

Identification of Biomarkers and Study of Mechanisms Related to Metastatic of Osteosarcoma Based on Integrated Bioinformatics Analyses

Guangzhi Wu

China-Japan Union Hospital of Jilin University

Minglei Zhang (✉ minglei111666@hotmail.com)

China-Japan Union Hospital of Jilin University <https://orcid.org/0000-0002-3881-9519>

Research article

Keywords: Osteosarcoma, metastatic, lncRNA, mRNA

Posted Date: December 31st, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-136142/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background

Osteosarcoma (OS) is a serious threat to public health. Because of high morbidity and fairly complicated pathogenesis. The study aim to identify candidate biomarkers and research the molecular mechanisms correlated of patients with metastatic OS.

Methods

The GSE21257 was downloaded from Gene Expression Omnibus(GEO) database, and the differentially expressed RNAs (DERs) were identified and functional enriched analysis by statistical soft-ware in R. Subsequently, the co-expression modules and its clinical characteristics of OS were identified by weighted gene co-expression network analysis (WGCNA) Following, the KEGG pathways directly related to metastatic OS was to researched by the Comparative Toxicogenomics Database 2019 update (CTD). Finally, the “survival” package in R was used to survival analysis and the DERs were verified using another independent profiling GSE14827.

Results

A total of 1,464 DERs were classified including 702 up-regulated and 762 down-regulated. In addition, a total of 1248 DERs were obtained by WGCNA analysis, the blue modules is the highest negative correlation ($P=0$) and the turquoise modules is highest positive correlation ($P=3E-196$) among all correlations with OS metastatic. The lncRNA-mRNA co-expression network including 4 lncRNAs and 507 mRNAs, and the cytokine-cytokine receptor interaction and JAK-STAT signaling pathway were found significantly correlation with metastatic. Finally, the increased expression levels of IFNGR1, lower DLEU1 and DLEU2 related to better prognosis. Which were significantly consistent in the another independent profiling GSE14827.

Conclusions

A bioinformatics analysis related to the IFNGR1, DLEU1 and DLEU2 may as candidate biomarkers for metastatic OS.

Background

Osteosarcoma (OS) is initiated in the metaphysis of the long bones predominantly and is very common non-haematological cancer in the bones (1). The metastases accounts major cause which lead to death of the patients with OS (2), and approximately 15–20% of patients show clinically detectable metastases(3). The metastasis of patients in the malignant progression of OS is the main pathological problem, which greatly hinders the effectiveness of clinical treatment of OS and brings adverse results to OS patients (4, 5). To sum up, it is urgently need to understanding the molecular mechanisms and identifying key molecules associated with metastasis of OS.

Long noncoding RNAs (lncRNAs) are recognized as non-protein-coding transcripts greater than 200 nucleotides in length (6). It has increasingly been suggested that the principal roles of lncRNAs function in various biological processes, including growth, invasion, metastasis, chemoresistance etc. (7, 8). Recently, considerable researches have shown that lncRNAs play a crucial role in OS. We found that lncRNA ITGB2-AS1 was up-regulated in OS and leads to poor prognosis (9); lncRNA MALAT1 promotes the up-regulation of RET and the activation of the Akt pathway

by competitively combines with miR-129-5p, which results in OS promotion of metastasis(10). Highly expressed lncRNA SNHG12 regulates Notch2 expression by inducing miR-195-5p, thereby activating Notch signaling pathway in OS and promoting tumorigenesis and metastasis in OS (11).

Messenger RNAs (mRNAs) are transcribed, spliced, and polyadenylated in the nucleus, then exported to the cytoplasm and to further form the cellular proteome (12). Related study shown that mRNAs inhibits OS metastasis by regulating related proteins. Such as miR-223-3p can inhibit cell invasion, migration, growth and proliferation in OS (13). PRDX1 was extracted and measured in OS patients, and found that the lowed expression of PRDX1 suppressed the progression and metastasis of OS cells(14).

Although previous preliminary analyzed of genome-wide expression profiling GSE21257 in OS and identified the importance of macrophages in controlling metastases. But the exact mechanism and antimetastatic functions are still unknown(15). Forethermore, the functions of numerous lncRNAs and mRNAs remain not elucidated, especially their clinical roles and molecular mechanisms in metastatic OS(16, 17). Therefore, it is necessary to conduct a comprehensive assessment of the functions of lncRNAs and mRNAs in OS.

In this study, bioinformatics analysis of gene expression profiles was used to further research the factors associated with the metastatic in OS. We have studied the potential molecular mechanism of metastatic in OS and screened the significant RNAs. It may be considered as the biomarkers for the treatment of OS patients with metastatic and provide a molecular basis for the early diagnosis, prognosis and treatment of metastatic OS.

Methods

Microarray data and data preprocessing

The GSE21257(18) was downloaded from the Gene Expression Omnibus (GEO, <https://www.ncbi.nlm.nih.gov/>) (19), the OS patients who developed metastases ($n = 34$) and not develop metastases ($n = 19$). Each of the dataset was analyzed by Illumina human-6 v2.0 expression BeadChip. The Limma package (Version 3.4.1, <https://bioconductor.org/packages/release/bioc/html/limma.html>) (20) in R software and quantile method were used for log2 conversion and the Microarray raw data were normalized. The lncRNA and mRNA in GSE21257 datasets were re-annotated using the information (Transcript

ID, Ref Seq ID and location) supported by the detection platform GPL10295 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GPL10295>) and the HUGO Gene Nomenclature Committee (HGNC) (<http://www.genenames.org/>) (19) which have the 4338 lncRNAs and 19218 mRNAs.

Screening of significant differentially expressed RNAs (DERs)

The Limma package (Version 3.4.1, <https://bioconductor.org/packages/release/bioc/html/limma.html>) (20) in R was used to performed DERs and calculated the FDR values and Fold Change values between metastasis and non-metastasis OS samples. The DERs was identified with FDR (false discovery rate) < 0.05 and $|\log_2$ fold change (FC)| > 0.263 were as the significance cut-off criteria. The hierarchical clustering analyses of DERs were performed using pheatmap package (Version 3.4.1, <https://cran.r-project.org/package=pheatmap>) (21) in R, and were presented by two-way hierarchical clustering heatmaps. A $P < 0.05$ as the significance cut-off criteria (22, 23).

Functional enrichment analyses

The online tool DAVID (Version 6.8; <https://david.ncifcrf.gov/>) (24, 25) was employed to identify the Gene Ontology (GO) biological processes and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways that significantly associated with the mRNAs in OS samples. A $P < 0.05$ was considered to achieve significant enrichment and have statistical significance.

Weighed gene co-expression network analysis (WGCNA)

To identify modules associated with OS and its clinical characteristics by using the WGCNA (Version1.61, <https://cran.r-project.org/web/packages/WGCNA/>) (26, 27) package in R (28, 29). In this study, all RNAs were detected and performed network construction and division. Based on the clinical information of OS samples to calculated the correlation between each partitioning module and each clinical indicator, with the minimal module size of 50 and the merge cut height of 0.99. Functional annotation was conducted for each stable module using the WGCNA package(30), with a Fold enrichment > 1. Additionally, DERs was performed for each module with a P-value and false discovery rate (FDR) of < 0.05. A cluster dendrogram among modules and a genes adjacency heatmap between modules were generated.

lncRNA-mRNA co-expression network construction

The co-expression analysis of lncRNAs and mRNAs was conducted based on the Pearson correlation coefficient (PCC) of the Cor function in the R language (Version 3.4.1, <http://77.66.12.57/R-help/cor.test.html>). Their expression levels was conduct a network visualization display via Cytoscape (Version 3.6.1, <http://www.cytoscape.org/>) (31). Then, Gene Set Enrichment Analysis (GSEA, <http://software.broadinstitute.org/gsea/index.jsp>) (32) in R was used to performe KEGG pathway enrichment analysis on optimal mRNAs in the lncRNA-mRNA network. The enrichment score (ES), normalised enrichment score (NES) and nominal P value were used in this analysis, and the NES absolute value increases and the P value decrease, suggesting a higher degree of enrichment and a higher significance of the result (33). A $P < 0.05$ was considered to screen KEGG pathways that were significantly enriched in the relevant mRNAs

OS relevant KEGG pathway network construction

The Comparative Toxicogenomics Database 2019 update (CTD, <http://ctd.mdibl.org/>)(34) was used to research the KEGG pathways directly related to OS with the keywords of “osteosarcoma”. Obtained the overlapping pathways by compared with the RNAs of significantly participated in pathways from the co-expression network, and to constructed a OS relevant pathway network.

