

Construction of an mRNA-miRNA-lncRNA network prognostic for triple-negative breast cancer

Yuan Huang^{1,*}, Xiaowei Wang^{2,*}, Yiran Zheng³, Wei Chen⁴, Yabing Zheng¹, Guangliang Li¹, Weiyang Lou⁵, Xiaojia Wang¹

¹Department of Breast Medical Oncology, The Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital), Institute of Basic Medicine and Cancer (IBMC), Chinese Academy of Sciences, Zhejiang Province, Hangzhou 310022, China

²Department of Colorectal Surgery, Sir Run Run Shaw Hospital, School of Medicine, Zhejiang University, Zhejiang Province, Hangzhou 310016, China

³School of Pharmaceutical Sciences, Soochow University, Jiangsu Province, Suzhou 215123, China

⁴Siligery Corporation, Zhejiang Province, Hangzhou 310012, China

⁵Department of Breast Surgery, First Affiliated Hospital, School of Medicine, Zhejiang University, Zhejiang Province, Hangzhou 310003, China

*Equal contribution

Correspondence to: Yuan Huang, Weiyang Lou, Xiaojia Wang; **email:** huangyuan@zicc.org.cn, 11718264@zju.edu.cn, wangxj@zjcc.org.cn

Keywords: triple-negative breast cancer, biomarker, competing endogenous RNA, prognosis

Received: May 4, 2020

Accepted: November 13, 2020

Published: January 3, 2021

Copyright: © 2020 Huang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

The aim of this study was to establish a novel competing endogenous RNA (ceRNA) network able to predict prognosis in patients with triple-negative breast cancer (TNBC). Differential gene expression analysis was performed using the GEO2R tool. Enrichr and STRING were used to conduct protein-protein interaction and pathway enrichment analyses, respectively. Upstream lncRNAs and miRNAs were identified using miRNet and mirTarBase, respectively. Prognostic values, expression, and correlational relationships of mRNAs, lncRNAs, and miRNAs were examined using GEPIA, starBase, and Kaplan-Meier plotter. In total, 860 upregulated and 622 downregulated differentially expressed mRNAs were identified in TNBC. Ten overexpressed and two underexpressed hub genes were screened. Next, 10 key miRNAs upstream of these key hub genes were predicted, of which six upregulated miRNAs were significantly associated with poor prognosis and four downregulated miRNAs were associated with good prognosis in TNBC. NEAT1 and MAL2 were selected as key lncRNAs. An mRNA-miRNA-lncRNA network in TNBC was constructed. Thus, we successfully established a novel mRNA-miRNA-lncRNA regulatory network, each component of which is prognostic for TNBC.

INTRODUCTION

Triple-negative breast cancer (TNBC) is immunohistochemically defined as estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and human epidermal growth factor receptor 2 (HER2) nonamplified breast cancer. TNBC composes nearly

15% of all breast cancers [1]. TNBC is aggressive and has the worst prognosis among all breast cancer subtypes. Despite significant advances in breast cancer treatment, TNBC still has a higher relapse rate, shorter overall survival (OS), and limited therapeutic options compared with other subtypes. The median OS for advanced TNBC is about 1 year [2–4]. The molecular

mechanisms involved in the occurrence and development of TNBC are still not clear. Thus, it is imperative to determine the mechanisms of TNBC to uncover valid treatment avenues and new prognostic biomarkers.

Numerous studies have revealed that competitive endogenous RNAs (ceRNAs) are vital molecules that regulate a variety of pathological processes [5–7]. The hypothesis of ceRNA, proposed by Salmena et al. [7], suggests a distinct molecular regulatory mechanism for posttranscriptional regulation. The key noncoding RNAs in this hypothesis are miRNAs, which are usually negative regulators of gene expression. ceRNAs can crosstalk by competing for microRNA binding and form a regulatory network across the transcriptome. Recently, it has been reported that the lncRNA-miRNA-mRNA ceRNA network might be a crucial factor in carcinogenesis and cancer development [8–10]. Through microRNA response elements (MREs), lncRNAs act as ceRNAs that sponge miRNA, thereby influencing gene expression of targeted mRNAs. However, the clinical significance of an lncRNA-miRNA-mRNA ceRNA network in TNBC has not been investigated.

In this study, we examined differentially expressed mRNAs (DE-mRNAs) in TNBC tissues, other breast cancer subtype tissues, and normal tissues by mining two Gene Expression Omnibus (GEO) data sets (GSE45827 and GSE6519). Functional enrichment analysis was conducted for these common DE-mRNAs. Next, we performed a protein-protein interaction (PPI) analysis to identify the hub genes. Combining the results of the expression analysis and prognosis analysis for hub genes in breast cancer, we selected 10 upregulated genes and two downregulated genes for further study. We then identified potential upstream lncRNAs and miRNAs and assessed the expression and prognostic value of these lncRNAs and miRNAs in breast cancer. The correlations between mRNAs, miRNAs, and lncRNAs were investigated according to the ceRNA hypothesis. Thus, we established a novel lncRNA-miRNA-mRNA ceRNA network, the components of which can be used to predict prognosis or serve as treatment targets in TNBC.

RESULTS

Identification of candidate DE-mRNAs in TNBC

We screened the gene expression microarrays related to TNBC in the GEO database, and the GSE45827 and GSE65194 data sets were ultimately included. Next, we identified differentially expressed genes between TNBC tissues, tissues of other subtypes of breast cancer, and

normal breast tissues using GEO2R ($|\log_2FC| > 1$ and P value < 0.05). DE-mRNAs in those data sets are shown as volcano plots in Figure 1A. In the GSE45827 data set, 4760 overexpressed mRNAs and 2422 underexpressed mRNAs were identified in the TNBC tissues relative to normal control samples. In addition, 1225 overexpressed and 1269 underexpressed genes were found in TNBC tissues relative to other breast cancer subtype tissues. In the GSE65194 data set, 4776 mRNAs were overexpressed and 2611 mRNAs were underexpressed in TNBC tissues compared with normal control samples. In addition, 1222 genes were overexpressed and 1244 genes were underexpressed in TNBC tissues compared with other breast cancer subtype tissues. After overlapping the genes, 860 upregulated (Figure 1B) and 622 downregulated (Figure 1C) common genes were identified. These DE-miRNAs are listed in Supplementary Table 1 and were chosen for further study.

GO functional and KEGG pathway enrichment analysis

We performed Gene Ontology (GO) functional annotation and pathway enrichment analysis using the Enrichr database to examine the functions of the DE-mRNAs. GO analysis was conducted as three sublevels: cellular component (CC), biological process (BP), and molecular function (MF). For pathway enrichment, KEGG's cell signaling pathway was conducted. As shown in Figure 2, overexpressed DE-mRNAs were dramatically enriched in terms of cell division and proliferation, such as centromere complex assembly, mitotic sister chromatid segregation, and DNA replication in the BP category; centromeric region, chromosome, and spindle in the CC group; and DNA helicase activity, DNA replication origin binding, and ATPase activity in the MF category. Besides cell cycle and DNA replication, signaling pathways related to cell death appeared in the top 10 enriched KEGG pathways for upregulated DE-mRNAs, including the p53 signaling pathway, mismatch repair, cellular senescence, and ferroptosis. For downregulated DE-mRNAs, the enriched GO functions included regulation of cellular response to growth factor stimulus, insulin-like growth factor receptor signaling pathway, and phosphate ion homeostasis in the BP category; vesicle, platelet granule membrane, and contractile actin filament bundle in the CC category; and insulin-like growth factor receptor binding, insulin receptor binding, oxidoreductase activity, and transmembrane transporter activity in the MF category. Subsequently, KEGG pathway analysis indicated that these downregulated DE-mRNAs were considerably enriched in several pathways associated with many cancers, such as prostate cancer and small-cell lung cancer, and proteoglycans.

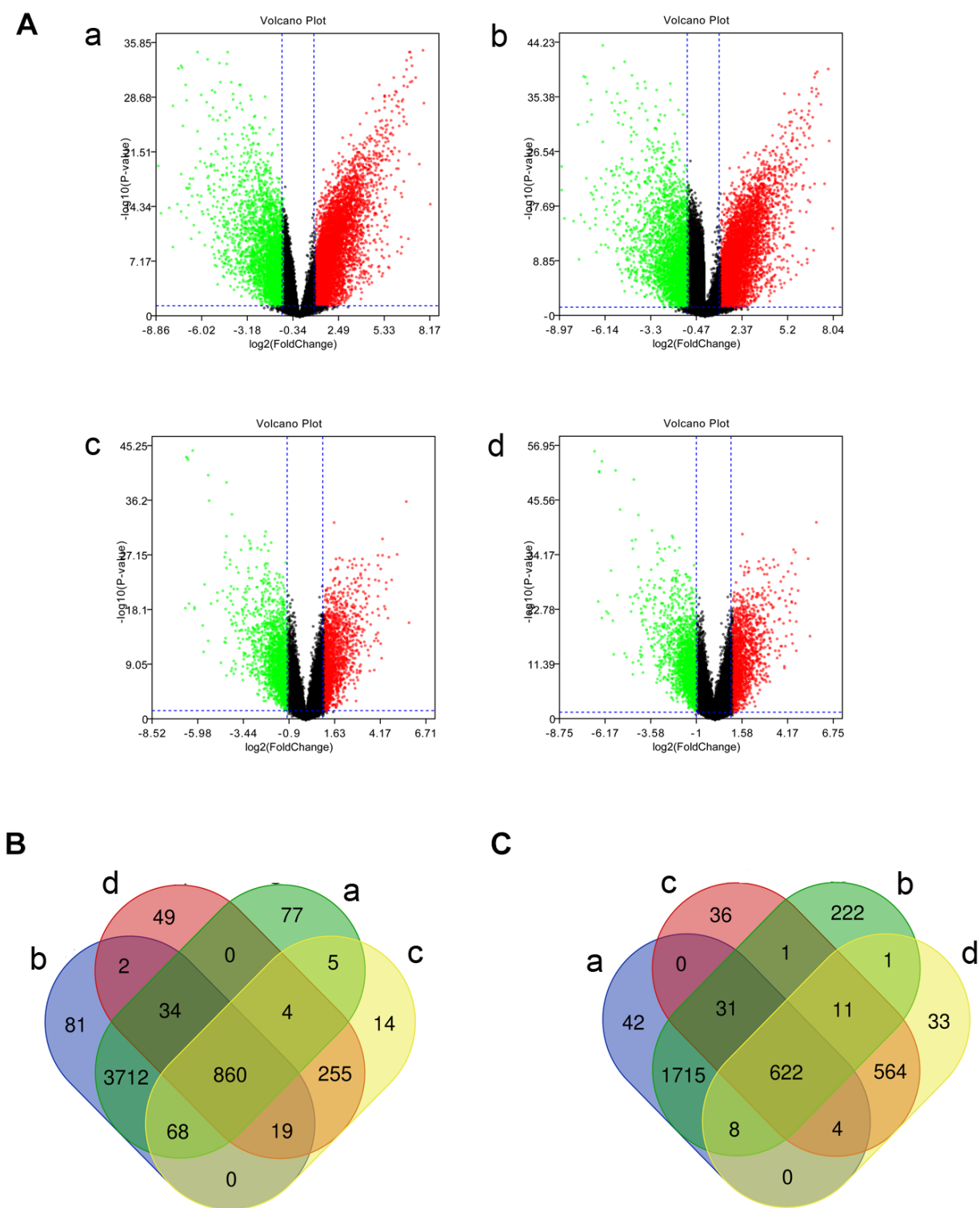


Figure 1. Identified differentially expressed mRNAs (DE-mRNAs) among triple-negative breast cancer (TNBC) tissues, tissues of other types of breast cancer, and normal samples in two Gene Expression Omnibus data sets. (A) The volcano plots of DE-mRNAs in the GSE45827 and GSE65194 data sets. The x-axis stands for \log_2 (fold change) of gene expression, and y-axis represents log-transformed P value. The red dots and green dots indicate the significantly overexpressed and underexpressed genes, respectively. The black dots indicate genes with no significant differential expression. $|\log_2FC| > 1$ and P value < 0.05 were the cutoff criteria. **(B)** The intersection of upregulated DE-mRNAs. **(C)** The intersection of downregulated DE-mRNAs. a: TNBC compared with normal samples in GSE45827; b: TNBC compared with normal samples in GSE65194; c: TNBC compared with tissues of other subtypes of breast cancer in GSE45827; d: TNBC compared with tissues of other subtypes of breast cancer in GSE65194.

