Comprehensive Analysis of lncRNA and miRNA Regulatory Network Reveals Potential Prognostic Non-coding RNA Involved in Breast Cancer Progression

Sheng Gao[1,2†], Xun Lu[3†], Jingjing Ma[2], Qian Zhou[2], RanRan Tang[4], Ziyi Fu[4], Fengliang Wang[2], Mingming Lv[2]* and Cheng Lu[1,2]*

[1] The First Clinical Medicine College, Nanjing University of Chinese Medicine, Nanjing, China, [2] Department of Breast, Nanjing Maternity and Child Health Care Hospital, Women's Hospital of Nanjing Medical University, Nanjing, China, [3] Milken School of Public Health, George Washington University, Washington, DC, United States, [4] Nanjing Maternal and Child Health Institute, Nanjing Maternity and Child Health Care Hospital, Women's Hospital of Nanjing Medical University, Nanjing, China

Breast cancer is one of the most common malignant tumors in women and is the second leading cause of cancer deaths among women. The tumorigenesis and progression of breast cancer are not well understood. The existing researches have indicated that non-coding RNAs, which mainly include long non-coding RNA (lncRNA) and microRNA (miRNA), have gradually become important regulators of breast cancer. We aimed to screen the differential expression of miRNA and lncRNA in the different breast cancer stages and identify the key non-coding RNA using TCGA data. Based on series test of cluster (STC) analysis, bioinformatics analysis, and negatively correlated relationships, 122 lncRNAs, 67 miRNAs, and 119 mRNAs were selected to construct the regulatory network of lncRNA and miRNA. It was shown that the miR-93/20b/106a/106b family was at the center of the regulatory network. Furthermore, 6 miRNAs, 10 lncRNAs, and 15 mRNAs were significantly associated with the overall survival (OS, log-rank $P < 0.05$) of patients with breast cancer. Overexpressed miR-93 in MCF-7 breast cancer cells was associated with suppressed expression of multiple lncRNAs, and these downregulated lncRNAs (MESTIT1, LOC100128164, and DNMBP-AS1) were significantly associated with poor overall survival in breast cancer patients. Therefore, the miR-93/20b/106a/106b family at the core of the regulatory network discovered by our analysis above may be extremely important for the regulation of lncRNA expression and the progression of breast cancer. The identified key miRNA and lncRNA will enhance the understanding of molecular mechanisms of breast cancer progression. Targeting these key non-coding RNA may provide new therapeutic strategies for breast cancer treatment and may prevent the progression of breast cancer from an early stage to an advanced stage.

Keywords: breast cancer, microRNA, long non-coding RNA, series test of cluster, overall survival

Abbreviations: TCGA, the Cancer Genome Atlas; OS, overall survival; STC, series test of cluster; AJCC, American Joint Committee on Cancer.

INTRODUCTION

Breast cancer is one of the most common malignant tumors in women and is the primary reason of cancer deaths among women around the world after lung cancer (Siegel et al., 2018). The 5-year survival rate of patients with early local breast cancer is over 90%, while that of patients with metastatic advanced breast cancer is only 27% (DeSantis et al., 2017), indicating that breast cancer progression, with the formation of distant advanced metastasis, accounts for this poor prognosis. Consequently, the breast cancer patients' mortality rate is mainly attributed by the advanced patients. Intensive efforts have been made to deepen our understanding of the initiation and progression of breast cancer. However, the molecular mechanisms underlying the tumorigenesis and progression of breast cancer are still not well understood. In particular, it is important to determine how to effectively monitor the progression of the breast cancer from early to late stage.

Most factors involved in the progress of breast cancer are related to changes in the expression of specific genes. Previous researches have indicated that non-coding RNA is an important regulator of gene expression (An and Song, 2011; Aalto and Pasquinelli, 2012). Long non-coding RNA (lncRNA) and microRNA (miRNA) are the two most important regulators (Taft et al., 2010; Chan and Tay, 2018). MiRNA, a small non-coding RNA, is approximately 21–24 nucleotides long and mostly repress the mRNA expression efficiently by base-pairing to the $3'$UTR (Finnegan and Pasquinelli, 2013). It has been demonstrated that miRNA associated with proliferation (Johnnidis et al., 2008), differentiation (Becker et al., 2018), cell development (Saultz et al., 2018), and apoptosis (Xu et al., 2016) is linked to all stages of cancers (Gibori et al., 2018; Karimi Mazraehshah et al., 2018; Sun et al., 2018). LncRNA represents an expanding class of ncRNA, consisting of all ncRNA longer than 200 nucleotides (Ponting et al., 2009; Blythe et al., 2016). It has been proved that abnormal expression of lncRNA was related to human diseases, including cancers. Several specific miRNA and lncRNA were associated with the different stages of breast cancer and hold great promise as biomarkers and therapeutic targets (Wang et al., 2017; Adhami et al., 2018).

Non-coding RNAs show a time- or tissue-specific expression pattern and play important roles in a variety of biological processes (Derrien et al., 2012; Li et al., 2015). Therefore, they may display different patterns in breast cancer development and progression. The tumor-node-metastasis classification (TNM classification) of malignant tumors is the most common system used to evaluate the treatment and prognosis of cancer and represents different stages of cancer progression (Greene, 2004; Singletary and Connolly, 2006).

In the present study, differentially expressed miRNA and lncRNA of breast cancer different stages were acquired by the Cancer Genome Atlas (TCGA). TCGA, a landmark cancer genomics program, molecularly characterized over 20,000 primary cancer and matched normal samples spanning 33 cancer types, with the goal of improving the understanding of the molecular basis of cancer and advancing the ability to diagnose, treat, and prevent cancer. Based on series

test of cluster (STC) analysis, the expression patterns of miRNA/lncRNA were identified in breast cancer progression. We also constructed the regulatory network of lncRNA and miRNA to determine the key lncRNA and miRNA in breast cancer progression. Finally, the potential prognostic miRNA, lncRNA, and mRNA involved in breast cancer were revealed. The identified key lncRNA and miRNA in our work would enhance the understanding of mechanism of progression of breast cancer, which may serve as the potential therapeutic targets.

MATERIALS AND METHODS

## RNA Expression Data and Sample Grouping

The RNA expression data and clinical information of breast cancer patients were collected from the TCGA database. Patients with the following conditions are excluded: (i) patients suffering from other malignancies besides breast cancer; (ii) male breast cancer cases; (iii) patients without complete staging information; and (iv) overall survival (OS) > 2,000 days. To seek the abnormally expressed miRNA/lncRNA in the different stages of breast cancer, all patients in TCGA were eventually enrolled in stage I, stage II, and stage III and IV groups.

Screening of Differentially Expressed Non-coding RNA

Eighty-four adjacent non-tumor tissues were set as normal controls. The differentially expressed RNA between the adjacent normal tissues and stage I, stage II, and stage III and IV cancer tissues was screened. The process of screening deregulated genes is shown in **Figure 1**. The differentially expressed non-coding RNAs were also screened between the normal tissues and cancer tissues concluding all the stages.

STC Analysis

The experiment was carried out on the basis of previous studies (Luan et al., 2017). When the diseases progressed from one stage to the next or when organisms were subjected to environmental stimulation, the gene expression changes exhibited different trend characteristics. STC analysis could reveal the gene expression's trend characteristics. The most representative genes with the same change characteristic could be confirmed in the disease progression and were gathered into the different changing-trend profile. We set adjacent non-tumor tissues, stages I, II, and III and IV, as different points of breast cancer development. The differential expression genes at a logical sequence according to RVM (random variance model) corrective ANOVA were selected. The raw expression values were converted into log2 ratio. The expression model profiles are related to the actual or the expected number of genes assigned to each model profile. Significant profiles have higher probability than expected by Fisher's exact test and multiple comparison test.

FIGURE 1 | Flow-chart of data analysis.
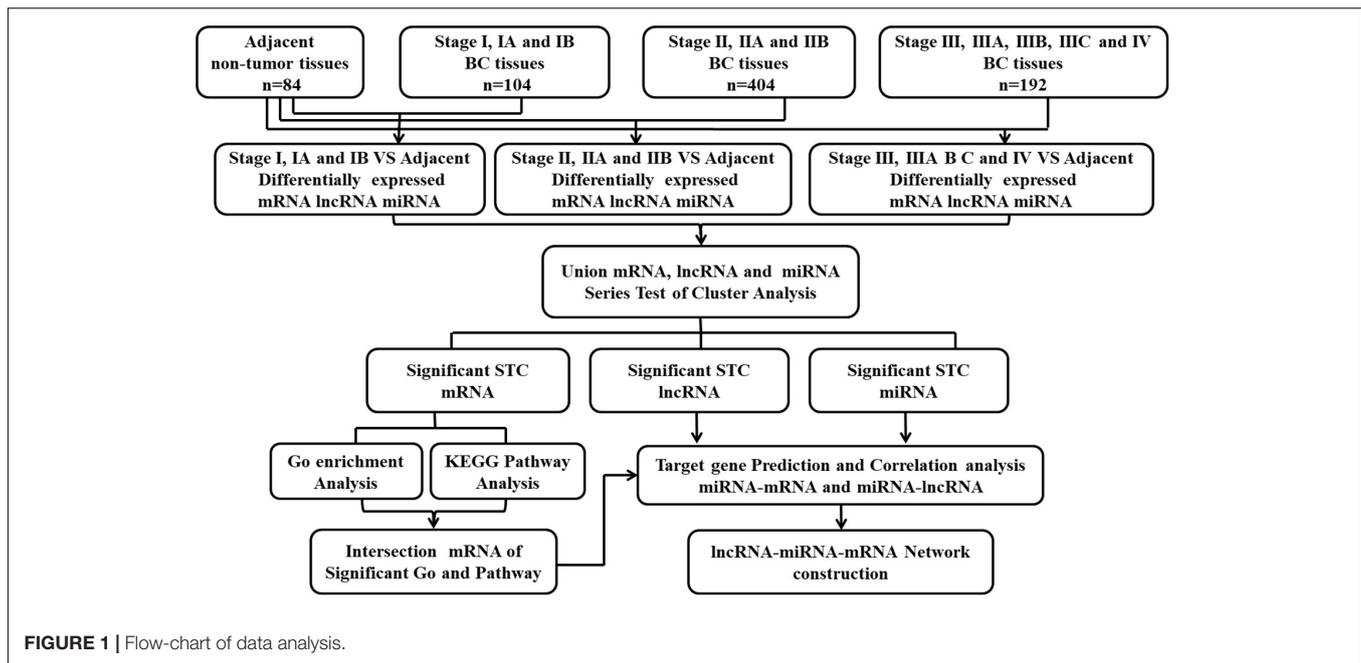
Bioinformatic Analysis of Differentially Expressed Genes