Survival analysis and DERs verification.

The R package “survival” (<http://bioconductor.org/packages/survivalr/>) (35) was used to identify the prognosis-related lncRNAs and mRNAs by COX regression univariate analysis. The Kaplan-Meier survival analysis was used to study the associations between the expression levels and the survival prognosis. The survival curves of samples with low expression of the RNAs and high expression of the RNAs were compared using the log-rank test. $P < 0.05$ indicated statistically significant differences.

In addition, to verify the reliability of the important RNAs of the metastatic and non-metastatic OS samples. Another independent profiling of GSE14827 containing 27 OS samples (18 non-metastases and 9 metastases) was downloaded from GEO (<https://www.ncbi.nlm.nih.gov/>) using the information supported by the platform GPL570 [HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array.

Results

Data preprocessing and DERs screening

After data preprocessing, a total of 13758 mRNAs and 58 lncRNAs were detected. We analyzed the DERs of GSE21257 and obtained 1,464 DERs, including 702 up-regulated and 762 down-regulated in OS samples who metastasis ($n = 34$) compared with non-metastasis ($n = 19$) when $p < 0.05$ and $|\log_2FC| > 0.263$ as the cutoff criteria. We identified all the DERs which were shown in the volcano map according to the value of $|\log_2FC|$ and displayed the DERs on a heatmap (Figure.1A). The expression values of differentially expressed lncRNAs and mRNAs were two-way hierarchically clustered, and the color contrast indicated that there was significantly difference in the expression levels between the non-metastatic and metastatic OS samples (Figure.1B).

GO and KEGG pathway enrichment analyses

A total of 16 significant related GO biological processes and 12 KEGG signaling pathways with $P < 0.001$ were identified for OS patients (Table 1). We found that immune response (GO:0006955, $P = 2.33E-16$), inflammatory response (GO:0006954, $P = 2.02E-12$) and interferon gamma mediated signaling pathway

(GO:0060333, $P = 1.58E-11$) were the three most significant pathways in 16 significant related GO biological processes. Meanwhile, cell adhesion molecules (hsa04514, $P = 3.76E-09$), antigen processing and presentation (hsa04612, $P = 1.13E-08$), lysosome (hsa04142, $P = 2.61E-06$) were the three most significant pathways in 12 KEGG signaling pathways were obtained.

Table 1

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis for the differentially expressed RNAs (DERs).

Category	Term	Count	P-Value	FDR	Gene
Biology Process	GO:0006955 ~ immune response	88	2.33E-16	4.22E-13	IL27RA, AQP9, CD8A, TLR4, IL15, TNFSF12, HLA-DMB, HLA-DMA, C1QC, CXCL12, CXCL10, MBP, CD96, HAMP, IL1B, SEMA3C, LTB, CIITA, C5AR1, CMKLR1.....
	GO:0006954 ~ inflammatory response	74	2.02E-12	3.81E-09	S100A8, AIF1, LY86, S100A9, PRDX5, TLR4, IL15, TLR5, CXCR3, CXCL12, TLR7, TLR8, CXCL10, MYD88, CXCR4, AOAH, CXCR6, IL1B, ADAM8, SYK.....
	GO:0060333 ~ interferon-gamma-mediated signaling pathway	27	1.58E-11	2.99E-08	HLA-DQB1, HLA-DQB2, IFI30, OAS1, CD44, FCGR1A, HLA-DPB1, IFNGR1, CIITA, HCK,HLA-A, HLA-B, HLA-E, STAT1, TRIM22, PRKCD, TRIM21, HLA-DQA1, HLA-F, NCAM1.....
	GO:0008283 ~ cell proliferation	68	1.69E-10	3.18E-07	PTGES3, MORF4L1, STIL, COPS2, USPL1, LY86, ENPEP, CDC16, MCM10, FES, GLI2, PTEN, LARP1, DAB2, SBDS, PICALM, RRN3, CREG1, YAP1, MYC.....
	GO:0006935 ~ chemotaxis	30	9.51E-08	1.79E-04	C3AR1, CCL2, CYSLTR1, CXCL9, FPR1, CCL8, CXCR3, FES, CCL5, CXCL12, CXCL10, DOCK2, CCL23, CCL20, CXCR4, CXCR6, RALA, RNASE2, C5AR1, CMKLR1.....
	GO:0006915 ~ apoptotic process	81	6.20E-07	1.17E-03	ITGB3BP, IER3, S100A8, LY86, SNCA, S100A9, PRDX5, GJA1, RASSF7, CXCR3, TNFSF12, PTEN, SMNDC1, PDCD2, CASP6, DAB2, GHITM, MYD88, CXCR4, MAP3K8.....
	GO:0019882 ~ antigen processing and presentation	18	8.64E-07	1.63E-03	HLA-DQB1, CD8A, RAB5B, HLA-A, FCGRT, HLA-B, CTSS, HLA-DMB, HLA-E, HLA-DQA1, CD74, CTSL, CD209, ULBP2, HLA-DPA1, HLA-DPB1, CTSH, HLA-DRA
	GO:0034341 ~ response to interferon-gamma	12	8.64E-07	1.63E-03	CIITA, SLC11A1, KYNU, CD86, BST2, IFITM1, IFITM2, IFITM3, CXCL16, SNCA, UBD, TRIM21
	GO:0050900 ~ leukocyte migration	27	3.93E-06	7.41E-03	GLG1, ITGAL, C3AR1, ATP1B3, GRB2, FPR1, ITGB2, ITGAM, CD74, SLC7A7, CD48, CD44, PROCR, CD2, PIK3CA, FCER1G, INPP5D, YES1, PTPN6, C5AR1.....
	GO:0071222 ~ cellular response to lipopolysaccharide	25	9.71E-06	1.83E-02	HAVCR2, IL6, CCL2, TNF, TNFSF4, CEBPB, LITAF, KLRK1, AXL, TLR4, TLR5, ABCA1, CXCL10, EDNRB, CD86, TNFRSF1B, CCL20, HAMP, CXCL16, SERPINE1...
	GO:0031295 ~ T cell costimulation	20	1.03E-05	1.95E-02	HLA-DQB1, PTPN6, HLA-DQB2, GRB2, EFNB2, CTLA4, KLRK1, TNFRSF14, VAV1, HLA-DQA1, PDCD1LG2, CD86, TNFSF13B, ICOS, MAP3K8, PIK3CA, HLA-DPA1, HLA-DPB1, YES1, HLA-DRA
	GO:0071407 ~ cellular response to organic cyclic compound	17	1.19E-05	2.25E-02	CEBPA, TNF, SMAD9, MSR1, CEBPB, CCL2, SMAD5, HCFC1, STAT1, CCL5, GLI2, CCNB1, P2RY13, CYBA, P2RY6, IL1B, RAP1B
	GO:0006952 ~ defense response	18	1.68E-05	3.17E-02	CYSLTR1, RNASE6, CXCL9, IL32, HLA-B, PTPRCAP, CXCL12, WAS, CD74, SP140, CD48, KCNN4, LILRA2, TIAL1, LILRB3, TAP1, TFF3, MX1
	GO:0009615 ~ response to virus	24	1.93E-05	3.64E-02	TNF, TNFSF4, IFITM1, FGR, BST2, IFITM2, IFITM3, CCL8, OAS1, PIM2, CCL5, TRIM22, CXCL12, TLR8, NPC2, IRAK3, MYD88, CXCR4, PRKRA, TPT1.....
	GO:0045429 ~ positive regulation of nitric oxide biosynthetic process	14	2.19E-05	4.12E-02	P2RX4, IL6, TNF, AIF1, KLRK1, PKD2, SMAD3, IL1B, TLR4, ITGB2, TLR5, DDAH2, KLF4, SOD2
	GO:0031663 ~ lipopolysaccharide-mediated signaling pathway	12	2.44E-05	4.60E-02	TNF, MTDH, CCL2, MYD88, LY96, HCK, IL1B, TLR4, CCL5, CD6, CD14, PTAFR
KEGG Pathway	hsa04514:Cell adhesion molecules (CAMs)	40	3.76E-09	2.13E-07	HLA-DQB1, GLG1, CLDN7, ITGAL, CD8A, ITGB2, CDH2, NEO1, SDC2, HLA-DMB, SDC4, HLA-DMA, ICOS, CLDN14, ITGAM, ALCAM, ITGB7, CD2, HLA-DPB1, HLA-DOA.....
	hsa04612:Antigen processing and presentation	27	1.13E-08	5.34E-07	HLA-DQB1, TNF, CD8A, LGMN, IFI30, HSPA1A, HLA-DMB, HLA-DMA, CANX, CD74, TAP1, HSPA4, HLA-DPB1, HLA-DOA, HLA-DOB, CIITA, HLA-A, CTSS, HLA-B, HLA-E.....
	hsa04142:Lysosome	31	2.61E-06	5.27E-05	LITAF, LGMN, HEXB, CLTC, ASAH1, SLC11A1, NAGPA, AP1S2, LAPTM5, GNPTAB, IDS, TPP1, GALNS, GALC, ATP6V0D1, CTSZ, LAPTM4A, LIPA, CTSS, MANBA.....
	hsa04110:Cell cycle	28	9.77E-05	1.45E-03	E2F2, E2F3, PKMYT1, TTK, ANAPC11, CDC16, RBX1, BUB1, CCNA2, MYC, BUB3, STAG1, CDC7, ANAPC4, SKP2, SMAD3, RB1, MCM3, ATM, CDC25B, CCNB1.....
	hsa04380:Osteoclast differentiation	29	1.03E-04	1.38E-03	IL1R1, TNF, GRB2, SPI1, FHL2, BTK, LILRA2, FCGR1A, PPP3CB, PIK3CA, IL1B, IFNGR1, TYROBP, SYK, CSF1R, NFATC1, BLNK, NCF1, NCF4, PIK3CD.....
	hsa04062:Chemokine signaling pathway	36	2.25E-04	2.65E-03	CCL2, GNAI3, FGR, GRB2, GNAI1, CXCL9, CCL8, CXCR3, GNG12, CCL5, CXCL12, CXCL10, DOCK2, CCL23, CCL20, CXCR4, CXCR6, PIK3CA, PLCB1, PLCB2.....