Establishment of PPI network and detection of hub genes

To explore the protein interaction networks, we used the STRING database to construct the PPI networks of the identified DE-mRNAs, as shown in Figure 3A, 3B. Based on the node degree, the top 20 hub genes in the downregulated and upregulated DE-mRNAs were identified using Cytoscape software and are listed in Table 1. To visualize better, the interactions of the top 20 hub genes were reconstructed and are presented in Figure 3C, 3D. The top 10 of these 20 hub genes were selected for further analyses.

Analysis of expression and prognosis of hub genes in breast cancer

To further examine the expression of the hub genes in TNBC, the expression of the top 10 upregulated and top

10 downregulated hub genes was analyzed using the GEPIA database. All of the 10 upregulated hub genes (CDK1, CCNB1, CCNA2, CDC20, TOP2A, CCNB2, MAD2L1, BUB1, KIF11, and RRM2) were dramatically upregulated in breast cancer, whereas five of the 10 downregulated hub genes (ESR1, IGF1, PDGFRB, PXN, and ZEB1) were significantly downregulated in breast cancer. The ability of the hub genes to predict prognosis in TNBC was evaluated using the Kaplan-Meier (KM) plotter database. The 10 upregulated hub genes correlated significantly with poor disease outcome. Of the downregulated hub genes, only ESR1 and IGF1 were correlated with favorable prognosis in TNBC. Expression boxplots and survival curves are shown separately in Figure 4. Thus, these 10 overexpressed and two under-expressed hub genes were selected for further investigation. In addition, the expression of the 12 key genes was verified in TNBC samples from The Cancer Genome Atlas (TCGA) (Supplementary Table 2).

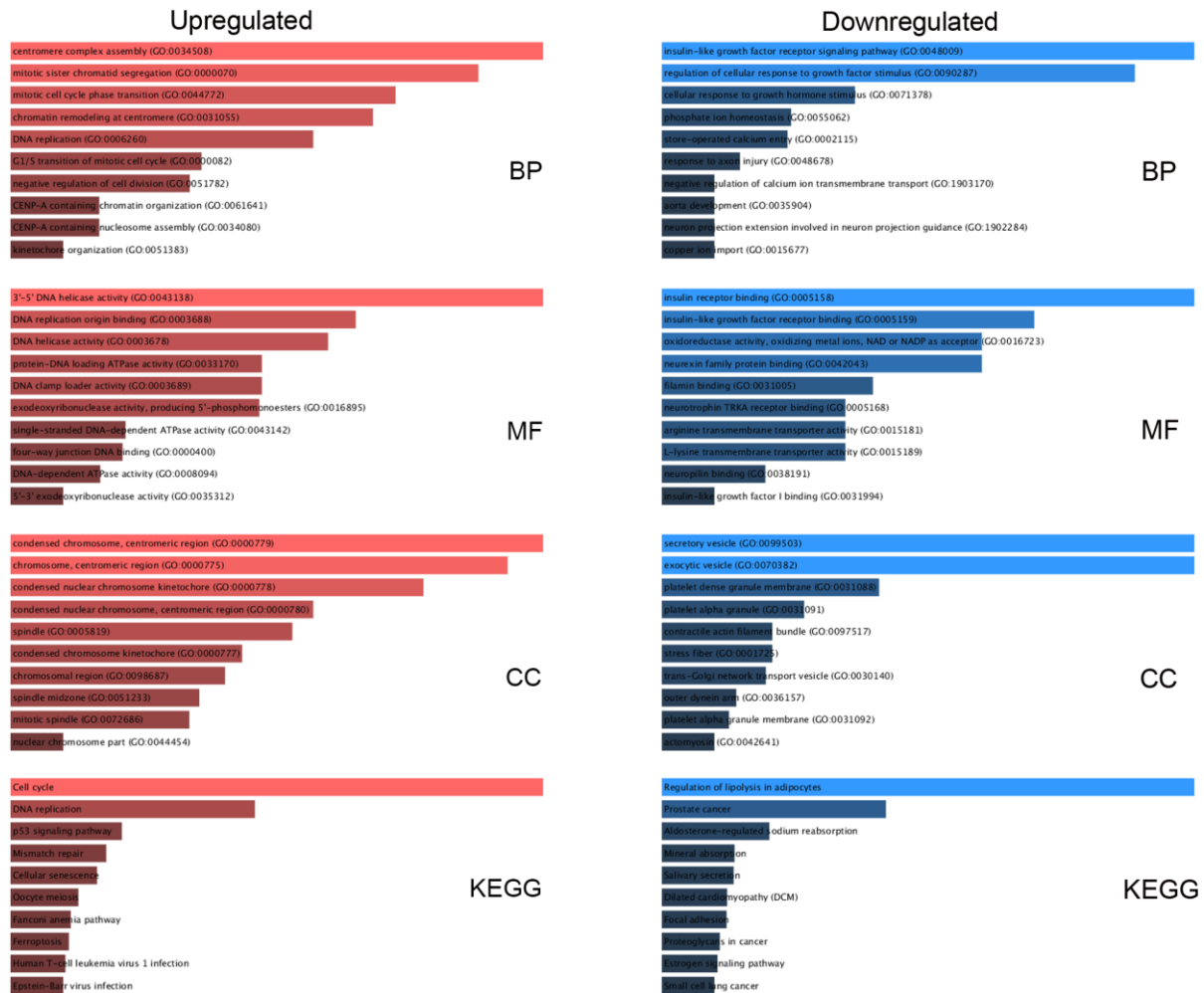


Figure 2. Functional enrichment analysis for key DE-mRNAs. The top 10 enriched molecular function (MF), cellular component (CC), biological process (BP), and KEGG pathways of the upregulated and downregulated significant DE-mRNAs.

Validation and prediction of upstream key miRNAs

Key miRNAs that regulate the 12 identified hub genes were predicted using miRTarBase. In view of the credibility of the predicted results, only microRNA–target gene interactions predicted by reporter assay were selected. As presented in Figure 5 and Supplementary Table 3, 11 miRNAs that could possibly modulate five of the key upregulated genes (CCNA2, MAD2L1, CDK1, RRM2, and CCNB1) and 32 miRNAs that could potentially modulate the two key downregulated genes (ESR1 and IGF1) were identified. Upstream potential miRNAs of five key genes (CDC20, TOP2A, MAD2L1, BUB1, and KIF11) were not included. We then evaluated the expression pattern and prognostic value of

the predicted miRNAs in breast cancer patients using the starBase and KM plotter databases according to the inverse regulation between miRNA and mRNA. Consequently, for upregulated hub genes, four miRNAs (hsa-let-7b-5p, hsa-miR-10b-3p, hsa-let-7a-5p, and hsa-miR-410-3p) were not only downregulated but also linked to favorable disease outcomes in patients with breast cancer. For downregulated hub genes, six miRNAs (hsa-miR-19a-3p, hsa-let-7e-5p, hsa-miR-130b-3p, hsa-miR-98-5p, hsa-miR-18b-5p, and hsa-miR-222-3p) were significantly upregulated and correlated with poor prognosis ($P < 0.05$). The expression boxplots and survival curves of the 10 key miRNAs are provided in Supplementary Figures 1, 2, respectively.

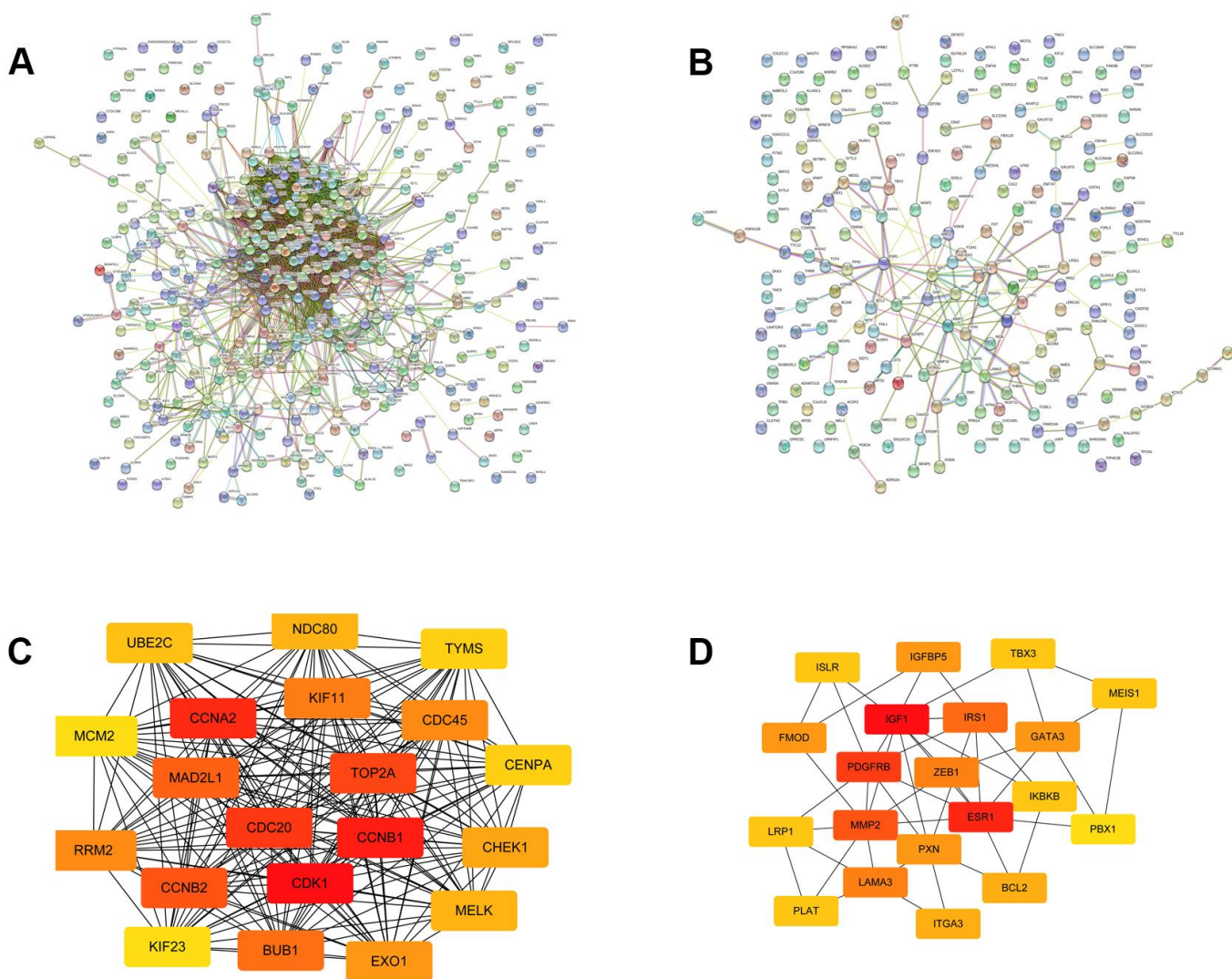


Figure 3. The top 20 hub genes selected from the PPI networks. (A) The PPI network of the upregulated significant DE-mRNAs. **(B)** The PPI network of the downregulated significant DE-mRNAs. **(C)** The top 20 hub genes of the upregulated significant DE-mRNAs. **(D)** The top 20 hub genes of the downregulated significant DE-mRNAs.

