Combined bioinformatics and machine learning methodologies reveal prognosis-related ceRNA network and propose *ABCA8*, *CAT*, and *CXCL12* as independent protective factors against osteosarcoma

*Jiaqi Fan[1,A,C,E,F], *Jianhong Liao[2,A–D], Yuwen Huang[3,B,D]

[1] Department of Orthopaedics, Capital Medical University Affiliated FuXing Hospital, Beijing, China
[2] Department of Otolaryngology Head and Neck Surgery, Beijing Tongren Hospital, Capital Medical University, China
[3] Department of Gastrointestinal and Hepatology, Beijing Youan Hospital, Capital Medical University, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation;
D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2024;33(8):857–868

Address for correspondence
Jiaqi Fan
E-mail: fanjiaqi1985@mail.ccmu.edu.cn

Funding sources
None declared

Conflict of interest
None declared

Acknowledgements
The authors would like to express their sincere gratitude to the Biotrainee team (www.biotrainee.com) for their valuable guidance regarding code analysis.

*Jiaqi Fan and Jianhong Liao contributed equally to this work.

Received on December 16, 2022
Reviewed on July 30, 2023
Accepted on September 20, 2023

Published online on February 5, 2024

Cite as
Fan J, Liao J, Huang Y. Combined bioinformatics and machine learning methodologies reveal prognosis-related ceRNA network and propose *ABCA8*, *CAT*, and *CXCL12* as independent protective factors against osteosarcoma. *Adv Clin Exp Med*. 2024;33(8):857–868. doi:10.17219/acem/172663

DOI
10.17219/acem/172663

Abstract

Background. Aberrant circular RNA (circRNA) acts as an oncogene or suppressor during neoplasm initiation and development. However, the functions of most circRNAs in osteosarcoma (OS) remain unclear.

Objectives. We aimed to investigate the expression, molecular functions and mechanisms underlying circRNAs in OS.

Materials and methods. Network interaction, pathway enrichment and regression analyses were performed to determine differentially expressed (DE) circRNAs, microRNAs (miRNAs) and messenger RNAs (mRNAs). We constructed competitive endogenous RcodeNA (ceRNA) networks and integrated patient clinical data to analyze the relationship between the networks and prognosis. The circRNA, miRNA and mRNA data were retrieved from Gene Expression Omnibus (GEO) microarray datasets. A circRNA-miRNA-mRNA interaction network was established and visualized using miRNet. Protein interactions were investigated using STRING and Cytoscape, and hub genes were identified using the MCODE plug-in. Gene Ontology, Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome pathway analyses were performed to determine the DEmRNAs. LIMMA and RobustRankAggreg were used to screen for DERNAs. Node genes in the interaction network were analyzed using least absolute shrinkage and selection operator (LASSO) and Cox regression to obtain OS-related ceRNA networks.

Results. We identified 9 DEcircRNAs, 243 DEmiRNAs and 211 DEmRNAs. We found that a ceRNA subnetwork, based on 1 circRNA, 1 miRNA and 8 mRNAs, was closely associated with OS prognosis. Integrating the proportional hazards model and survival analysis revealed 3 independent protective factors: adenosine triphosphate (ATP)-binding cassette sub-family A member 8 (*ABCA8*), catalase (*CAT*) and C-X-C motif chemokine ligand 12 (*CXCL12*).

Conclusions. Our study provides novel insights into circRNA-related ceRNA networks and identifies potential prognostic biomarkers of OS.

Key words: prognosis, microRNA, osteosarcoma, circular RNA, competitive endogenous RNA

Background

Osteosarcoma (OS) is a common malignant bone tumor that occurs in children and adolescents,[1] with its occurrence rate reaching a 2nd peak in individuals aged over 60.[2,3] With the emergence of neoadjuvant chemotherapy, patient survival rates have greatly improved over the past 40 years. However, because chemotherapy causes high treatment resistance in OS, the 5-year overall survival rate is still lower than 80%.[2,4] Therefore, identifying molecular markers is important for OS prognosis and treatment.

Circular ribonucleic acids (circRNAs) belong to a class of circular noncoding RNAs with continuous and covalently closed structures that modulate protein expression by acting as microRNAs (miRNAs) or protein inhibitor sponges.[5] With the discovery of circRNA functions, they are increasingly becoming important in understanding disease pathology and treatment. Their biological functions exhibit 5 characteristics: sponge effect, rolling circle translation, circRNA-derived pseudogenes, post-transcriptional regulation, and splicing interference.[6] In recent years, increasing focus has been directed toward understanding the relationship between circRNA and tumor occurrence and metastasis. Multiple circRNAs have been implicated in the development, migration, invasion, and metastasis of OS.[7–9] As such, identifying more unknown circRNAs that may be involved in cancerization and cancer development is critical, especially in malignant tumors such as OS.

The miRNAs are small noncoding RNAs that are 20–24 nucleotides in length.[10] Similar to classic noncoding RNAs, many miRNAs play vital roles in almost all aspects of tumorigenesis and tumor occurrence and development, including invasion, angiogenesis, proliferation, and apoptosis.[11–13] An in-depth understanding of their roles in the development of diseases, especially tumors, is required. Furthermore, miRNAs are attractive tools and potential targets for developing new treatment modalities.[14,15]

Overall, regulatory networks based on the composition of circRNAs, miRNAs and messenger RNAs (mRNAs) play a crucial role in tumor development and prognosis. Nonetheless, further research is imperative to thoroughly analyze the competing endogenous RNA (ceRNA) foundation of OS. In the present study, the Gene Expression Omnibus (GEO) database (www.ncbi.nlm.nih.gov/geo) was investigated, and prognosis-related ceRNA networks were constructed by validating the clinical information from the TARGET (Therapeutically Applicable Research to Generate Effective Treatments; www.cancer.gov/ccg/research/genome-sequencing/target) database. The results may provide valuable insights that could facilitate diagnosis, monitoring and prediction of prognosis and circRNA research in primary OS patients (Fig. 1).

Objectives

We investigated the GEO database to screen for RNAs associated with OS occurrence. By validating clinical information in the TARGET database, we constructed a prognosis-related ceRNA network, which may further enable OS diagnosis, monitoring and prognosis, and facilitate research on circRNAs in primary OS (Fig. 1).

Materials and methods

Study design

The GEO and TARGET databases were used to identify differentially expressed miRNAs (DEmiRNAs) and DEmRNAs associated with OS prognosis. A ceRNA network and protein-protein interaction (PPI) network were constructed to discern the interactions between mRNAs. Statistical and survival analyses were performed using the Cox model.

Setting

Data acquisition and processing

According to the data upload description file, the included datasets encompass gene sequencing results from tumor tissue and paracancerous tissue. We downloaded 6 expression profiles from the GEO database: 4 mRNA datasets (GSE28425 (GPL13376, 7 OS and 4 control), GSE99671 (GPL20148, 18 OS and 18 control), GSE36004 (GPL6102, 7 OS and 4 control), and GSE126209 (GPL20301, 6 OS and 5 control)), 1 miRNA expression profile (GSE28425 (GPL8227, 7 OS and 4 control)) and 1 circRNA expression profile (GSE140256 (GPL27741, 3 OS and 3 control)). In relation to prognostic outcomes, we used a GSE79181 profile (GPL15497, 25 patients with OS) and the TARGET database (172 available patients with OS) to identify OS prognosis-associated miRNA and mRNA. The basic traits of these 7 microarray datasets (GSE28425, GSE36004, GSE99671, GSE126209, GSE140256, GSE79181, and TARGET) are listed in Supplementary Table 1. Ethical approval or informed consent was not required for this data as they are publicly available in the GEO and TARGET databases.

