

# Integrated analysis of lncRNA-miRNA-mRNA ceRNA network and the potential prognosis indicators in sarcomas

**Lu Gao**

the general of western theater command

**Yu Zhao**

the general hospital of western theater command

**Xuelei Ma**

Sichuan university

**Ling Zhang** (✉ [zhangling1718@126.com](mailto:zhangling1718@126.com))

general hospital of western theater command

---

## Research article

**Keywords:** sarcomas, ceRNA network, overall survival, prognosis indicators

**Posted Date:** December 8th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-21667/v2>

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

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on March 2nd, 2021. See the published version at <https://doi.org/10.1186/s12920-021-00918-x>.

# Abstract

**Background:** Competitive endogenous RNA (ceRNA) has been demonstrated to play crucial roles in multiple biological processes and in the development of neoplasms that potentially serve as diagnostic and prognosis markers as well as therapeutic targets.

**Methods:** In this work, we identified differentially expressed mRNAs (DEGs), lncRNAs (DELs) and miRNAs (DEMs) in sarcoma by comparing the genes expression profiles between sarcoma samples and normal tissue samples in Gene Expression Omnibus (GEO) datasets. Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analyses were applied to investigate the primary functions of the overlapped DEGs. Then, lncRNA-miRNA and miRNA-mRNA interactions were predicted, and the ceRNA regulatory network was constructed in Cytoscape. In addition, the protein-protein interaction (PPI) network was constructed and survival analysis was performed.

**Results:** A total of 1296 DEGs were identified in sarcoma samples by combining the GO and KEGG pathway enrichment analyses, 338 DELs were discovered after the probes were reannotated, and 36 DEMs were ascertained through intersecting two different expression miRNAs sets. Further, through target gene prediction, a lncRNA-miRNA-mRNA ceRNA network that contained 113 mRNAs, 69 lncRNAs and 29 miRNAs was constructed. The PPI network identified the six most significant hub proteins. Survival analysis revealed that seven mRNAs, four miRNAs and one lncRNA were associated with overall survival of sarcoma patients.

**Conclusions:** Overall, we constructed a ceRNA network in sarcomas, which likely provides insights for further research on the molecular mechanism and potential prognosis biomarkers.

## Background

Sarcomas are rare types of malignant neoplasms of mesenchymal origin and often spread to other tissues in the body [1]. Sarcomas can be classified into more than 100 distinct subtypes, which are typically divided into two heterogeneous group, including bone sarcomas and soft tissue sarcomas [2, 3]. Despite the use of extensive surgical resection, intensive and multiagent chemotherapy, radiotherapy, and emerging treatment strategies consisting of molecular targeting agents, immune checkpoint inhibitors, and adoptive T-cell therapy, the overall 5-year survival rates have not significantly improved over the past few decades for sarcomas [2, 4, 5]. Unfortunately, sarcoma patients accompanied with metastatic spread show a lower 5-year overall survival rate [6]. Further investigations into the pathogenesis and identification of the novel prognostic biomarkers that facilitate the improvement of the therapy and prognosis of sarcomas are urgently needed.

Although noncoding RNAs (ncRNAs), such as long noncoding RNAs (lncRNAs) and microRNAs (miRNAs), function as key regulators of gene expression, their involvement in various human diseases is being gradually revealed, and the multilayered regulatory networks formed by cross-linked ncRNAs and mRNAs seemly provide new insights into their regulatory mechanism with regards to both physiology and

pathology [7, 8]. The ceRNA network has been suggested to be involved in essential biological processes and play crucial roles in the initiation and development of neoplasms and to potentially serve as diagnostic and prognosis markers or therapeutic targets [9]. miRNAs are small RNAs of 19–25 nucleotides in length that can guide the posttranscriptional repression of protein-coding genes by binding to their mRNAs [10]. The lncRNAs that harbor a miRNA response element can compete with other RNA transcripts and thus theoretically function as competing endogenous RNAs (ceRNAs) [11]. That a single miRNA can regulate multiple target RNAs that contain the specific miRNA response element and these RNAs can be regulated by multiple miRNAs lay the foundation for the construction of a ceRNA network [12]. Recent studies have shown novel roles of the ceRNA network in lung cancer, breast cancer, gastric cancer, esophageal adenocarcinoma, esophageal squamous cell carcinoma, cholangiocarcinoma, ovarian cancer, and uterine corpus endometrial carcinoma [13-21]. In addition, a previous study has constructed a ceRNA network and predicted the prognosis of soft tissue sarcoma recurrence [22]. However, the ceRNA network based on sarcoma and normal tissue samples has not been reported.

In the present study, the genes expression profiles of sarcomas in the Gene Expression Omnibus (GEO) database were integrated, and the differentially expressed genes (DEGs), lncRNAs (DELs) and miRNAs (DEMs) were identified. Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment analyses were performed to screen out significant functional groups of DEGs. Thereafter, a ceRNA network consisting of lncRNA-miRNA-mRNA was constructed. In addition, a protein-protein interaction (PPI) network analysis for the DEGs that were included in the ceRNA network and prognostic analysis based on the information of sarcoma patients in the TCGA database were conducted to explore the effects of these transcripts on sarcomas.

## Methods

### Acquisition of available datasets and lncRNA annotation

Microarray data including high-throughput genes or miRNA expression profiling of sarcoma patients and normal samples were acquired from the GEO database ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)). The Affymetrix Human Genome U133 Plus 2.0 Array was reannotated according to the latest release of the updated gene database to filter lncRNA probes [23]. The process of lncRNA annotation was as follows: gene transcription sequences beginning with NM and XM in the RefSeq database were selected as a mRNA database, and ncRNAs with a length greater than 200 nucleotides that had gene transcription sequences beginning with NR and XR were selected from RefSeq, Ensembl and NONCODE2016 as a lncRNA database. Sequence alignments between 54675 probes that were downloaded from the annotation file of the Affymetrix Human Genome U133 Plus 2.0 Array and the mRNA database were performed by using BLAST, and the E value was set as  $10^{-5}$ . Probes that match the respective gene were included in the mRNA probe set, and others were compared with the lncRNA database. The E value was set as  $10^{-5}$ , and probes that match the respective gene were included in the lncRNA probe set, which was use for the following analysis.