Table 1. The top 20 hub genes in the PPI networks.

Overexpressed gene		Underexpressed gene	
Gene symbol	Degree	Gene symbol	Degree
CDK1	113	ESR1	13
CCNB1	108	IGF1	10
CCNA2	105	PDGFRB	8
CDC20	104	PXN	7
TOP2A	103	ZEB1	6
CCNB2	102	MMP2	6
MAD2L1	99	IRS1	6
BUB1	96	BCL2	5
KIF11	94	PPID	4
RRM2	89	ITGA3	3
CDC45	89	RUNX1T1	3
EXO1	86	NEGR1	3
CHEK1	85	PDGFD	3
MELK	84	MEIS1	3
NDC80	84	PBX1	3
UBE2C	83	FRS2	2
TYMS	82	ALCAM	2
CENPA	82	TOX3	2
KIF23	80	WISP2	2
MCM2	80	RTN1	2

Validation and prediction of upstream key lncRNAs

Studies have demonstrated that lncRNAs suppress miRNA expression by acting as miRNA sponges [11–13]. As a result, we predicted the important upstream lncRNAs that could possibly bind to the 10 vital miRNAs (hsa-let-7b-5p, hsa-miR-10b-3p, hsa-let-7a-5p, hsa-miR-410-3p, hsa-miR-19a-3p, hsa-let-7e-5p, hsa-miR-130b-3p, hsa-miR-98-5p, hsa-miR-18b-5p, and hsa-miR-222-3p) using the online miRNet database. In total, 374 lncRNA-miRNA pairs were identified (Supplementary Table 4). According to the ceRNA hypothesis, expression of these lncRNAs was evaluated using the GEPIA database. When compared with normal controls, nine lncRNAs (AC018766.4, CROCCP2, CTD-3092A11.2, LINC00342, RP11-553L6.5, XXbac-B461K10.4, NEAT1, RP11-228B15.4, and RP11-311C24.1) that target to upregulated miRNAs showed lower expression in breast cancer, whereas five lncRNAs (HOTAIR, LINC00467, RECQL4, LINC00665, and MAL2) that target to downregulated miRNAs were significantly upregulated in breast cancer (Supplementary Figure 3). Subsequent survival analysis

using the KM plotter database revealed that patients with low NEAT1 expression and high MAL2 expression had an unfavorable prognosis. Thus, NEAT1 and MAL2 were recognized as key lncRNAs (Figure 6).

Establishment of the mRNA-miRNA-lncRNA regulatory prognostic network in TNBC

We applied bioinformatics to develop a key lncRNA-miRNA-mRNA ceRNA network in TNBC. As depicted in Figure 7, the network includes 11 miRNA-mRNA pairs (miR-19a-3p-ESR1, miR-18b-5p-ESR1, miR-222-3p-ESR1, let-7e-5p-IGF1, miR-130b-3p-IGF1, miR-98-5p-IGF1, miR-18b-5p-IGF1, let-7b-5p-CCNA2, miR-10b-3p-CCNA2, let-7a-5p-RRM2, and miR-410-3p-CCNB1), five miRNA-lncRNA pairs (let-7e-5p-NEAT1, miR-98-5p-NEAT1, let-7a-5p-MAL2, miR-410-3p-MAL2, and let-7b-5p-MAL2), and four mRNA-lncRNA pairs (IGF1-NEAT1, RRM2-MAL2, CCNB1-MAL2, and CCNA2-MAL2). Based on the ceRNA hypothesis, lncRNA works as a ceRNA to compete for shared miRNA and sequester miRNA away from mRNA. LncRNA has an inverse co-expression

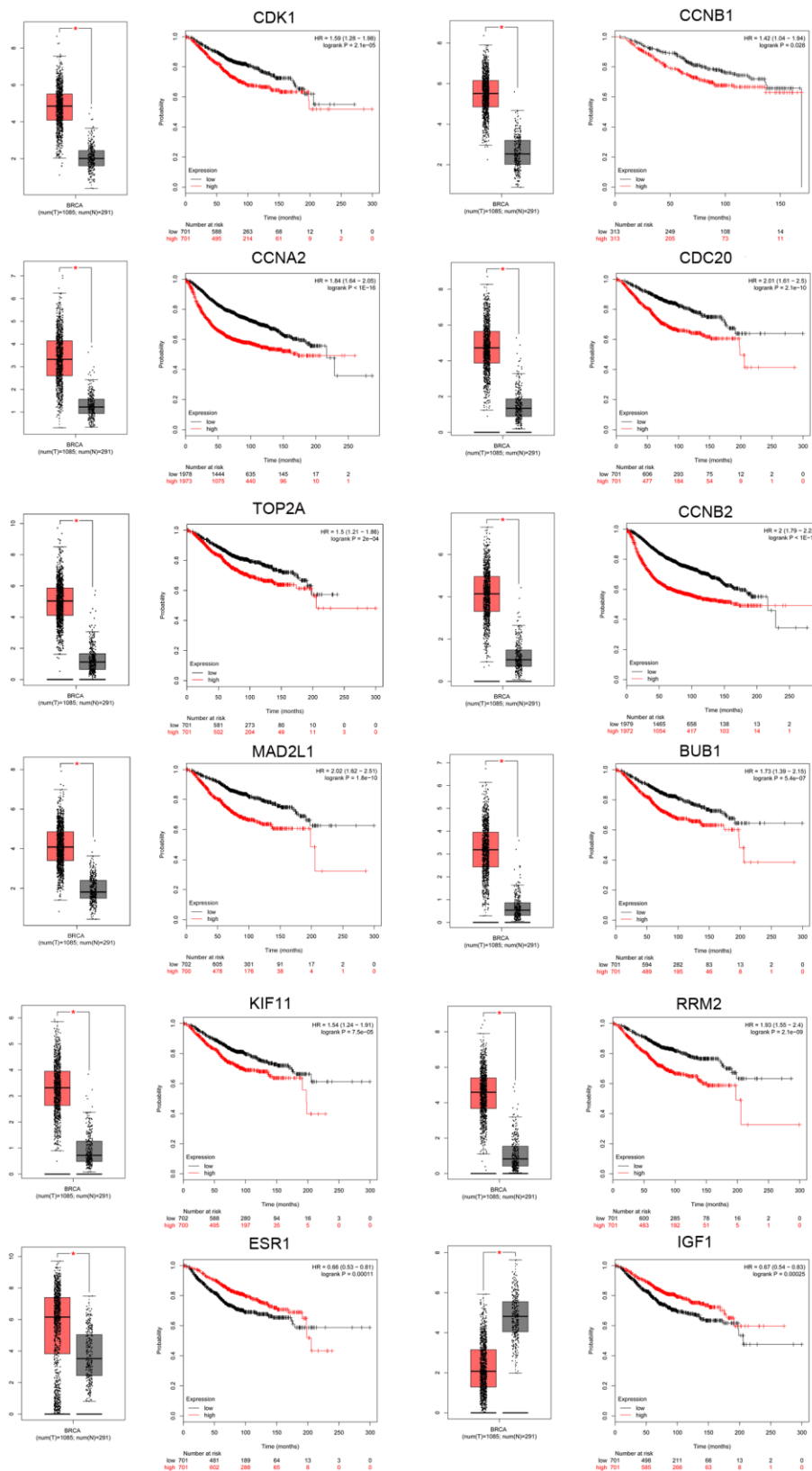


Figure 4. Screening of key genes in TNBC. Key genes were identified from the top 10 hub genes of the significant dysregulated DE-mRNAs by merging the prognosis and expression analyses using Kaplan Meier-plotter and GEPIA databases. Expression boxplots and survival curves (overall survival [OS]) of 12 key genes, including 10 upregulated hub genes (CDK1, CCNB1, CCNA2, CDC20, TOP2A, CCNB2, MAD2L1, BUB1, KIF11, and RRM2) and two downregulated hub genes (ESR1 and IGF1) in TNBC are presented.

relationship with miRNA but a positive co-expression relationship with mRNA. Therefore, the correlations between mRNA-lncRNA, miRNA-lncRNA, and mRNA-miRNA pairs in the constructed network were

assessed using the starBase database, and results are shown in Table 2. Except for one miRNA-lncRNA pair (let-7e-5p-NEAT1), the other pairs were all fitted with the ceRNA mechanism. By considering the three levels,

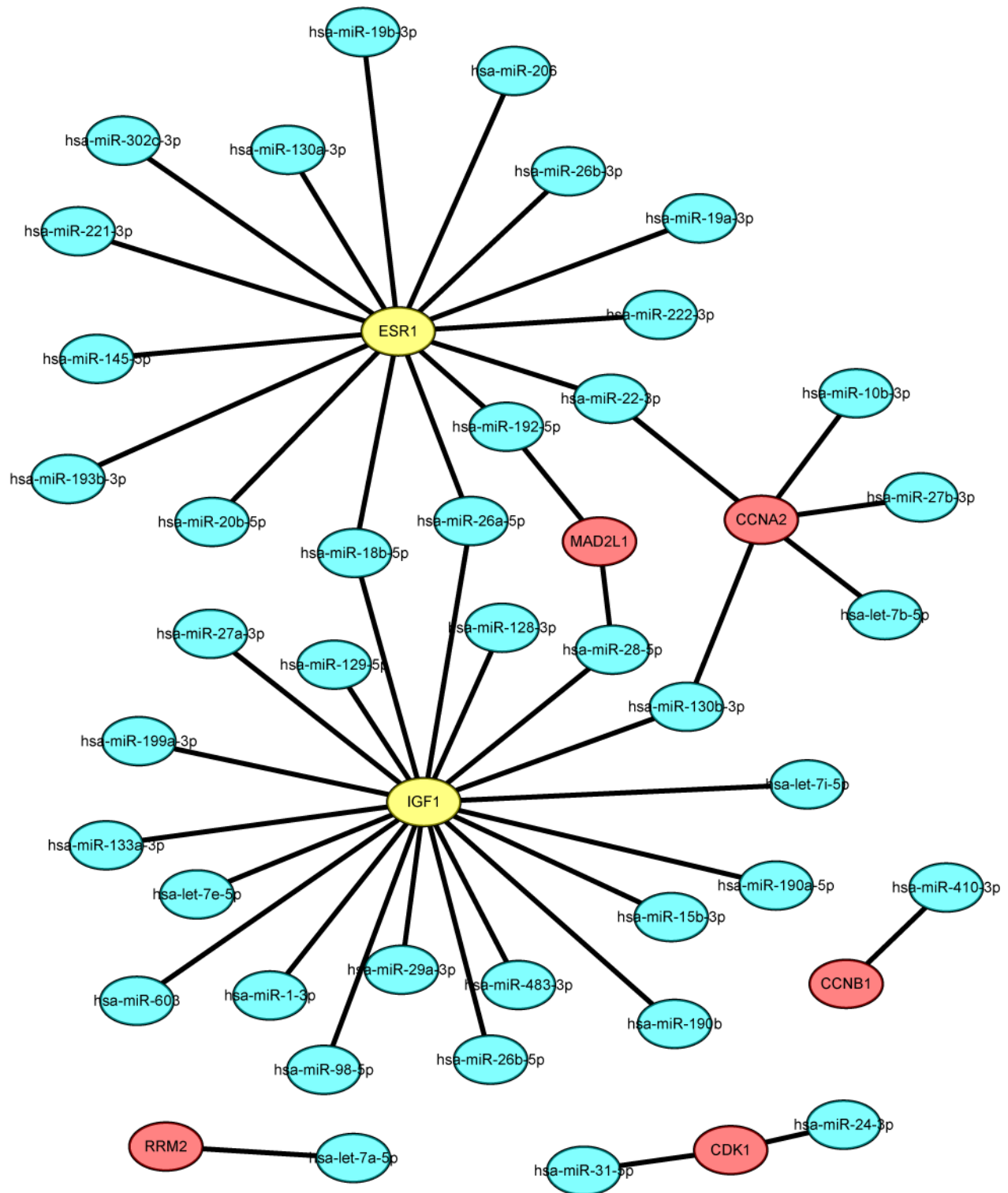


Figure 5. The mRNA-miRNA network established by Cytoscape software.

a novel mRNA-miRNA-lncRNA triple subnetwork, including four mRNA-miRNA-lncRNA axes (IGF1-miR-98-5p-NEAT1, RRM2-let-7a-5p-MAL2, CCNB1-miR-410-3p-MAL2, and CCNA2-let-7b-5p-MAL2), was ultimately constructed and possessed significant prognostic value in TNBC.