Category	Term	Count	P-Value	FDR	Gene
	hsa04610:Complement and coagulation cascades	18	4.18E-04	4.72E-03	C3AR1, A2M, C5AR1, F13A1, BDKRB1, SERPING1, C1QC, PLAUR, C1QA, C1QB, CD46, SERPINE1, TFPI, SERPINA1, C2, CFI, CFD, PLAUI
	hsa04670:Leukocyte transendothelial migration	25	4.59E-04	4.99E-03	ACTB, ITGAL, CLDN7, GNAI3, ROCK2, NCF1, GNAI1, NCF4, SIPA1, PIK3CD, CD99, ITGB2, VAV1, CXCL12, ITGAM, CLDN14, CLDN23, ACTG1, CYBA, CXCR4.....
	hsa04060:Cytokine-cytokine receptor interaction	41	1.40E-03	1.36E-02	IL1R1, CCL2, TNF, BMPR2, CXCL9, CCL8, CXCR3, IL15, TNFSF12, CCL5, CXCL12, IL11, CXCL10, TNFRSF1B, CCL23, CCL20, CXCR4, IL10RA, CXCR6, IL1B.....
	hsa04640:Hematopoietic cell lineage	19	2.52E-03	2.21E-02	IL1R1, IL6, TNF, IL2RA, CD8A, ITGAM, IL11, CD9, CD38, CD37, CD44, FCGR1A, CD33, CD2, IL1B, CD14, CSF2RA, HLA-DRA, CSF1R
	hsa04611:Platelet activation	25	2.74E-03	2.33E-02	ACTB, GNAI3, TBXAS1, ROCK2, GNAI1, PIK3CD, PRKCI, PPP1CC, APBB1IP, ITPR1, BTK, ACTG1, VAMP8, MAPK13, PPP1R12A, FCER1G, PIK3CA, RAP1B, FCGR2A, SNAP23.....
	hsa04064:NF-kappa B signaling pathway	18	6.06E-03	4.66E-02	IL1R1, TNF, LY96, TLR4, BCL2L1, CXCL12, ATM, BTK, CCL13, MYD88, TNFSF13B, IL1B, TNFAIP3, LTB, CD14, PLAUI, SYK, BLNK

WGCNA

WGCNA was performed to identify metastasis OS associated modules and genes on the obtained all DERs. Firstly, we explored the value of power ranging from 1 to 30. Following definition of the adjacent function, the correlation coefficient reached 0.9 for the first time, which the power value of 6 was selected. Under a power value of 6, the mean connectivity degree of the RNAs was 1, which conformed to the small world property in a scale-free network (Figure.2A). A total of 10 gene modules were identified (Figure.2B), and the relationship between the gene modules and the clinical features (age, gender, histological, grade, metastases) was analyzed (Figure. 2C). The results shown that the stabilities of the 7 modules were assessed ($P<0.05$) (Table 2) and a total of 1248 DERs are distributed and displayed among the various modules with the black, blue, brown, green, grey, pink, red, turquoise, yellow (Figure.2D). Based on the clinical information in OS, the correlation between each module and the clinical factors was analyzed. Among the 7 stable modules, the blue and turquoise modules correlated significantly with metastatic (Figure.2C), the blue modules is the highest negative correlation ($P=0$) and the turquoise modules is highest positive correlation ($P=3E-196$) among all correlations with metastatic. Functional annotation for each stable modules revealed that the lncRNAs in the blue (1 lncRNAs and 495 mRNAs) (Enrichment fold = 3.201) and turquoise (5 lncRNAs and 538 mRNAs) (Enrichment fold = 1.626) modules were predominantly enriched in metastasis. The heatmap of expression changes of RNAs was shown (Figure.2E and 2F). The RNAs of blue modules were down-regulated significantly in metastatic OS samples, and the RNAs of turquoise modules were up-regulated significantly in metastatic OS samples.

Table 2

Differentially expressed RNAs (DERs) in modules.