Variables and data measurement

Identification of circRNA, miRNA, and mRNA

We selected datasets that were investigated using cancerous compared to pan-cancerous tissues through next-generation sequencing. All the datasets were normalized or $\log_2$-transformed. The circRNA and miRNA were identified using linear models for microarray data in LIMMA

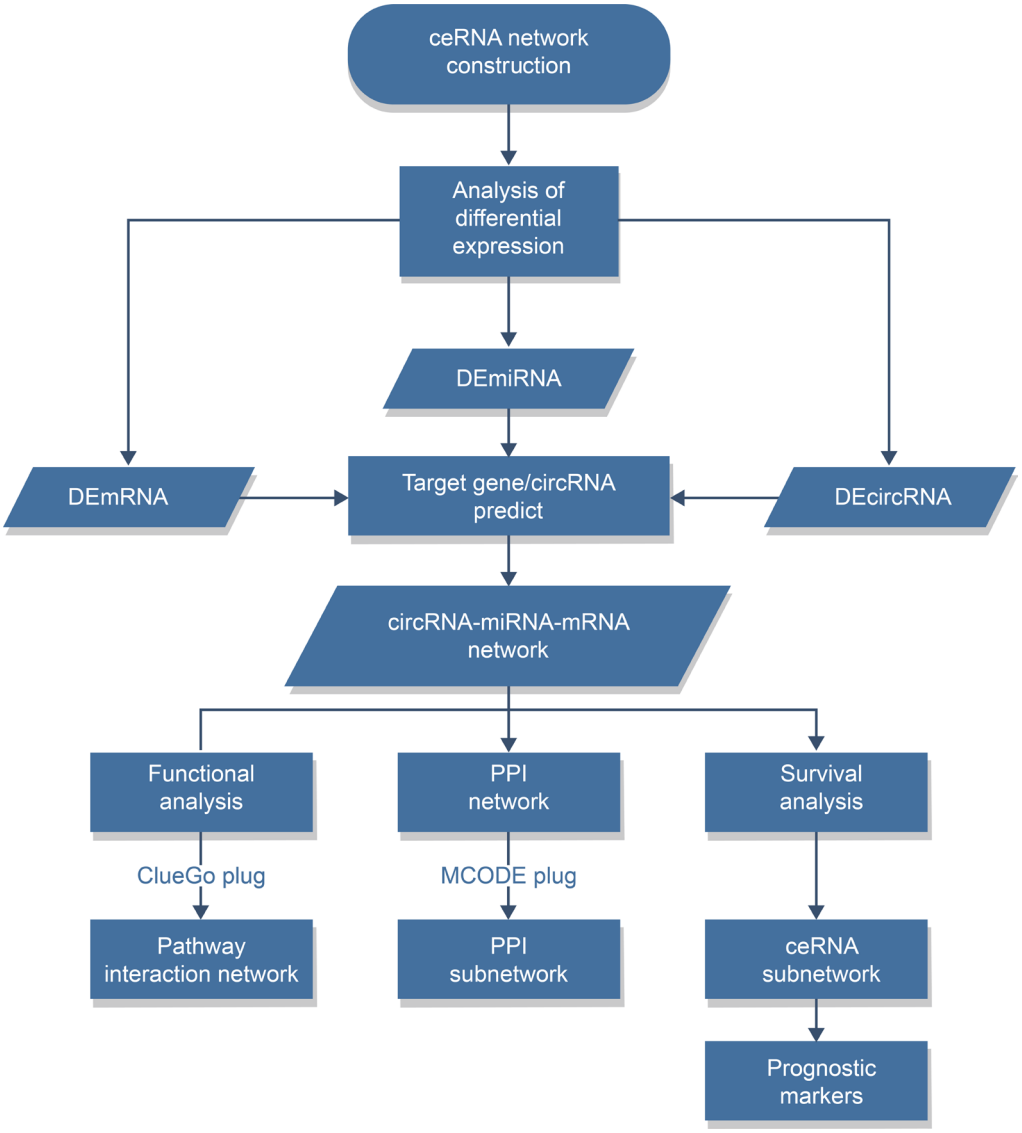

Fig. 1. Flowchart of the competitive endogenous RNA network analysis. circRNA, differentially expressed circular RNAs; miRNA, differentially expressed microRNAs; mRNA, differentially expressed messenger RNAs; PPI network, protein-protein interaction network; ceRNA network, competitive endogenous RNA network

v. 3.42.2, an R package that processes normalized data and analyzes the differential expression of genes (https://bioconductor.org/packages/release/bioc/html/limma.html). Key mRNAs with common differential expression in the 4 datasets were screened using the RobustRankAggreg package v. 1.1 (https://cran.r-project.org/web/packages/RobustRankAggreg/index.html). A p-value < 0.05 and a $\log_2$ fold change (FC) >1 were considered the cutoff values in the differential expression of genes.

Construction of a ceRNA network

Using miRNet (www.mirnet.ca), a network visual analytics platform based on 11 different miRNA databases, we predicted target genes and circRNAs. Interaction networks were visualized using Cytoscape v. 3.7.1 (http://cytoscape.org/).

Construction of a PPI network and module analysis

To discern the interactions between mRNA involved in the ceRNA network, we established a PPI network using an online database search tool for the retrieval of interacting genes (STRING; https://string-db.org/). Cytoscape was used for visualization. The molecular complex detection (MCODE) plug-in was used to extract the subnetwork from the PPI network.

Functional and pathway enrichment analyses

To elucidate the mRNA-associated mechanisms in OS, Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome pathway enrichment analyses of mRNAs were performed using ClueGO v. 2.5.6, based on Cytoscape. A p-value <0.05 indicated statistical significance in enrichment.

Statistical methods and quantitative variables

Statistical and survival analyses

The GSE79181 miRNA expression profile (25 samples), which is based on GPL15497 (TaqMan array human miRNA cards (A+B Card Set v3); Applied Biosystems, Foster City, USA), was downloaded from the GEO database. The mRNA expression profile and clinical information of 172 primary OS samples were downloaded from the TARGET database. Of these, 63 patients with mRNA expression profiles were selected as the discovery dataset, whereas 109 patients with basic information were used as the testing dataset. We applied the proportional hazard model (least absolute shrinkage and selection operator (LASSO) and multivariate Cox regression analyses) and performed multivariate Cox regression analysis between the time-to-survival and RNA expression in the ceRNA network using the data of 63 patients from the TARGET database and 25 patients from GSE79181. To verify the prediction efficiency of the Cox model, we used the index of concordance (C-index), time-dependent receiver-operating characteristic (ROC) curve, Kaplan–Meier (K–M) survival curves, and log-rank tests to compare the survival rates between high- and low-risk groups. The R v. 3.1-11 packages survival ROC v. 1.0.3 and Survminer v. 0.4.6 (www.cran.r-project.org/web/packages/survivalROC/survivalROC.pdf, https://cran.r-project.org/web/packages/survminer/index.html) were used to analyze and visualize the overall survival of circRNA, miRNA and mRNA. Genes absent from half of the samples were not considered during these analyses.

Results

Identification of circRNAs, miRNAs and mRNAs

We identified 9 circRNAs (Supplementary Table 2), 243 miRNAs and 211 mRNAs in the datasets. Of the 243 miRNAs, 128 were upregulated and 115 were downregulated (Fig. 2A). Three of the 9 circRNAs were upregulated and the remaining 6 were downregulated (Fig. 2B). Lastly, of the 211 mRNAs, 103 were upregulated and 108 were downregulated (Fig. 2C).