The Gene Ontology database and KEGG (Kyoto Encyclopedia of Genes and Genomes) were used to analyze the functions of the differential expression of genes in significant patterns of the STC analysis. Two-sided Fisher's exact test and χ^2 test were used to classify GO enrichment and select the significant KEGG pathway, and FDR (false discovery rate) was calculated to correct the p-value. Genes that were upregulated and downregulated were analyzed. We were interested in GO biological processes and KEGG pathways with statistical significance. FDR < 0.05 was considered statistically significant.

Construction of lncRNA and miRNA Regulatory Network

LncRNA can act as a miRNA decoy or sponge by complementary sequence and can indirectly regulate the expression and activity of mRNA. Based on this theory, the lncRNA and miRNA regulatory network was constructed. The differentially expressed miRNAs with significant patterns from the STC analysis were selected. MRNAs targeted by these miRNAs were predicted by TargetScan and miRanda. The predicted lncRNA targeted by these miRNAs were determined by miRanda. Then, we combined predicted lncRNA/mRNA and the differentially expressed lncRNA/mRNA to choose the intersecting lncRNA and mRNA. Finally, according to the theory that miRNA negatively regulates lncRNA and mRNA, the miRNA–lncRNA and miRNA–mRNA relationships of negative correlation were screened, and the lncRNA and miRNA regulatory network was constructed.

Cell Culture and Transfection

Human breast cancer cell line MCF-7 was purchased from American Type Culture Collection and maintained in RPMI 1640 containing 10% FBS and 1% penicillin/streptomycin (Gibico) in a humidified atmosphere of 5% CO_2 at 37°C. Cells were transfected with mimic/inhibitor for miR-93-5p or control (RiboBio) mixed with Opti-MEM medium using Lipofectamine 2000 (Invitrogen).

Total RNA Extraction and Real-Time PCR

MCF-7 cells were transiently transfected with miR-93-5p mimic/inhibitor or control, and after 48 h of transfection, total RNA was isolated using TRIzol reagent (Invitrogen). cDNA was synthesized using oligodT primers or miR-93-5p/U6-specific primers (RiboBio) with a Reverse transcription kit (Thermo fisher). Real-time PCR was performed in triplicate using TaqMan Universal PCR Master Mix (Applied Biosystems) in the ViiA7 system (Applied Biosystems). The primers of miRNA-93-5P were designed and provided by RiboBio.

Statistical Analysis

The associations between breast cancer patients' overall survival time and the specific lncRNA/mRNA expression levels were analyzed by using Kaplan–Meier method, the Cox proportional hazard regression models, and log-rank test.

RESULTS

Differential Expression of miRNA and lncRNA in the Different Stages

The mRNA, miRNA, and lncRNA expression data of 1,097 breast cancer from TCGA were collected and obtained. According to the clinical information given in the TCGA database, we removed 397 unsuitable cases, and 700 cases were included in this study. To seek the deregulated lncRNA and miRNA in the different

stages of breast cancer, 700 cases in TCGA were divided into three groups (stage I, stage II, and stages III and IV) according to their diagnosed stages that were based on the recommendations of the AJCC. There were 104 cases presented with stage I, 404 cases presented with stage II, and 192 cases presented with stages III and IV. In addition, 84 adjacent non-tumor tissues were set as normal controls to screen the differentially expressed RNA. The process of screening the deregulated genes and that of the bioinformatics analysis is shown in **Figure 1**. We identified differentially expressed 193 lncRNA and 148 miRNA between stage I and the adjacent normal tissues, 208 lncRNA and 166 miRNA between stage II and the adjacent normal tissues, 207 lncRNA and 163 miRNA between stages III and IV, and the adjacent normal tissues (absolute fold change > 2, FDR < 0.05, and $P < 0.05$). In addition, we identified differentially expressed 204 lncRNA and 162 miRNA between the normal tissues and cancer tissues concluding all the stages (**Supplementary Table 1**).
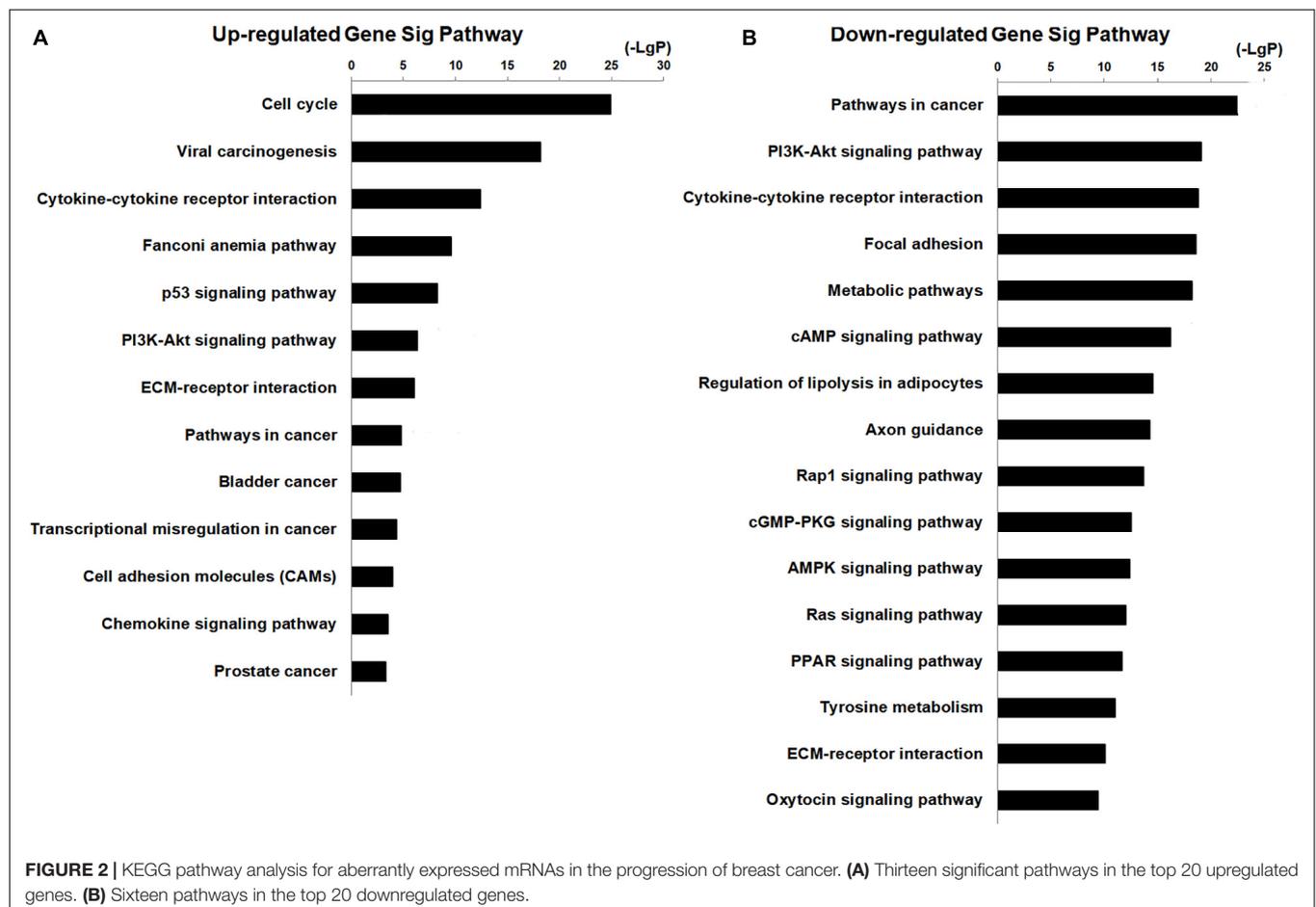
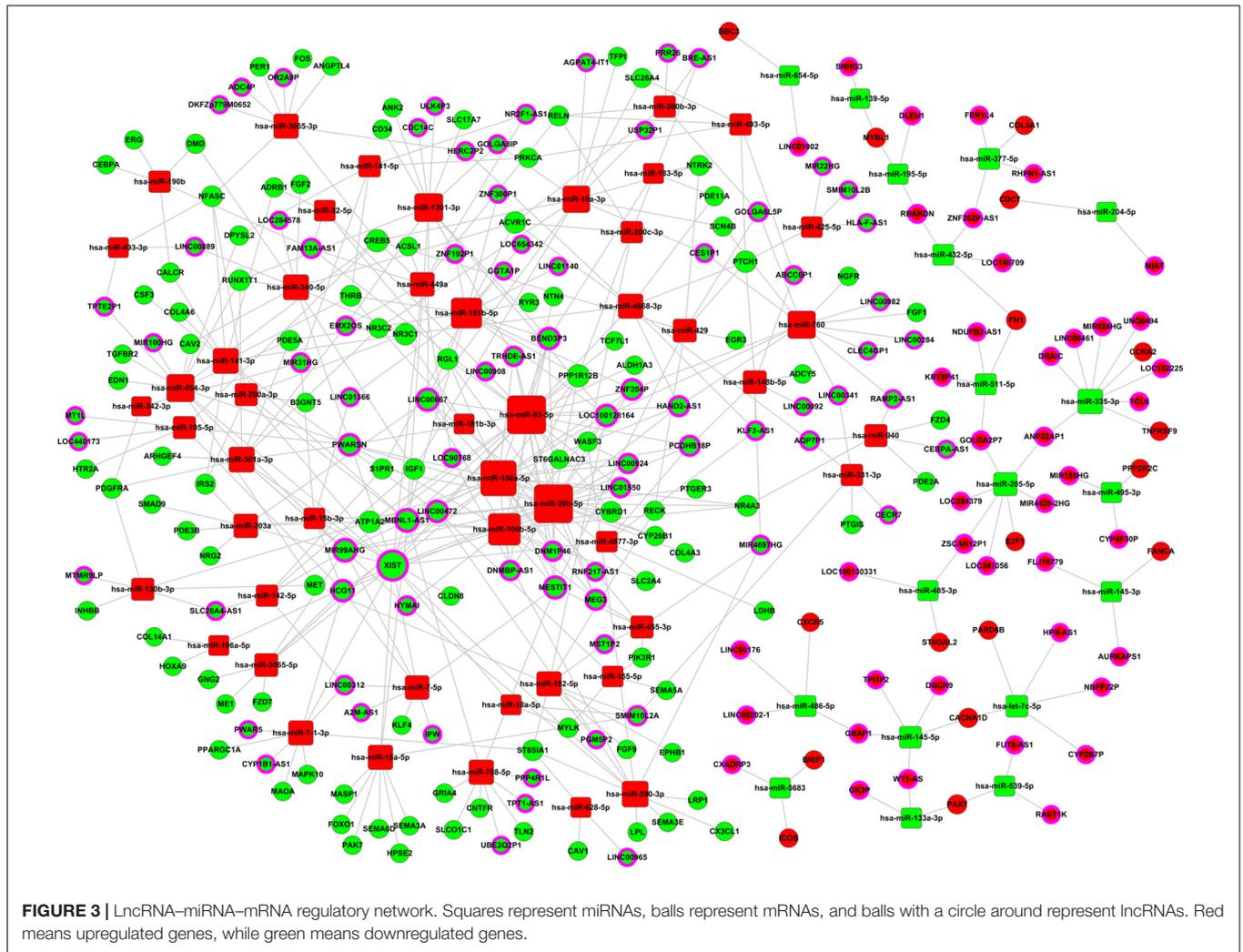## MiRNA/LncRNA Expression Model Analysis With STC Analysis