Color	Total RNA	DERs(lncRNA)	Enrichment fold[95%CI]	Phyper
black	66	9	0.699[0.305–1.414]	4.304E-01
blue	751	496(1)	3.201[2.804–3.654]	2.200E-16
brown	677	16	0.121[0.0686–0.199]	2.200E-16
green	321	21	0.335[0.204–0.524]	3.630E-08
grey	2016	105	0.267[0.215–0.328]	2.200E-16
magenta	59	-	-	-
pink	62	11	0.909[0.431–1.748]	8.740E-01
red	68	1	0.0754[0.00188-0.435]	1.095E-04
turquoise	1712	543(5)	1.626[1.448–1.825]	2.741E-16
yellow	667	46	0.354[0.255–0.480]	2.802E-14

LncRNA-mRNA co-expression network construction

A total of 4 lncRNA, 507 mRNA, 1001 connection sides and 511 nodes were obtained in the established co-expression network by using OS associated gene expression value from turquoise and blue panel (Figure. 3). The GSEA analyses for the mRNAs regulatory network revealed that 9 KEGG pathways (antigen processing and presentation, cell adhesion molecules (CAMs), cytokine-cytokine receptor interaction, natural killer cell mediated cytotoxicity, ribosome, complement and coagulation cascades, chemokine signaling pathway, FC-epsilon-RI signaling pathway and JAK-STAT signaling pathway) were obviously related with the metastatic OS (Table 3).

Table 3
Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway that significantly correlate with mRNAs in lncRNA-mRNA co-expression

Name	Count	ES	NES	Np-val	FDR q-val	Gene
KEGG_ANTIGEN_PROCESSING_AND_PRESENTATION	12	-0.7075	-2.7271	0	0	CTSS,TAP1,CD74,HLA-F,HLA-B,HLA-I DMB,HLA-DQA1,HLA-DRA
KEGG_CELL_ADHESION_MOLECULES_CAMS	13	-0.8004	-2.8820	0	0	ITGAL,HLA-A,PECAM1,CD2,HLA-F,HL DMA,HLA-DMB,HLA-DQA1,HLA-DRA
*KEGG_CYTOKINE_CYTOKINE_RECEPTOR_INTERACTION	8	-0.7214	-2.5474	0	0	IFNGR1,CSF2RA,IL10RA,TNFSF13B,
KEGG_NATURAL_KILLER_CELL_MEDIATED_CYTOTOXICITY	10	-0.7309	-2.6253	0	0	ITGAL,HLA-A,VAV1,FCER1G,TNF,LCP
KEGG_RIBOSOME	11	0.8440	2.7165	0	0	RPLP1,RPL10A,RPL7,RPL7A,RPL23,
KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	4	-0.8603	-2.1242	0	0.001493	C2,C1QB,C1QA,C1QC
KEGG_CHEMOKINE_SIGNALING_PATHWAY	10	-0.5484	-2.1075	0	0.001823	VAV1,NCF1,HCK,DOCK2,WAS,CXCR3
KEGG_FC_EPSILON_RI_SIGNALING_PATHWAY	6	-0.6587	-1.8502	0.011682	0.016056	VAV1,FCER1G,TNF,LCP2,BTK,INPP5C
*KEGG_JAK_STAT_SIGNALING_PATHWAY	5	-0.6813	-1.7668	0.0133	0.02635	STAT1,IFNGR1,STAT5A,CSF2RA,IL10
The "*" and bold labeled pathway entries are pathways that overlap the osteosarcoma-related pathways in the Comparative Toxicogenomics Database 2019 enrichment score; NES:normalised enrichment score; Np-value: Nominal p-value						

OS relevant KEGG pathway network construction

A total 127 KEGG pathways were screened to be related with OS by searching CTD database. Cytokine-cytokine receptor interaction and JAK-STAT signaling pathway were the overlapping pathways by compared with enriched 9 KEGG pathways. Moreover, a total of 3 lncRNA (DLEU1, DLEU2 and HCP5) and 10 mRNAs (IFNGR1, STAT5A, STAT1, CAF2RA, IL10RA, CCL5, CCL8, TNFSF1B, TNFSF13B and CXCL10) which had intersections with the overlapping pathways were gained to construct random lncRNA-mRNA pairs (Fig. 4).

Survival analysis and DERs verification

According to the expression levels of 13 RNAs, the OS samples were classified into high-expression and low-expression group. The Kaplan-Meier survival analysis showed that samples with high expression levels of IFNGR1 had better prognosis ($P = 3.823e-02$), while reduced the expression levels of DLEU1 ($p = 4.345e-02$) and DLEU2 ($P = 2.933e-02$) related to better prognosis (Fig. 5).

The expression level of IFNGR1, DLEU1 and DLEU2 of the GSE21257 was significantly consistent in the metastatic and non-metastatic OS samples of another independent profiling GSE14827 (Fig. 6).

Discussion

OS is a highly aggressive bone tumor, which early systemic metastasis in young people, and results in poor survival for patients with OS (36). Therefore, it is very important to develop new therapeutic strategies and to treat the metastatic OS patients. However, the molecular mechanisms underlying the metastasis of OS are still very urgently to understand. This study aimed to investigate the metastatic processes of OS and to identify potential biomarkers of metastatic OS.

The gene expression profile of GSE21257 using bioinformatics analysis and shown that there were 762 up-regulated and 702 down-regulated DERs which may be potential biomarkers of metastatic OS. Furthermore, The results of GO term and KEGG enrichment analysis were demonstrated that the DERs previously been associated with the cell adhesion molecules (CAMs), antigen processing and presentation, lysosome (2). The functional enrichment analysis suggested that aberrant regulation of RNAs may contribute to metastasis of OS. The DERs identified in this study may be associated with metastatic processes of OS. They can be used as a criterion for malignancy phenotype of bone cancer and prediction of metastatic potentiality of OS (37). Furthermore, the WCGNA results showed the blue module (1 lncRNAs and 495 mRNAs) and the turquoise module (5 lncRNAs and 529 mRNAs) have significantly relevance with metastatic OS. The highly stable blue and turquoise modules were mainly involved in cellular immune responses, cell adhesion and the cell cycle, and indicating their likely association with metastasis OS pathogenesis. In addition, 13 RNAs with cytokine-cytokine receptor interaction and JAK-STAT signaling pathway RNAs-pathway co-expression network was constructed. The Kaplan-Meier curve analysis was performed to explore the associations between overall survival and these 13 essential RNAs in patients with OS. Consequently, IFNGR1, DLEU1 and DLEU2 were found to be significantly associated with metastatic OS ($P < 0.05$). The three prognostic RNAs have all been reported to be associated with tumor, suggesting the reliability of the methods we used in this study.

To our surprise, IFNGR1 expression level was increased in metastasis OS group (Figure. 5). Since IFNGR1 have been identified in 1996, IFNGR1 plays a important role in the immune response against mycobacteria (38). IFNGR1 was also found to be mainly mutated in non-small-cell lung cancer and found that interferon associated IL2-STAT5 pathways were significantly enriched (39). Importantly, IFNGR1 plays an important role in the cytokine-cytokine receptor interaction and JAK-STAT signaling pathway simultaneously, which are reportedly involved in up-regulated to against OS. So it may be involved immune response for OS and as a protective factor in the metastasis of OS (40).

lncRNA DLEU1 was acted as an oncogene in various types of cancer. DLEU1 was overexpressed and promoted the cell proliferation, migration and invasion in OS cells. Moreover, DLEU1 interacts directly directly with miR-671-5p and regulating the expression of DDX5 in OS (41). Similarity, in our study showed that

DLEU1 was up-regulated in samples with metastasis OS. Thus, the OS genes may identified as significantly gene in metastasis OS(42). DLEU2 is the host gene of miR-15a, the study was indicated the transcription of DLEU2 was repressed by hypoxia. These findings indicated that miR-15a may be a valuable target for the treatment of OS(43), suggesting the carcinogenic role of DLEU2 in OS. DLEU1 and DLEU2 elicited OS cell metastasis by inhibiting the activation of the JAK-STAT signaling pathway, thus, they are considered as potential target genes for cancer treatment.

