioConductor", we identified 1033 DEGs in U2OS cells before and after THZ2 treatment, of which 800 were up-regulated genes and 233 were down-regulated genes. The volcano map and heat map produced by the "ggplot2" software package show that DEGs have a significant influence on the difference between the 2 groups of cells (Figs. 1 and 2).

3.2. GO enrichment analysis

The GO enrichment analysis of DEGs was carried out using DAVID database, and the results showed that DEGs were enriched in biological processes. We used the "ggplot2" software package to make a bubble chart to observe its trends more intuitively (Fig. 3). In Figure 3, we can observe that the cell component is the main enrichment area of DEGs, mainly in the nucleus, cytoplasm, and nucleoplasm. In addition, in molecular function analysis, DEGs are mainly enriched in the process of protein binding. In biological processes analysis, changes in DEGs can also be observed in transcription and regulation using DNA as a template.

3.3. Screening of module genes and KEGG enrichment analysis

Considering that too many DEGs will have an impact on the results, we screened the "MCODE" plugin (Degree cutoff >2, node score cutoff >2, K score >2) through the molecular complex detection in Cytoscape, and selected 29 key ones Molecular complex (module) (Table 1). For the module results, use the "Gluego" plugin to perform KEGG analysis to screen out the pathways concentrated in the module. The P-value $<.05$ is the admission standard. The results are shown in Table 2. The module genes are mainly enriched in the ribosome biogenesis in eukaryotes pathway (Fig. 4). The relationship between ribosomes and DNA and its transcription is very close, and the results of KEGG analysis and GO enrichment analysis correspond to each other.

3.4. PPI network construction

Use search tool for the retrieval of interacting genes database (version 11.0) to evaluate the interaction between DEGs and integrate PPI networks. Using Cytoscape software for visual

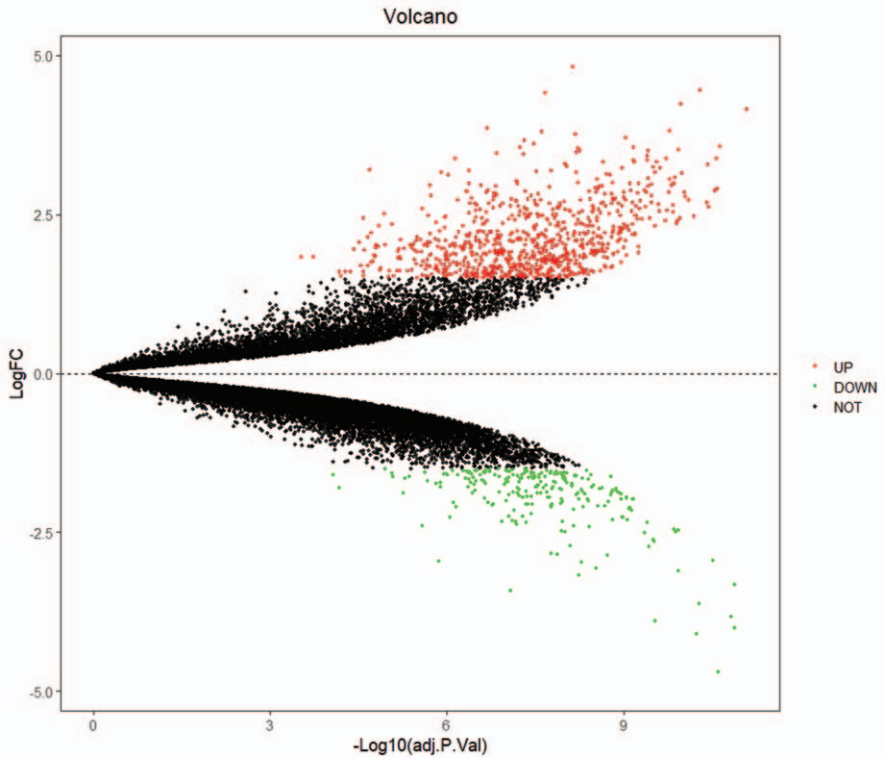

Figure 2. Volcano map of all DEGs, screening criteria: *P*-value < .01 and logFC| > 1.5. DEGs = differentially expressed genes, FC = fold change.

adjustment, there are a total of 878 points and 8524 edges in the entire PPI network (Fig. 5).

3.5. Screening of Hub genes

We sorted according to the number of Degrees and screened out the top 30 genes with differences, as shown in Figure 6. Then, it was compared with the related genes in the Ribosome biogenesis in eukaryotes pathway from KEGG (Table 3). As a result, 4 Hub genes were identified.