STC is used to reveal the expression tendency of the genes. To further assess the profile pattern of differentially expressed miRNA/lncRNA among the different stage of breast cancer, we used the union of differentially expressed miRNA/lncRNA in the breast cancer different stage, and STC analysis was performed. There were 26 possible model (**Supplementary Figure 1**) profiles to enrich the expression tendency of the miRNA/lncRNA. We identified five expression patterns of aberrantly expressed miRNA (profiles 2, 4, 5, 22, and 25) and five expression patterns of aberrantly expressed lncRNA (profiles 1, 2, 5, 22, and 25) that were significant ($P < 0.05$, **Supplementary Figure 2**). The differential expression of 224 lncRNA and 172 miRNA in the significantly changing trend profiles might have vital roles in breast cancer progression. We also identified eight patterns of aberrantly expressed mRNA (profiles 1, 2, 4, 5, 7, 22, 25, and 26) that were significant ($P < 0.05$) for the following analysis.

GO Enrichment and KEGG Pathway Analysis

The potential functions of non-coding RNA were identified by GO and KEGG pathways analysis. There were 17 terms related to breast cancer in the top 20 GO analysis terms (upregulated genes), and 8 of the terms included cell division, cell proliferation, cell cycle, and DNA repair (**Supplementary Figure 3**). There were 15 terms related to cancer in the top 20 GO analysis terms (downregulated genes), and 9 of the terms included signal transduction, cell adhesion, angiogenesis, apoptosis, and cell proliferation (**Supplementary Figure 3**). These GO analyses

FIGURE 2 | KEGG pathway analysis for aberrantly expressed mRNAs in the progression of breast cancer. **(A)** Thirteen significant pathways in the top 20 upregulated genes. **(B)** Sixteen pathways in the top 20 downregulated genes.

FIGURE 3 | LncRNA–miRNA–mRNA regulatory network. Squares represent miRNAs, balls represent mRNAs, and balls with a circle around represent lncRNAs. Red means upregulated genes, while green means downregulated genes.
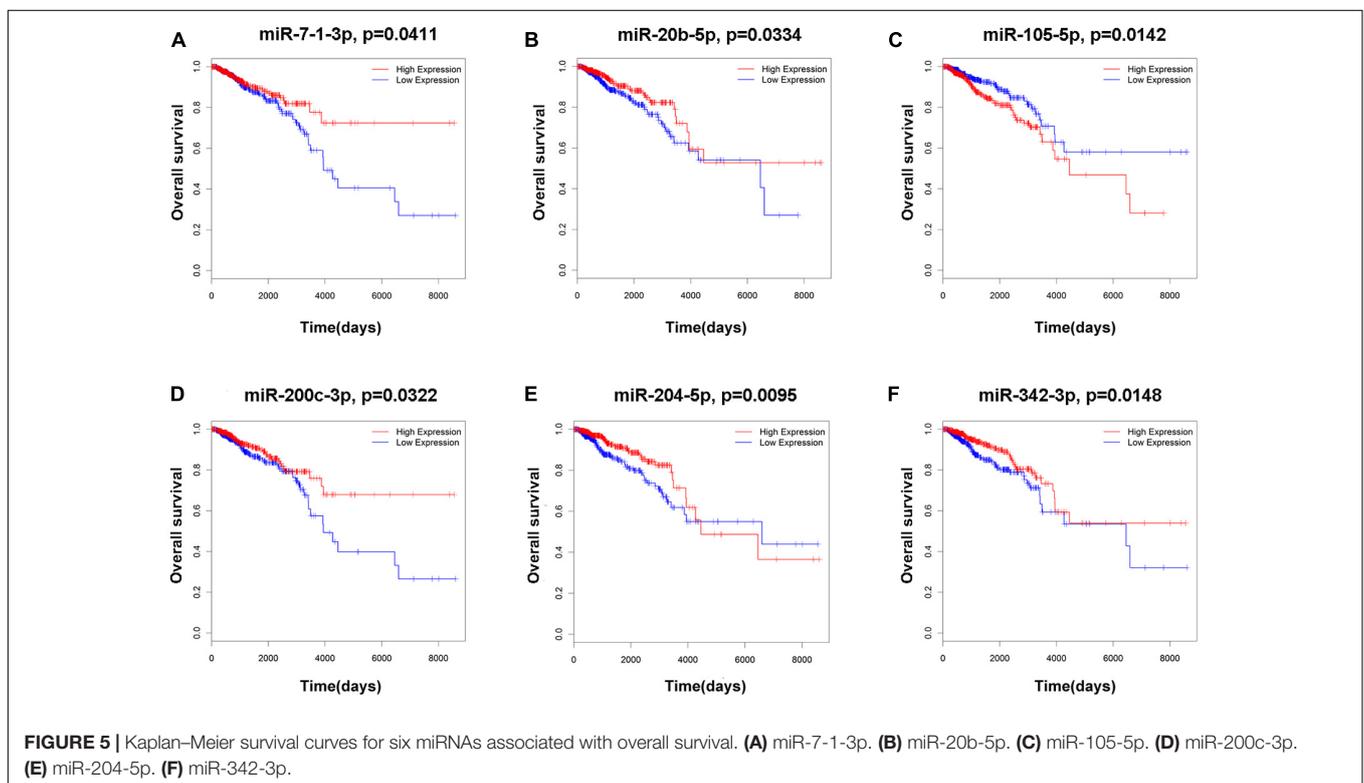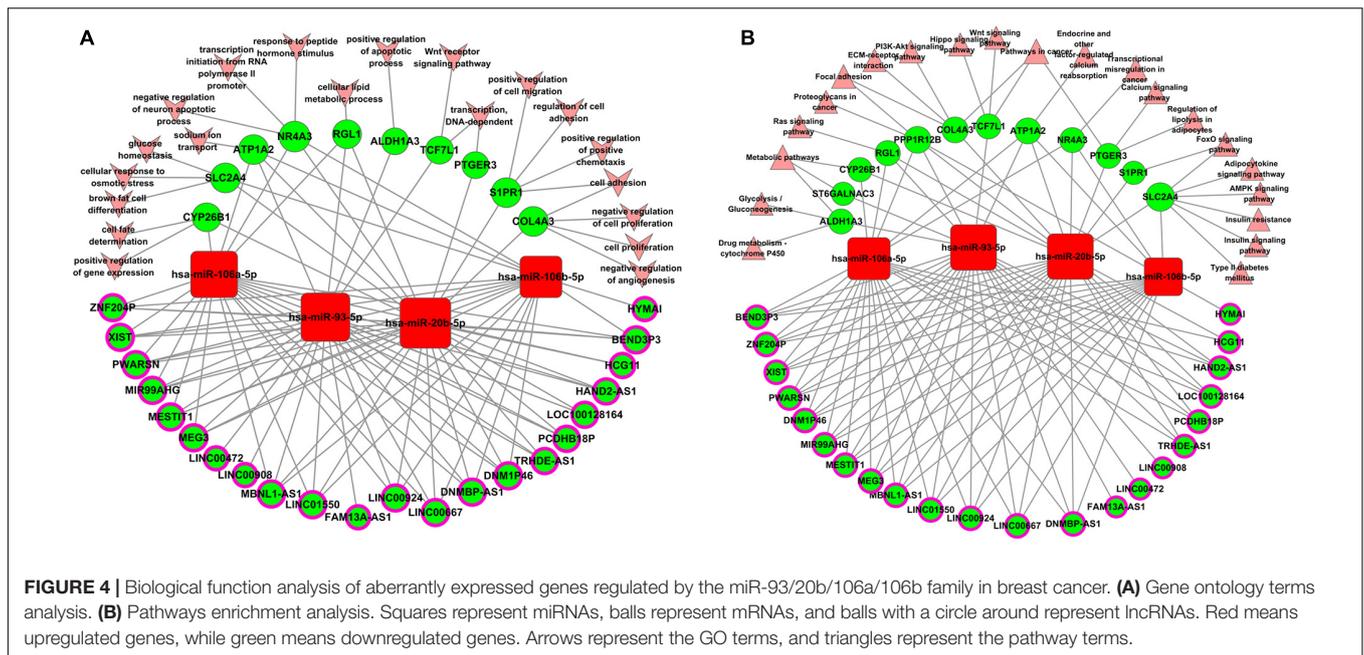
indicated that aberrantly expressed genes were mostly related to cell proliferation, cell apoptosis, cell cycle, cell adhesion, and angiogenesis, which were related to the biological behavior of cancer cells. Pathway analysis indicated that upregulated genes are related to 63 pathways and that there were 13 pathways in the top 20 pathways related to breast cancer (**Figure 2A**). The downregulated genes are enriched in 144 pathways, and there were 16 pathways in the top 20 pathways related to cancer (**Figure 2B**). The potential signaling pathways have been identified in breast cancer progression, including cell cycle pathway, P53 signaling pathway, PI3K–AKT signaling pathways, focal adhesion pathways, pathways in cancer, metabolic pathway, cAMP pathway, AMPK signal pathway, and Ras signal pathway (**Figure 2**). These findings suggested that the aberrant expression genes were mainly enriched in cancer cells' biological behavior. At the same time, the results also suggested that downregulated genes were concerned with the regulation of miRNA in cancer. To further narrow the possible target genes, we combined the genes in the GO and Pathway analysis and chose 637 intersecting genes with significant difference (FDR < 0.05) to perform the following analysis.