Conclusion

In conclusion, the increased expression levels of IFNGR1 may be as a protective factor in metastatic OS. Meanwhile, high DLEU1 and DLEU2 expression predicts poor overall survival of OS patients with metastasis, and they were significantly consistent in the metastatic and non-metastatic OS samples of another independent profiling GSE14827. The IFNGR1, DLEU1 and DLEU2 all played important roles in OS with metastasis by influencing cytokine-cytokine receptor interaction and JAK-STAT signaling pathway. Therefore, our study revealed that the IFNGR1, DLEU1 and DLEU2 might provide a potential new therapeutic strategy for metastatic OS treatment.

Abbreviations

Osteosarcoma (OS)

Gene Expression Omnibus (GEO)

Differentially expressed RNAs(DERs)

weighted gene co-expression network analysis (WGCNA)

Comparative Toxicogenomics Database (CTD)

Long noncoding RNAs (lncRNAs)

Messenger RNAs (mRNAs)

fold change (FC)

Kyoto Encyclopedia of Genes and Genomes (KEGG)

false discovery rate (FDR)

cell adhesion molecules(CAMs)

Declarations

Ethics approval and consent to participate: Not applicable

Consent for publication: Not applicable

Availability of data and material

The raw data were collected and analyzed by the Authors, and are not ready to share their data because the data have not been published.

Competing interests: The authors declare no conflict of interest.

Funding: none

Authors' contributions

GZW and MLZ participated in the design of this study, and they both performed the statistical analysis. GZW carried out the study and collected important background information. MLZ drafted the manuscript. All authors read and approved the final manuscript.

Acknowledgements: none

References

1. Zhang J, Lan Q, Lin JJWJoSO. Identification of key gene modules for human osteosarcoma by co-expression analysis. 2018;16(1):89.
2. Wang S, Zheng S, Hu K, Sun H, Zhang J, Rong G, et al. A predictive model to estimate the pretest probability of metastasis in patients with osteosarcoma. 2017;96(3):e5909.
3. S F, S S, M M, F B, A L, P R, et al. Neoadjuvant chemotherapy with high-dose Ifosfamide, high-dose methotrexate, cisplatin, and doxorubicin for patients with localized osteosarcoma of the extremity: a joint study by the Italian and Scandinavian Sarcoma Groups. 2005;23(34):8845.
4. Ragland BD, Bell WC, Lopez RR, Siegal GP. Cytogenetics and Molecular Biology of Osteosarcoma. Laboratory Investigation. 2002;82(4):365-73.
5. Wang H, Zeng X, Oliver P, Le LP, Chen J, Chen L, et al. MDM2 oncogene as a target for cancer therapy: An antisense approach. 1999;15(4):653.

6. Andreia S, Marc B, George CJC. The Clinical Relevance of Long Non-Coding RNAs in Cancer. 2015;7(4):2169-82.
7. Liz J, Esteller MJBBA. lncRNAs and microRNAs with a role in cancer development. 2016;1859(1):169-76.
8. Long-Fei, Rong, Chuan-Bin, Xiao-Gang, Dong, Wang, et al. Screening of lymph nodes metastasis associated lncRNAs in colorectal cancer patients. 2014;;20(25):8139-50.
9. Dai J, Xu LJ, Han GD, Jiang HT, Sun HLJAC, Nanomedicine, Biotechnology. Down-regulation of long non-coding RNA ITGB2- AS1 inhibits osteosarcoma proliferation and metastasis by repressing Wnt/ β -catenin signalling and predicts favourable prognosis. 2018:1-8.
10. Ji Q, Zhang L, Liu X, Zhou L, Wang W, Han Z, et al. Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ[[sol]]PTBP2 complex. 2014;111(4):736.
11. Zhou S, Yu L, Xiong M, Dai GJB, Communications BR. LncRNA SNHG12 promotes tumorigenesis and metastasis in osteosarcoma by upregulating Notch2 by sponging miR-195-5p. 2018;495(2):1822.
12. Bicknell AA, Ricci EPJBST. When mRNA translation meets decay. 2017;45(2):339.
13. Ji Q, Xu X, Song Q, Xu Y, Tai Y, Goodman SB, et al. miR-223-3p inhibits human osteosarcoma metastasis and progression by directly targeting CDH6. 2018;26(5).
14. Wang Y, Liu M, Yang P, Peng HJMSMIMJoE, Research C. Peroxiredoxin 1 (PRDX1) Suppresses Progressions and Metastasis of Osteosarcoma and Fibrosarcoma of Bone. 2018;24:4113-20.
15. Buddingh EP, Kuijjer ML, Duim RAJ, Horst B, Konstantin A, Ola M, et al. Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. 2011;17(8):2110-9.
16. Lin X, Yao Z, Zhang Y, Li D, Hu F, Liao Y, et al. Deep RNA sequencing reveals the dynamic regulation of miRNA, lncRNAs, and mRNAs in osteosarcoma tumorigenesis and pulmonary metastasis. 2018;9(7):772-.
17. Xie J, Lin D, Lee DH, Akunowicz J, Hansen M, Miller C, et al. Copy number analysis identifies tumor suppressive lncRNAs in human osteosarcoma. 2017;50(3).
18. Buddingh EP, Kuijjer ML, Duim RA, Bürger H, Agelopoulos K, Myklebost O, et al. Tumor-infiltrating macrophages are associated with metastasis suppression in high-grade osteosarcoma: a rationale for treatment with macrophage activating agents. 2011;17(8):2110-9.
19. Tanya B, Troup DB, Wilhite SE, Pierre L, Carlos E, Kim IF, et al. NCBI GEO: archive for functional genomics data sets--10 years on. 2013;39(Database issue):1005-10.
20. Grigoriev IV, Nordberg H, Shabalov I, Aerts A, Cantor M, Goodstein D, et al. The genome portal of the Department of Energy Joint Genome Institute. Nucleic Acids Res 40:D26-D32. 2014;42(Database issue):26-31.
21. Wang L, Cao C, Ma Q, Zeng Q, Wang H, Cheng Z, et al. RNA-seq analyses of multiple meristems of soybean: novel and alternative transcripts, evolutionary and functional implications. 2014;14(1):169.
22. Szekely GJ, Rizzo MLJJoC. Hierarchical Clustering via Joint Between-Within Distances: Extending Ward's Minimum Variance Method. 2005;22(2):151-83.
23. Kandemir-Cavas C, Nasibov E, editors. Alternative hierarchical clustering approach in construction of phylogenetic trees. Biomedical Engineering Meeting; 2009.
24. Da WH, Sherman BT, Lempicki RAJNP. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. 2009;4(1):44.
25. Wei HD, Sherman BT, Lempicki RA, %J Nucleic Acids Research. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. 2009;37(1):1.
26. Langfelder P, Horvath SJB. WGCNA: an R package for weighted correlation network analysis. 2008;9(1):559.
27. Cao J, Zhang SJB. A Bayesian Extension of the Hypergeometric Test for Functional Enrichment Analysis. 2014;70(1):84-94.
28. Oldham M, Konopka G, K, Langfelder P, Kato T, Horvath S, Geschwind DJNN. Functional organization of the transcriptome in human brain. 2008;11(11):1271.
29. Qi L, Changning L, Xiongying Y, Shuli K, Ruoyu M, Hui X, et al. Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. 2011;39(9):3864-78.
30. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics. 2008;9:559.
31. Shannon, Research PJG. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. 2003;13(11):2498.
32. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. 2005;102(43):15545-50.
33. Zhao Q, Fan CJBMG. A novel risk score system for assessment of ovarian cancer based on co-expression network analysis and expression level of five lncRNAs. 2019;20(1).
34. Davis AP, Grondin CJ, Johnson RJ, Sciaky D, McMorran R, Wiegiers J, et al. The comparative toxicogenomics database: update 2019. 2018;47(D1):D948-D54.
35. Wang P, Wang Y, Hang B, Zou X, Mao J-HJO. A novel gene expression-based prognostic scoring system to predict survival in gastric cancer. 2016;7(34):55343.
36. Wang JS, Wang YG, Zhong YS, Li XD, Du SX, Xie P, et al. Identification of co-expression modules and pathways correlated with osteosarcoma and its metastasis. 2019;17(1):46.
37. You MR, Liu XP, Xiu-Chun YUJJoER. Expression of Cell Adhesion Molecules in Osteosarcoma and Their Relationship with the Invasion and Metastasis of Sarcoma. 2004.