4. Discussion

Osteosarcoma is one of the most common primary bone malignancies. The current standard treatment for osteosarcoma is to surgically remove the tumor to make it have a suitable

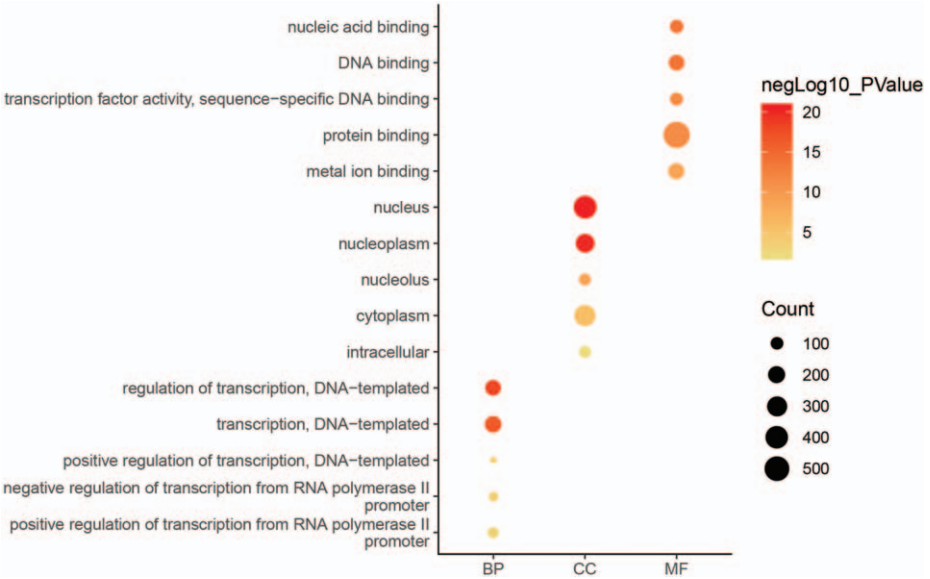

Figure 3. GO functional enrichment analysis results from up and down regulate DEGs. DEGs = differentially expressed genes.

Table 1

Key molecular complexes by "MCODE" module in Cytoscape.

Module name
DHX32, DDX10, LTV1, CEBPZ, UTP15, WDR43, GNL3, ZNF622, RRS1, NIFK, BRIX1, PPAN, IMP3, WDR3, UTP23, TFB2M, FTSJ2, EMG1, MPHOSPH10, NGDN, RPP38, ABT1, RIOK2, TRMT1L, WDR75, RRP1B, POLR1B, TEX10, BYSL

ABT1 = activator of basal transcription 1, BYSL = bystin like, BRIX1 = biogenesis of ribosomes BRX1, CEBPZ = CCAAT enhancer binding protein zeta, DHX32 = DEAH-box helicase 32 (Putative), DDX10 = DEAD-box helicase 10, EMG1 = essential for mitotic growth 1, FTSJ2 = mitochondrial RRNA methyltransferase 2, NGDN = neuroguidin, GNL3 = G protein nucleolar 3, IMP3 = U3 SnoRNP protein 3 homolog, LTV1 = LTV1 ribosome biogenesis factor, MPHOSPH10 = M-phase phosphoprotein 10, NIFK = nucleolar protein interacting with the FHA domain of MKI67, POLR1B = RNA polymerase I subunit B, PPAN = peter pan homolog, RIOK2 = RIO kinase 2, RRS1 = ribosome biogenesis regulator 1 homolog, RRP1B = ribosomal RNA processing 1B, RPP38 = ribonuclease P/MRP subunit P38, TFB2M = transcription factor B2, mitochondrial, TEX10 = testis expressed 10, TRMT1L = TRNA methyltransferase 1 Like, UTP15 = U3 small nucleolar RNA-associated protein 15 homolog, UTP23 = UTP23 small subunit processome component, WDR75 = WD repeat domain 75, WDR43 = WD repeat domain 43, WDR3 = WD repeat domain 3. ZNF622 = zinc finger protein 622.

margin, and to cooperate with chemotherapy before and after surgery. This therapy has now entered a bottleneck. Traditional chemotherapeutic drugs have limited efficacy and it is difficult to achieve breakthrough progress. At present, the research and development of specific molecular targeted drugs for osteosarcoma has become a hot spot at home and abroad.[29] However, due to the heterogeneity of tumors, the development of new targeted drugs is extremely difficult. At present, targeted therapies for receptor activator of NF-kB ligand, glycoprotein neuromedin B, and disialylganglioside have made preliminary progress, and PI3K/mTOR inhibitors have also been shown to have certain research potential in the treatment of osteosarcoma.[12–16] The biological characteristics of osteosarcoma are still being extensively studied. Finding the key genes and pathways for the occurrence and development of osteosarcoma can greatly promote the diagnosis and help the treatment and prognosis evaluation.