Prediction of the miRNA Targets and Construction of the lncRNA and miRNA Regulatory Network

Firstly, 172 differentially expressed miRNA (in five significant patterns in the above STC analysis) were selected to further predict their targeted mRNA by miRanda and TargetScan. MiRNA may interact with lncRNA through miRNA recognition

TABLE 1 | miRNA-93/20b/106a/106b family and related lncRNAs/mRNAs in lncRNA-miRNA-mRNA regulatory network.

miRNA	LncRNA	mRNA
miR-20b-5p miR-93-5p miR-106a-5p mir-106b-5p	BEND3P3, DNM1P46, DNMBP-AS1, HAND2-AS1, HCG11, LINC00667, LINC00924, LINC01550, LOC100128164, MBNL1-AS1, MEG3, MESTIT1, MIR99AHG, PCDHB18P, PWARSN, TRHDE-AS1, XIST, ZNF204P, FAM13A-AS1, LINC00472, LINC00908, LINC00924	ATP1A2, COL4A3, CYBRD1, CYP26B1, NR4A3, PPP1R12B, PTGER3, RGL1, S1PR1, SLC2A4, ALDH1A3, NTN4, ST6GALNAC3, TCF7L1

FIGURE 4 | Biological function analysis of aberrantly expressed genes regulated by the miR-93/20b/106a/106b family in breast cancer. **(A)** Gene ontology terms analysis. **(B)** Pathways enrichment analysis. Squares represent miRNAs, balls represent mRNAs, and balls with a circle around represent lncRNAs. Red means upregulated genes, while green means downregulated genes. Arrows represent the GO terms, and triangles represent the pathway terms.

FIGURE 5 | Kaplan–Meier survival curves for six miRNAs associated with overall survival. **(A)** miR-7-1-3p. **(B)** miR-20b-5p. **(C)** miR-105-5p. **(D)** miR-200c-3p. **(E)** miR-204-5p. **(F)** miR-342-3p.

elements. We predicted the lncRNA targeted by these miRNAs by miRanda, and the results show that these 156 miRNAs target 200 lncRNAs. Then, we chose the intersecting lncRNA and mRNA between the predicted target mRNA/lncRNA, and we chose differential expression of 220 lncRNA and 637 mRNA from above. Finally, we selected these differentially expressed RNAs that were negatively regulated between the lncRNA/mRNA

and the miRNA according to the theory that miRNA negatively regulates lncRNA and mRNA. Based on the above data, 122 lncRNAs, 67 miRNAs, and 119 mRNAs were selected to construct the regulatory network of lncRNA and miRNA (**Figure 3**). Degree refers to the degree of association or connectivity between genes and other genes in gene interaction network, which is the most basic characteristic of a gene-to-gene interaction.

FIGURE 6 | Kaplan–Meier survival curves for 10 lncRNAs associated with overall survival. **(A)** DNMBP-AS1. **(B)** HPN-AS1. **(C)** LINC00461. **(D)** LINC00472. **(E)** LINC00667. **(F)** LOC284578. **(G)** LOC100128164. **(H)** MESTIT1. **(I)** RHPN1-AS1. **(J)** SLC26A4-AS1.
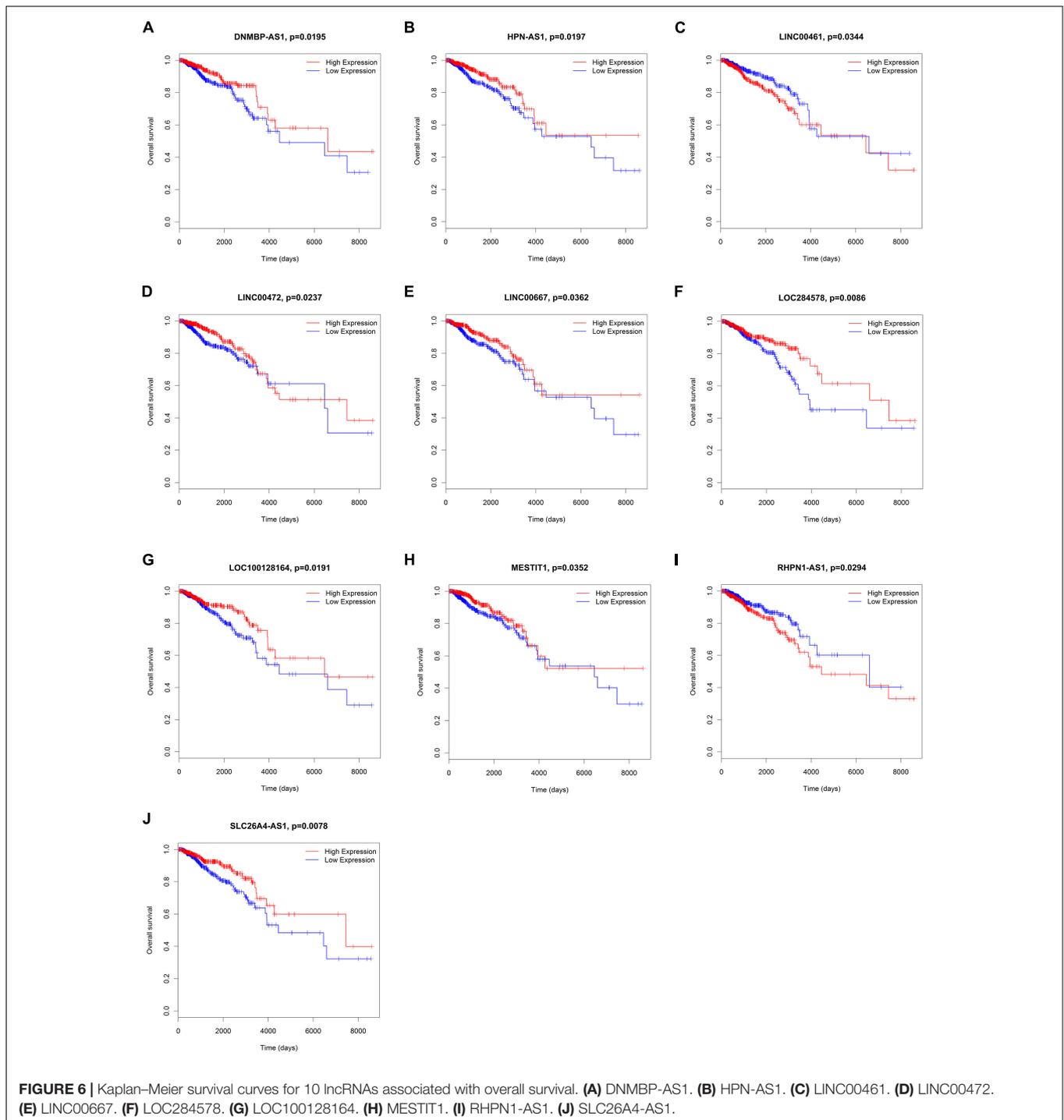
The genes with the highest degree value in gene interaction network have the greatest influence on the overall structure of the network and other gene relationships. The top 15 key lncRNAs and miRNAs that may target these lncRNA are shown in **Supplementary Table 2**. The top 15 key miRNA and their target lncRNA/mRNA in the regulatory network in breast cancer progression are shown in **Supplementary Table 3**. We found that in the regulatory network, miR-301a-3p (Zheng et al., 2018),