38. Boudjemaa S, Dainese L, Héritier S, Masserot C, Hachemane S, Casanova JL, et al. Disseminated BCG osteomyelitis related to STAT 1 gene deficiency mimicking a metastatic neuroblastoma. 2016.
39. Chen H, Chong W, Teng C, Yao Y, Wang X, Li X. The immune response-related mutational signatures and driver genes in non-small-cell lung cancer. *Cancer science*. 2019;110(8):2348-56.
40. Nava P, Koch S, Laukoetter MGL, Lee WY, Kolegraff K, Capaldo CT, Beeman N, et al. Interferon-gamma regulates intestinal epithelial homeostasis through converging beta-catenin signaling pathways. 2010;32(3):392-402.
41. Chen X, Zhang C, Wang X. Long noncoding RNA DLEU1 aggravates osteosarcoma carcinogenesis via regulating the miR-671-5p/DDX5 axis. *Artificial cells, nanomedicine, and biotechnology*. 2019;47(1):3322-8.
42. Guo Y, Bai M, Lin L, Huang J, An Y, Liang L, et al. LncRNA DLEU2 aggravates the progression of hepatocellular carcinoma through binding to EZH2. *Biomedicine & Pharmacotherapy*. 2019;118:109272.
43. Leng J, Song Q, Zhao Y, Wang Z, Ji J, O. miR-15a represses cancer cell migration and invasion under conditions of hypoxia by targeting and downregulating Bcl-2 expression in human osteosarcoma cells. 2018;52(4).

Figures

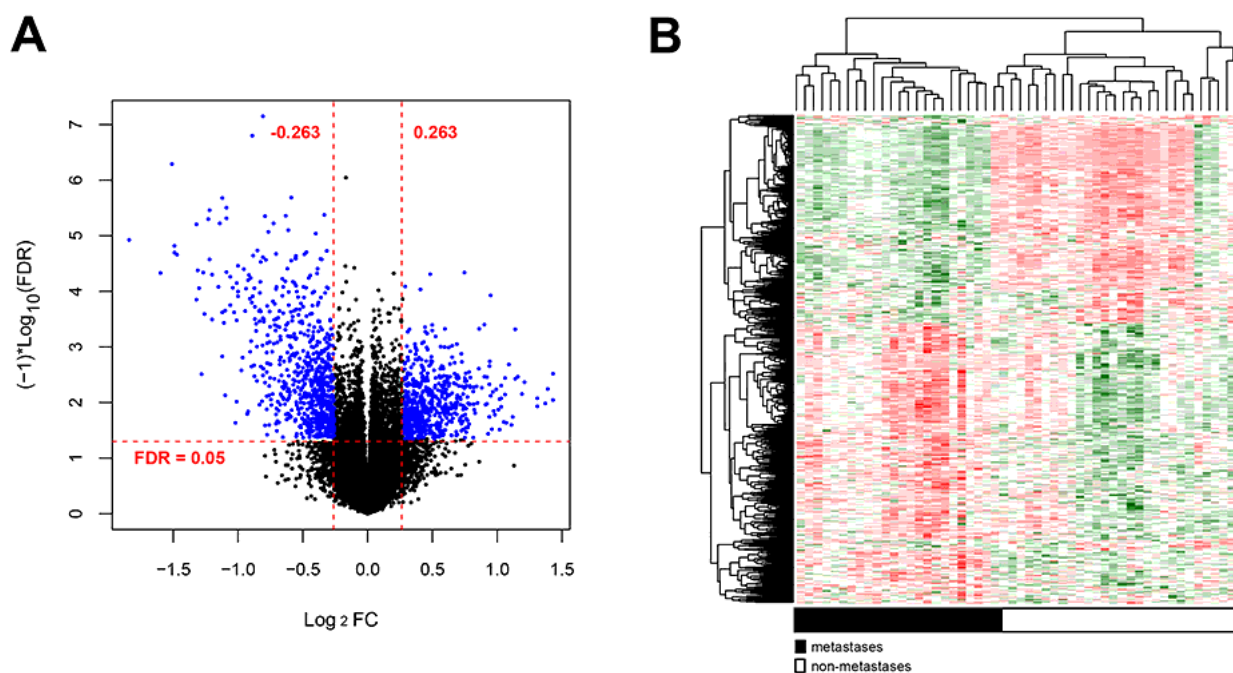


Figure 1

The hierarchically clustering analysis of screened DEGs. A: log2FC-log10 (FDR) volcano map for GSE21257 using the significantly DEGs. Blue dots indicate significant DEGs; The red horizontal dashed line indicates FDR<0.05; Two red vertical lines indicate |Log2FC|>0.263. B: Two-way hierarchically clustered heatmap for GSE21257 using the DEGs. Red up-regulated DEGs; green down-regulated DEGs.