THZ2 is a new type of CDK7-specific inhibitor that inhibits the activity of CDK7 through covalent binding and selectively inhibits superenhancer-related genes, especially oncogenes. Studies in vivo and in vitro experiments have verified that THZ2 can inhibit the phosphorylation level of CDK7 substrate CDK2T160, inactivate the survival signal pathway EGFR/RAS/AKT/ERK, and increase the level of intracellular ROS, which has an impact on the growth of gastric cancer cells.[30] THZ2 can also effectively inhibit the growth of triple negative breast cancer cells.[31,32] For osteosarcoma cells, THZ2 affects its PI3K-AKT signaling and MAPK signaling pathways, and has certain effects on cell apoptosis and cell migration.[22] In this study, we used the changes in cell gene expression before and after THZ2 treatment to infer which genes play a key role in the entire treatment process, and then look for new targets for osteosarcoma targeted therapy.

After software analysis, we identified 1033 DEGs, of which 800 were up-regulated genes and 233 were down-regulated genes. GO analysis shows that DEGs are concentrated in the nucleus, nucleoplasm and cytoplasmic cell components, and are reflected in the process of transcription and regulation of DNA as a template and protein binding. Through the screening of key molecular complexes, we identified 29 module genes represented by DEAH-box helicase 32 (Putative), DEAD-box helicase 10, and LTV1 ribosome biogenesis factor. After KEGG analysis of the module, it was found that the module gene was concentrated in the pathway Ribosome biogenesis in eukaryotes. This result coincides with the result of GO analysis. Pelletier et al[33] retrospectively proposed that changes in the number and quality of ribosomes lead to deviations in translation patterns, which may ultimately promote the onset of cancer, and inhibiting ribosome biogenesis may be a potential treatment avenue of cancer. Next, we compared the DEGs contained in the above pathways with the top 30 genes in all DEGs and identified 4 Hub genes: essential for mitotic growth 1 (EMG1), U3 SnoRNP protein 3 homolog (IMP3), U3 small nucleolar RNA-associated protein 15 homolog (UTP15), and WD repeat domain 3 (WDR3).

We searched the GeneCards database for information about Hub genes. The 4 Hub genes are all related to the biogenesis of

Table 2

KEGG pathway analysis of module genes associated with osteosarcoma.

Pathway ID	Term	Count	*P*-value
hsa03008	Ribosome biogenesis in eukaryotes	10	1.85E-18

KEGG = kyoto encyclopedia of genes and genomes.

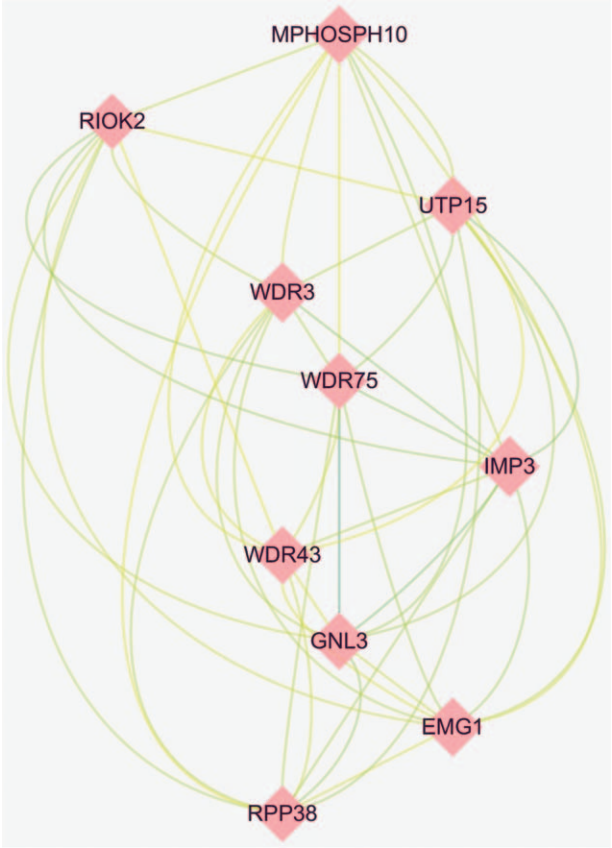

Figure 4. The module genes in the ribosome biogenesis in eukaryotes pathway.