DISCUSSION

Breast cancer is associated with high mortality in women, and incidence is increasing worldwide. TNBC is the most fatal subtype of breast cancer. In addition, the therapeutic options for TNBC are limited.

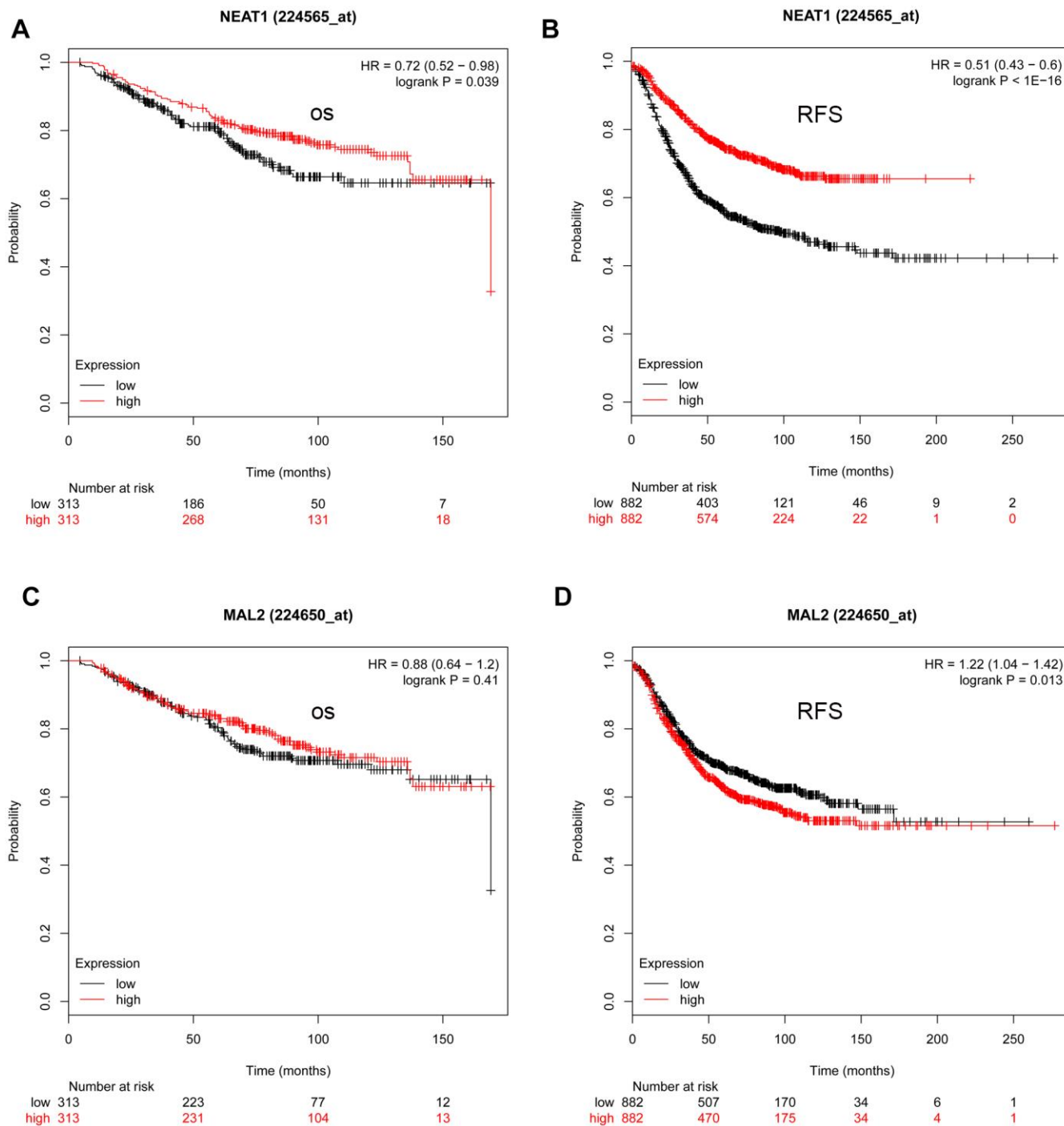


Figure 6. The prognostic values of NEAT1 and MAL2 in TNBC determined by the Kaplan-Meier plotter. (A) The prognostic value (overall survival [OS]) of NEAT1 in TNBC. (B) The prognostic value (relapse-free survival [RFS]) of NEAT1 in TNBC. (C) The prognostic value (OS) of MAL2 in TNBC. (D) The prognostic value (RFS) of MAL2 in TNBC.

Chemotherapy remains the primary therapy for advanced TNBC. Because it is believed that development of TNBC is regulated by sophisticated signaling networks, revealing the specific molecular mechanism of TNBC could result in development of effective treatments and novel prognostic biomarkers.

Recently, a novel mechanistic function of lncRNAs, known as ceRNAs, has been reported. lncRNAs sponge miRNAs to regulate their expression, thereby contributing to various pathological processes, including the development of cancer [12–14]. For example, Zhang et al. reported that a downregulated lncRNA, MT1JP, affected the progression of gastric cancer by competitively binding to miR-92a-3p and regulating FBXW7 expression [15]. A study by Wang et al. determined that the lncRNA HOXD-AS1 can bind to miR-130a-3p and inhibit SOX4 degradation, thus activating MMP2 and EZH2 expression and promoting the metastasis of hepatocellular carcinoma [16]. Wu et al. observed that, in papillary thyroid carcinoma, the lncRNA SNHG15 can serve as a ceRNA and thus

modulate YAP1-Hippo signal transduction by sponging miR-200a-3p [17]. Luan et al. found that there is crosstalk between the lncRNA XLOC_006390 and miR-338-3p and miR-331-3p expression, which aggravates cervical cancer [18]. Studies have also suggested that ceRNA influences the pathogenesis of breast cancer. Zhao et al. suggested that the lncRNA TUSC8 may affect epithelial-mesenchymal transition (EMT)-associated protein levels by functioning as a ceRNA of myosin regulatory light chain interacting protein (MYLIP) as it binds miR-190b-5p, leading to the suppression of breast cancer progression [19]. Lu et al. found that lncARAP1-AS promotes tumorigenesis by enhancing the proliferative and migratory abilities of breast cancer cells by modulating the miR-2110/HDAC2/PLIN1 axis [20]. However, there are different breast cancer subtypes, including HER2-positive breast cancer, luminal A and B breast cancer, and TNBC, and comprehensive studies on the ceRNA networks in each breast cancer subtype are limited, particularly regarding the identification of their specific molecular mechanisms.

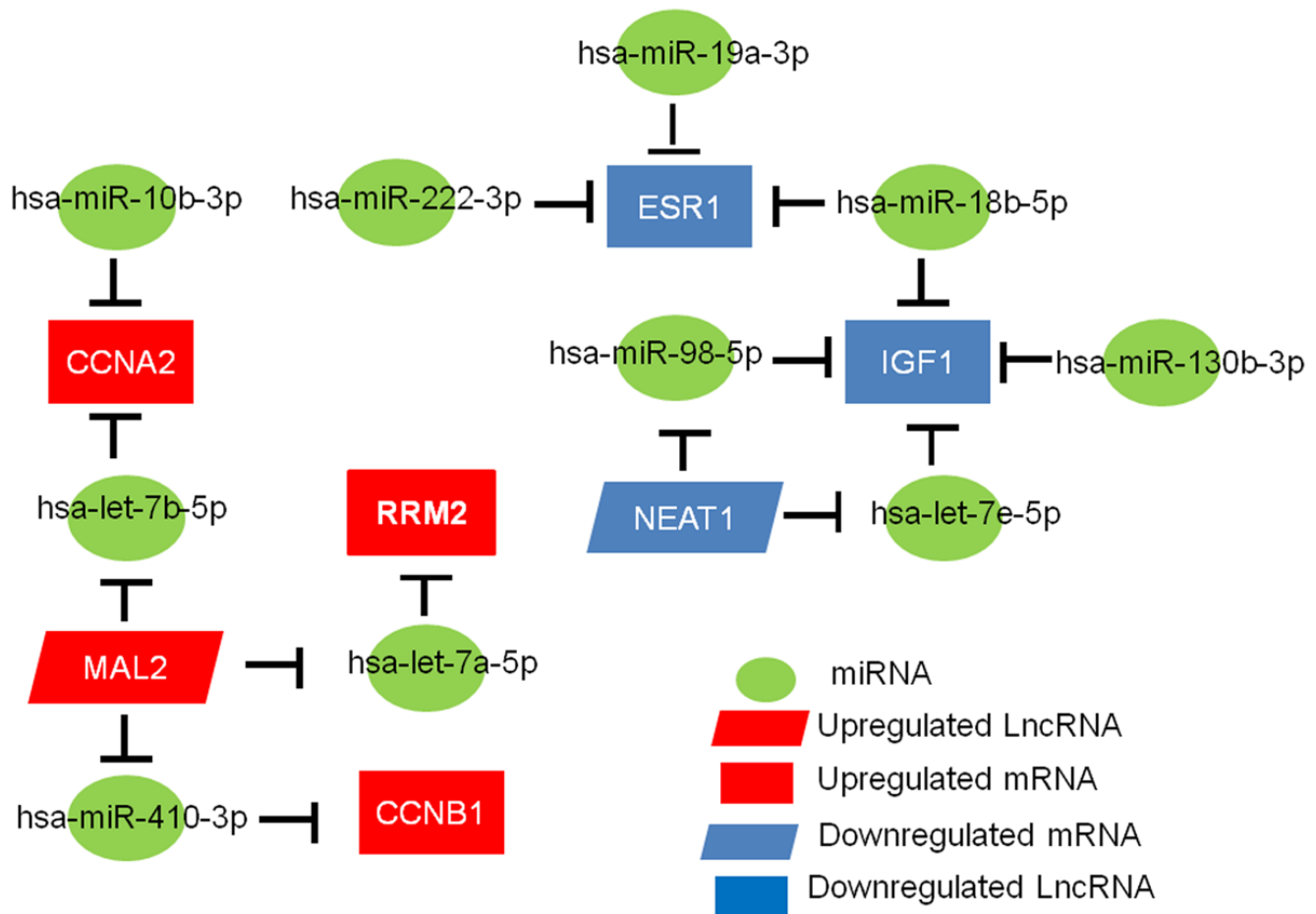


Figure 7. The mRNA-miRNA-lncRNA competing endogenous RNA (ceRNA) network related to the prognosis of TNBC.