## Identification of DEGs, DELs and DEMs

DEGs, DELs and DEMs between sarcoma patients and normal muscle tissue samples were determined by using the limma package in the R software (version 3.3.2) [24, 25].  $P < 0.05$ , the false discovery rate (FDR)  $< 0.05$  and fold change (FC)  $> 3$  was used as the screening criterion for DEGs and DELs, while  $P < 0.05$ ,  $FDR < 0.05$  and  $FC > 2$  was used for DEMs. Hierarchical clustering was performed using EPCLUST to represent the relationships among the samples based upon the expression data matrix of DEGs, DELs and DEMs, respectively.

## GO and KEGG pathway enrichment analysis of DEGs

GO enrichment analysis of DEGs was performed to identify the potential functional genes according to the Gene Ontology database, which can organize genes into hierarchical categories and uncover the gene regulatory network on the basis of biological processes and molecular function [26]. KEGG pathway analysis of DEGs was carried out to ascertain the potential functions of these genes that participated in the pathways based on the KEGG database [27-29]. Fisher's exact test was used to classify the GO category and identify the significant genes through the P-value and FDR [30].

## Target genes and lncRNA prediction of DEMs

Three algorithms, miRanda (<http://www.microrna.org/>), Targetscan (<http://www.targetscan.org/>) and miRWalk (<http://129.206.7.150/>) were used to inspect interactions between DEMs that were identified above and their target genes. The target genes that overlapped with genes that were identified by GO and KEGG pathway enrichment analyses were considered as the predicted results. MiRanda and PITA ([https://genie.weizmann.ac.il/pubs/mir07/mir07\\_exe.html](https://genie.weizmann.ac.il/pubs/mir07/mir07_exe.html)) were used to predict target lncRNAs, and those that overlapped with DELs between sarcoma samples and muscle tissue samples were used for the construction of the ceRNA network.

## Construction of the ceRNA network

Based on the ceRNA theory that lncRNA can serve as an endogenous "sponge" to regulate the expression of mRNA by sinking miRNA [31], upregulated or downregulated miRNAs, and lncRNAs or mRNAs that are inverse relationships with miRNAs in the intersection lncRNA-miRNA and miRNA-mRNA interaction pairs were chosen to construct the lncRNA-miRNA-mRNA ceRNA network [32]. The ceRNA networks were constructed and visualized using Cytoscape software (version 2.8.2) [33]. The flow chart, shown in Figure 1, represents the overall process of ceRNA network construction.

## PPI network analysis

The PPI network of between the encoded proteins of the DEGs that were included in the ceRNA network was constituted by using the STRING database version 11.0 (<https://string-db.org/>). Cytoscape software (version 2.8.2) was used to visualize the PPI network with a combined score of protein pairs  $> 0.4$  as the cut-off value [34].

## Survival analysis

High-throughput experimental data with survival profiles of sarcoma patients was extracted from the TCGA database. The associations between the expression levels of mRNAs, miRNAs and lncRNAs that were included in the ceRNA network and overall survival of sarcoma patients were evaluated by using log rank test, with P-values less than 0.05 recognized as statistically significant.

## Results

### Available datasets of sarcomas

There were four datasets (GSE55625, GSE31045, GSE17674 and GSE18546) obtained from GEO, whereas the GSE55625 and GSE31045 datasets, with multiple zero and negative expression values, were excluded. The GSE17674 dataset contains Ewing sarcoma patients and 18 normal muscle tissue samples and is based on the array platforms of Affymetrix Human Genome U133 Plus 2.0 Array, from which we extracted probes for lncRNA annotation. The GSE18546 dataset contains two subtypes of bone sarcoma, Ewing sarcoma and synovial sarcoma, and 5 muscle tissue samples. Five Ewing sarcoma samples and 10 synovial sarcoma samples were allocated to two case-control groups, respectively.

### Identification of DEGs, DELs and DEMs in sarcomas

DEGs and DELs were identified in 32 Ewing sarcoma patients compared with 18 normal muscle tissue samples in the GSE17674 dataset. DEMs were identified in 10 synovial sarcoma samples and 5 Ewing sarcoma samples compared with muscle tissue samples in the GSE 18546 dataset. A total of 3415 mRNAs (2554 upregulated and 861 downregulated genes) and 338 lncRNAs (234 upregulated and 104 downregulated lncRNAs) were differentially expressed in sarcoma patients of the GSE17674 dataset (Table S1; Table S2).

There were two different subtypes of sarcomas, Ewing sarcoma and synovial sarcoma, in the GSE18546 dataset, and two different expression miRNAs sets were identified. Fifty-two miRNAs (39 upregulated and 13 downregulated miRNAs) were differentially expressed in Ewing sarcoma in comparison to skeletal muscle. In addition, there were 145 (109 upregulated and 36 downregulated miRNAs) DEMs between synovial sarcoma and skeletal muscle. We created a Venn diagram intersecting the two DEM sets, and 26 upregulated miRNAs and 10 downregulated miRNAs were identified (Figure S1; Table S3). Hierarchical clustering of the identified DEGs, DELs and DEMs was displayed as a heatmap (Figure 2).

### GO and KEGG enrichment analysis of DEGs

GO enrichment analysis and KEGG pathways analysis of 3415 DEGs were performed to identify the potential functional genes. The upregulated and down-regulated genes were analyzed, respectively. The top 25 significantly enriched upregulated DEGs are presented in Figure 3A, including transcription, DNA-templated (GO:0006351;  $P=2.64 \times 10^{-104}$ ), cell division (GO:0051301;  $P=2.57 \times 10^{-82}$ ) and regulation of

transcription, DNA-templated (GO:0006355;  $P=7.99\times 10^{-75}$ ). The most significantly enriched downregulated DEGs are presented in Figure 3B, including muscle contraction (GO:0006936;  $P=1.89\times 10^{-57}$ ), muscle filament sliding (GO:0030049;  $P=5.78\times 10^{-46}$ ) and sarcomere organization (GO:0045214;  $P=8.40\times 10^{-38}$ ). The main pathways for both the upregulated genes and the downregulated genes are metabolic pathways (Figure 3C, D). The significant GO terms and pathways of up- and downregulated DEGs are presented in Table S4 and Table S5.

Finally, 1296 intersecting DEGs were extracted from the significantly enriched genes in GO and KEGG pathway analyses involved in up- and down-regulated genes (Table S6).

### **Target genes and lncRNAs of DEMs**

In this study, we have identified 36 sarcoma-associated miRNAs, and we focused on whether these miRNAs would target the 1296 DEGs identified by GO and KEGG pathway analyses and 338 DELs between sarcoma samples and muscle tissue samples. Based on the predicted targets of DEMs, 448 miRNA-mRNA interactor pairs (including 34 miRNAs and 269 mRNAs) and 454 lncRNA-miRNA interaction pairs (including 117 lncRNAs and 36 lncRNAs) were obtained.