Construction of the ceRNA network

Based on miRNet, we screened 8 targeted circRNAs from 8,975 miRNA/mRNA using overlapping circRNA. We identified 119 of 13,473 targeted mRNAs based on 65,227 pairs of interactions between miRNAs and mRNAs. Using Cytoscape, we constructed a circRNA-miRNA-mRNA network based on the circRNA–DEmiRNA and DEmiRNA–DEmRNA interactions. Accordingly, we identified 6 circRNAs, 92 miRNAs and 119 mRNAs in the ceRNA network (Supplementary Fig. 1A). Of the 119 mRNAs, 69 were upregulated and 50 were downregulated. Using MCODE in Cytoscape, we recognized 2 subnetworks in which hsa-mir-20a-5p and Ras association domain family member 2 (*RASSF2*) were identified as key nodes, indicating that they may play a critical role in the OS development (Supplementary Fig. 1B,C).

Construction of the PPI network

We constructed a PPI network with the 119 mRNAs that participated in the ceRNA network (Supplementary Fig. 2A). Using MCODE, we identified 4 significant modules. The 1st module consisted of 6 target genes: *CD14*, *IGSF6*, *C1QA*, *GIMAP4*, *C1QB*, and *FGL2* (Supplementary Fig. 2B). The 2nd module consisted of 5 target genes: *SEC61G*, *RPL7*, *MRPL13*, *RPL6*, and *RPS28* (Supplementary Fig. 2C). The 3rd module comprised 4 target genes: *CDH2*, *TF*, *APOE*, and *SERPINA1* (Supplementary Fig. 2D), while the 4th module consisted of 3 target genes: *PTGES3*, *PTGS1*, and *TBXAS1* (Supplementary Fig. 2E). Notably, we identified *CD14*, *SEC61G*, and *PTGES3* as the key genes in the 1st, 2nd and 4th modules, respectively.

Gene Ontology and pathway enrichment analyses

We identified 67 functional enrichment terms from GO, including 56 biological processes and 11 cellular components. The top 5 biological processes were "xenobiotic metabolic process," "response to progesterone," "antigen processing and presentation of exogenous peptide antigen via major histocompatibility complex class I," "response to electrical stimulus," and "regulation of telomerase activity." The top 5 cellular components were "cytosolic large ribosomal subunit," "tertiary granule lumen," "ficolin-1-rich granule," "ficolin-1-rich granule lumen," and "luminal side of membrane" (Fig. 3). Nineteen KEGG and Reactome pathways were notably enriched. Major pathways enriched in OS were the "signaling by erb-b2 receptor tyrosine kinase 4 (*ERBB4*)," "nuclear signaling by *ERBB4*" and "gelatin degradation by matrix metalloproteinases (MMP) 1, 2, 3, 7, 8, 9, 12, and 13" (Fig. 4).

Statistical and survival analyses

Twenty-one RNAs were derived from the ceRNA network, with coef/se(coef) <0.01 and $p > 0.05$. To find optimal prognostic biomarkers among 3 specific RNAs, we performed a multivariate Cox regression analysis of the RNA expression profiles, clinical traits of 63 patients and other clinical information for 109 patients (Fig. 5A). The predictive potency of adenosine triphosphate (ATP)-binding cassette sub-family A member 8 (*ABCA8*), C-X-C motif

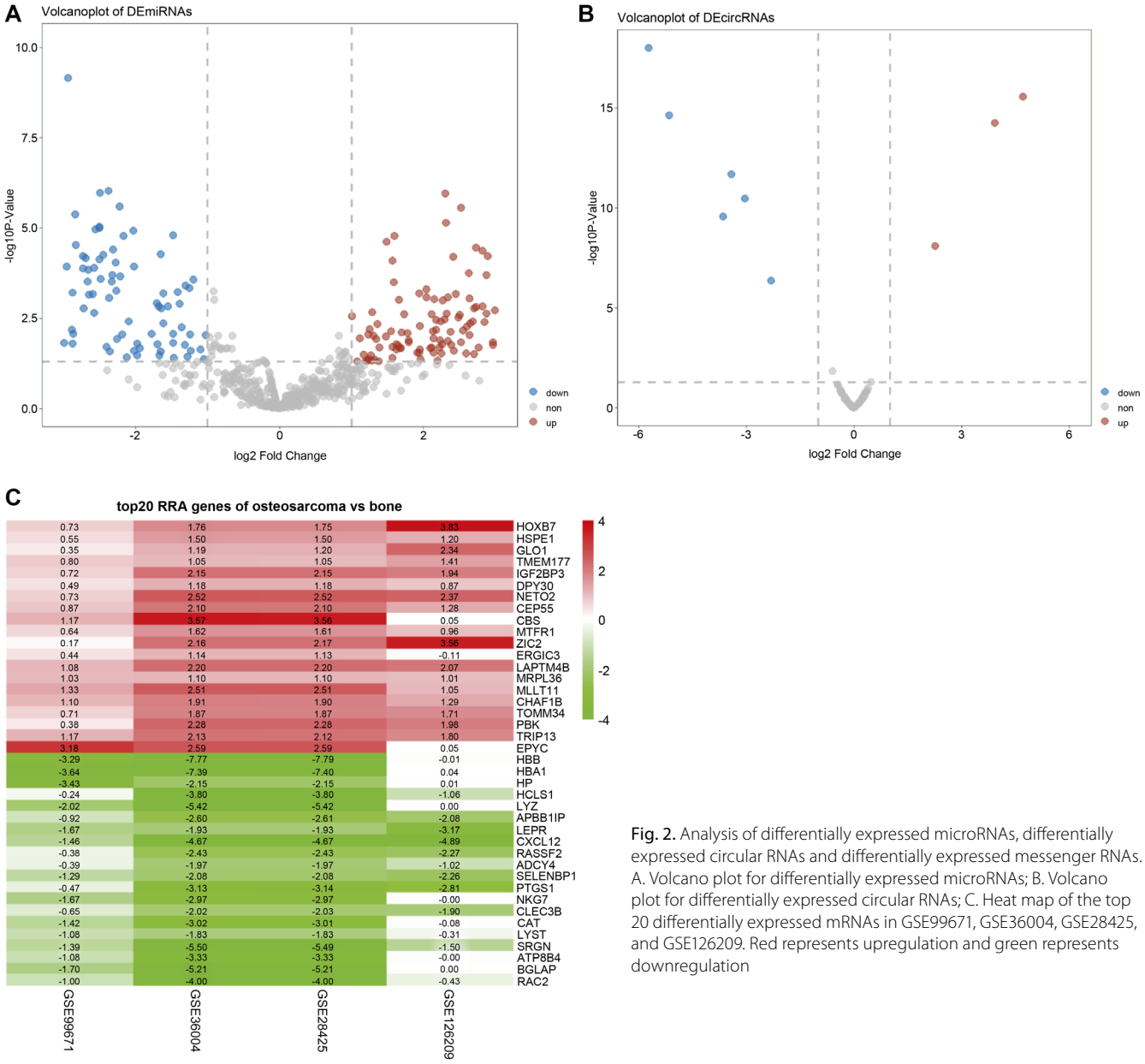

Fig. 2. Analysis of differentially expressed microRNAs, differentially expressed circular RNAs and differentially expressed messenger RNAs. A. Volcano plot for differentially expressed microRNAs; B. Volcano plot for differentially expressed circular RNAs; C. Heat map of the top 20 differentially expressed mRNAs in GSE99671, GSE36004, GSE28425, and GSE126209. Red represents upregulation and green represents downregulation