Table 2. The correlation between mRNA, miRNA, and lncRNA according to the starBase database.

miRNA	mRNA	R	P value
hsa-miR-19a-3p	ESR1	-0.426	5.47e-49
hsa-miR-18b-5p	ESR1	-0.276	1.85e-20
hsa-miR-222-3p	ESR1	-0.306	6.71e-25
hsa-let-7e-5p	IGF1	-0.058	5.80e-2
hsa-miR-130b-3p	IGF1	-0.304	1.25e-24
hsa-miR-98-5p	IGF1	-0.182	1.58e-9
hsa-miR-18b-5p	IGF1	-0.092	2.45e-3
hsa-let-7b-5p	CCNA2	-0.239	1.62e-15
hsa-miR-10b-3p	CCNA2	-0.301	3.88e-24
hsa-let-7a-5p	RRM2	-0.306	5.49e-25
hsa-miR-410-3p	CCNB1	-0.189	3.49e-10
miRNA	lncRNA	R	P value
hsa-let-7e-5p	NEAT1	0.057	5.85e-02
hsa-miR-98-5p	NEAT1	-0.198	5.24e-11
hsa-let-7a-5p	MAL2	-0.066	2.96e-02
hsa-miR-410-3p	MAL2	-0.144	2.01e-06
hsa-let-7b-5p	MAL2	-0.007	8.28e-1
mRNA	lncRNA	R	P value
IGF1	NEAT1	0.142	2.04e-06
RRM2	MAL2	0.139	3.77e-06
CCNB1	MAL2	0.223	7.46e-14
CCNA2	MAL2	0.207	4.11e-12

In this study, we identified 1482 DE-mRNAs in TNBC and then constructed a unique lncRNA-miRNA-mRNA network based on the ceRNA hypothesis. To our knowledge, we are the first to identify the specific ceRNA network involved in TNBC using a stepwise reverse prediction from mRNA to lncRNA, instead of using the lncRNA-miRNA-mRNA pattern. In addition, each component in this network is significantly associated with breast cancer prognosis, providing potential therapeutic targets and prognostic biomarkers.

First, we screened for significantly dysregulated mRNAs in TNBC tissues compared with normal control tissue and non-TNBC breast cancer tissue in two GEO data sets (GSE36259 and GSE42568). A total of 860 upregulated mRNAs and 622 downregulated mRNAs were selected as specific DE-mRNAs in TNBC. GO analysis [21] indicated that these DE-mRNAs are significantly enriched in some factors of cell division, proliferation, and response, such as centromere complex assembly [22], mitotic sister chromatid segregation and

DNA replication [23], insulin-like growth factor receptor signaling pathway [24], and transmembrane transporter activity [25, 26]. Subsequently, KEGG pathway analysis indicated that these dysregulated DE-mRNAs were remarkably enriched in damage repair, cell death, and cancer-related pathways, including the p53 signaling pathway [27, 28], mismatch repair [29, 30], cellular senescence [31], ferroptosis [32, 33], prostate cancer, and small-cell lung cancer. Therefore, these DE-mRNAs identified through the intersection of two GEO data sets were found to be closely associated with the pathogenesis of TNBC.

To explore the potential association among these identified DE-mRNAs, two PPI networks were established via the STRING database and exhibited a variety of interactions among the DE-mRNAs, especially in the upregulated group. Genes that have a greater node degree in the PPI network typically play vital roles. Thus, the hub genes in the two PPI networks we screened had node degree calculated by Cytoscape

software. In addition, the top 10 overexpressed and underexpressed hub genes were chosen for survival and expression analyses to identify key genes in TNBC. All 10 upregulated hub genes (CDK1, CCNB1, CCNA2, CDC20, TOP2A, CCNB2, MAD2L1, BUB1, KIF11, and RRM2) and two downregulated hub genes (ESR1 and IGF1) were found to affect progression of TNBC (ie, they are key genes). Interestingly, based on previous reports, in addition to breast cancer, most of these hub genes are also related to other cancers. For example, in pancreatic ductal adenocarcinoma, overexpression of CDK1 indicates poor prognosis [34], and CCNA2 is considered a possible biomarker for progression (eg, growth and apoptosis) of colorectal cancer [35]. Evidence indicates that CCNB1 can predict the prognosis of ER-positive breast cancer, as well as the efficacy of hormonal therapy [36]. ESR1 mutation results in acquired endocrine resistance in breast cancer [37].

We initially predicted the miRNAs of the hub genes as described previously based on the ceRNA mechanism. After the expression and survival analyses, we identified 10 key miRNAs, among which six upregulated miRNAs (hsa-let-7e-5p, hsa-miR-19a-3p, hsa-miR-130b-3p, hsa-miR-18b-5p, hsa-miR-98-5p, and hsa-miR-222-3p) were significantly associated with poor prognosis and four downregulated miRNAs (hsa-let-7b-5p, hsa-miR-10b-3p, hsa-let-7a-5p, and hsa-miR-410-3p) were associated with better prognosis in breast cancer. Aberrant expression patterns of these key miRNAs contribute to the progression and prognosis of many cancers. For instance, the upregulation of MiR-98-5p enhances the progression of non-small-cell lung cancer, as well as breast cancer [38, 39]. In addition, the upregulation of miR-18a-5p enhances the ability of lung adenocarcinoma cells to proliferate via the miR-18b-5p/VMA21 axis [40]. The miR-18b-5p/DOCK4 axis inhibits the EMT and migratory capacity of breast cancer cells [41]. Let-7b-5p suppresses the motility and proliferation of squamous cell carcinoma cells [42]. miR-10b-3p expression facilitates the pathogenesis of liver cancer by interacting with CMTM5 [43]. Yang et al. conducted bioinformatics analysis and reported that miR-10b-3p can act as a prognostic biomarker in colorectal cancer [44]. MiR-130b-5p was also found to contribute to the occurrence of pancreatic ductal adenocarcinoma [45].

In our study, we conducted further analysis to identify upstream lncRNAs for the key miRNAs. By combining survival analysis and expression validation, only two lncRNAs (NEAT1 and MAL2) were selected as the key lncRNAs. Overexpression of NEAT1 has been observed in various tumors and has been associated with tumorigenesis and poor prognosis [46–48]. Li et al. implicated

the ER α -NEAT1-FOXN3/NEAT1/SIN3A-GATA3 axis in the metastasis of breast cancer [49]. The oncogenic effects of NEAT1 have been reported to be influenced by the CDC5L-AGRN transcriptional regulation circuit in prostate cancer [50]. Some studies have shown that MAL2 functions as an oncogene in various malignant tumors [51]. Bhandari et al. demonstrated that MAL2 is elevated in breast cancer tissues and that the upregulation of MAL2 was related to the lowest OS rate in the TCGA cohort, suggesting that MAL2 could be an oncogene for breast cancer [52].

This study has some limitations. After identification of candidate DE-mRNAs in TNBC tissues, the screening of key mRNAs, miRNAs, and lncRNAs was performed according to their expression and prognostic values among the whole breast cancer group instead of the TNBC subgroup. Expressions of 12 key genes was further confirmed in TNBC samples from TCGA. The limited TNBC samples in both the GEO and TCGA data sets limited our ability to perform statistically significant analysis of the expression and prognosis of key genes in TNBC. Therefore, we did not stratify the tumor samples and predicted them based on all breast cancers. In addition, both the expression and prognostic values were considered during screening, which could have resulted in our missing some valuable molecules.

In conclusion, we successfully established a novel mRNA-miRNA-lncRNA regulatory network in TNBC using comprehensive bioinformatics analysis. Each component of the network has significant prognostic predictive value for TNBC. Co-expression analysis for all of the RNA pairs in the network revealed that four mRNA-miRNA-lncRNA axes (IGF1-miR-98-5p-NEAT1, RRM2-let-7a-5p-MAL2, CCNB1-miR-410-3p-MAL2, and CCNA2-let-7b-5p-MAL2) act as key ceRNA subnetworks and can be used to predict prognosis or serve as treatment targets in TNBC. Experimental validation of the network should be conducted in future trials.

MATERIALS AND METHODS

Gene expression profile data

To compare genome-wide gene expression profiles among TNBC tissues, non-TNBC tissues, and normal tissues, data sets from the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) database were searched according to our selection criteria. Data sets containing the mRNA expression profiling on triple-negative primary breast cancer (TNPBC) tissues, non-TNPBC tissues such as luminal and HER2-positive breast cancer, and normal tissues were exclusively enrolled. Each of the three groups contains no less than 10 samples. Details of the

data sets, including the summary and overall design, were further evaluated. Finally, the GSE45827 (including 41 TNBC samples, 89 non-TNBC samples, and 11 normal tissue samples) and GSE65194 (including 55 TNBC samples, 98 non-TNBC samples, and 11 normal tissue samples) data sets were chosen for subsequent study. To enhance the reliability of screening results, the TNBC TCGA data sets were added as validation data sets.

Differential expression screening

Upregulated or downregulated DE-mRNAs in the two selected data sets between TNBC tissues and the other two groups (non-TNBC tissues and normal tissues) were assessed using the online analytic tool GEO2R (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>), and the cutoff criteria were set as $|\log_2FC| > 1$ and P value < 0.05 when analyzing differential expression. Next, Venn diagrams were constructed using VENNY 2.1.0 (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>). In all four differential expression analyses, the commonly dysregulated DE-mRNAs were redefined as the significant DE-mRNAs and were chosen for further study.

The clinical and RNA sequencing raw data from 113 normal samples and 1109 breast cancer tumor samples were obtained from TCGA database, and 149 TNBC tissues and 12 paired paracancer tissues were selected. The raw data were first standardized using the method of $\log_2(x + 1)$. Then, the normalized processing of data was conducted using the normalize Between Array function from R package LIMMA. Subsequently, the LIMMA package in R (version 3.4.1) was used to identify DE-mRNAs between TNBC tissues and normal breast tissues in TCGA data sets, and the cutoff criteria were set as adjusted $P < 0.05$ and $|\log_2FC| > 1$.

Functional enrichment analysis

For the commonly identified DE-mRNAs, GO functional annotation and analysis of biological pathways based on the KEGG pathway database were conducted using Enrichr, a comprehensive gene set enrichment analysis (<http://amp.pharm.mssm.edu/Enrichr/>). The top 10 enriched GO items and pathways were visualized and downloaded directly from the webpage.

Construction of a PPI network and screening for hub genes

To assess the interactions between DE-mRNAs, the PPI interaction networks were explored using the Search Tool for the Retrieval of Interacting Genes (STRING)

database (<https://string-db.org/>) [53]. The interactors with a combined confidence score ≥ 0.4 were used to generate the PPI network, and the comprehensive interaction pair information of these differentially expressed genes was downloaded from STRING. Based on the connection degree, the hub genes in the PPI networks were subsequently selected using the Cytoscape plugin, CytoHubba [54]. Finally, the top 20 genes of the common DE-mRNAs were found in Cytoscape (Version 3.6.1), identified as hub genes, and ranked by node degree.

Gene expression analysis

The Gene Expression Profiling Interactive Analysis (GEPIA; <http://gepia.cancer-pku.cn/detail.php>) database contains RNA sequencing expression data from 8587 normal samples and 9736 tumors obtained from Genotype-Tissue Expression data set projects and TCGA [55]. We used the GEPIA database, which contains 1085 breast cancer samples and 291 normal controls, to identify DE-mRNAs and lncRNAs in TNBC. $P < 0.05$ was chosen as the cutoff value for statistical significance.

Survival analysis

Prognostic values of mRNAs, lncRNAs, and miRNAs in TNBC were analyzed using the Kaplan-Meier (KM) plotter database (<http://kmplot.com>) [56], a commonly used website tool to simultaneously integrate gene expression and clinical data retrieved from the GEO, EGA, and TCGA. Subsequently, the potential of these biomarkers to determine the prognosis of cancer was examined. First, the mRNAs, lncRNAs, and miRNAs were entered into the database. Next, KM graphs were generated using the KM plotter database. The log-rank P value and hazard ratio (HR) with 95% confidence intervals were calculated. Log-rank $P < 0.05$ was considered statistically significant.

Prediction of miRNA

To predict the upstream miRNAs of key DE-mRNAs, we obtained miRNA-mRNA interactions from miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw>) in which the interactions between miRNA targets were confirmed using reporter assay, quantitative polymerase chain reaction, next-generation sequencing, western blot, and microarray experiments [57]. To identify more reliable candidate results, only the miRNA-target interactions verified by reporter assay were collected for further study. Then, the expression of the target miRNA in breast cancer and normal tissues was exported using starBase v3.0, an online platform to investigate the differential expression

analysis of miRNAs data from the TCGA [58]. Statistical significance was set at $P < 0.05$.