### **ceRNA network**

According to the ceRNA theory, we use the shared miRNA as a junction, that is, upregulated miRNAs, accompany with downregulated lncRNAs and mRNAs, and downregulated miRNAs, accompany with upregulated lncRNAs and mRNAs in the miRNA-mRNA and miRNA-lncRNA interaction pairs, to construct a lncRNA-miRNA-mRNA ceRNA network. Finally, the ceRNA network consisted of 1440 interactions, including 69 lncRNAs, 29 miRNAs and 113 mRNAs (Figure 4; Table S7).

### **PPI network analysis**

To further explore the most significant clusters of DEGs in the ceRNA network, we conducted PPI network analysis using the STRING database version 11.0 and visualization by Cytoscape. The most significant hub upregulated proteins were IGF1, PRKCB and GNAI3, and the most significant hub downregulated proteins were AR, CYCS and PPP1CB in the PPI network (Figure 5; Table S8).

### **Survival of miRNAs, lncRNAs and mRNAs**

Furthermore, we performed the survival analysis based on the miRNAs, mRNAs and lncRNAs that were involved in the ceRNA network. The results show that three miRNAs, three mRNAs and one lncRNA were significantly associated with the overall survival of sarcoma patients ( $P<0.05$ ). Among them, high expression levels of SMARCC1, SRSF10, PRPF38A, JARID2, GNAI3, miR-301a-3p, miR-106b-5p, miRNA-130b-3p, miR-423-3p and LINC01296 and low expression levels of ARF3 and PRKCB were associated with shorter overall survival in sarcomas (Figure 6).

## **Discussion**

Increasing evidence shows that lncRNAs and miRNAs were differentially expressed and implicated in series of molecular processes, including differentiation, proliferation, metastasis, and transcriptional regulation, in sarcomas [35, 36]. However, the entire regulatory network that links the functions of coding and noncoding RNAs has not been extensively discussed. In the present study, bioinformatics analysis was utilized to integrate available sequencing data sets of sarcomas, and 1296 DEGs were identified in sarcoma samples by combining the GO and KEGG pathway enrichment analyses. A total of 338 DELs were discovered after the probes were reannotated, and 36 DEMs were ascertained through intersecting two different expression miRNAs sets. Further, 448 miRNA–mRNA interactions and 454 miRNA–lncRNA interactions were obtained through target gene prediction, and then, we constructed a lncRNA-miRNA-mRNA ceRNA network that contained 29 miRNAs, 69 lncRNAs and 113 mRNAs.

The ceRNA network identified in our study provided useful clues for further study. According to the DEGs in the ceRNA network, we constructed a PPI network, which showed the upregulated hub nodes, including IGF1, PRKCB and GNAI3, and the downregulated hub nodes, including AR, CYCS and PPP1CB. In addition, there were twelve RNAs in the ceRNA network associated with the prognosis of sarcomas based on the TCGA database. Among the seven survival-associated mRNAs, the high expression levels of SMARCC1, SRSF10, PRPF38A, JARID2 and GNAI3 were significantly associated with shorter overall survival in sarcomas ( $P=0.0018$ ,  $P=0.037$ ,  $P=0.0058$ ,  $P=0.0093$ , and  $P=0.0234$ , respectively). However, the high expressions of ARF3 and PRKCB were significantly associated with longer overall survival ( $P=0.0018$  and  $P=0.0162$ ), indicating that ARF3 and PRKCB overexpression could be positive prognostic factors in sarcoma patients. In these overall survival-associated mRNAs, GNAI3 and PRKCB, with the highest degree of connectivity, were hubs and tended to be essential [37]. Moreover, the PPI network showed that GNAI3 and PRKCB interacted with each other. The protein encoded by PRKCB is one of the PKC family members that has been reported to be involved in many different cellular functions. Surdez et al.'s study showed PRKCB is strongly overexpressed in Ewing sarcoma. PRKCB inhibition significantly increased apoptosis in Ewing sarcoma cells and prevented tumor growth in vivo [38]. GNAI3 encodes an alpha subunit of guanine nucleotide-binding proteins and is involved in various transmembrane signaling pathways. In addition, their function and role in the prognosis of sarcoma patients remain to be further defined. In addition, in the ceRNA network, both GNAI3 and ARF3 can be regulated by miR-133b and miR-133a-3p, which are at the center of the regulatory network and interact with lncRNAs including NONHSAT001973.2, NONHSAT203034.1, SNHG12 and ZNF37BP. Recent research has demonstrated that SNHG12 was significantly overexpressed in osteosarcoma, and high expression of SNHG12 tended to lead to a poor prognosis of osteosarcoma patients. Researchers further confirmed that SNHG12 promoted tumorigenesis and metastasis by activating the Notch signaling pathway, wherein it functioned as a ceRNA, modulating the expression of Notch2 by competing with miR-195-5p [39]. Therefore, the interaction of SNHG12, miR-133a-3p/miR-133b and GNAI3 in sarcoma was worthwhile of exploration. Additionally, ARF3 could also be regulated by miR-378a-3p and miR-422a, and among them, miR-378a-3p targeted LINC01296. It has been revealed that the high expressions were significantly associated with shorter overall survival of sarcoma patients ( $P=0.0297$ ) through our survival analysis.

The high expressions of miRNA-301a-3p, miRNA-106b-5p, miRNA-130b-3p, and miRNA-423-3p were significantly associated with a shorter overall survival of sarcoma patients ( $P < 0.0001$ ,  $P = 0.0046$ ,  $P = 0.0128$ , and  $P = 0.0147$ , respectively). The overexpression of miR-301a was revealed in Ewing sarcoma cell lines. Additionally, the transfection of anti-miR-301a inhibited the proliferation and cell cycle progression of Ewing sarcoma cells [40]. In the ceRNA network, the NEAT1/miR-301a-3p, XIST/miR-301a-3p, miR-301a-3p/NR3C2, and miR-301a-3p/AR interactions were identified. XIST was significantly upregulated in osteosarcoma tissues and cell lines, and increased XIST expression was associated with the poor overall survival of patients [41]. Nakatani et al. have defined miR-130b as an independent predictor of risk for disease progression and survival by microarray analysis [42]. Up until now, there is no study that reports the association of miRNA-423-3p and miRNA-106b-5p with sarcomas. The NEAT1/miR-301a-3p/AR, XIST/miR-301a-3p/AR, NEAT1/miR-301a-3p/NR3C2, and XIST/miR-301a-3p/AR axes predicted by the construction of the ceRNA network may provide a more precise research direction for exploring biological mechanisms extending from this ceRNA network.