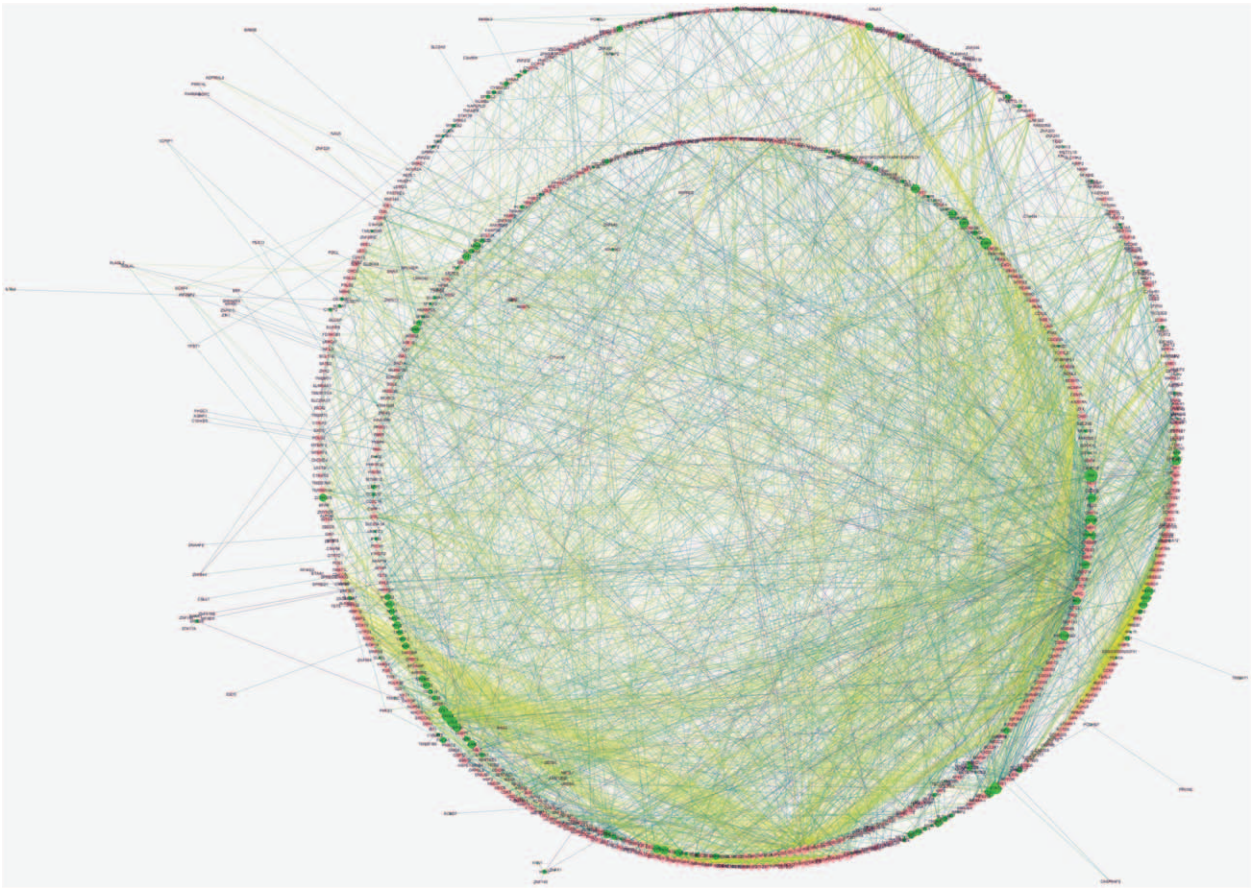

Figure 5. PPI network constructed by STRING database for the DEGs, the red diamonds represent the up-regulated DEGs and the green balls represent the down-regulated DEGs. DEGs = differentially expressed genes, STRING = search tool for the retrieval of interacting genes.

ribosomal RNA (Table 4). The WDR3 family is also involved in cell cycle processes, signal transduction, cell apoptosis and gene regulation. EMG1 is an essential and conserved eukaryotic protein, which plays a key role in ribosomal biogenesis as a ribosome assembly factor. This protein can methylate pseudouridine in 18S rRNA. The specific process is to recruit small ribosomal subunit protein RPS19 to mature ribosomes and remove snR57, which is the snoRNA component of snoRNP and is responsible for yeast 18S rRNA 2'-OH ribose methylation in 1570.[34] Studies have shown that the decrease of EMG1 level will cause the decrease of 18S rRNA level, which will affect the protein synthesis level and cell proliferation rate.[35] However, its effect on tumor cells is not yet clear, but we can guess its expression changes in levels are also closely related to the proliferation rate of cancer cells, which is worthy of further research.