miR-93-5p (Ni et al., 2018), miR-454-3p (Ren et al., 2019), miR-181b-5p (Sochor et al., 2014), XIST (Salvador et al., 2013), LINC00472 (Wang et al., 2019), MIR31HG (Augoff et al., 2012), MEG3 (Mondal et al., 2015), MIR155HG (Ghafouri-Fard et al., 2020), and LINC00667 (Zhu et al., 2020) have been reported in breast cancer. This indicates that our results are consistent with the results of existing research. Some non-coding RNAs that have not been reported in breast cancer need further research to
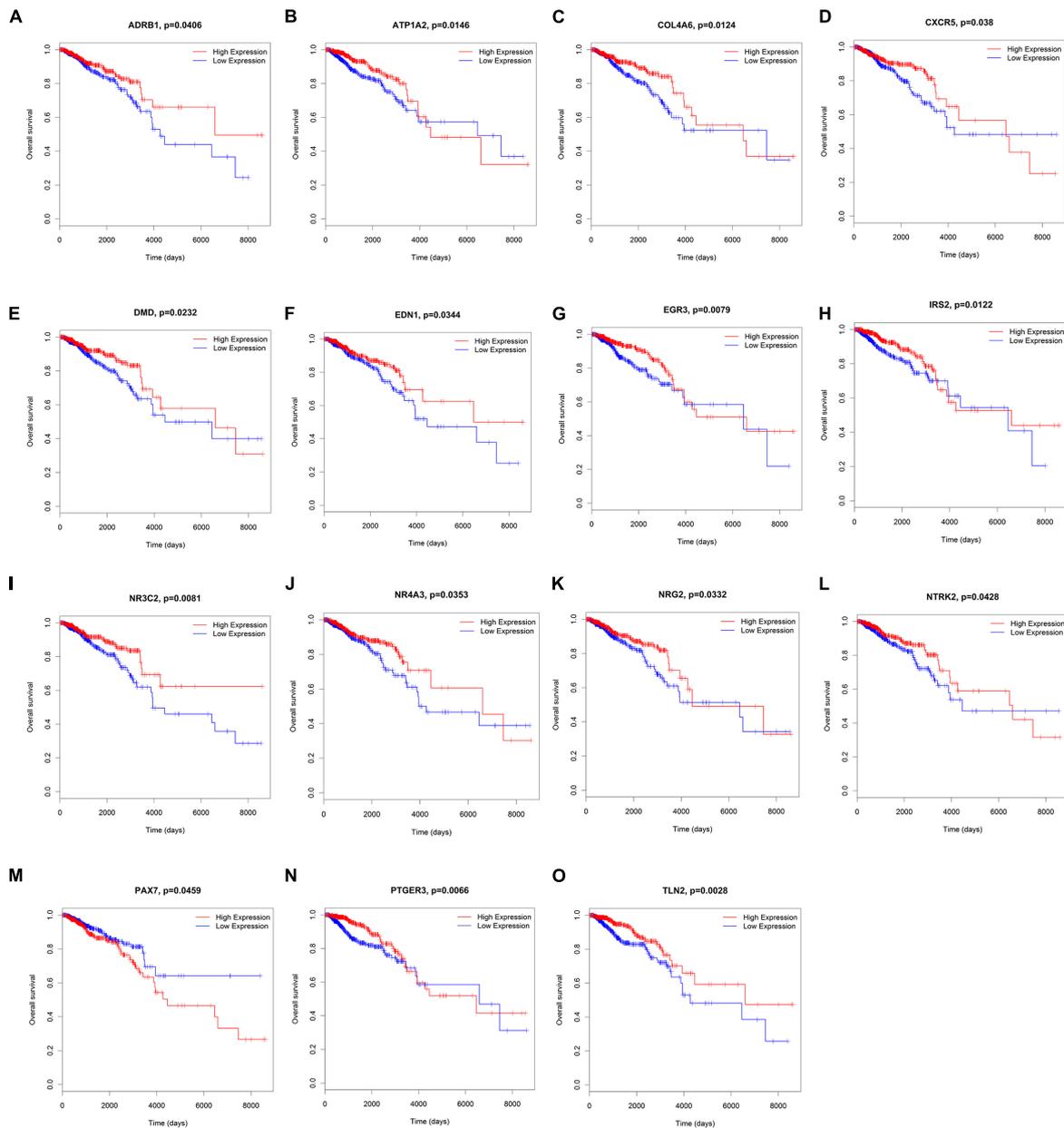
FIGURE 7 | Kaplan–Meier survival curves for 15 mRNAs associated with overall survival. **(A)** ADRB1, **(B)** ATP1A2, **(C)** COL4A6, **(D)** CXCR5, **(E)** DMD, **(F)** EDN1, **(G)** EGR3, **(H)** IRS2, **(I)** NR3C2, **(J)** NR4A3, **(K)** NRG2, **(L)** NTRK2, **(M)** PAX7, **(N)** PTGER3, **(O)** TLN2.

reveal their functions and mechanisms. It is worth noting that the miR-93/20b/106a/106b family is located at the center of the regulatory network.

Potential Biological Functions of the Aberrantly Expressed miR-93/20b/106a/106b Family in Breast Cancer

MiR-106b, miR-106a, miR-93, and miR-20b share the same seed sequence "AAAGUGC," which is encoded in two genetic clusters and has similar physiological functions in a variety of

diseases (**Supplementary Figure 4A**). They all show a profile 22 pattern in the five aberrantly expressed patterns of miRNA (**Supplementary Figure 4B**). MiR-20b and miR-106a are encoded by the miR-106a-363 cluster located in chromosome X. MiR-93 and miR-106b were encoded by the miR-106b-25 cluster located in chromosome 7. The relative expression of miR-93 was the highest (**Supplementary Figure 4C**). In view of the importance of the miRNA-93/20b/106a/106b family and the related lncRNA, which were separately displayed, further GO and KEGG pathway analyses were conducted. As shown in **Table 1**, all 22 lncRNAs and 14 mRNAs in the miRNA-93/20b/106a/106b family are in the regulatory network. The GO and KEGG pathway analyses

TABLE 2 | lncRNA/miRNA/mRNAs that were significantly related to overall survival in breast cancer progression.

Gene_id	Gene type	Expression	Hazard ratio	p-Value	Relationship to overall survival
DMD	mRNA	Down	0.6371	0.0232	Positive
NRG2	mRNA	Down	0.6567	0.0332	Positive
NTRK2	mRNA	Down	0.6715	0.0428	Positive
PAX7	mRNA	Up	1.4916	0.0459	Negative
PTGER3	mRNA	Down	0.5840	0.0066	Positive
TLN2	mRNA	Down	0.5486	0.0028	Positive
ADRB1	mRNA	Down	0.6668	0.0406	Positive
ATP1A2	mRNA	Down	0.6134	0.0146	Positive
COL4A6	mRNA	Down	0.6097	0.0124	Positive
CXCR5	mRNA	Up	0.6629	0.0380	Positive
EDN1	mRNA	Down	0.6562	0.0344	Positive
EGR3	mRNA	Down	0.5929	0.0079	Positive
IRS2	mRNA	Down	0.6115	0.0122	Positive
NR3C2	mRNA	Down	0.5871	0.0081	Positive
NR4A3	mRNA	Down	0.6588	0.0353	Positive
miR-200c-3p	miRNA	Up	0.6471	0.0322	Positive
miR-204-5p	miRNA	Down	0.5939	0.0095	Positive
miR-20b-5p	miRNA	Up	0.6506	0.0334	Positive
miR-7-1-3p	miRNA	Up	0.6590	0.0411	Positive
miR-105-5p	miRNA	Up	1.6361	0.0142	Negative
miR-342-3p	miRNA	Up	0.6136	0.0148	Positive
DNMBP-AS1	lncRNA	Down	0.6268	0.0195	Positive
HPN-AS1	lncRNA	Up	0.6263	0.0197	Positive
LINC00461	lncRNA	Up	1.5183	0.0344	Negative
LINC00472	lncRNA	Down	0.6394	0.0237	Positive
LINC00667	lncRNA	Down	0.6600	0.0362	Positive
LOC100128164	lncRNA	Down	0.6267	0.0191	Positive
LOC284578	lncRNA	Down	0.5961	0.0086	Positive
MESTIT1	lncRNA	Down	0.6527	0.0352	Positive
RHPN1-AS1	lncRNA	Up	1.5462	0.0294	Negative
SLC26A4-AS1	lncRNA	Down	0.5895	0.0078	Positive

show that the miR-93/20b/106a/106b family members are mainly involved in development, ion transport, and metabolism. The go and pathway terms that related to cancer were displayed in **Figure 4**.