There were several limitations in the present study. Sarcomas contains multiple distinct subtypes, and only two subtypes were involved in our study. More RNA sequencing datasets of all subtypes of sarcomas are required for the construction of a ceRNA network of sarcomas or a particular subtype. Furthermore, regulatory pathways of the ceRNA network are very complex, and ceRNA activity is influenced by multiple factors, such as the abundance and subcellular location of ceRNA components, miRNA/ceRNA affinity, RNA editing and RBPs [9]. This study only investigates the putative interactions of lncRNAs, miRNAs and genes, which require further validation.

## Conclusion

In conclusion, our work successfully constructed a ceRNA network by bioinformatics analysis based on the GEO database, providing a comprehensive resource for investigating the ceRNA regulation in sarcomas. Importantly, we screened out candidate prognostic biomarkers that are involved in the ceRNA network, which may exhibit important roles in the therapeutic target and prognosis analysis in sarcoma patients.

## Declarations

### Abbreviations

GEO: Gene Expression Omnibus; DEGs: Differentially expressed genes; DELs: Differentially expressed lncRNAs; DEMs: Differentially expressed miRNAs; GO: gene ontology; KEGG: Kyoto encyclopedia of genes and genomes; PPI: Protein–protein interaction; FDR: false discovery rate; TCGA: The Cancer Genome Atlas.

### Declarations

### Ethics approval and consent to participate

Not applicable.

### **Consent for publication**

Not applicable.

### **Availability of data and materials**

The analyzed data sets generated during the present study are available from the corresponding author (Email: zhangling1718@126.com) on reasonable request.

### **Funding**

This study was supported by grants from the Science and Technology Program of Sichuan Province (2018JY0165). We thank all of the participants in this study for their contributions. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### **Authors' contributions**

LG and LZ conceived and designed the study. LG and YZ wrote the manuscript. LG, YZ and XM were responsible for the data visualization, and analysis. XM and LZ reviewed and revised the manuscript. All authors read and approved the final manuscript.

### **Competing interests**

The authors have no conflicts of interest to declare.

### **Acknowledgements**

Not applicable.

## **References**

1. Ressing M, Wardelmann E, Hohenberger P, Jakob J, Kasper B, Emrich K, Eberle A, Blettner M, Zeissig SR: Strengthening health data on a rare and heterogeneous disease: sarcoma incidence and histological subtypes in Germany. *BMC public health* 2018, 18(1):235.
2. Alavi SN, Florou V, Tinoco G, Trent JC, Wilky BA: A precision medicine approach in sarcoma: identification of patients who may benefit from early use of pazopanib. *Discovery medicine* 2018, 25(137):131-144.