IMP3 is a Vg1-RBP/VERA homologous gene, originally called KOC, and was discovered due to its highly elevated expression in pancreatic cancer.[36] The IMPs are also known as insulin-like growth factor 2 messenger RNA insulin-like growth factor 2 binding proteins, which are highly conserved cancer-fetal RNA binding proteins that regulate RNA processing in terms of localization, translation and stability process. During embryogenesis, IMPs are expressed in large quantities, and they are believed to play an important role in cell migration, metabolism and stem cell renewal. Except for IMP2 after birth, IMP1, and IMP3 are both in a low-level expression state.[37] However, in tumor cells, we can observe a substantial increase in the expression levels of these 3 proteins. They are believed to not only promote the occurrence and development of tumors, but also play a key role in maintaining the establishment and stability of tumor levels.[38] In osteosarcoma, some studies speculate that the expression of IMP3 is related to its vascular metastasis, and the inhibition of IMP3 has broad application prospects in tumor vascular targeted therapy.[39]

Utp15 is a U3 snoRNA-related protein of the small subunit process group, which is essential for the biogenesis of 18S rRNA.[40] In yeast, the loss of any U3 protein will cause 18S rRNA biogenesis defects and disrupt the formation of 40S small ribosomal subunits.[40] Studies have shown that p53 mediates the lack of UTP15, which affects the angiogenesis pathway.[41] At present, the effect of UTP15 on tumor cells is not clear, but because of its close association with tumor suppressor gene p53, we can guess that UTP15 may have an effect on p53 expression or other ways to affect tumor cells.

WDR3 is a nuclear protein composed of 10 WD repeat units. It is located on chromosome 1p12–p13, which is often changed in malignant tumors and solid tumors.[42] UTP12, a yeast homologue of WDR3, is a component of the pre-rRNA processing complex, which is essential for rRNA processing and the synthesis of small ribosomal subunits.[40] The study by McMahon et al showed that the uncontrolled expression of

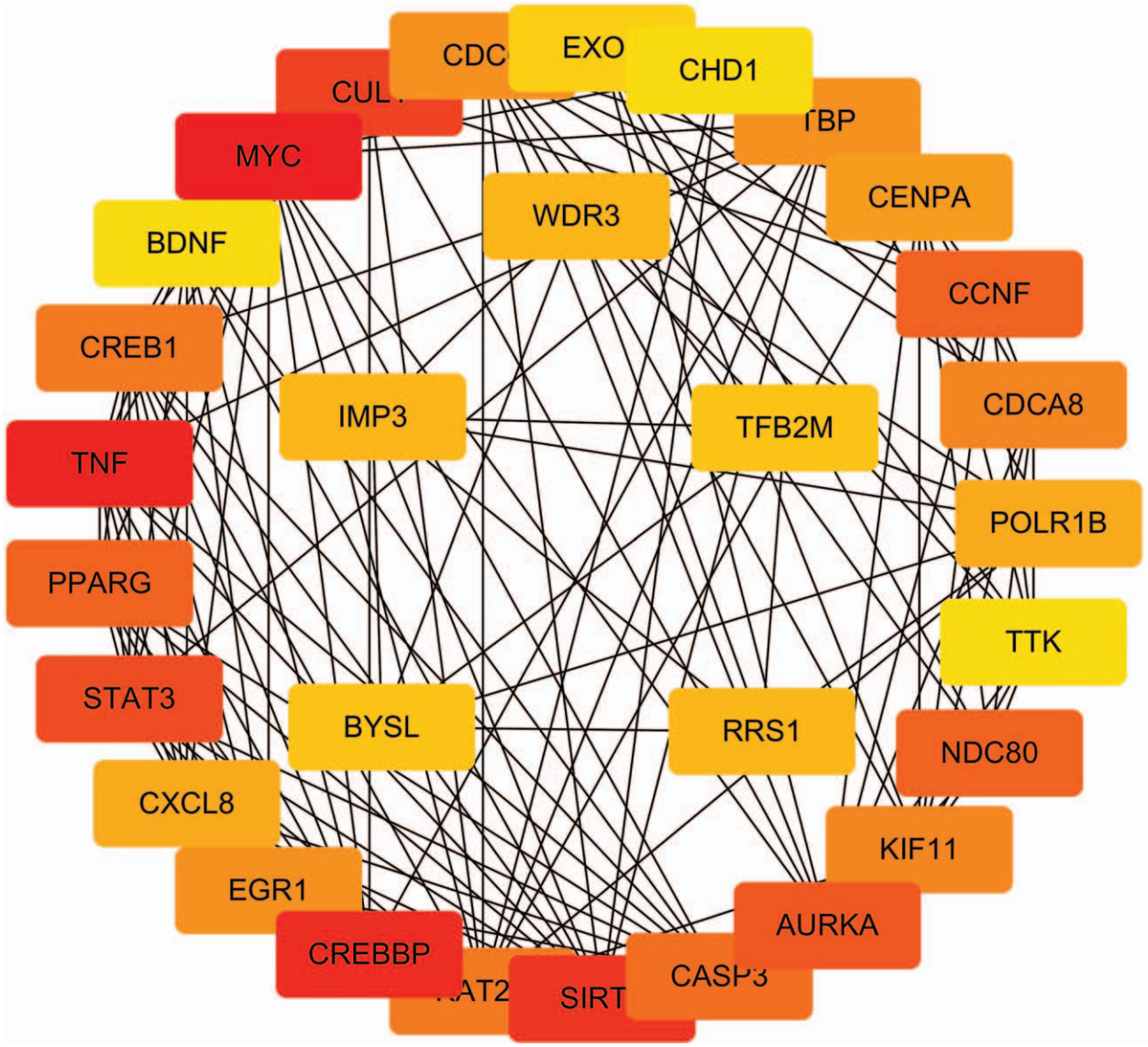

Figure 6. The top 30 genes in degree score.