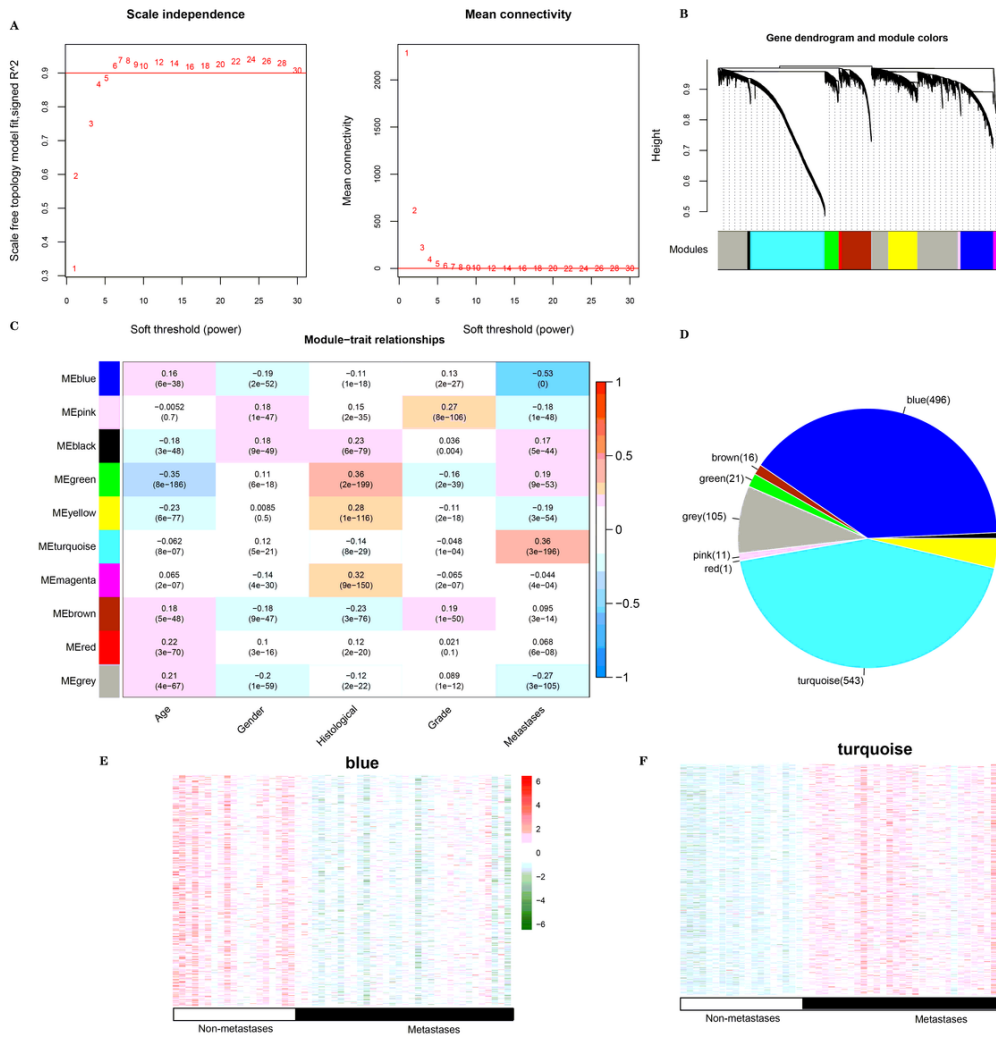


Figure 2

Weighed gene co-expression network analysis (WGCNA) was used to identify modules associated with OS and its clinical characteristics. A: Analysis of network topology for various soft-thresholding powers. (left) Selection diagram of the adjacency matrix weighting parameter power. The horizontal and vertical axes represent the weight parameter power and the square values (R^2) of correlation coefficients between $\log(k)$ and $\log p(k)$ in the corresponding network, respectively. The red line is the standard when the square of the correlation coefficient reaches 0.9; (right) The mean connectivity of RNAs under different values of power. The red line represents the mean connectivity of 1 when the power is 6; B: Module partition trees of datasets. Modules are indicated by different colors; C: Correlation heatmap between each module and clinical factors; the horizontal and vertical axes separately represent clinical factors and modules; the change of color from green to red indicates the change of correlation from negative to positive; The numbers in grids represent correlation coefficients, and the numbers in parentheses represent significant P-values. ME, module eigengene; D: The pie chart of significantly DEGs distribution in modules results: Each part represents the number refers to the number of DEGs; E: Heatmap of DEGs expression levels in blue and turquoise modules. Red up-regulated DEGs; green down-regulated DEGs

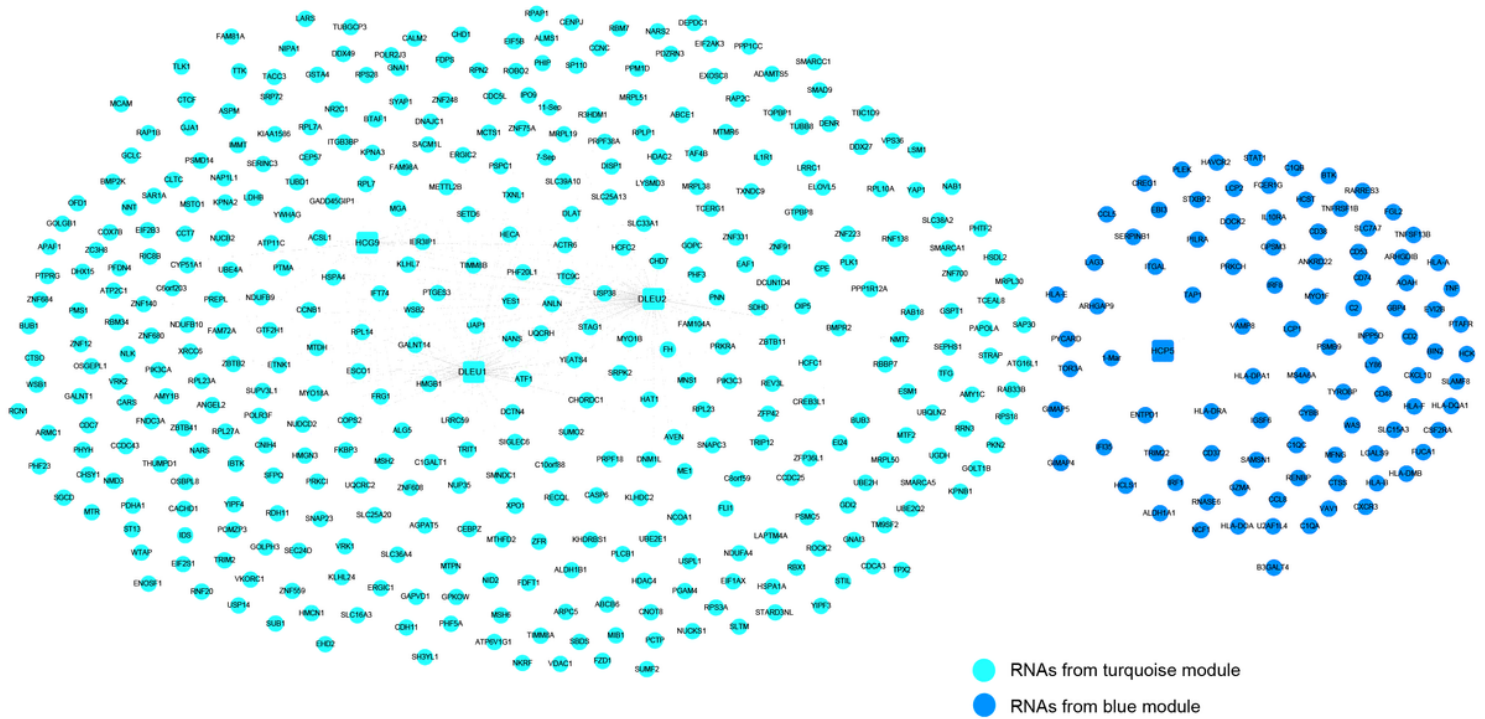


Figure 3

The heatmap of lncRNA-mRNA co-expression network of based on the turquoise and blue modules. The circles represent lncRNAs, square represent mRNAs respectively. The color of the node represents the color from the modules.