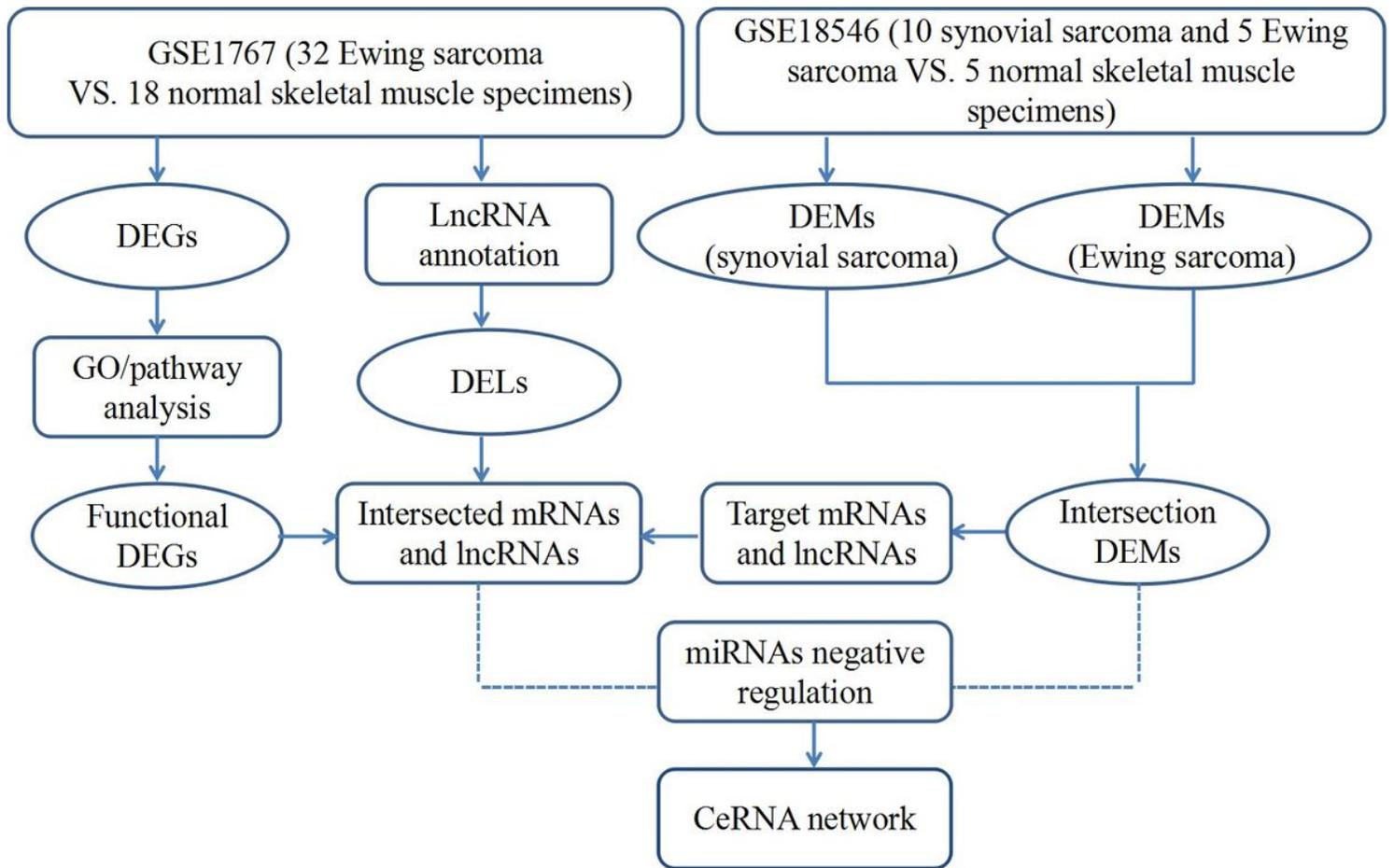
3. Doyle LA: Sarcoma classification: an update based on the 2013 World Health Organization Classification of Tumors of Soft Tissue and Bone. *Cancer* 2014, 120(12):1763-1774.
4. Miwa S, Yamamoto N, Hayashi K, Takeuchi A, Igarashi K, Tsuchiya H: Therapeutic Targets for Bone and Soft-Tissue Sarcomas. *International journal of molecular sciences* 2019, 20(1).
5. Gaspar N, Hawkins DS, Dirksen U, Lewis IJ, Ferrari S, Le Deley MC, Kovar H, Grimer R, Whelan J, Claude L *et al*: Ewing Sarcoma: Current Management and Future Approaches Through Collaboration. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2015, 33(27):3036-3046.
6. Meyer WH: Improving Outcomes for Patients With Pulmonary Metastatic Ewing Sarcoma. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2019:Jco1902522.
7. Cava C, Bertoli G: Portrait of Tissue-Specific Coexpression Networks of Noncoding RNAs (miRNA and lncRNA) and mRNAs in Normal Tissues. 2019, 2019:9029351.
8. Zhu KP, Zhang CL, Ma XL, Hu JP, Cai T, Zhang L: Analyzing the Interactions of mRNAs and ncRNAs to Predict Competing Endogenous RNA Networks in Osteosarcoma Chemo-Resistance. *Molecular therapy : the journal of the American Society of Gene Therapy* 2019, 27(3):518-530.
9. Qi X, Zhang DH, Wu N, Xiao JH, Wang X, Ma W: ceRNA in cancer: possible functions and clinical implications. *Journal of medical genetics* 2015, 52(10):710-718.
10. Li L, Peng M, Xue W, Fan Z, Wang T, Lian J, Zhai Y, Lian W, Qin D, Zhao J: Integrated analysis of dysregulated long non-coding RNAs/microRNAs/mRNAs in metastasis of lung adenocarcinoma. 2018, 16(1):372.
11. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP: A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell* 2011, 146(3):353-358.
12. Plotnikova O, Baranova A, Skoblov M: Comprehensive Analysis of Human microRNA-mRNA Interactome. *Frontiers in genetics* 2019, 10:933.
13. Wang XW, Guo QQ, Wei Y, Ren KM, Zheng FS, Tang J, Zhang HY, Zhao JG: Construction of a competing endogenous RNA network using differentially expressed lncRNAs, miRNAs and mRNAs in nonsmall cell lung cancer. *Oncology reports* 2019, 42(6):2402-2415.
14. Wang JJ, Huang YQ, Song W, Li YF, Wang H, Wang WJ, Huang M: Comprehensive analysis of the lncRNA-associated competing endogenous RNA network in breast cancer. *Oncology reports* 2019, 42(6):2572-2582.
15. Le K, Guo H, Zhang Q, Huang X, Xu M, Huang Z, Yi P: Gene and lncRNA co-expression network analysis reveals novel ceRNA network for triple-negative breast cancer. *Scientific reports* 2019, 9(1):15122.
16. Pan H, Guo C, Pan J, Guo D, Song S, Zhou Y, Xu D: Construction of a Competitive Endogenous RNA Network and Identification of Potential Regulatory Axis in Gastric Cancer. *Frontiers in oncology* 2019, 9:912.
17. Yu Y, Chen X, Cang S: Cancer-related long noncoding RNAs show aberrant expression profiles and competing endogenous RNA potential in esophageal adenocarcinoma. *Oncology letters* 2019,

18(5):4798-4808.

18. Li CY, Zhang WW, Xiang JL, Wang XH, Wang JL, Li J: Integrated analysis highlights multiple long non-coding RNAs and their potential roles in the progression of human esophageal squamous cell carcinoma. *Oncology reports* 2019, 42(6):2583-2599.
19. Xu F, Zhao Y, Qin G, Huan Y, Li L, Gao W: Comprehensive analysis of competing endogenous RNA networks associated with cholangiocarcinoma. *Experimental and therapeutic medicine* 2019, 18(5):4103-4112.
20. Li G, Han L, Ren F, Zhang R, Qin G: Prognostic value of the tumor-specific ceRNA network in epithelial ovarian cancer. *Journal of cellular physiology* 2019, 234(12):22071-22081.
21. Ouyang D, Li R, Li Y, Zhu X: Construction of a Competitive Endogenous RNA Network in Uterine Corpus Endometrial Carcinoma. *Medical science monitor : international medical journal of experimental and clinical research* 2019, 25:7998-8010.
22. Huang R, Meng T, Chen R, Yan P, Zhang J, Hu P, Zhu X, Yin H, Song D, Huang Z: The construction and analysis of tumor-infiltrating immune cell and ceRNA networks in recurrent soft tissue sarcoma. *Aging* 2019, 11(22):10116-10143.
23. Chalifa-Caspi V, Yanai I, Ophir R, Rosen N, Shmoish M, Benjamin-Rodrig H, Shklar M, Stein TI, Shmueli O, Safran M *et al*: GeneAnnot: comprehensive two-way linking between oligonucleotide array probesets and GeneCards genes. *Bioinformatics (Oxford, England)* 2004, 20(9):1457-1458.
24. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK: limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic acids research* 2015, 43(7):e47.
25. Zou JB, Chai HB, Zhang XF, Guo DY, Tai J, Wang Y, Liang YL, Wang F, Cheng JX, Wang J *et al*: Reconstruction of the lncRNA-miRNA-mRNA network based on competitive endogenous RNA reveal functional lncRNAs in Cerebral Infarction. *Scientific reports* 2019, 9(1):12176.
26. The Gene Ontology (GO) project in 2006. *Nucleic acids research* 2006, 34(Database issue):D322-326.
27. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M: The KEGG resource for deciphering the genome. *Nucleic acids research* 2004, 32(Database issue):D277-280.
28. Yi M, Horton JD, Cohen JC, Hobbs HH, Stephens RM: WholePathwayScope: a comprehensive pathway-based analysis tool for high-throughput data. *BMC bioinformatics* 2006, 7:30.
29. Draghici S, Khatri P, Tarca AL, Amin K, Done A, Voichita C, Georgescu C, Romero R: A systems biology approach for pathway level analysis. *Genome research* 2007, 17(10):1537-1545.
30. Wei D: A multigene support vector machine predictor for metastasis of cutaneous melanoma. *Molecular medicine reports* 2018, 17(2):2907-2914.
31. Guo LL, Song CH, Wang P, Dai LP, Zhang JY, Wang KJ: Competing endogenous RNA networks and gastric cancer. *World journal of gastroenterology* 2015, 21(41):11680-11687.
32. Xu Y, Chen J, Yang Z, Xu L: Identification of RNA Expression Profiles in Thyroid Cancer to Construct a Competing Endogenous RNA (ceRNA) Network of mRNAs, Long Noncoding RNAs (lncRNAs), and