Survival Prediction of Non-coding RNA and mRNA

To elucidate the relationship between these key genes (non-coding RNA and mRNA) and the outcome of breast cancer patients, univariate Cox proportional hazards regression model was used to validate their prognostic values. MiR-200c-3p, miR-204-5p, miR-20b-5p, miR-7-1-3p, miR-105-5p, miR-342-3p, DNMBP-AS1, HPN-AS1, LINC00461, LINC00472, LINC00667, LOC100128164, LOC284578, MESTIT1, RHPN1-AS1, and SLC26A4-AS1 were significantly associated with breast cancer patients' overall survival (log-rank $P < 0.05$, **Figures 5**, **6**). Additionally, we found 15 mRNAs that were significantly related to overall survival (**Figure 7**). The expression of these 6 miRNAs, 10 lncRNAs, and 15

mRNAs is shown in **Table 2**. In addition, the relationship between these miRNA, lncRNA, and mRNA expression levels and prognosis in breast cancer patients is also shown in **Table 2**. CXCR5, miR-200c-3p, miR-20b-5p, miR-7-1-3p, miR-342-3p, and HPN-AS1 were upregulated genes, while these high expression genes were significantly associated with breast cancer patients' better overall survival. Among these non-coding RNAs, miR-20b-5p, MESTIT1, LOC100128164, LINC00667, and DNMBP-AS1 were related to the miR-93/20b/106a/106b family.
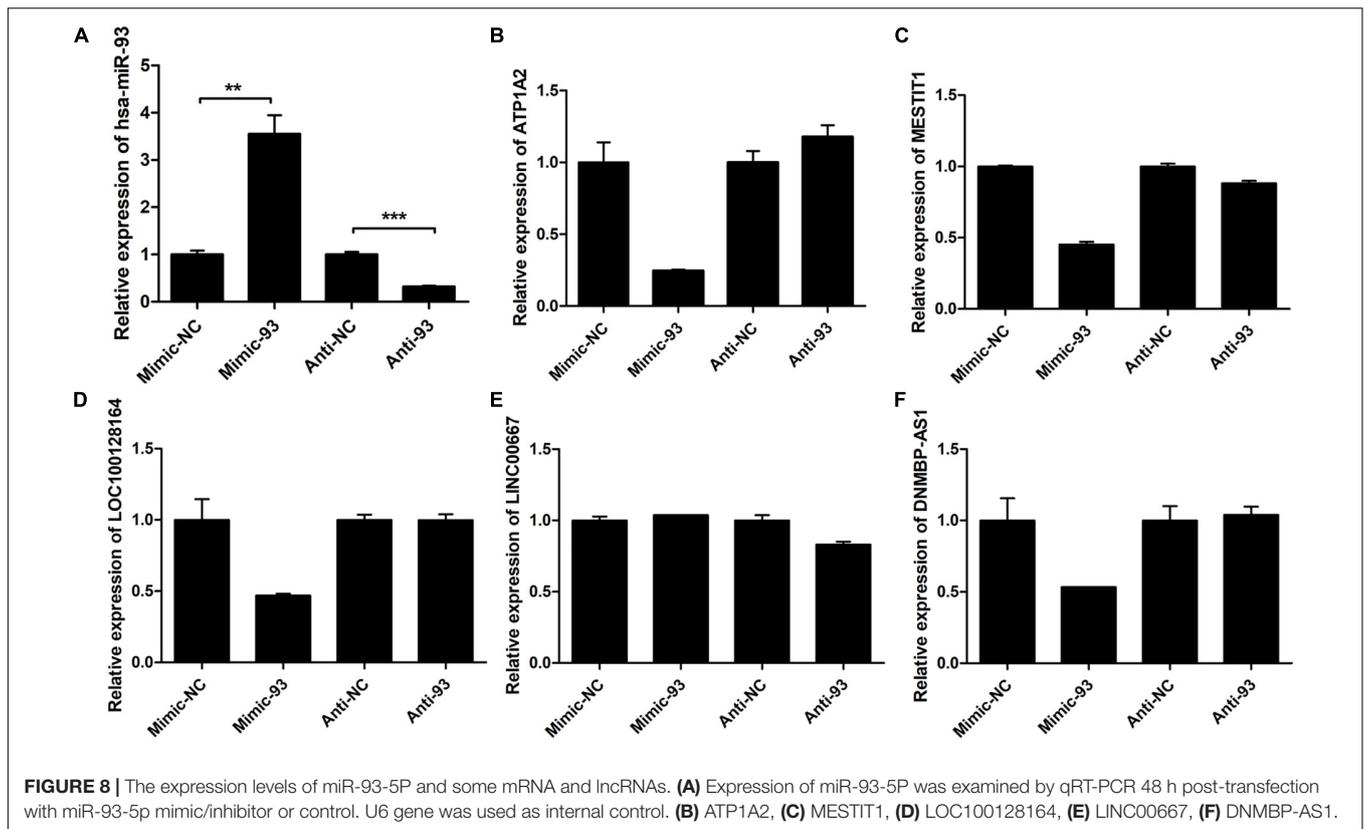
Experimental Verification of Possible Regulatory Relationships Between miR-93 and Some lncRNA/mRNA

Some central lncRNAs and miRNAs predicted here may have vital roles in breast cancer progression. Therefore, we further used GEO data to analyze the expression levels of four miRNAs in the miRNA-93/20b/106a/106b family at the core of the regulatory network. The results showed that miRNA-93/20b/106a/106b were significantly and highly expressed in tumor tissues in one or more GEO data (**Supplementary Figure 5**). Among them, the expression of miR-93 is the most stable and relatively significant. Next, MCF-7 cells were transiently transfected with miR-93-5p mimic/inhibitor or control to validate the miRNA-93/20b/106a/106b family regulatory network. The expression levels of miR-93-5P were examined by qRT-PCR after transfection with miR-93-5p mimic/inhibitor or control (**Figure 8A**). It was found that the expression levels of ATP1A2, MESTIT1, LOC100128164, and DNMBP-AS1 were all downregulated by overexpressing miR-93 in MCF-7 cells. Therefore, the miRNA-93/20b/106a/106b family suppresses multiple lncRNA expression (MESTIT1, LOC100128164, and DNMBP-AS1). However, these gene expression levels did not significantly change after treatment with miR-93-5p inhibitor (**Figures 8B–F**). The above results can indicate that overexpressed miR-93 in breast cancer tissues is associated with the repression of multiple lncRNAs, and these downregulated lncRNAs (MESTIT1, LOC100128164, and DNMBP-AS1) were significantly associated with poor overall survival in breast cancer patients.

In this study, we found low expression mRNAs/lncRNAs were significantly associated with cancer patients' better overall survival and high expression mRNAs were significantly associated with cancer patients' worse overall survival except for CXCR5 and HPN-AS1. However, upregulated miRNAs (miR-200c-3p, miR-20b-5p, miR-7-1-3p, and miR-342-3p) were significantly associated with breast cancer patients' better overall survival. It is certain that some lncRNAs suppressed by the highly expressed miRNA-93/20b/106a/106b family are not conducive to the prognosis of patients, and the regulation between miRNAs and lncRNAs may be more complicated.

DISCUSSION

Breast cancer is one of the most common malignant tumors in women and is the primary reason of cancer deaths among women around the world. The miRNA and lncRNA expression

FIGURE 8 | The expression levels of miR-93-5P and some mRNA and lncRNAs. **(A)** Expression of miR-93-5P was examined by qRT-PCR 48 h post-transfection with miR-93-5p mimic/inhibitor or control. U6 gene was used as internal control. **(B)** ATP1A2, **(C)** MESTIT1, **(D)** LOC100128164, **(E)** LINC00667, **(F)** DNMBP-AS1.

profiles were analyzed in a variety of human cancer using high-throughput microarrays or sequencing. The biological functions of non-coding RNA in breast cancer have been revealed, including tumorigenesis, proliferation, apoptosis, and invasion. In previous studies, the differentially expressed miRNA and lncRNA were mostly screened from the different subtypes of breast cancer tissues, cancerous and normal breast tissues, and breast cancer drug-resistant cells (Van Schooneveld et al., 2012; Lv et al., 2014; Sun et al., 2014; Chen et al., 2015). Most studies focus on the function and mechanism of single non-coding RNA. Recently, evidence has begun to accumulate that these two ncRNAs can interact with each other and can influence the gene expression in breast cancer (Zhang H.Y. et al., 2017; Lu et al., 2018; Zhao et al., 2018). Other studies have reported that miR-510, PVT1, CCAT1, and Linc00861 may play important roles in breast invasive carcinoma (Zhang Y. et al., 2017). MiR-19a and lncRNA DLEU1 might be co-expressed to regulate the expression of ER in breast cancer (Wu et al., 2015).