WDR3 in cancer cells would disrupt the signaling pathway between ribosome biogenesis and p53 activation. Increasing the expression of WDR3 would promote cell proliferation, especially in p53 mutants or nonexpressed ones in tumor cells.[43] Therefore, the study of targeted molecular therapy of WDR3 related genes has great research value for the treatment and prognosis of tumors caused by p53 mutation or nonexpression.

Table 3

Associated genes in KEGG pathway and degree top 30 gene.

Ribosome biogenesis in eukaryotes	Degree top 30 genes
EMG1, GNL3, IMP3, MPHOSPH10, RIOK2, RPP38, UTP15, WDR3, WDR43, WDR75	TBP, PPARG, EXO1, CASP3, EMG1, CENPA, TNF, CCNF, CDC6, CREB1, TFB2M, UTP15, KIF11, CDCA8, KAT2B, RRS1, CHD1, MYC, IMP3, CUL1, WDR3, SIRT1, AURKA, BYSL, CXCL8, CREBBP, POLR1B, EGR1, STAT3, NDC80

AURKA = aurora kinase A, BYSL = bystin like, EMG1 = essential for mitotic growth 1, CASP3 = caspase 3, CCNF = cyclin F, CDCA8 = cell division cycle associated 8, CDC6 = cell division cycle 6, CHD1 = chromodomain helicase DNA binding protein 1, CREB1 = CAMP responsive element binding protein 1, CREBBP = CREB binding protein, CENPA = centromere protein A, CXCL8 = C-X-C motif chemokine ligand 8, CUL1 = cullin 1, EGR1 = early growth response 1, EXO1 = exonuclease 1, GNL3 = G protein nucleolar 3, IMP3 = U3 SnoRNP protein 3 homolog, KAT2B = lysine acetyltransferase 2B, KEGG = kyoto encyclopedia of genes and genomes, KIF11 = kinesin family member 11, MYC = MYC proto-oncogene, BHLH transcription factor, NDC80 = NDC80 kinetochore complex component, PPARG = peroxisome proliferator activated receptor gamma, POLR1B = RNA polymerase I subunit B, RPP38 = ribonuclease P/MRP subunit P38, RRS1 = ribosome biogenesis regulator 1 homolog, RIOK2 = RIO kinase 2, STAT3 = signal transducer and activator of transcription 3, SIRT1 = sirtuin 1, TNF = tumor necrosis factor, TFB2M = transcription factor B2, mitochondrial, UTP15 = U3 small nucleolar RNA-associated protein 15 homolog, WDR43 = WD repeat domain 43, WDR75 = WD repeat domain 75, WDR3 = WD repeat domain 3.

Table 4

Function of 4 hub genes.

Gene symbol	Full name	Function
EMG1	Essential for mitotic growth 1	An essential, conserved eukaryotic protein that methylates pseudouridine in 18S rRNA. The related protein in yeast is a component of the small subunit processome and is essential for biogenesis of the ribosomal 40S subunit. Alternative splicing results in multiple transcript variants.
IMP3	U3 SnoRNP protein 3 homolog	Component of the 60 to 80S U3 small nucleolar ribonucleoprotein (U3 snoRNP). Required for the early cleavages during pre-18S ribosomal RNA processing.
UTP15	U3 small nucleolar RNA-Associated protein 15 homolog	Ribosome biogenesis factor. Involved in nucleolar processing of pre-18S ribosomal RNA. Required for optimal preribosomal RNA transcription by RNA polymerase I.
WDR3	WD repeat domain 3	A protein coding gene. Proteins belonging to the WD repeat family are involved in a variety of cellular processes, including cell cycle progression, signal transduction, apoptosis, and gene regulation.

UTP15 = UTP15 small subunit processome component

5. Conclusion

Regarding the treatment of osteosarcoma, surgical resection combined with drug chemotherapy is still the standard treatment method.[8] In recent years, various new therapies have been researched and applied to the treatment of osteosarcoma, but most of them have not made breakthrough progress.[44,45] Our research helps to further understand the molecular mechanism of the occurrence and development of osteosarcoma. Hub genes EMG1, IMP3, UTP15 and WDR3 participate in the occurrence and development of osteosarcoma cells through the pathway of Ribosome biogenesis in eukaryotes. This discovery will promote the research and development of future molecular targeted therapy of osteosarcoma, and provide a possible and effective research approach for improving the survival rate and prognosis of patients with osteosarcoma.