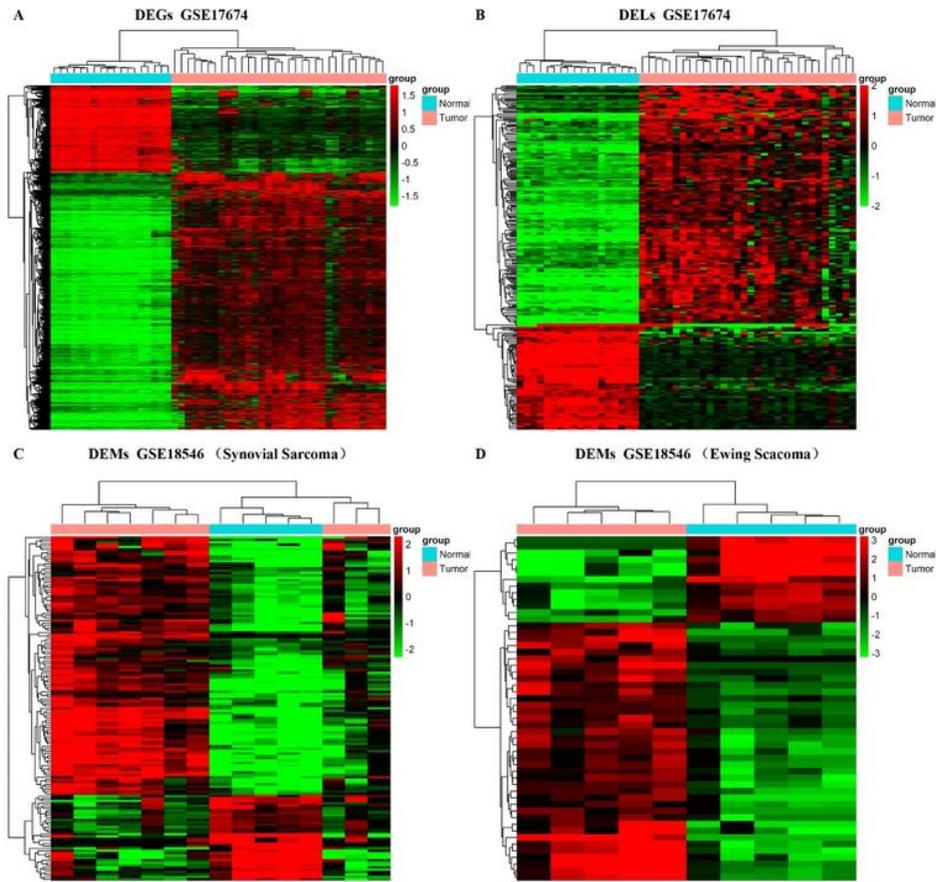
- microRNAs (miRNAs). *Medical science monitor : international medical journal of experimental and clinical research* 2019, 25:1140-1154.
33. Demchak B, Hull T, Reich M, Liefeld T, Smoot M, Ideker T, Mesirov JP: Cytoscape: the network visualization tool for GenomeSpace workflows. *F1000Research* 2014, 3:151.
  34. Bhushan R, Rani A, Ali A, Singh VK, Dubey PK: Bioinformatics enrichment analysis of genes and pathways related to maternal type 1 diabetes associated with adverse fetal outcomes. *Journal of diabetes and its complications* 2020, 34(5):107556.
  35. Ma L, Sun X, Kuai W, Hu J, Yuan Y, Feng W, Lu X: LncRNA SOX2 overlapping transcript acts as a miRNA sponge to promote the proliferation and invasion of Ewing's sarcoma. *American journal of translational research* 2019, 11(6):3841-3849.
  36. Li M, Chen H, Zhao Y, Gao S, Cheng C: H19 Functions as a ceRNA in Promoting Metastasis Through Decreasing miR-200s Activity in Osteosarcoma. *DNA and cell biology* 2016, 35(5):235-240.
  37. He X, Zhang J: Why do hubs tend to be essential in protein networks? *PLoS genetics* 2006, 2(6):e88.
  38. Surdez D, Benetkiewicz M, Perrin V, Han ZY, Pierron G, Ballet S, Lamoureux F, Redini F, Decouvelaere AV, Daudigeos-Dubus E *et al*: Targeting the EWSR1-FLI1 oncogene-induced protein kinase PKC-beta abolishes ewing sarcoma growth. *Cancer research* 2012, 72(17):4494-4503.
  39. Zhou S, Yu L, Xiong M, Dai G: LncRNA SNHG12 promotes tumorigenesis and metastasis in osteosarcoma by upregulating Notch2 by sponging miR-195-5p. *Biochemical and biophysical research communications* 2018, 495(2):1822-1832.
  40. Kawano M, Tanaka K, Itonaga I, Iwasaki T, Tsumura H: MicroRNA-301a promotes cell proliferation via PTEN targeting in Ewing's sarcoma cells. *International journal of oncology* 2016, 48(4):1531-1540.
  41. Li GL, Wu YX, Li YM, Li J: High expression of long non-coding RNA XIST in osteosarcoma is associated with cell proliferation and poor prognosis. *European review for medical and pharmacological sciences* 2017, 21(12):2829-2834.
  42. Nakatani F, Ferracin M, Manara MC, Ventura S, Del Monaco V, Ferrari S, Alberghini M, Grilli A, Knuutila S, Schaefer KL *et al*: miR-34a predicts survival of Ewing's sarcoma patients and directly influences cell chemo-sensitivity and malignancy. *The Journal of pathology* 2012, 226(5):796-805.