In our research, we successfully obtained differentially expressed miRNA and lncRNA in the different stages of breast cancer by bioinformatics analysis based on large numbers of samples and clinical information from the TCGA database. Moreover, we analyzed the abnormal expression patterns of miRNA and lncRNA by STC analysis. A total of 67 miRNAs, 122 lncRNAs, and 119 mRNAs were selected to construct the non-coding regulatory network. The results showed that the miR-93/20b/106a/106b family was in the center of the

regulatory network, which may help to determine the key lncRNA and target genes that participated in the different stages of breast cancer. MiR-20b, miR-106a, miR-93, and miR-106b were overexpressed in tumor tissue, serum, or plasma samples from patients with cancer, including breast cancer (Fang et al., 2012; Deng et al., 2014; Zheng et al., 2015; Li M. et al., 2018). It has been demonstrated that the miR-93/20b/106a/106b family plays important roles in breast cancer and other cancers by regulating their target genes. Notch, MAPK, PI3K/Akt, and Hippo pathways are also involved in the regulation of miR-93/20b/106a/106b family (Shen et al., 2016; Li et al., 2017; Guarnieri et al., 2018; Li L. et al., 2018).

We also observed that 6 miRNAs (miR-7-1-3p, miR-20b-5p, miR-105-5p, miR-200c-3p, miR-204-5p, and miR-342-3p) and 10 lncRNA (DNMBP-AS1, HPN-AS1, LINC00461, LINC00472, LINC00667, LOC100128164, LOC284578, MESTIT1, RHPN1-AS1, and SLC26A4-AS1) can predict overall survival of patients with breast cancer. In existing research, highly expressed genes were significantly associated with cancer patients' worse overall survival, and genes with low expression were significantly associated with better overall survival. In this study, we found that upregulated miRNAs (miR-200c-3p, miR-20b-5p, miR-7-1-3p, and miR-342-3p) were significantly associated with breast cancer patients' better overall survival.

Previous findings also showed that miRNA was a universal mechanism of post-transcriptional gene regulation that controlled precise gene expression. MiRNA is the regulator of gene expression, and accumulating evidence has suggested

that miRNAs share a key role in breast cancer progression through the regulation of their target genes. More importantly, a single miRNA may regulate multiple target genes (Buchan and Parker, 2007; Wu and Pfeffer, 2016); therefore, multiple target genes jointly determine the function of miRNA in cancer patients' overall survival. For these reasons, some high expression of miRNAs is associated with breast cancer patients' better overall survival. Besides, compared to mRNA and lncRNA, the regulation of miRNA expression is complicated. During transcription, miRNAs were transcribed by RNA Polymerase II as long primary transcripts characterized by hairpin structures (pri-miRNAs). After transcription, the enzymes Drosha and Dicer process pri-miRNAs and pre-miRNAs to generate mature miRNAs (Feng et al., 2011; Iorio and Croce, 2012). Therefore, our results suggest that lncRNA/mRNA may be more suitable as prognostic biomarkers than miRNAs. The paradox of relationship between miRNA expression with breast cancer prognosis may be important in the progression of breast cancer, and it is more important to clarify this phenomenon in the prevention and treatment of breast cancer.

CONCLUSION

In conclusion, we screened abnormal expression of miRNA and lncRNA using TCGA data. STC analysis, lncRNA–miRNA–mRNA regulatory network, GEO data, and cell experiments showed that key miRNA and lncRNA performed vital roles in breast cancer progression. However, these non-coding RNA related functions and mechanisms still need to be further developed, especially these miRNA and lncRNA with prognostic value. Targeting these key non-coding RNAs may provide new therapeutic options for patients with breast cancer and may prevent the development of breast cancer from the early stages to an advanced disease.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga.

AUTHOR CONTRIBUTIONS

ML and CL constructed the main conceptual ideas and proposed the research. SG, XL, JM, and QZ conceived the technical details, performed the analysis, and prepared the manuscript. RT, ZF, and FW collected and processed the data. All authors contributed to the article and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.621809/full#supplementary-material

REFERENCES

Aalto, A. P., and Pasquinelli, A. E. (2012). Small non-coding RNAs mount a silent revolution in gene expression. *Curr. Opin. Cell. Biol.* 24, 333–340. doi: 10.1016/j.ceb.2012.03.006

Adhami, M., Haghdoost, A. A., Sadeghi, B., and Malekpour Afshar, R. (2018). Candidate miRNAs in human breast cancer biomarkers: a systematic review. *Breast Cancer* 25, 198–205. doi: 10.1007/s12282-017-0814-8

An, S., and Song, J. J. (2011). The coded functions of noncoding RNAs for gene regulation. *Mol. Cells.* 31, 491–496. doi: 10.1007/s10059-011-1004-8

Augoff, K., Mccue, B., Plow, E. F., and Sossey-Alaoui, K. (2012). miR-31 and its host gene lncRNA LOC554202 are regulated by promoter hypermethylation in triple-negative breast cancer. *Mol. Cancer* 11:5. doi: 10.1186/1476-4598-11-5

Becker, W., Nagarkatti, M., and Nagarkatti, P. S. (2018). miR-466a targeting of TGF-beta2 contributes to FoxP3(+) regulatory T cell differentiation in a murine model of allogeneic transplantation. *Front. Immunol.* 9:688. doi: 10.3389/fimmu.2018.00688

Blythe, A. J., Fox, A. H., and Bond, C. S. (2016). The ins and outs of lncRNA structure: how, why and what comes next? *Biochim. Biophys. Acta* 1859, 46–58. doi: 10.1016/j.bbagrm.2015.08.009

Buchan, J. R., and Parker, R. (2007). Molecular biology. the two faces of miRNA. *Science* 318, 1877–1878. doi: 10.1126/science.1152623

Chan, J. J., and Tay, Y. (2018). Noncoding RNA:RNA regulatory networks in cancer. *Int. J. Mol. Sci.* 19:1310. doi: 10.3390/ijms19051310

Chen, C., Li, Z., Yang, Y., Xiang, T., Song, W., and Liu, S. (2015). Microarray expression profiling of dysregulated long non-coding RNAs in triple-negative breast cancer. *Cancer Biol. Ther.* 16, 856–865. doi: 10.1080/15384047.2015.1040957

Deng, Z. Q., Qian, J., Liu, F. Q., Lin, J., Shao, R., Yin, J. Y., et al. (2014). Expression level of miR-93 in formalin-fixed paraffin-embedded tissues of breast cancer patients. *Genet. Test Mol. Biomarkers* 18, 366–370. doi: 10.1089/gtmb.2013.0440

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789. doi: 10.1101/gr.132159.111

DeSantis, C. E., Ma, J., Goding Sauer, A., Newman, L. A., and Jemal, A. (2017). Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J. Clin.* 67, 439–448. doi: 10.3322/caac.21412

Fang, L., Du, W. W., Yang, W., Rutnam, Z. J., Peng, C., Li, H., et al. (2012). MiR-93 enhances angiogenesis and metastasis by targeting LATS2. *Cell Cycle* 11, 4352–4365. doi: 10.4161/cc.22670

Feng, B., Chen, S., Mcarthur, K., Wu, Y., Sen, S., Ding, Q., et al. (2011). miR-146a-Mediated extracellular matrix protein production in chronic diabetes complications. *Diabetes* 60, 2975–2984. doi: 10.2337/db11-0478

Finnegan, E. F., and Pasquinelli, A. E. (2013). MicroRNA biogenesis: regulating the regulators. *Crit. Rev. Biochem. Mol. Biol.* 48, 51–68. doi: 10.3109/10409238.2012.738643

Ghafouri-Fard, S., Taherian-Esfahani, Z., Dashti, S., Kholghi Oskooei, V., Taheri, M., and Samsami, M. (2020). Gene expression of indoleamine and tryptophan dioxygenases and three long non-coding RNAs in breast cancer. *Exp. Mol. Pathol.* 114:104415. doi: 10.1016/j.yexmp.2020.104415

Gibori, H., Eliyahu, S., Krivitsky, A., Ben-Shushan, D., Epshtein, Y., Tiram, G., et al. (2018). Amphiphilic nanocarrier-induced modulation of PLK1 and miR-34a leads to improved therapeutic response in pancreatic cancer. *Nat. Commun.* 9:16. doi: 10.1038/s41467-017-02283-9

Greene, F. L. (2004). TNM: our language of cancer. *CA Cancer J. Clin.* 54, 129–130.