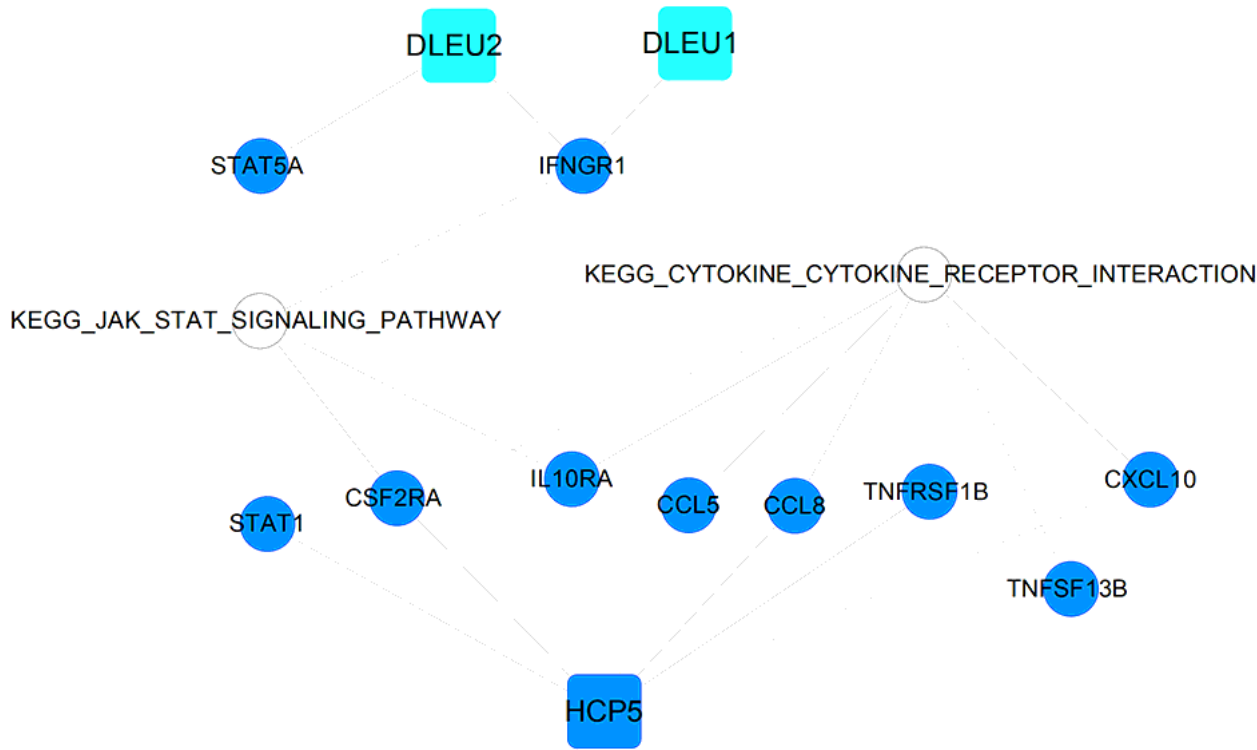


Figure 4

OS relevant KEGG pathway network. The square represent lncRNAs, circles represent mRNAs respectively. The color of the node represents the color from the turquoise and blue modules, and the white circular node represents the KEGG signaling pathway associated with OS.

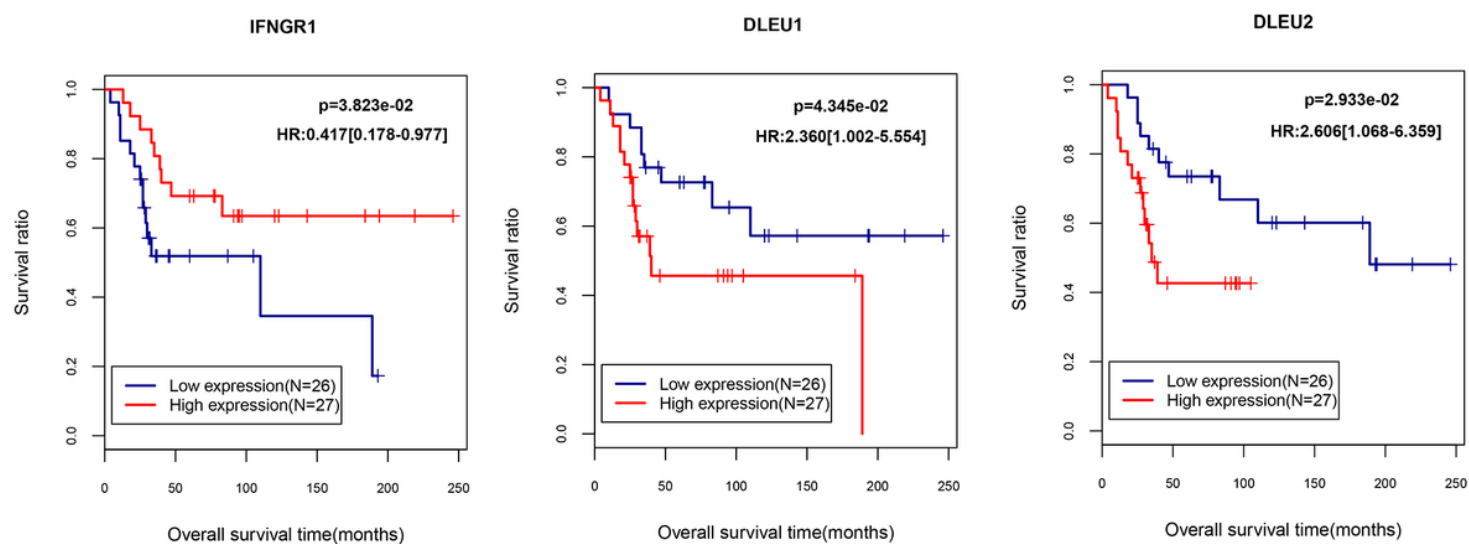


Figure 5

The Kaplan-Meier survival analysis for 13 RNAs. Increased expression levels of IFNGR1, while lower DLEU1 and DLEU2 related to better prognosis.

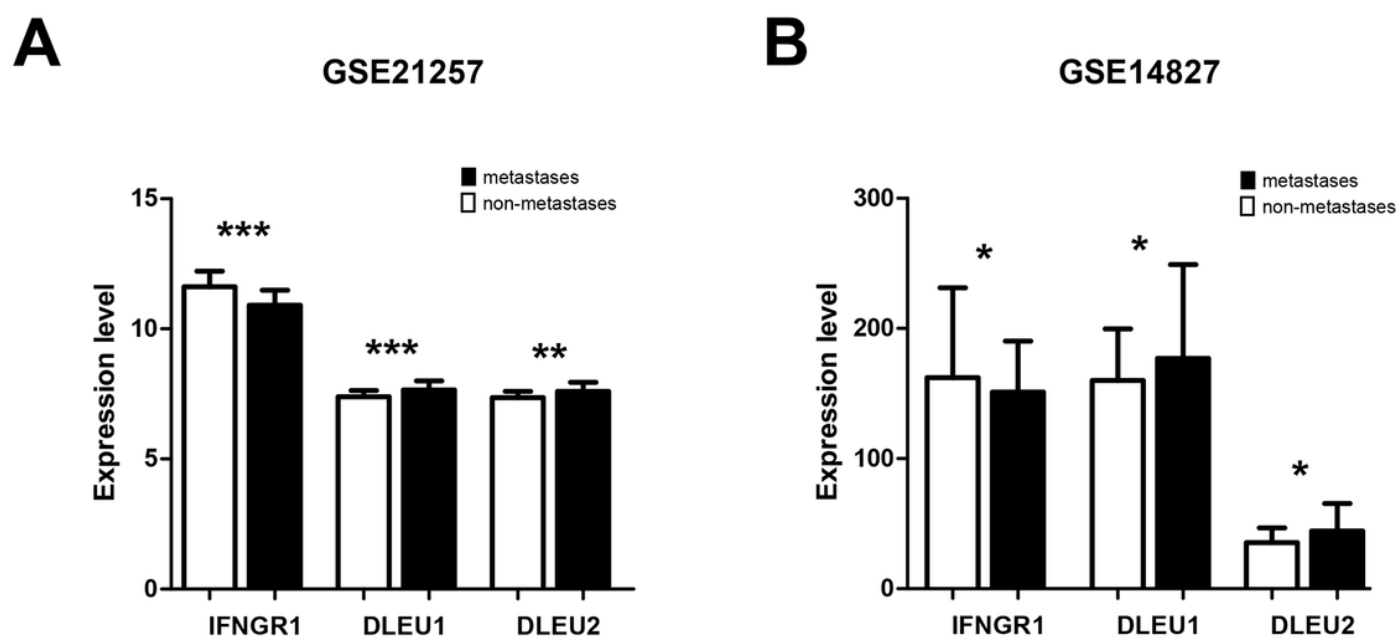


Figure 6

The histograms of Kaplan-Meier survival analysis using GSE14827 for validation. A: The expression level of IFNGR1, DLEU1 and DLEU2 in GSE21257; B: The expression level of IFNGR1, DLEU1 and DLEU2 in GSE14827.