## Figures



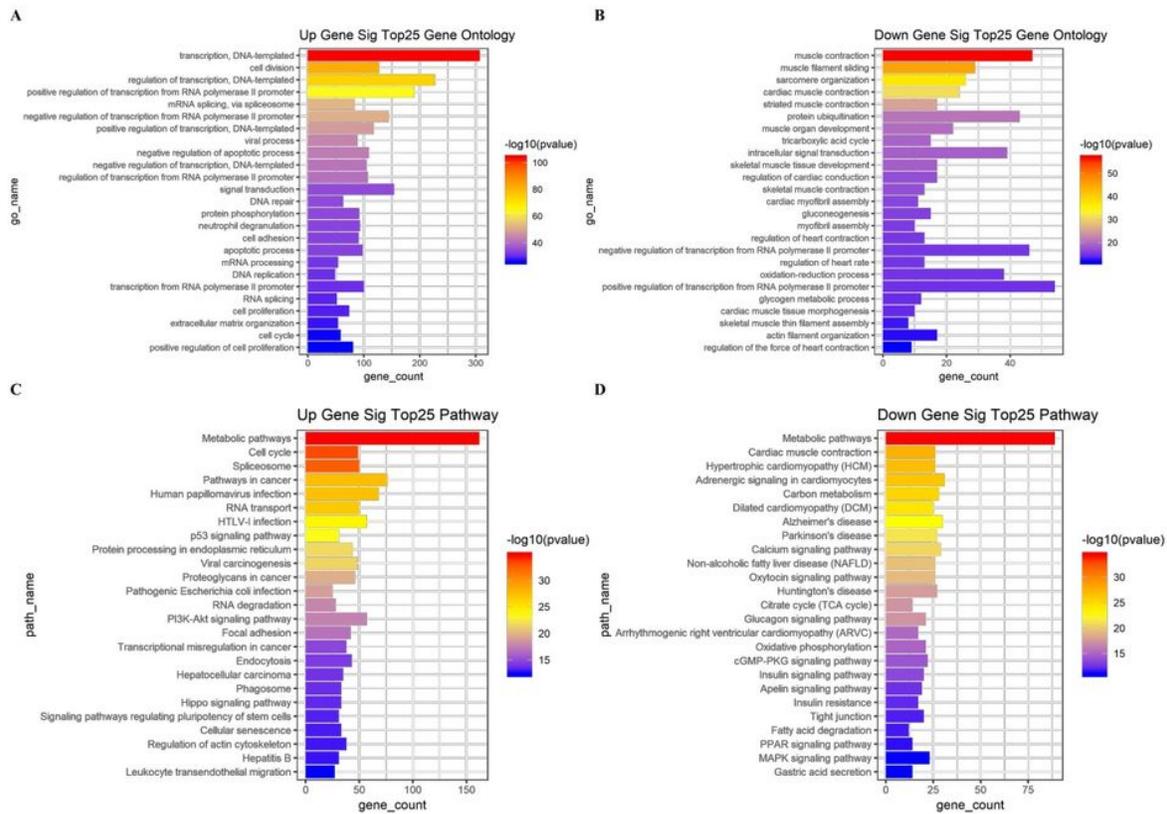
**Figure 1**

Flow chart of constructing the ceRNA network.



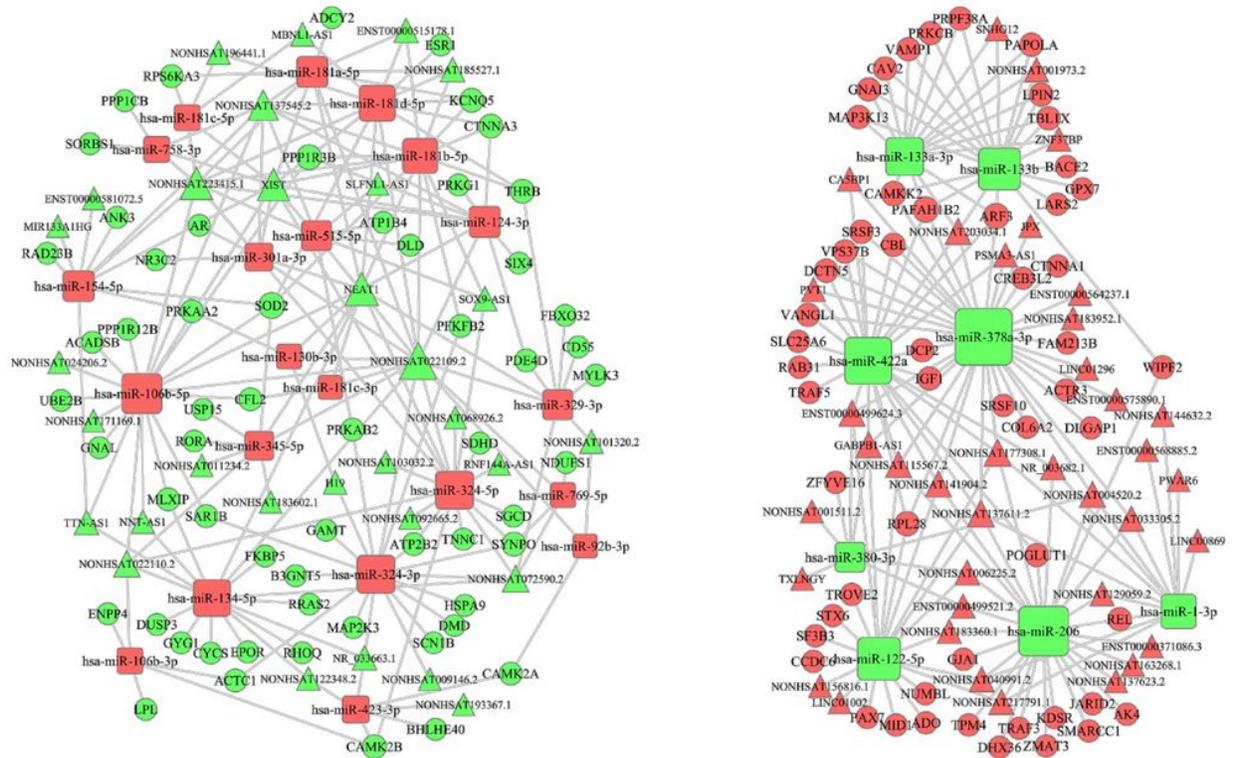
**Figure 2**

The heatmap of differentially expressed mRNAs, lncRNAs, and miRNAs between sarcoma samples and skeletal muscle samples. The rows represent specific DERs, and the columns represent each sample. Blue represents the expression levels of DERs at a comparatively low level and red represents a high level. A. DEGs; B. DELs; and C. DEMs between synovial sarcoma samples and skeletal muscle samples; D. DEMs between Ewing sarcoma samples and skeletal muscle samples.