chemokine ligand 12 (*CXCL12*) and catalase (*CAT*) in OS was mutually independent (Fig. 5B). We calculated the correlation and variance inflation factor of these RNAs. The variance inflation factor was <1.1 and the correlation of each pair was <0.5, indicating no multicollinearity between independent variables (Supplementary Table 3). We made proportional hazard assumptions for the model (Supplementary Table 4 and Supplementary Fig. 3), with results showing that the model and the factors are proportional over time. Linear assumption showed that 3 predictors were linearly correlated with hazard function (Supplementary Fig. 3). Utilizing the same TARGET patient data, we noticed that the area under the curve (AUC) of 5-year overall survival for the module was 0.9, which is close to the C-index of 0.87, thus demonstrating a stable predictive effect (Fig. 5C). We grouped the discovery dataset using the 5-year risk score and used the K–M survival curves and log-rank tests to compare survival rates between the high- (n = 22) and low-risk (n = 39) groups (Fig. 5D). To investigate associations among the risk scores, OS and biomarkers, we ranked patients according to their risk scores and displayed their clinical information and gene expression on the same abscissa (Fig. 6). The high-risk group corresponded to a poor prognosis, with the 3 genes demonstrating great consistency with patient outcomes. To verify the reliability of this model, we calculated the C-index of *ABCA8*, *CXCL12* and *CAT*, and confirmed the reliability of the prediction (Supplementary Table 3). The RobustRankAggreg analysis for the 3 mRNAs is shown in Supplementary Table 5.

We applied a K–M one-way survival analysis for patients with OS and grouped them by median values. Of the 8 circRNAs, a low expression of hsa_circFADS2_007 was associated with improved survival (Supplementary Fig. 4A), and a high expression of hsa-mir-335-5p correlated with improved prognosis (Supplementary Fig. 4B).

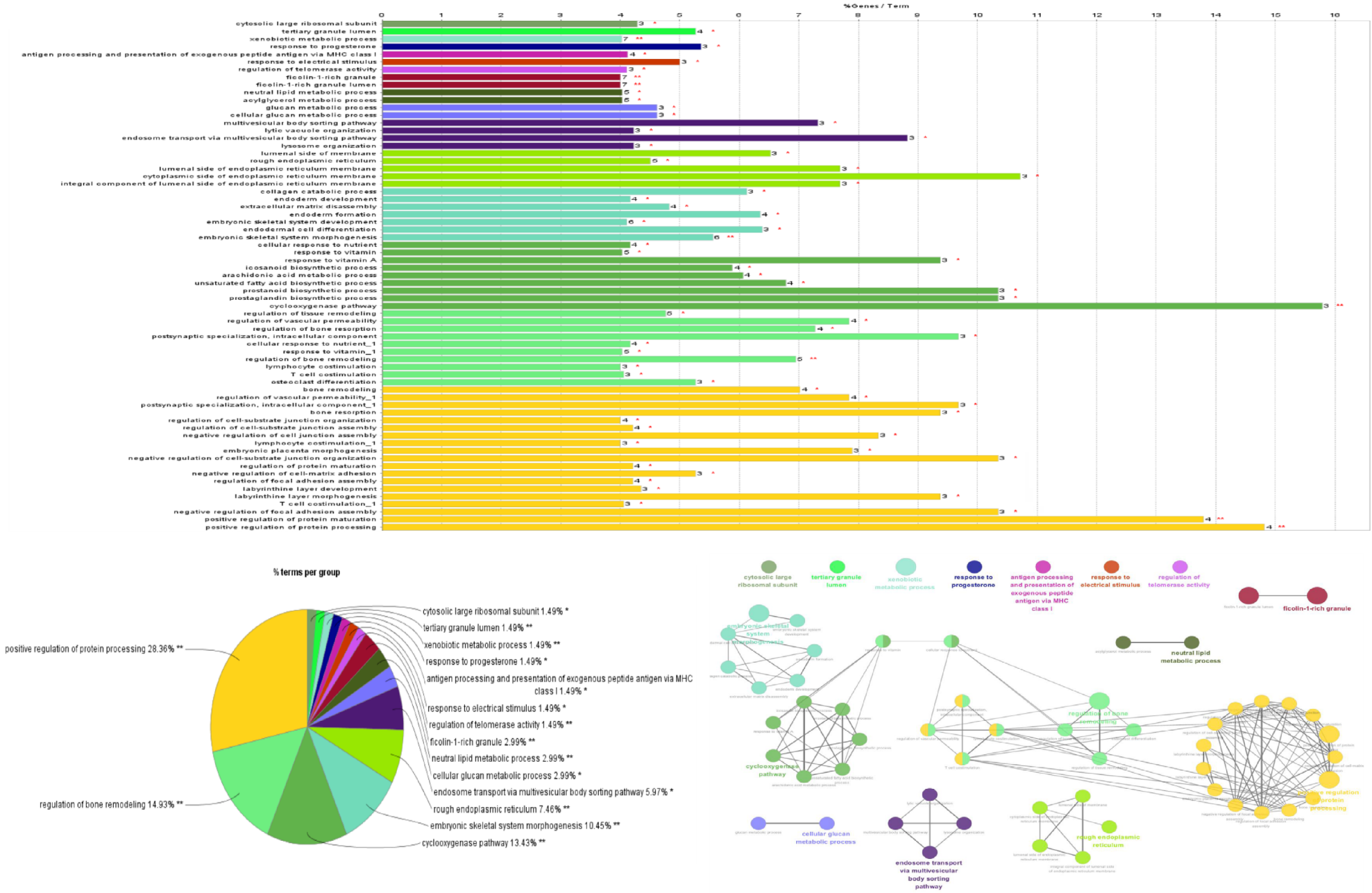

Fig. 3. Functional enrichment analyses of the target genes involved in the competitive endogenous RNA network of osteosarcoma. Bar graphs show the proportion of genes enriched in the gene ontology analyses. The pie chart shows the proportion of pathways in each group