Prediction of lncRNA

The miRNet database, which incorporates data from 11 integrated microRNA databases, was used to determine the upstream lncRNAs of miRNA [59]. Selection criteria were “target type-lncRNAs” and “Organism-H.sapiens.” Expression of the candidate target lncRNAs was further assessed using the GEPIA database.

Correlation analysis

After miRNA-mRNA, mRNA-lncRNA, and miRNA-lncRNA pairs were obtained, the starBase database (<http://starbase.sysu.edu.cn/>), which contains ncRNA data sets, was used to investigate the correlations [58] between the pairs in invasive breast carcinoma. $P < 0.05$ was set as the cutoff for statistical significance.

Data availability

Additional data for this work can be obtained from the corresponding author upon request.

AUTHOR CONTRIBUTIONS

YH, XWW, YRZ, WC, XJW, and WYL were involved in study design, data analysis, and manuscript writing; YBZ and GLL reviewed the manuscript. All the authors read and approved the final version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

FUNDING

This work was financially supported by the National Natural Science Foundation of China (Grant No. NSFC-81502618), Zhejiang Provincial Natural Science Foundation of China (Grant No. LQ16H160012), and Zhejiang Provincial Research Center for Cancer Intelligent Diagnosis and Molecular Technology (JBZX-202003).

REFERENCES

1. Foulkes WD, Smith IE, Reis-Filho JS. Triple-negative breast cancer. *N Engl J Med*. 2010; 363:1938–48. <https://doi.org/10.1056/NEJMra1001389> PMID:21067385
2. Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P, Narod SA. Triple-negative breast cancer: clinical features and patterns of recurrence. *Clin Cancer Res*. 2007; 13:4429–34. <https://doi.org/10.1158/1078-0432.CCR-06-3045> PMID:17671126
3. Kassam F, Enright K, Dent R, Dranitsaris G, Myers J, Flynn C, Fralick M, Kumar R, Clemons M. Survival outcomes for patients with metastatic triple-negative breast cancer: implications for clinical practice and trial design. *Clin Breast Cancer*. 2009; 9:29–33. <https://doi.org/10.3816/CBC.2009.n.005> PMID:19299237
4. Kumar P, Aggarwal R. An overview of triple-negative breast cancer. *Arch Gynecol Obstet*. 2016; 293:247–69. <https://doi.org/10.1007/s00404-015-3859-y> PMID:26341644
5. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature*. 2014; 505:344–52. <https://doi.org/10.1038/nature12986> PMID:24429633
6. Smillie CL, Sirey T, Ponting CP. Complexities of post-transcriptional regulation and the modeling of ceRNA crosstalk. *Crit Rev Biochem Mol Biol*. 2018; 53:231–45. <https://doi.org/10.1080/10409238.2018.1447542> PMID:29569941
7. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the rosetta stone of a hidden RNA language? *Cell*. 2011; 146:353–58. <https://doi.org/10.1016/j.cell.2011.07.014> PMID:21802130
8. Jiang T, Guo J, Hu Z, Zhao M, Gu Z, Miao S. Identification of potential prostate cancer-related pseudogenes based on competitive endogenous RNA network hypothesis. *Med Sci Monit*. 2018; 24:4213–39. <https://doi.org/10.12659/MSM.910886> PMID:29923546
9. Liu Y, Xie D, He Z, Zheng L. Integrated analysis reveals five potential ceRNA biomarkers in human lung adenocarcinoma. *PeerJ*. 2019; 7:e6694. <https://doi.org/10.7717/peerj.6694> PMID:31106044
10. Wang H, Xu D, Huang H, Cui Y, Li C, Zhang C, Guo S, Zhang L, Xu X, Xu J, Lu J, Wang L, Li K. Detection of dysregulated competing endogenous RNA modules associated with clear cell kidney carcinoma. *Mol Med Rep*. 2018; 18:1963–72. <https://doi.org/10.3892/mmr.2018.9189> PMID:29956728
11. Bak RO, Mikkelsen JG. miRNA sponges: soaking up miRNAs for regulation of gene expression. *Wiley Interdiscip Rev RNA*. 2014; 5:317–33. <https://doi.org/10.1002/wrna.1213> PMID:24375960

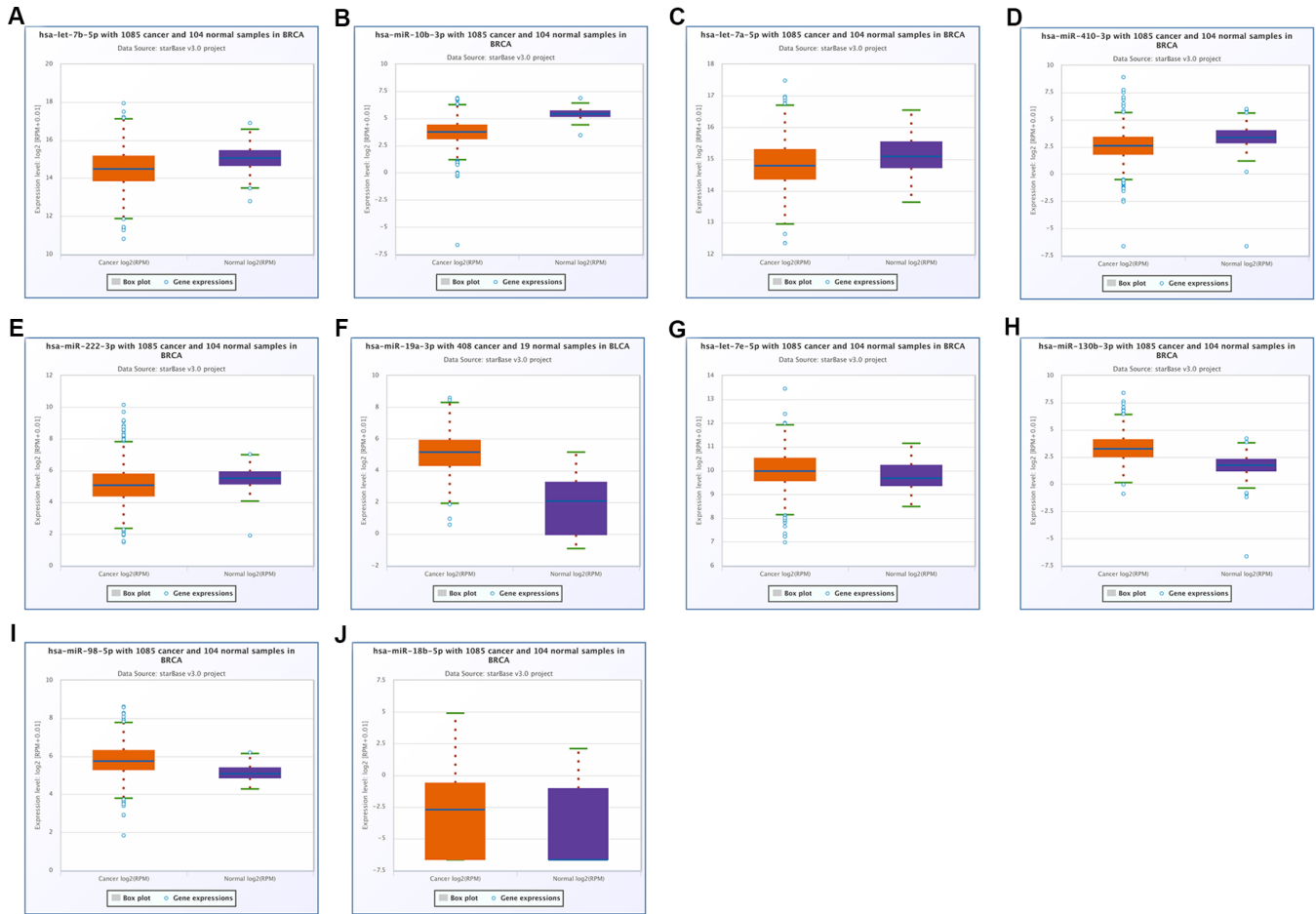
12. Chan JJ, Tay Y. Noncoding RNA:RNA regulatory networks in cancer. *Int J Mol Sci.* 2018; 19:1310. <https://doi.org/10.3390/ijms19051310> PMID:[29702599](https://pubmed.ncbi.nlm.nih.gov/29702599/)
13. Kondo Y, Shinjo K, Katsushima K. Long non-coding RNAs as an epigenetic regulator in human cancers. *Cancer Sci.* 2017; 108:1927–33. <https://doi.org/10.1111/cas.13342> PMID:[28776911](https://pubmed.ncbi.nlm.nih.gov/28776911/)
14. Tam C, Wong JH, Tsui SK, Zuo T, Chan TF, Ng TB. LncRNAs with miRNAs in regulation of gastric, liver, and colorectal cancers: updates in recent years. *Appl Microbiol Biotechnol.* 2019; 103:4649–77. <https://doi.org/10.1007/s00253-019-09837-5> PMID:[31062053](https://pubmed.ncbi.nlm.nih.gov/31062053/)
15. Zhang G, Li S, Lu J, Ge Y, Wang Q, Ma G, Zhao Q, Wu D, Gong W, Du M, Chu H, Wang M, Zhang A, Zhang Z. LncRNA MT1JP functions as a ceRNA in regulating FBXW7 through competitively binding to miR-92a-3p in gastric cancer. *Mol Cancer.* 2018; 17:87. <https://doi.org/10.1186/s12943-018-0829-6> PMID:[29720189](https://pubmed.ncbi.nlm.nih.gov/29720189/)
16. Wang H, Huo X, Yang XR, He J, Cheng L, Wang N, Deng X, Jin H, Wang N, Wang C, Zhao F, Fang J, Yao M, et al. STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4. *Mol Cancer.* 2017; 16:136. <https://doi.org/10.1186/s12943-017-0680-1> PMID:[28810927](https://pubmed.ncbi.nlm.nih.gov/28810927/)
17. Wu DM, Wang S, Wen X, Han XR, Wang YJ, Shen M, Fan SH, Zhang ZF, Shan Q, Li MQ, Hu B, Lu J, Chen GQ, Zheng YL. LncRNA SNHG15 acts as a ceRNA to regulate YAP1-hippo signaling pathway by sponging miR-200a-3p in papillary thyroid carcinoma. *Cell Death Dis.* 2018; 9:947. <https://doi.org/10.1038/s41419-018-0975-1> PMID:[30237435](https://pubmed.ncbi.nlm.nih.gov/30237435/)
18. Luan X, Wang Y. LncRNA XLOC_006390 facilitates cervical cancer tumorigenesis and metastasis as a ceRNA against miR-331-3p and miR-338-3p. *J Gynecol Oncol.* 2018; 29:e95. <https://doi.org/10.3802/jgo.2018.29.e95> PMID:[30207103](https://pubmed.ncbi.nlm.nih.gov/30207103/)
19. Zhao L, Zhou Y, Zhao Y, Li Q, Zhou J, Mao Y. Long non-coding RNA TUSC8 inhibits breast cancer growth and metastasis via miR-190b-5p/MYLIP axis. *Aging (Albany NY).* 2020; 12:2974–91. <https://doi.org/10.18632/aging.102791> PMID:[32039833](https://pubmed.ncbi.nlm.nih.gov/32039833/)
20. Lu C, Wang X, Zhao X, Xin Y, Liu C. Long non-coding RNA ARAP1-AS1 accelerates cell proliferation and migration in breast cancer through miR-2110/HDAC2/PLIN1 axis. *Biosci Rep.* 2020; 40:BSR20191764. <https://doi.org/10.1042/BSR20191764> PMID:[32110804](https://pubmed.ncbi.nlm.nih.gov/32110804/)
21. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, et al. Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat Genet.* 2000; 25:25–29. <https://doi.org/10.1038/75556> PMID:[10802651](https://pubmed.ncbi.nlm.nih.gov/10802651/)
22. Schalch T, Steiner FA. Structure of centromere chromatin: from nucleosome to chromosomal architecture. *Chromosoma.* 2017; 126:443–55. <https://doi.org/10.1007/s00412-016-0620-7> PMID:[27858158](https://pubmed.ncbi.nlm.nih.gov/27858158/)
23. Hillers KJ, Jantsch V, Martinez-Perez E, Yanowitz JL. Meiosis. *WormBook.* 2017; 2017:1–43. <https://doi.org/10.1895/wormbook.1.178.1> PMID:[26694509](https://pubmed.ncbi.nlm.nih.gov/26694509/)
24. Hakuno F, Takahashi SI. IGF1 receptor signaling pathways. *J Mol Endocrinol.* 2018; 61:T69–86. <https://doi.org/10.1530/JME-17-0311> PMID:[29535161](https://pubmed.ncbi.nlm.nih.gov/29535161/)
25. Diallinas G. Dissection of transporter function: from genetics to structure. *Trends Genet.* 2016; 32:576–90. <https://doi.org/10.1016/j.tig.2016.06.003> PMID:[27430403](https://pubmed.ncbi.nlm.nih.gov/27430403/)
26. Tosolini M, Pengo P, Tecilla P. Biological activity of trans-membrane anion carriers. *Curr Med Chem.* 2018; 25:3560–76. <https://doi.org/10.2174/092986732566618030911322> PMID:[29521206](https://pubmed.ncbi.nlm.nih.gov/29521206/)
27. Mello SS, Attardi LD. Deciphering p53 signaling in tumor suppression. *Curr Opin Cell Biol.* 2018; 51:65–72. <https://doi.org/10.1016/j.ceb.2017.11.005> PMID:[29195118](https://pubmed.ncbi.nlm.nih.gov/29195118/)
28. Bourgeois B, Madl T. Regulation of cellular senescence via the FOXO4-p53 axis. *FEBS Lett.* 2018; 592:2083–97. <https://doi.org/10.1002/1873-3468.13057> PMID:[29683489](https://pubmed.ncbi.nlm.nih.gov/29683489/)
29. Fishel R. Mismatch repair. *J Biol Chem.* 2015; 290:26395–403. <https://doi.org/10.1074/jbc.R115.660142> PMID:[26354434](https://pubmed.ncbi.nlm.nih.gov/26354434/)
30. Li Z, Pearlman AH, Hsieh P. DNA mismatch repair and the DNA damage response. *DNA Repair (Amst).* 2016; 38:94–101. <https://doi.org/10.1016/j.dnarep.2015.11.019> PMID:[26704428](https://pubmed.ncbi.nlm.nih.gov/26704428/)
31. Hernandez-Segura A, Nehme J, Demaria M. Hallmarks of cellular senescence. *Trends Cell Biol.* 2018; 28:436–53.