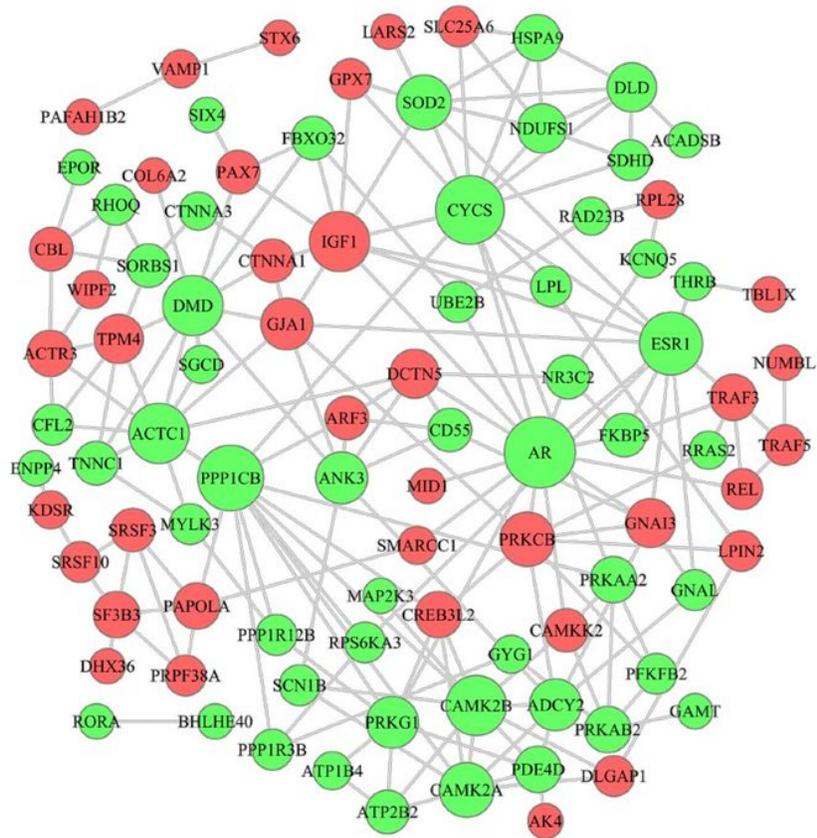
**Figure 3**

Top 25 enriched DEGs from the GO and KEGG pathways analyses. The bar plot shows the enrichment scores of GO and KEGG pathways. Different colors of bars represent different P values.



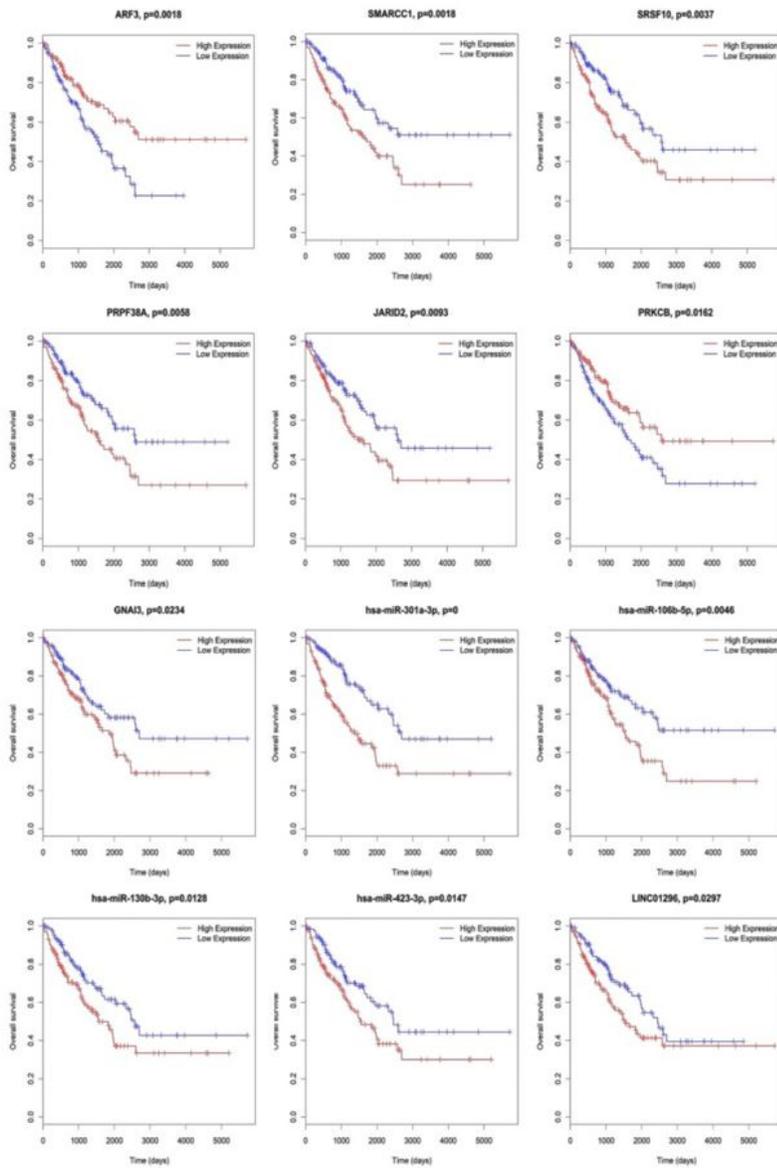
**Figure 4**

The lncRNA-miRNA-mRNA ceRNA network. Rectangles, circles and triangles represent miRNAs, mRNAs and lncRNAs, respectively. Red and green represent upregulated and downregulated genes in sarcomas. Gray lines indicate interactions between miRNA-mRNA and miRNA-lncRNA.



**Figure 5**

PPI network of DEGs created by STRING. Red and blue nodes represent upregulated and downregulated proteins, respectively. The edge between nodes represents the interaction between two connected proteins.



**Figure 6**

Kaplan-Meier survival curves of DEGs, DELs and DEMs associated with overall survival in sarcomas.

## Supplementary Files

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

- FigureS1.tif
- SupTable1.xlsx
- SupTable2.xlsx
- SupTable3.xlsx
- SupTable4.xlsx
- SupTable5.xlsx
- SupTable6.xlsx
- SupTable7.xls
- SupTable8.xlsx