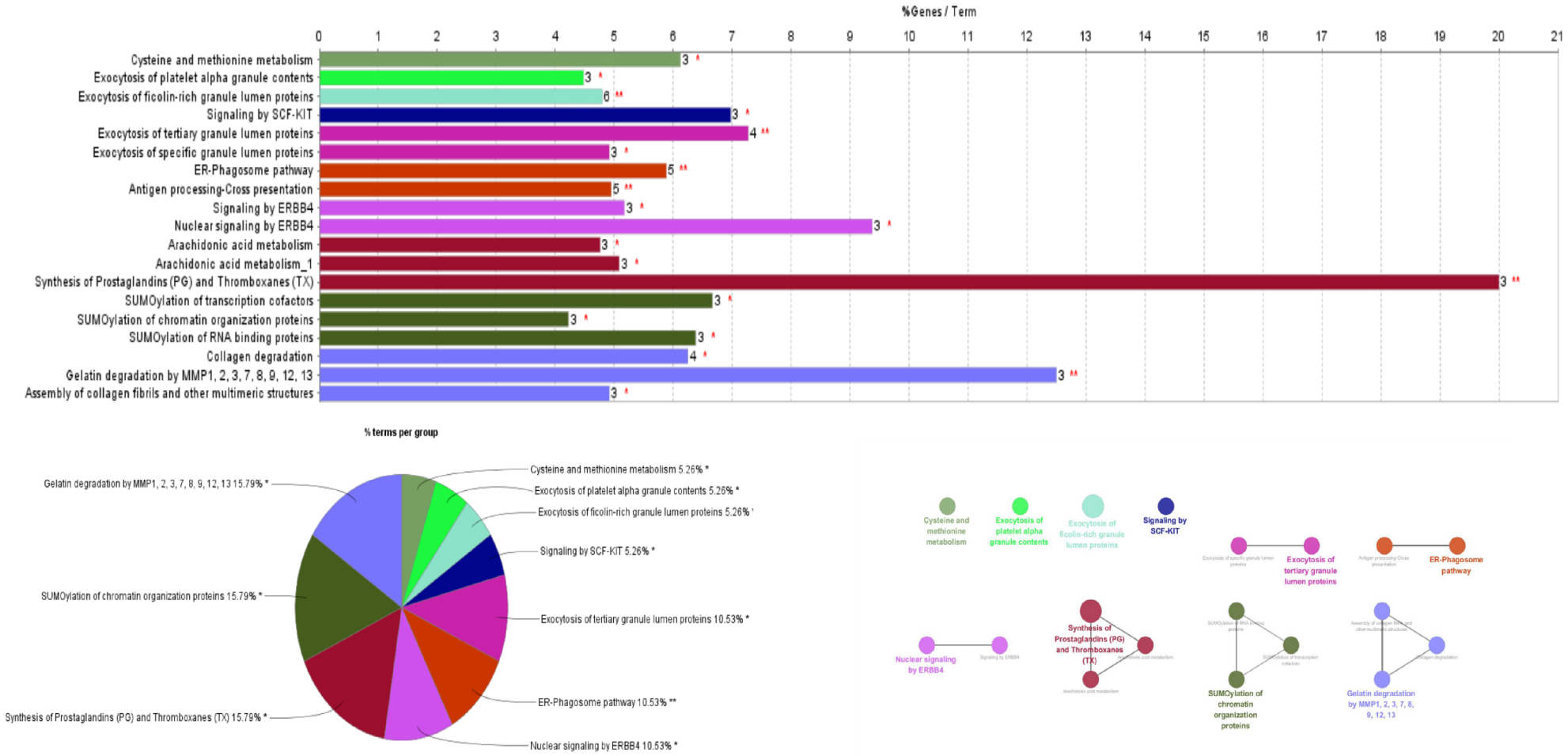

Fig. 4. Pathway enrichment analyses of the target genes involved in the competitive endogenous RNA network of osteosarcoma. Bar graphs show the proportion of genes enriched in the pathway analyses. The pie chart shows the proportion of pathways in each group

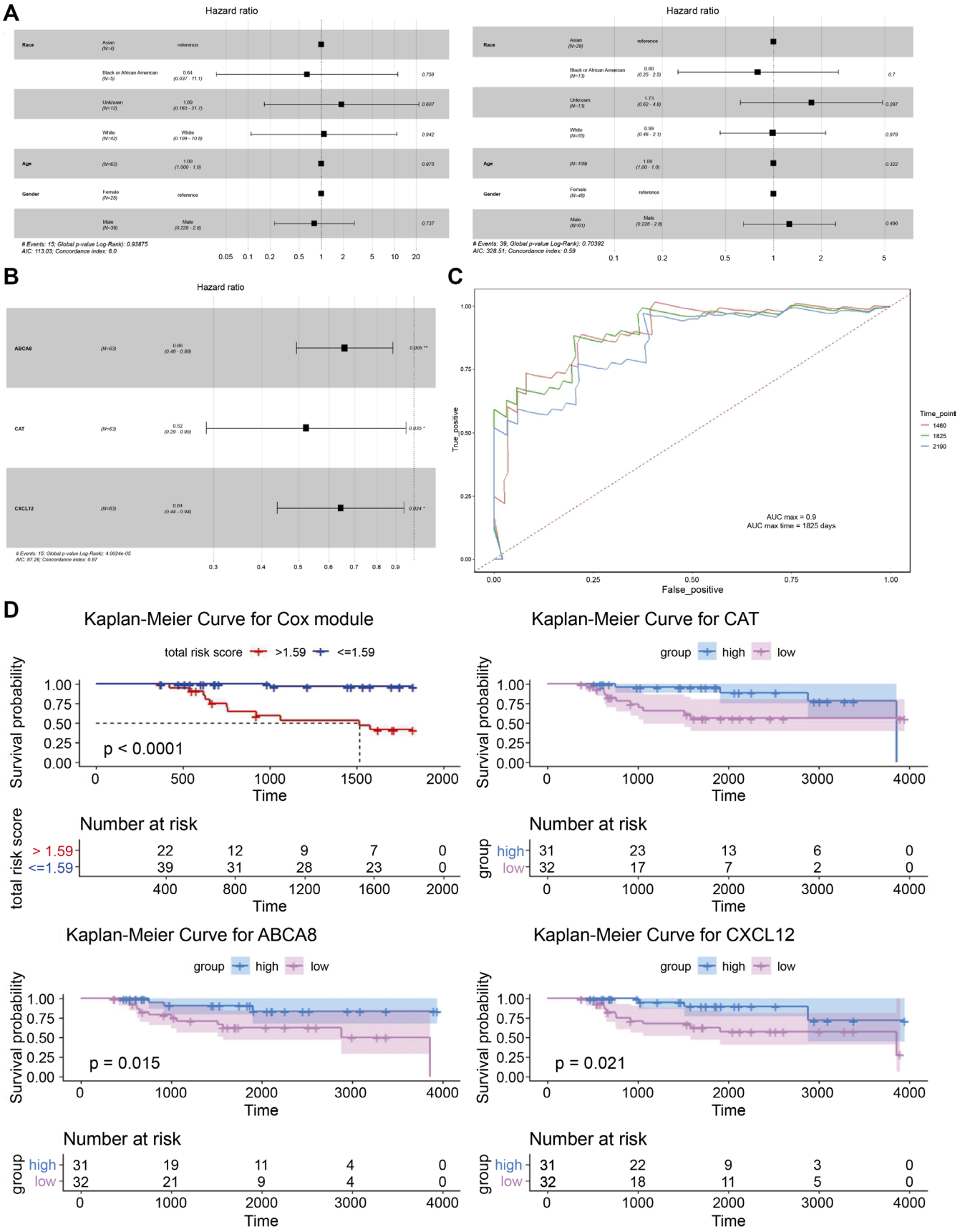

Fig. 5. Multivariate Cox regression analysis. A. Forest plot of risk factors in testing (left) and discovery (right) datasets; B. Visualization of Cox regression; C. The receiver operating characteristic (ROC) curve for 4-, 5- and 6-year survival; D. Kaplan–Meier survival analysis of the module and RNAs