- <https://doi.org/10.1016/j.tcb.2018.02.001>
PMID:29477613
32. Cao JY, Dixon SJ. Mechanisms of ferroptosis. *Cell Mol Life Sci.* 2016; 73:2195–209.
<https://doi.org/10.1007/s00018-016-2194-1>
PMID:27048822
33. Xie Y, Hou W, Song X, Yu Y, Huang J, Sun X, Kang R, Tang D. Ferroptosis: process and function. *Cell Death Differ.* 2016; 23:369–79.
<https://doi.org/10.1038/cdd.2015.158> PMID:26794443
34. Dong S, Huang F, Zhang H, Chen Q. Overexpression of BUB1B, CCNA2, CDC20, and CDK1 in tumor tissues predicts poor survival in pancreatic ductal adenocarcinoma. *Biosci Rep.* 2019; 39:BSR20182306.
<https://doi.org/10.1042/BSR20182306>
PMID:30765611
35. Gan Y, Li Y, Li T, Shu G, Yin G. CCNA2 acts as a novel biomarker in regulating the growth and apoptosis of colorectal cancer. *Cancer Manag Res.* 2018; 10:5113–24.
<https://doi.org/10.2147/CMAR.S176833>
PMID:30464611
36. Ding K, Li W, Zou Z, Zou X, Wang C. CCNB1 is a prognostic biomarker for ER+ breast cancer. *Med Hypotheses.* 2014; 83:359–64.
<https://doi.org/10.1016/j.mehy.2014.06.013>
PMID:25044212
37. Jeselsohn R, Buchwalter G, De Angelis C, Brown M, Schiff R. ESR1 mutations—a mechanism for acquired endocrine resistance in breast cancer. *Nat Rev Clin Oncol.* 2015; 12:573–83.
<https://doi.org/10.1038/nrclinonc.2015.117>
PMID:26122181
38. Wu F, Mo Q, Wan X, Dan J, Hu H. NEAT1/hsa-mir-98-5p/MAPK6 axis is involved in non-small-cell lung cancer development. *J Cell Biochem.* 2019; 120:2836–46.
<https://doi.org/10.1002/jcb.26442> PMID:29095526
39. Shi XY, Wang H, Wang W, Gu YH. MiR-98-5p regulates proliferation and metastasis of MCF-7 breast cancer cells by targeting Gab2. *Eur Rev Med Pharmacol Sci.* 2019; 23:2847–55.
https://doi.org/10.26355/eurrev_201904_17562
PMID:31002135
40. Xue M, Tao W, Yu S, Yan Z, Peng Q, Jiang F, Gao X. lncRNA ZFPM2-AS1 promotes proliferation via miR-18b-5p/VMA21 axis in lung adenocarcinoma. *J Cell Biochem.* 2020; 121:313–21.
<https://doi.org/10.1002/jcb.29176> PMID:31297866
41. Wang YY, Yan L, Yang S, Xu HN, Chen TT, Dong ZY, Chen SL, Wang WR, Yang QL, Chen CJ. Long noncoding RNA AC073284.4 suppresses epithelial-mesenchymal transition by sponging miR-18b-5p in paclitaxel-resistant breast cancer cells. *J Cell Physiol.* 2019; 234:23202–15.
<https://doi.org/10.1002/jcp.28887> PMID:31215650
42. Zheng S, Liu Q, Ma R, Tan D, Shen T, Zhang X, Lu X. Let-7b-5p inhibits proliferation and motility in squamous cell carcinoma cells through negative modulation of KIAA1377. *Cell Biol Int.* 2019; 43:634–41.
<https://doi.org/10.1002/cbin.11136> PMID:30958603
43. Guan L, Ji D, Liang N, Li S, Sun B. Up-regulation of miR-10b-3p promotes the progression of hepatocellular carcinoma cells via targeting CMTM5. *J Cell Mol Med.* 2018; 22:3434–41.
<https://doi.org/10.1111/jcmm.13620> PMID:29691981
44. Yang G, Zhang Y, Yang J. A five-microRNA signature as prognostic biomarker in colorectal cancer by bioinformatics analysis. *Front Oncol.* 2019; 9:1207.
<https://doi.org/10.3389/fonc.2019.01207>
PMID:31799184
45. Fukuhisa H, Seki N, Idichi T, Kurahara H, Yamada Y, Toda H, Kita Y, Kawasaki Y, Tanoue K, Mataka Y, Maemura K, Natsugoe S. Gene regulation by antitumor miR-130b-5p in pancreatic ductal adenocarcinoma: the clinical significance of oncogenic EPS8. *J Hum Genet.* 2019; 64:521–34.
<https://doi.org/10.1038/s10038-019-0584-6>
PMID:30858505
46. Yu X, Li Z, Zheng H, Chan MT, Wu WK. NEAT1: a novel cancer-related long non-coding RNA. *Cell Prolif.* 2017; 50:e12329.
<https://doi.org/10.1111/cpr.12329> PMID:28105699
47. Klec C, Prinz F, Pichler M. Involvement of the long noncoding RNA NEAT1 in carcinogenesis. *Mol Oncol.* 2019; 13:46–60.
<https://doi.org/10.1002/1878-0261.12404>
PMID:30430751
48. Ghafouri-Fard S, Taheri M. Nuclear enriched abundant transcript 1 (NEAT1): a long non-coding RNA with diverse functions in tumorigenesis. *Biomed Pharmacother.* 2019; 111:51–59.
<https://doi.org/10.1016/j.biopha.2018.12.070>
PMID:30576934
49. Li W, Zhang Z, Liu X, Cheng X, Zhang Y, Han X, Zhang Y, Liu S, Yang J, Xu B, He L, Sun L, Liang J, Shang Y. The FOXN3-NEAT1-SIN3A repressor complex promotes progression of hormonally responsive breast cancer. *J Clin Invest.* 2017; 127:3421–40.
<https://doi.org/10.1172/JCI94233>
PMID:28805661
50. Li X, Wang X, Song W, Xu H, Huang R, Wang Y, Zhao W, Xiao Z, Yang X. Oncogenic properties of NEAT1 in prostate cancer cells depend on the CDC5L-AGRN