Guarnieri, A. L., Towers, C. G., Drasin, D. J., Oliphant, M. U. J., Andrysik, Z., Hotz, T. J., et al. (2018). The miR-106b-25 cluster mediates breast tumor initiation through activation of NOTCH1 via direct repression of NEDD4L. *Oncogene* 37, 3879–3893. doi: 10.1038/s41388-018-0239-7

Iorio, M. V., and Croce, C. M. (2012). Causes and consequences of microRNA dysregulation. *Cancer J.* 18, 215–222. doi: 10.1097/PPO.0b013e3182 50c001

Johnnidis, J. B., Harris, M. H., Wheeler, R. T., Stehling-Sun, S., Lam, M. H., Kirak, O., et al. (2008). Regulation of progenitor cell proliferation and granulocyte function by microRNA-223. *Nature* 451, 1125–1129. doi: 10.1038/nature 06607

Karimi Mazraehshah, M., Tavangar, S. M., Saidijam, M., Amini, R., Bahreini, F., Karimi Dermani, F., et al. (2018). Anticancer effects of miR-200c in colorectal cancer through BMI1. *J. Cell. Biochem.* 119, 10005–10012. doi: 10.1002/jcb.27330

Li, F., Xiao, Y., Huang, F., Deng, W., Zhao, H., Shi, X., et al. (2015). Spatiotemporal-specific lncRNAs in the brain, colon, liver and lung of macaque during development. *Mol. Biosyst.* 11, 3253–3263. doi: 10.1039/c5mb00474h

Li, L., Zhao, J., Huang, S., Wang, Y., Zhu, L., Cao, Y., et al. (2018). MiR-93-5p promotes gastric cancer-cell progression via inactivation of the Hippo signaling pathway. *Gene* 641, 240–247. doi: 10.1016/j.gene.2017.09.071

Li, M., Zhou, Y., Xia, T., Zhou, X., Huang, Z., Zhang, H., et al. (2018). Circulating microRNAs from the miR-106a-363 cluster on chromosome X as novel diagnostic biomarkers for breast cancer. *Breast Cancer Res. Treat.* 170, 257–270. doi: 10.1007/s10549-018-4757-3

Li, N., Miao, Y., Shan, Y., Liu, B., Li, Y., Zhao, L., et al. (2017). MiR-106b and miR-93 regulate cell progression by suppression of PTEN via PI3K/Akt pathway in breast cancer. *Cell. Death Dis.* 8:e2796. doi: 10.1038/cddis.2017.119

Lu, G., Li, Y., Ma, Y., Lu, J., Chen, Y., Jiang, Q., et al. (2018). Long noncoding RNA LINC00511 contributes to breast cancer tumourigenesis and stemness by inducing the miR-185-3p/E2F1/Nanog axis. *J. Exp. Clin. Cancer Res.* 37:289. doi: 10.1186/s13046-018-0945-6

Luan, S., Li, P., and Yi, T. (2017). Series test of cluster and network analysis for lupus nephritis, before and after IFN-K-immunosuppressive therapy. *Nephrology (Carlton)* 23, 997–1006. doi: 10.1111/nep.13159

Lv, J., Xia, K., Xu, P., Sun, E., Ma, J., Gao, S., et al. (2014). miRNA expression patterns in chemoresistant breast cancer tissues. *Biomed. Pharmacother.* 68, 935–942. doi: 10.1016/j.biopha.2014.09.011

Mondal, T., Subhash, S., Vaid, R., Enroth, S., Uday, S., Reinius, B., et al. (2015). MEG3 long noncoding RNA regulates the TGF-beta pathway genes through formation of RNA-DNA triplex structures. *Nat. Commun.* 6:7743. doi: 10.1038/ncomms8743

Ni, Q., Stevic, I., Pan, C., Muller, V., Oliveira-Ferrer, L., Pantel, K., et al. (2018). Different signatures of miR-16, miR-30b and miR-93 in exosomes from breast cancer and DCIS patients. *Sci. Rep.* 8:12974. doi: 10.1038/s41598-018-31108-y

Ponting, C. P., Oliver, P. L., and Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell* 136, 629–641. doi: 10.1016/j.cell.2009.02.006

Ren, L., Chen, H., Song, J., Chen, X., Lin, C., Zhang, X., et al. (2019). MiR-454-3p-mediated Wnt/beta-catenin signaling antagonists suppression promotes breast cancer metastasis. *Theranostics* 9, 449–465. doi: 10.7150/thno.29055

Salvador, M. A., Wicinski, J., Cabaud, O., Toiron, Y., Finetti, P., Josselin, E., et al. (2013). The histone deacetylase inhibitor abexinostat induces cancer stem cells differentiation in breast cancer with low Xist expression. *Clin. Cancer Res.* 19, 6520–6531. doi: 10.1158/1078-0432.CCR-13-0877

Saultz, J. N., Freud, A. G., and Mundy-Bosse, B. L. (2018). MicroRNA regulation of natural killer cell development and function in leukemia. *Mol. Immunol.* 115, 12–20. doi: 10.1016/j.molimm.2018.07.022

Shen, C. T., Qiu, Z. L., Song, H. J., Wei, W. J., and Luo, Q. Y. (2016). miRNA-106a directly targeting RARB associates with the expression of Na(+)/I(-) symporter in thyroid cancer by regulating MAPK signaling pathway. *J. Exp. Clin. Cancer Res.* 35:101. doi: 10.1186/s13046-016-0377-0

Siegel, R. L., Miller, K. D., and Jemal, A. (2018). Cancer statistics, 2018. *CA Cancer J. Clin.* 68, 7–30. doi: 10.3322/caac.21442

Singletary, S. E., and Connolly, J. L. (2006). Breast cancer staging: working with the sixth edition of the AJCC Cancer Staging Manual. *CA Cancer J. Clin.* 56, 37–47. quiz 50-31,

Sochor, M., Basova, P., Pesta, M., Dusilkova, N., Bartos, J., Burda, P., et al. (2014). Oncogenic microRNAs: miR-155, miR-19a, miR-181b, and miR-24 enable monitoring of early breast cancer in serum. *BMC Cancer* 14:448. doi: 10.1186/1471-2407-14-448

Sun, E. H., Zhou, Q., Liu, K. S., Wei, W., Wang, C. M., Liu, X. F., et al. (2014). Screening miRNAs related to different subtypes of breast cancer with miRNAs microarray. *Eur. Rev. Med. Pharmacol. Sci.* 18, 2783–2788.

Sun, J., Cai, X., Yung, M. M., Zhou, W., Li, J., Zhang, Y., et al. (2018). miR-137 mediates the functional link between c-Myc and EZH2 that regulates cisplatin resistance in ovarian cancer. *Oncogene* 38, 564–580. doi: 10.1038/s41388-018-0459-x

Taft, R. J., Pang, K. C., Mercer, T. R., Dinger, M., and Mattick, J. S. (2010). Non-coding RNAs: regulators of disease. *J. Pathol.* 220, 126–139. doi: 10.1002/path.2638

Van Schooneveld, E., Wouters, M. C., Van Der Auwera, I., Peeters, D. J., Wildiers, H., Van Dam, P. A., et al. (2012). Expression profiling of cancerous and normal breast tissues identifies microRNAs that are differentially expressed in serum from patients with (metastatic) breast cancer and healthy volunteers. *Breast Cancer Res.* 14:R34. doi: 10.1186/bcr3127

Wang, J., Song, C., Tang, H., Zhang, C., Tang, J., Li, X., et al. (2017). miR-629-3p may serve as a novel biomarker and potential therapeutic target for lung metastases of triple-negative breast cancer. *Breast Cancer Res.* 19:72. doi: 10.1186/s13058-017-0865-y

Wang, Z., Katsaros, D., Biglia, N., Shen, Y., Loo, L., Yu, X., et al. (2019). ERalpha upregulates the expression of long non-coding RNA LINC00472 which suppresses the phosphorylation of NF-kappaB in breast cancer. *Breast Cancer Res. Treat.* 175, 353–368. doi: 10.1007/s10549-018-05108-5

Wu, Q., Guo, L., Jiang, F., Li, L., Li, Z., and Chen, F. (2015). Analysis of the miRNA-mRNA-lncRNA networks in ER+ and ER- breast cancer cell lines. *J. Cell. Mol. Med.* 19, 2874–2887. doi: 10.1111/jcmm.12681

Wu, Z. H., and Pfeffer, L. M. (2016). MicroRNA regulation of F-box proteins and its role in cancer. *Semin. Cancer Biol.* 36, 80–87. doi: 10.1016/j.semcancer.2015.09.016

Xu, Z., Bu, Y., Chitnis, N., Koumenis, C., Fuchs, S. Y., and Diehl, J. A. (2016). miR-216b regulation of c-Jun mediates GADD153/CHOP-dependent apoptosis. *Nat. Commun.* 7:11422. doi: 10.1038/ncomms11422

Zhang, H. Y., Liang, F., Zhang, J. W., Wang, F., Wang, L., and Kang, X. G. (2017). Effects of long noncoding RNA-ROR on tamoxifen resistance of breast cancer cells by regulating microRNA-205. *Cancer Chemother. Pharmacol.* 79, 327–337. doi: 10.1007/s00280-016-3208-2

Zhang, Y., Li, Y., Wang, Q., Zhang, X., Wang, D., Tang, H. C., et al. (2017). Identification of an lncRNAmiRNAmRNA interaction mechanism in breast cancer based on bioinformatic analysis. *Mol. Med. Rep.* 16, 5113–5120. doi: 10.3892/mmr.2017.7304

Zhao, W., Geng, D., Li, S., Chen, Z., and Sun, M. (2018). LncRNA HOTAIR influences cell growth, migration, invasion, and apoptosis via the miR-20a-5p/HMGA2 axis in breast cancer. *Cancer Med.* 7, 842–855. doi: 10.1002/cam4.1353

Zheng, J. Z., Huang, Y. N., Yao, L., Liu, Y. R., Liu, S., Hu, X., et al. (2018). Elevated miR-301a expression indicates a poor prognosis for breast cancer patients. *Sci. Rep.* 8:2225. doi: 10.1038/s41598-018-20680-y

Zheng, R., Pan, L., Gao, J., Ye, X., Chen, L., Zhang, X., et al. (2015). Prognostic value of miR-106b expression in breast cancer patients. *J. Surg. Res.* 195, 158–165. doi: 10.1016/j.jss.2014.12.035

Zhu, M., Lv, Q., Huang, H., Sun, C., Pang, D., and Wu, J. (2020). Identification of a four-long non-coding RNA signature in predicting breast cancer survival. *Oncol. Lett.* 19, 221–228. doi: 10.3892/ol.2019.11063