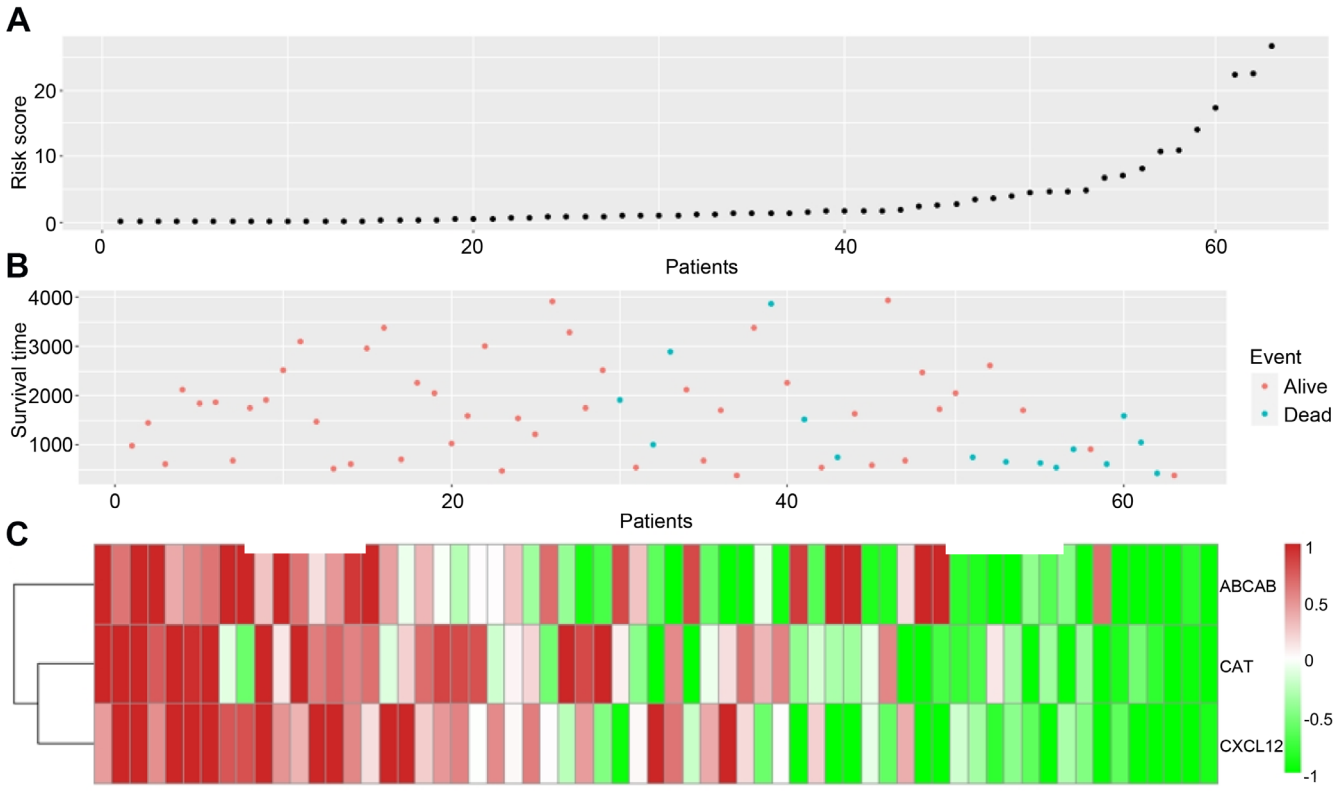

Fig. 6. Characteristics of the prognostic signature in the discovery dataset. A. Sorting of the patients in ascending order of risk score; B. Association of clinical characteristics with risk scores; C. Heatmap of *ABCA8*, *CAT* and *CSCL12* in high- and low-risk groups

Eight mRNAs, including 7 upregulated mRNAs (*ABCA8*, *ACSL5*, *FABP4*, *FGL2*, *LAPTM5*, *SLC38A2*, and *VNN2*) and 1 downregulated DEmRNA (*FADS1*), correlated with improved prognosis (Supplementary Fig. 4C,D). The high expression of *ABCA8*, *FABP4* and *VNN2* was significant for improved OS and event-free survival. We constructed a prognosis-related subnetwork based on 1 circRNA (hsa_circFADS2_007), 1 miRNA (hsa-mir-335-5p) and 8 mRNAs (*ABCA8*, *ACSL5*, *FABP4*, *FADS1*, *FGL2*, *LAPTM5*, *SLC38A2*, and *VNN2*). A high expression of all RNAs, except for hsa_circFADS2_007 and FADS1, was associated with improved OS prognosis (Fig. 7).

Discussion

Although OS is a rare malignant tumor, its characteristics, i.e., rapid invasion and ease of metastasis in vivo, lead to the development of metastases during treatment and follow-up in 30–40% of patients, making ideal treatment difficult.[4] Osteosarcoma treatment primarily involves surgery and chemotherapy; however, the prognosis for patients remains unsatisfactory.[16]

To elucidate the mechanisms underlying RNAs in OS, we constructed a ceRNA network based on differences between primary OS and controls. The GO enrichment analysis revealed that genes participating in the ceRNA network were mainly enriched in the positive regulation of protein processing and regulation of bone remodeling-related pathways, such as "bone remodeling," "regulation of vascular permeability" and "regulation of bone resorption." The ability to deregulate bone remodeling is one of the main characteristics of OS, inducing an increase in factors initially trapped in the bone matrix, such as transforming growth factor-β, to promote tumor progression.[17]

Exosomes are important for vascular permeability dysregulation and OS development, and a difference in miRNA content was confirmed between metastatic and nonmetastatic OS-derived exosomes.[18] The *MMPs* are overexpressed in OS and facilitate the survival, growth and metastasis of OS cells; thus, their expression is associated with a shorter survival time in OS patients.[19] SUMOylation is the post-translational modification of proteins whose disorders are believed to be related to malignant transformation in normal cells, cancer progression and the abnormal expression of oncogenes.[20] Prostaglandin promotes cell migration and invasion in different tumors, including OS, by increasing survivin expression and focal adhesion kinase phosphorylation and enhancing cell adhesion and migration.[21] Thus, ceRNA networks play critical roles in OS occurrence and progression.

To further investigate potential prognostic markers in the ceRNA network, we analyzed the correlation between 219 RNAs and OS or event-free survival. The hsa-mir-335-5p-related ceRNA subnetwork was significantly