Acknowledgments

We are grateful to Dr. Dong Yin (Department of Orthopedics, The People's Hospital of Guangxi Zhuang Autonomous Region) for his kindly assistance in all stages of the present study.

Author contributions

Conceptualization: Yang An, Guoyong Xu, Dong Yin.
Data curation: Yang An, Dong Yin.
Formal analysis: Xin Jin, Dong Yin.
Methodology: Yinan Liao, Ge Huang, Dong Yin.
Software: Yuanlin Wang, Chengxin Xie.
Visualization: Yuanlin Wang, Qinglong Li.
Writing – original draft: Yang An, Dong Yin.
Writing – review & editing: Dong Yin.

References

[1] Nagarajan R, Kamruzzaman A, Ness KK, et al. Twenty years of follow-up of survivors of childhood osteosarcoma: a report from the childhood cancer survivor study. Cancer 2011;117:625–34.
[2] Bielack SS, Kempf-Bielack B, Delling G, et al. Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. J Clin Oncol 2002;20:776–90.
[3] Zhang Y, Li J, Wang Y, Jing J, Li J. The roles of circular RNAs in osteosarcoma. Med Sci Monit 2019;25:6378–82.
[4] Jo VY, Fletcher CDM. WHO classification of soft tissue tumours: an update based on the 2013 (4th) edition. Pathology 2014;46:95–104.
[5] Musselman JR, Bergemann TL, Ross JA, et al. Case-parent analysis of variation in pubertal hormone genes and pediatric osteosarcoma: a Children's Oncology Group (COG) study. Int J Mol Epidemiol Genet 2012;3:286–93.
[6] Mirabello L, Pfeiffer R, Murphy G, et al. Height at diagnosis and birth-weight as risk factors for osteosarcoma. Cancer Causes Control 2011;22:899–908.
[7] Kansara M, Teng MW, Smyth MJ, Thomas DM. Translational biology of osteosarcoma. Nat Rev Cancer 2014;14:722–35.
[8] Longhi A, Errani C, De Paolis M, Mercuri M, Bacci G. Primary bone osteosarcoma in the pediatric age: state of the art. Cancer Treat Rev 2006;32:423–36.
[9] Tsuchiya H, Kanazawa Y, Abdel-Wanis ME, et al. Effect of timing of pulmonary metastases identification on prognosis of patients with osteosarcoma: the Japanese musculoskeletal oncology group study. J Clin Oncol 2002;20:3470–7.
[10] Kager L, Zoubek A, Pötschger U, et al. Primary metastatic osteosarcoma: presentation and outcome of patients treated on neoadjuvant cooperative osteosarcoma study group protocols. J Clin Oncol 2003;21:2011–8.
[11] Kawai A, Yonemori K, Takahashi S, Araki N, Ueda T. Systemic therapy for soft tissue sarcoma: proposals for the optimal use of pazopanib, trabectedin, and eribulin. Adv Ther 2017;34:1556–71.
[12] Beristain AG, Narala SR, Di Grappa MA, Khokha R. Homotypic RANK signaling differentially regulates proliferation, motility and cell survival in osteosarcoma and mammary epithelial cells. J Cell Sci 2012;125:943–55.
[13] Kolb EA, et al. Initial testing (stage 1) of glembatumumab vedotin (CDX-011) by the pediatric preclinical testing program. Pediatr Blood Cancer 2014;61:1816–21.
[14] Roth M, Barris DM, Piperdi S, et al. Targeting glycoprotein NMB with antibody-drug conjugate, glembatumumab vedotin, for the treatment of osteosarcoma. Pediatr Blood Cancer 2016;63:32–8.
[15] Roth M, Linkowski M, Tarim J, et al. Ganglioside GD2 as a therapeutic target for antibody-mediated therapy in patients with osteosarcoma. Cancer 2014;120:548–54.
[16] Poon VI, Roth M, Piperdi S, et al. Ganglioside GD2 expression is maintained upon recurrence in patients with osteosarcoma. Clin Sarcoma Res 2015;5:4–14.
[17] Jiang YY, Lin DC, Mayakonda A, et al. Targeting super-enhancer-associated oncogenes in oesophageal squamous cell carcinoma. Gut 2017;66:1358–68.
[18] Christensen CL, Kwiatkowski N, Abraham BJ, et al. Targeting transcriptional addictions in small cell lung cancer with a covalent CDK7 inhibitor. Cancer Cell 2014;26:909–22.
[19] Fisher RP. Secrets of a double agent: CDK7 in cell-cycle control and transcription. J Cell Sci 2005;118:5171–80.
[20] Drapkin R, Le Roy G, Cho H, Akoulitchev S, Reinberg D. Human cyclin-dependent kinase-activating kinase exists in three distinct complexes. Proc Natl Acad Sci - PNAS 1996;93:6488–93.
[21] Akhtar MS, Heidemann M, Tietjen JR, et al. TFIIH kinase places bivalent marks on the carboxy-terminal domain of RNA polymerase II. Mol Cell 2009;34:387–93.
[22] Zhang J, Liu W, Zou C, et al. Targeting super-enhancer-associated oncogenes in osteosarcoma with THZ2, a covalent CDK7 inhibitor. Clin Cancer Res 2020;26:2681–92.
[23] Barrett T, Wilhite SE, Ledoux P, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Res 2012;41:D991–5.