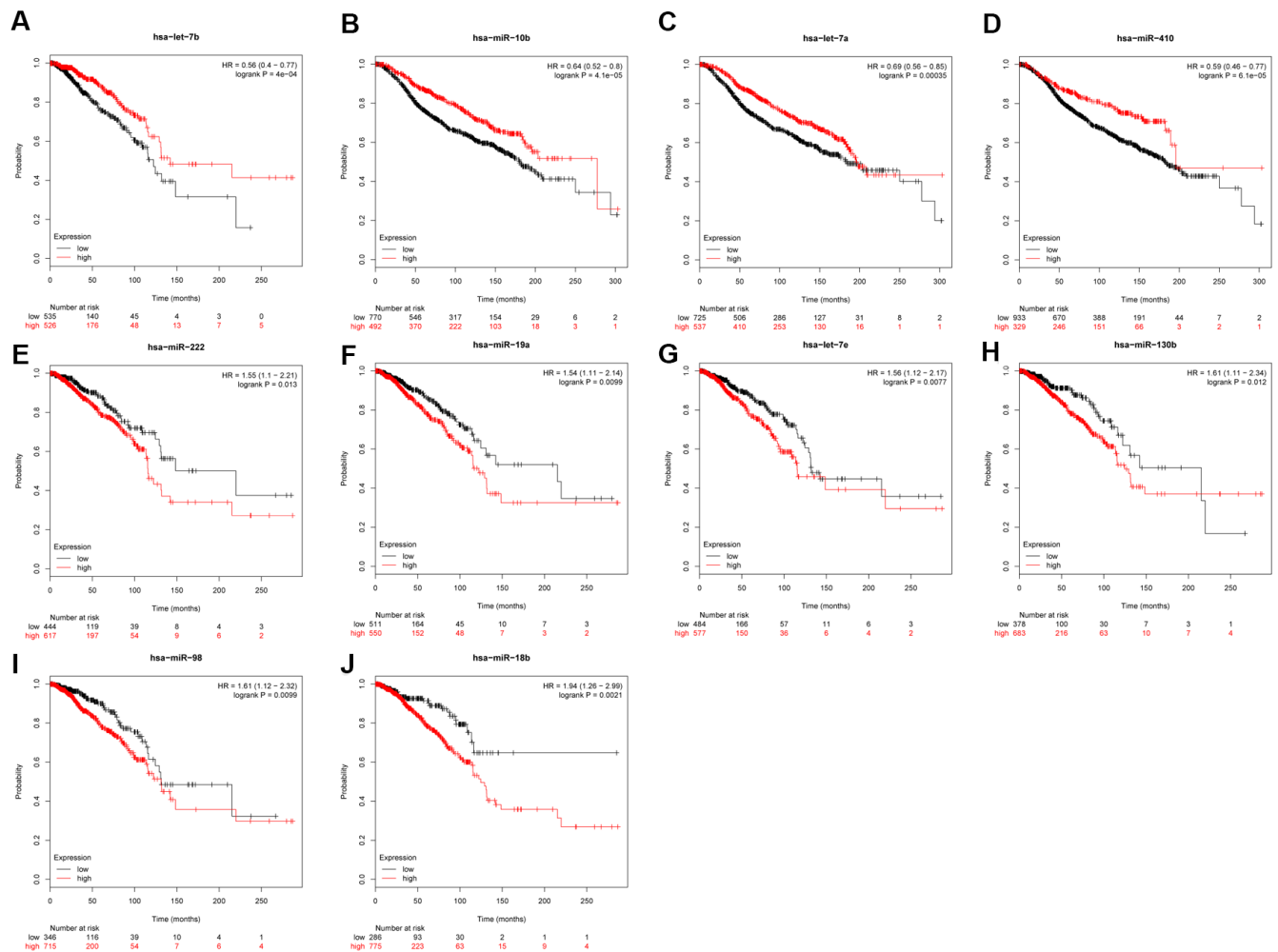
- transcriptional regulation circuit. *Cancer Res.* 2018; 78:4138–49.
<https://doi.org/10.1158/0008-5472.CAN-18-0688>
PMID:29871935
51. Byrne JA, Maleki S, Hardy JR, Gloss BS, Murali R, Scurry JP, Fanayan S, Emmanuel C, Hacker NF, Sutherland RL, Defazio A, O'Brien PM. MAL2 and tumor protein D52 (TPD52) are frequently overexpressed in ovarian carcinoma, but differentially associated with histological subtype and patient outcome. *BMC Cancer.* 2010; 10:497.
<https://doi.org/10.1186/1471-2407-10-497>
PMID:20846453
52. Bhandari A, Shen Y, Sindan N, Xia E, Gautam B, Lv S, Zhang X. MAL2 promotes proliferation, migration, and invasion through regulating epithelial-mesenchymal transition in breast cancer cell lines. *Biochem Biophys Res Commun.* 2018; 504:434–39.
<https://doi.org/10.1016/j.bbrc.2018.08.187>
PMID:30195491
53. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, Doerks T, Stark M, Muller J, Bork P, Jensen LJ, von Mering C. The STRING database in 2011: functional interaction networks of proteins, globally integrated and scored. *Nucleic Acids Res.* 2011; 39:D561–68.
<https://doi.org/10.1093/nar/gkq973> PMID:21045058
54. Lou W, Liu J, Ding B, Xu L, Fan W. Identification of chemoresistance-associated miRNAs in breast cancer. *Cancer Manag Res.* 2018; 10:4747–57.
<https://doi.org/10.2147/CMAR.S172722>
PMID:30425571
55. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 2017; 45:W98–102.
<https://doi.org/10.1093/nar/gkx247>
PMID:28407145
56. Nagy Á, Lánckzy A, Menyhárt O, Gyórfy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci Rep.* 2018; 8:9227.
<https://doi.org/10.1038/s41598-018-27521-y>
PMID:29907753
57. Chou CH, Shrestha S, Yang CD, Chang NW, Lin YL, Liao KW, Huang WC, Sun TH, Tu SJ, Lee WH, Chiew MY, Tai CS, Wei TY, et al. miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 2018; 46:D296–302.
<https://doi.org/10.1093/nar/gkx1067>
PMID:29126174
58. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-seq data. *Nucleic Acids Res.* 2014; 42:D92–97.
<https://doi.org/10.1093/nar/gkt1248>
PMID:24297251
59. Fan Y, Siklenka K, Arora SK, Ribeiro P, Kimmins S, Xia J. miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis. *Nucleic Acids Res.* 2016; 44:W135–41.
<https://doi.org/10.1093/nar/gkw288>
PMID:27105848

SUPPLEMENTARY MATERIALS

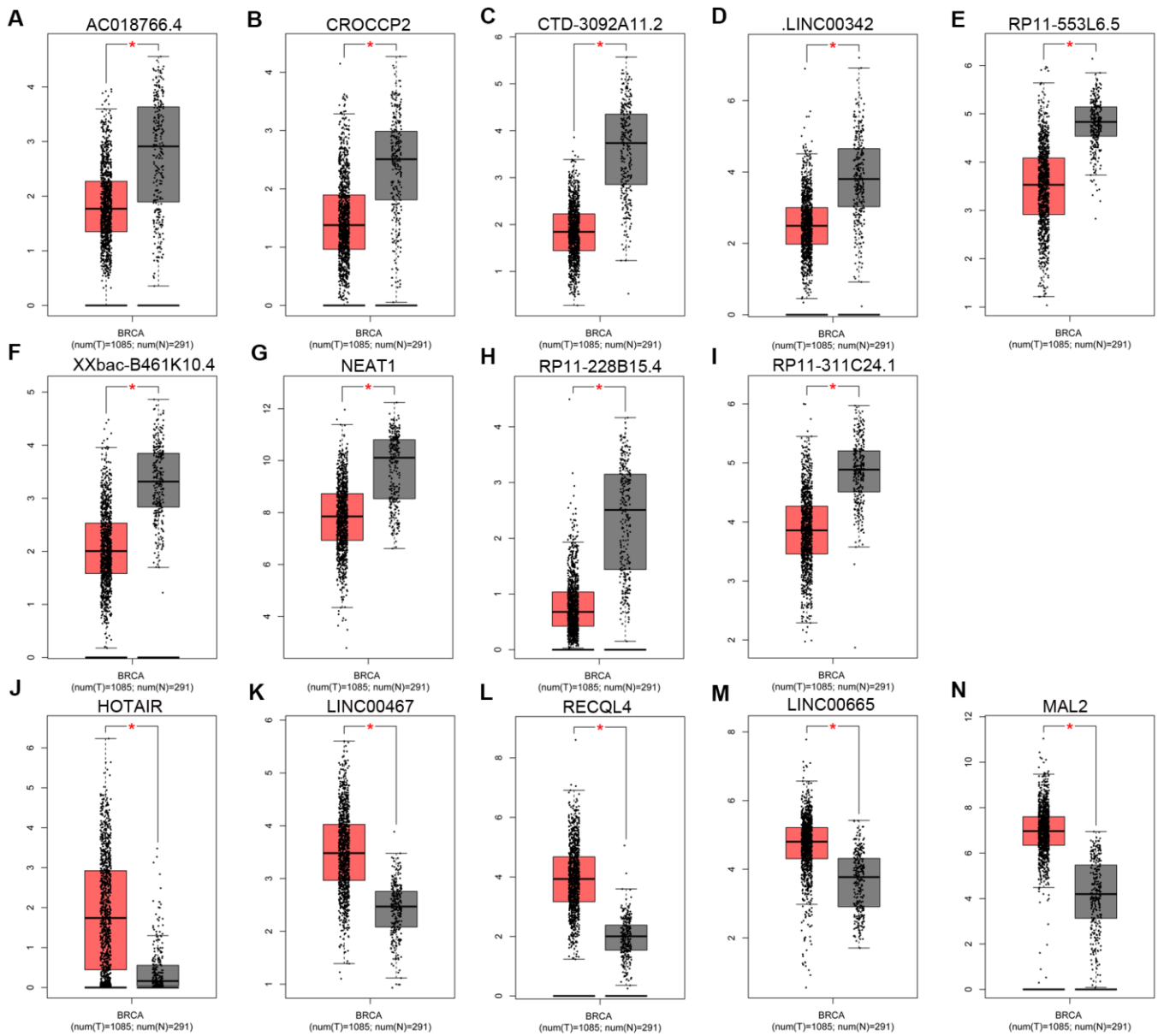
Supplementary Figures



Supplementary Figure 1. The expression of 10 key miRNAs in breast cancer. (A) hsa-let-7b-5p. (B) hsa-miR-10b-3p. (C) hsa-let-7a-5p. (D) hsa-miR-410-3p. (E) hsa-miR-222-3p. (F) hsa-miR-19a-3p. (G) hsa-let-7e-5p. (H) hsa-miR-130b-3p. (I) hsa-miR-98-5p. (J) hsa-miR-18-5p.



Supplementary Figure 2. The prognostic value (overall survival [OS]) of 10 key miRNAs in breast cancer. (A) hsa-let-7b-5p. (B) hsa-miR-10b-3p. (C) hsa-let-7a-5p. (D) hsa-miR-410-3p. (E) hsa-miR-222-3p. (F) hsa-miR-19a-3p. (G) hsa-let-7e-3p. (H) hsa-miR-130b-3p. (I) hsa-miR-98-5p. (J) hsa-miR-18-5p.



Supplementary Figure 3. The expression of key lncRNAs in breast cancer. (A) AC018766.4. **(B)** CROCCP2. **(C)** CTD-309A11.2. **(D)** LINC00342. **(E)** RP11-553L6.5. **(F)** XXbac-B461K10.4. **(G)** NEAT1. **(H)** RP11-228B15.4. **(I)** RP11-311C24.1. **(J)** HOTAIR. **(K)** LINC00467. **(L)** RECQL4. **(M)** LINC00665. **(N)** MAL2.

Supplementary Tables

Please browse Full Text version to see the data of Supplementary Tables 1, 4.

Supplementary Table 1. Common differentially expressed mRNAs in the GSE45827 and GSE65194 datasets.

Supplementary Table 2. The expression of 12 hub genes in TNBC from TCGA.

Gene symbol	logFC	<i>P</i>	Adjusted <i>P</i>
CDK1	3.72	1.51e-21	1.22e-18
CCNB1	2.83	4.95e-19	2.30e-16
CCNA2	3.65	6.00e-20	3.30e-17
CDC20	4.75	1.31e-26	6.04e-23
TOP2A	4.18	1.90e-19	9.41e-17
CCNB2	3.92	2.09e-23	2.81e-20
MAD2L1	2.51	3.93e-11	3.44e-09
BUB1	4.39	8.99e-23	1.01e-19
KIF11	3.09	6.10e-17	1.84e-14
RRM2	4.41	1.66e-26	6.71e-23
ESR1	-4.79	5.11e-14	8.00e-12
IGF1R	-1.34	0.00754	0.046499

Supplementary Table 3. The mRNA-miRNA pairs predicted by miRTarBase.

mRNA	miRNA
CCNA2	hsa-let-7b-5p
CCNA2	hsa-miR-10b-3p
CCNA2	hsa-miR-130b-3p
CCNA2	hsa-miR-27b-3p
CCNA2	hsa-miR-22-3p
MAD2L1	hsa-miR-192-5p
MAD2L1	hsa-miR-28-5p
CDK1	hsa-miR-31-5p
CDK1	hsa-miR-24-3p
RRM2	hsa-let-7a-5p
CCNB1	hsa-miR-410-3p
ESR1	hsa-miR-302c-3p
ESR1	hsa-miR-206
ESR1	hsa-miR-193b-3p
ESR1	hsa-miR-18b-5p
ESR1	hsa-miR-22-3p
ESR1	hsa-miR-19a-3p
ESR1	hsa-miR-19b-3p
ESR1	hsa-miR-20b-5p
ESR1	hsa-miR-221-3p
ESR1	hsa-miR-222-3p
ESR1	hsa-miR-130a-3p
ESR1	hsa-miR-26a-5p
ESR1	hsa-miR-145-5p
ESR1	hsa-miR-26b-3p
ESR1	hsa-miR-192-5p
IGF1	hsa-miR-1-3p
IGF1	hsa-miR-27a-3p
IGF1	hsa-miR-483-3p
IGF1	hsa-let-7e-5p
IGF1	hsa-miR-190b
IGF1	hsa-let-7i-5p
IGF1	hsa-miR-199a-3p
IGF1	hsa-miR-190a-5p
IGF1	hsa-miR-29a-3p
IGF1	hsa-miR-133a-3p
IGF1	hsa-miR-18b-5p
IGF1	hsa-miR-28-5p
IGF1	hsa-miR-603

IGF1	hsa-miR-15b-3p
IGF1	hsa-miR-130b-3p
IGF1	hsa-miR-98-5p
IGF1	hsa-miR-26b-5p
IGF1	hsa-miR-129-5p
IGF1	hsa-miR-128-3p
IGF1	hsa-miR-26a-5p

Supplementary Table 4. The lncRNA-miRNA pairs predicted by miRNet database.