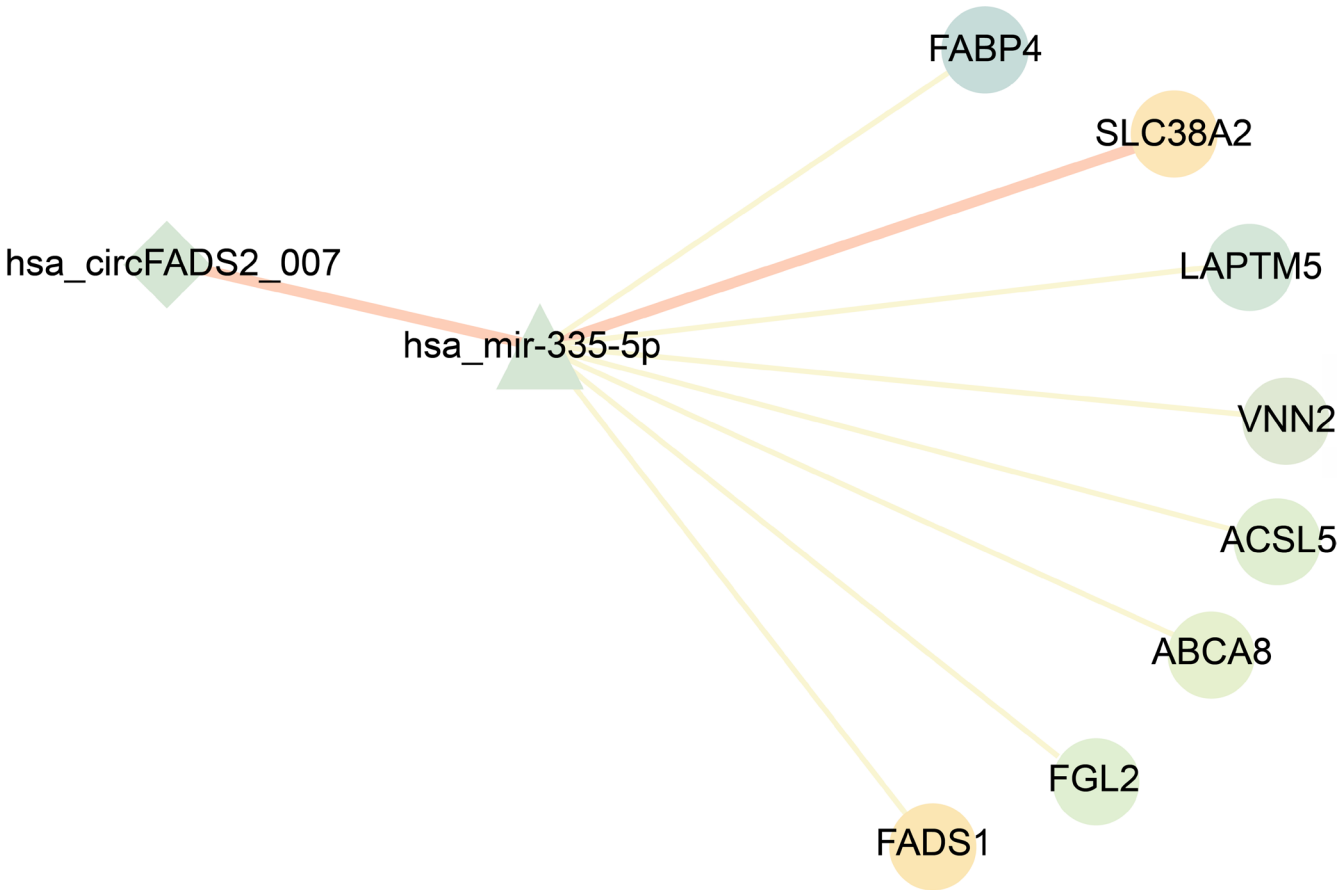

Fig. 7. Survival-related competitive endogenous RNA network. The light color indicates increased RNA expression, whereas the dark color indicates decreased RNA expression. Diamonds indicate circular RNAs, rectangles indicate microRNAs and circles indicate messenger RNAs

correlated with OS prognosis. In particular, we identified 1 circRNA (hsa_circFADS2_007), 1 miRNA (hsa-mir-335-5p) and 8 mRNAs (*ABCA8*, *ACSL5*, *FABP4*, *FADS1*, *FGL2*, *LAPTM5*, *SLC38A2*, and *VNN2*) in this subnetwork. Previous studies revealed that these RNAs were involved in tumor progression. A large-scale genetic study found that *FADS1* and *FADS2* (11q12. 2) fatty acid metabolism-related genes were associated with colorectal cancer risk.[22] Meanwhile, hsa-mir-335-5p affected the recurrence and prognosis of OS by regulating 2 potential modules.[23] *ABCA8* was reportedly underexpressed in 3 cases of pre-invasive breast cancer but not in invasive cases.[24] The high-level expression of the adenosine triphosphate (ATP)-binding cassette transporters for the "A" subfamily in patients with serous ovarian cancer was significantly associated with a poor survival rate.[25] In patients with colorectal cancer, *ACSL5* is closely related to tumor occurrence and prognosis.[26] Overexpression of miRNA-106a reportedly inhibits the effect of *VNN2* and leads to the invasion, proliferation and migration of OS cells.[21] Furthermore, the dysregulation of *FABP4*, *FGL2*, *LAPTM5*, and *SLC38A2* is identified in multiple cancers, such as colorectal, lung, ovarian, and breast.[27–30]

The high expression of *ABCA8*, *CAT* and *CXCL12* is significantly associated with survival time, and also serves as an independent protective factor in patients with OS. In our study, *CAT* was enriched in the "cyclooxygenase pathway" and "neutral lipid metabolic process," biological process (BP) pathways, and the "ficolin-1 rich granule" and "exocytosis of ficolin-rich granule lumen proteins," cellular component (CC) pathways, which are considered closely related to the hypoxia induction and abnormal proliferation of cells. The *CAT* plays a critical role in cellular resistance to oxidative stress, and its downregulation promotes the occurrence of tumors by increasing the level of reactive oxygen species (ROS) in transformed cells.[31] The Cox regression model showed that a high *CAT* expression in OS was associated with better prognoses. Analogously, a local increase in the CAT level is a feasible method to reverse or treat tumor cells and thus has attracted increasing attention.[32] Our results are consistent with those from previous studies showing that the hypoexpressive chemokine *CXCL12* controls metastasis and immune responses in OS, supporting that therapies targeting *CXCL12* have a potential for therapeutic intervention.[33] *CXCL12* reportedly induces the constitutive expression and activity of *CAT* by activating downstream p38, Akt and extracellular signal-regulated kinase, which are essential for protecting β cells from DNA damage caused by hydrogen peroxide (H_2O_2).[34] This may explain the consistency

in the expression of *CAT* and *CXCL2*. Moreover, *ABCA8*, *CAT* and *CXCL12* may complement the current prognostic gene signatures.

We focused on investigating the circRNAs, miRNAs and mRNAs between primary OS and normal bone or adjacent tissues to identify factors for improved tissue specificity. Although the circRNA-miRNA-mRNA network was built based on GEO data and verified using the TARGET database, the results require further experimental verification.

Limitations

We used reliable statistical methods to explore the dysregulated ceRNA network in OS. Given the current scarcity of studies on gene sequencing in OS, it is imperative that our findings be corroborated through larger sample sizes and additional experimentation.

Conclusions

Our findings provide new insights into the molecular mechanisms of OS and reveal potential targets for its diagnosis, monitoring and therapy. Using bioinformatics, we identified 3 independent protective factors in OS. No dataset suitable for Cox model validation was found in the public database to verify their applicability; hence, further experimentation is needed.

Supplementary data

The Supplementary materials are available at https://doi.org/10.5281/zenodo.10184102. The package includes the following files:

Supplementary Table 1. Basic traits of the seven microarray datasets from the GEO and TCGA.

Supplementary Table 2. Basic characteristics of the 9 differentially expressed circRNAs.

Supplementary Table 3. Index of concordance (C-index) and variance inflation factor (VIF) of *ABCA8*, *CXCL12* and *CAT*.

Supplementary Table 4. LASSO and Cox regression analysis of ceRNA with coef/se(coef) < 0.01 and p-value of proportional hazards assumption (PH) >0.05.

Supplementary Table 5. Robust rank aggregation analysis of *ABCA8*, *CXCL12* and *CAT*. LogFC in the 4 datasets and RRA score of the three RNAs.

Supplementary Fig. 1. Competitive endogenous RNA in osteosarcoma.

Supplementary Fig. 2. Protein–protein interaction (PPI) network of genes in the competitive endogenous RNA network.

Supplementary Fig. 3. Proportional hazards assumption (left) and linearity assumption (right).

Supplementary Fig. 4. Survival analysis for competitive endogenous RNA in osteosarcoma.

ORCID iDs

Jiaqi Fan https://orcid.org/0000-0003-0270-0900
Jianhong Liao https://orcid.org/0000-0003-2465-3850
Yuwen Huang https://orcid.org/0009-0007-3726-4340