[24] Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
[25] Gatto L, Breckels LM, Naake T, Gibb S. Visualization of proteomics data using R and bioconductor. Proteomics 2015;15:1375–89.
[26] Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44–57.
[27] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.
[28] Bindea G, Mlecnik B, Hackl H, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009;25:1091–3.
[29] Shaikh AB, Li F, Li M, et al. Present advances and future perspectives of molecular targeted therapy for osteosarcoma. Int J Mol Sci 2016;17:506.
[30] Huang JR, Qin WM, Wang K, et al. Cyclin-dependent kinase 7 inhibitor THZ2 inhibits the growth of human gastric cancer *in vitro* and *in vivo*. Am J Transl Res 2018;10:3664–76.
[31] Li B, Ni Chonghaile T, Fan Y, et al. Therapeutic rationale to target highly expressed CDK7 conferring poor outcomes in triple-negative breast cancer. Cancer Res 2017;77:3834–45.
[32] Wang Y, Zhang T, Kwiatkowski N, et al. CDK7-dependent transcriptional addiction in triple-negative breast cancer. Cell 2015;163:174–86.
[33] Pelletier J, Thomas G, Volarevic S. Ribosome biogenesis in cancer: new players and therapeutic avenues. Nat Rev Cancer 2018;18:51–63.
[34] Buchhaupt M, Meyer B, Kötter P, et al. Genetic evidence for 18S rRNA binding and an Rps19p assembly function of yeast nucleolar protein Nep1p. Mol Genet Genom 2006;276:273–84.
[35] Armistead J, Hemming R, Patel N, et al. Mutation of EMG1 causing Bowen–Conradi syndrome results in reduced cell proliferation rates concomitant with G2/M arrest and 18S rRNA processing delay. BBA Clin 2014;1:33–43.
[36] Müeller-Pillasch F, Lacher U, Wallrapp C, et al. Cloning of a gene highly overexpressed in cancer coding for a novel KH-domain containing protein. Oncogene 1997;14:2729–33.
[37] Degrauwe N, Suvà ML, Janiszewska M, et al. IMPs: an RNA-binding protein family that provides a link between stem cell maintenance in normal development and cancer. Genes Dev 2016;30:2459–74.
[38] Lederer M, Bley N, Schleifer C, Hüttelmaier S. The role of the oncofetal IGF2 mRNA-binding protein 3 (IGF2BP3) in cancer. Semin Cancer Biol 2014;29:3–12.
[39] Chen P, Wang SJ, Wang HB, et al. The distribution of IGF2 and IMP3 in osteosarcoma and its relationship with angiogenesis. J Mol Histol 2012;43:63–70.
[40] Dragon F, Gallagher JE, Compagnone-Post PA, et al. A large nucleolar U3 ribonucleoprotein required for 18S ribosomal RNA biogenesis. Nature 2002;417:967–70.
[41] Mouillesseaux K, Chen J. Mutation in UTP15 disrupts vascular patterning in a p53-dependent manner in zebrafish embryos. PloS one 2011;6:e25013.
[42] Claudio JO, Liew CC, Ma J, et al. Cloning and expression analysis of a novel WD repeat gene, WDR3, mapping to 1p12-p13. Genomics 1999;59:85–9.
[43] McMahon M, Ayllón V, Panov KI, O'Connor R. Ribosomal 18S RNA processing by the IGF-I-responsive WDR3 protein is integrated with p53 function in cancer cell proliferation. J Biol Chem 2010;285:18309–18.
[44] Isakoff MS, Bielack SS, Meltzer P, Gorlick R. Osteosarcoma: current treatment and a collaborative pathway to success. J Clin Oncol 2015;33:3029–35.
[45] Bishop MW, Janeway KA, Gorlick R. Future directions in the treatment of osteosarcoma. Curr Opin Pediatr 2016;28:26–33.