References

1. Gao F, Zuo Q, Jiang T, Song H, Zhou J. A newly synthesized oleanolic acid derivative inhibits the growth of osteosarcoma cells in vitro and in vivo by decreasing c-MYC-dependent glycolysis. *J Cell Biochem*. 2019;120(6):9264–9276. doi:10.1002/jcb.28202
2. Mirabello L, Troisi RJ, Savage SA. Osteosarcoma incidence and survival rates from 1973 to 2004: Data from the Surveillance, Epidemiology, and End Results Program. *Cancer*. 2009;115(7):1531–1543. doi:10.1002/cncr.24121
3. Ottaviani G, Jaffe N. The epidemiology of osteosarcoma. In: Jaffe N, Bruland OS, Bielack S, eds. *Pediatric and Adolescent Osteosarcoma*. Cancer Treatment and Research. Vol. 152. Boston, USA: Springer US; 2009:3–13. doi:10.1007/978-1-4419-0284-9_1
4. Smith MA, Altekruse SF, Adamson PC, Reaman GH, Seibel NL. Declining childhood and adolescent cancer mortality. *Cancer*. 2014;120(16):2497–2506. doi:10.1002/cncr.28748
5. Kristensen LS, Andersen MS, Stagsted LVW, Ebbesen KK, Hansen TB, Kjems J. The biogenesis, biology and characterization of circular RNAs. *Nat Rev Genet*. 2019;20(11):675–691. doi:10.1038/s41576-019-0158-7
6. Liu J, Liu T, Wang X, He A. Circles reshaping the RNA world: From waste to treasure. *Mol Cancer*. 2017;16(1):58. doi:10.1186/s12943-017-0630-y
7. Xi Y, Fowdur M, Liu Y, Wu H, He M, Zhao J. Differential expression and bioinformatics analysis of circRNA in osteosarcoma. *Biosci Rep*. 2019;39(5):BSR20181514. doi:10.1042/BSR20181514
8. Wu Y, Xie Z, Chen J, et al. Circular RNA circTADA2A promotes osteosarcoma progression and metastasis by sponging miR-203a-3p and regulating CREB3 expression. *Mol Cancer*. 2019;18(1):73. doi:10.1186/s12943-019-1007-1
9. Song YZ, Li JF. Circular RNA hsa_circ_0001564 regulates osteosarcoma proliferation and apoptosis by acting miRNA sponge. *Biochem Biophys Res Commun*. 2018;495(3):2369–2375. doi:10.1016/j.bbrc.2017.12.050
10. Esquela-Kerscher A, Slack FJ. Oncomirs: MicroRNAs with a role in cancer. *Nat Rev Cancer*. 2006;6(4):259–269. doi:10.1038/nrc1840
11. Lee YS, Dutta A. MicroRNAs in cancer. *Annu Rev Pathol Mech Dis*. 2009;4(1):199–227. doi:10.1146/annurev.pathol.4.110807.092222
12. Sandiford OA, Moore CA, Du J, et al. Human aging and cancer: Role of miRNA in tumor microenvironment. *Adv Exp Med Biol*. 2018;1056:137–152. doi:10.1007/978-3-319-74470-4_9
13. Saliminejad K, Khorram Khorshid HR, Soleymani Fard S, Ghaffari SH. An overview of microRNAs: Biology, functions, therapeutics, and analysis methods. *J Cell Physiol*. 2019;234(5):5451–5465. doi:10.1002/jcp.27486
14. Rupaimoole R, Slack FJ. MicroRNA therapeutics: Towards a new era for the management of cancer and other diseases. *Nat Rev Drug Discov*. 2017;16(3):203–222. doi:10.1038/nrd.2016.246
15. Mishra S, Yadav T, Rani V. Exploring miRNA based approaches in cancer diagnostics and therapeutics. *Crit Rev Oncol Hematol*. 2016;98:12–23. doi:10.1016/j.critrevonc.2015.10.003
16. Moore DD, Luu HH. Osteosarcoma. *Cancer Treat Res*. 2014;162:65–92. doi:10.1007/978-3-319-07323-1_4
17. Lamora A, Talbot J, Mullard M, Brounais-Le Royer B, Redini F, Verrecchia F. TGF-β signaling in bone remodeling and osteosarcoma progression. *J Clin Med*. 2016;5(11):96. doi:10.3390/jcm5110096
18. Chicón-Bosch M, Tirado OM. Exosomes in bone sarcomas: Key players in metastasis. *Cells*. 2020;9(1):241. doi:10.3390/cells9010241
19. Hadjimichael AC, Foukas AF, Savvidou OD, Mavrogenis AF, Psyrri AK, Papagelopoulos PJ. The anti-neoplastic effect of doxycycline in osteosarcoma as a metalloproteinase (MMP) inhibitor: A systematic review. *Clin Sarcoma Res*. 2020;10(1):7. doi:10.1186/s13569-020-00128-6
20. Vlachostergios PJ, Papandreou CN. The role of the small ubiquitin-related modifier (SUMO) pathway in prostate cancer. *Biomolecules*. 2012;2(2):240–255. doi:10.3390/biom2020240
21. Niu JC, Ma N, Liu W, Wang PJ. EP1 receptor is involved in prostaglandin E2-induced osteosarcoma growth. *Bosn J Basic Med Sci*. 2019;19(3):265–273. doi:10.17305/bjbms.2019.4177

22. Zhang B, Jia WH, Matsuda K, et al. Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. *Nat Genet*. 2014;46(6):533–542. doi:10.1038/ng.2985
23. Chen Y, Chen Q, Zou J, Zhang Y, Bi Z. Construction and analysis of a ceRNA-ceRNA network reveals two potential prognostic modules regulated by hsa-miR-335-5p in osteosarcoma. *Int J Mol Med*. 2018;42(3):1237–1246. doi:10.3892/ijmm.2018.3709
24. Sultan G. Towards the early detection of ductal carcinoma (a common type of breast cancer) using biomarkers linked to the PPAR(γ) signaling pathway. *Bioinformation*. 2019;15(11):799–805. doi:10.6026/97320630015799
25. Hedditch EL, Gao B, Russell AJ, et al. *ABCA* transporter gene expression and poor outcome in epithelial ovarian cancer. *J Nat Cancer Inst*. 2014;106(7):dju149. doi:10.1093/jnci/dju149
26. Hartmann F, Sparla D, Tute E, et al. Low acyl-CoA synthetase 5 expression in colorectal carcinomas is prognostic for early tumour recurrence. *Pathol Res Pract*. 2017;213(3):261–266. doi:10.1016/j.prp.2016.09.002
27. Chen Y, Huang T, Yang X, et al. MicroRNA-106a regulates the proliferation and invasion of human osteosarcoma cells by targeting VNN2. *Oncol Rep*. 2018;40(4):2251–2259. doi:10.3892/or.2018.6601
28. Zhang Y, Zhao X, Deng L, et al. High expression of *FABP4* and *FABP6* in patients with colorectal cancer. *World J Surg Onc*. 2019;17(1):171. doi:10.1186/s12957-019-1714-5
29. Nuylan M, Kawano T, Inazawa J, Inoue J. Down-regulation of *LAPTM5* in human cancer cells. *Oncotarget*. 2016;7(19):28320–28328. doi:10.18632/oncotarget.8614
30. Morotti M, Bridges E, Valli A, et al. Hypoxia-induced switch in SNAT2/SLC38A2 regulation generates endocrine resistance in breast cancer. *Proc Natl Acad Sci U S A*. 2019;116(25):12452–12461. doi:10.1073/pnas.1818521116
31. Winternitz MC, Meloy CR. On the occurrence of catalase inhuman tissues and its variations in diseases. *J Exp Med*. 1908;10(6):759–781. doi:10.1084/jem.10.6.759
32. Hyoudou K, Nishikawa M, Umeyama Y, Kobayashi Y, Yamashita F, Hashida M. Inhibition of metastatic tumor growth in mouse lung by repeated administration of polyethylene glycol-conjugated catalase. *Clin Cancer Res*. 2004;10(22):7685–7691. doi:10.1158/1078-0432.CCR-04-1020
33. Kim SY, Lee CH, Midura BV, et al. Inhibition of the CXCR4/CXCL12 chemokine pathway reduces the development of murine pulmonary metastases. *Clin Exp Metastasis*. 2008;25(3):201–211. doi:10.1007/s10585-007-9133-3
34. Dinić S, Grdović N, Uskoković A, et al. CXCL12 protects pancreatic β-cells from oxidative stress by a Nrf2-induced increase in catalase expression and activity. *Proc Jpn Acad Ser B Phys Biol Sci*. 2016;92(9):436–454. doi:10.2183/pjab.92.436