



# Identification of Key lncRNA–mRNA Pairs and Functional lncRNAs in Breast Cancer by Integrative Analysis of TCGA Data

Zhe Li<sup>1\*†</sup>, Zheng Qian<sup>1†</sup>, Fei Chen<sup>1</sup>, Shujun Jiang<sup>1</sup>, Lingjia Meng<sup>2</sup> and Jinzhong Chen<sup>3\*</sup>

<sup>1</sup> Department of Breast Surgery, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai, China, <sup>2</sup> Department of General Surgery, Putuo Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai, China, <sup>3</sup> State Key Laboratory of Genetic Engineering, School of Life Sciences, Fudan University, Shanghai, China

## OPEN ACCESS

### Edited by:

Ellen Jones,  
National Center for Toxicological  
Research (FDA), United States

### Reviewed by:

Vedbar Singh Khadka,  
University of Hawaii, United States  
Suilan Zheng,  
Purdue University, United States

### \*Correspondence:

Zhe Li  
lizhezsh@shutcm.edu.cn  
Jinzhong Chen  
kingbellchen@fudan.edu.cn

<sup>†</sup> These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Cancer Genetics and Oncogenomics,  
a section of the journal  
Frontiers in Genetics

Received: 14 May 2021

Accepted: 16 July 2021

Published: 20 August 2021

### Citation:

Li Z, Qian Z, Chen F, Jiang S,  
Meng L and Chen J (2021)  
Identification of Key lncRNA–mRNA  
Pairs and Functional lncRNAs  
in Breast Cancer by Integrative  
Analysis of TCGA Data.  
Front. Genet. 12:709514.  
doi: 10.3389/fgene.2021.709514

Long non-coding RNAs (lncRNAs) play an important role in many diseases and are involved in the post-transcriptional regulatory network of tumors. The purpose of this study is to mine new lncRNA–mRNA regulatory pairs and analyze the new mechanism of lncRNA involvement in breast cancer progression. Using breast cancer miRNA and mRNA expression profiling from The Cancer Genome Atlas (TCGA), we identified 59 differentially expressed lncRNAs, 88 differentially expressed miRNAs, and 1,465 differentially expressed mRNAs between breast cancer tissue and adjacent normal breast cancer. Whereafter, four candidate lncRNAs (FGF14-AS2, LINC01235, AC055854.1, and AC124798.1) were identified by the Kaplan–Meier (K–M) plotter. Furthermore, we screened the hub lncRNA (LINC01235) through univariate Cox analysis, multivariate Cox analysis, and qPCR validation, which was significantly correlated with breast cancer stage, ER status, and pathological N. Subsequently, 107 LINC01235-related mRNAs were obtained by combining differentially expressed miRNAs, differentially expressed mRNAs, and LINC01235 targeting miRNAs and mRNAs. The protein–protein interaction (PPI) network was established by Cytoscape software, and 53 key genes were screened. Function and pathway enrichment showed that LINC01235-related key genes might be involved in the process of cell differentiation, cell proliferation, and p53 signal pathway. In addition, LINC01235 has been confirmed to regulate the proliferation, migration, and invasion of MCF-7 cells in *in vitro* experiments. Furthermore, we screened three mRNAs (ESR1, ADRA2A, and DTL) associated with breast cancer drug resistance from key genes. Through RNA interference experiments *in vitro* and correlation analysis, we found that there was a negative feedback mechanism between LINC01235 and ESR1/ADRA2A. In conclusion, our results suggest that LINC01235–ESR1 and LINC01235–ADRA2A could serve as important co-expression pairs in the progression of breast cancer, and LINC01235 plays a key role as an independent prognostic factor in patients with breast cancer. The findings of this work greatly increase our understanding of the molecular regulatory mechanisms of lncRNA in breast cancer.

**Keywords:** LINC01235, proliferation, prognosis, breast cancer, invasion

## INTRODUCTION

Long non-coding RNAs (lncRNAs) are a class of RNA molecules that are over 200 nucleotides in length, which have no protein-coding function but have epigenetic regulation and other biological functions (Peng et al., 2017). Currently, mounting evidences have demonstrated that lncRNAs and mRNAs played critical roles in various physiological processes, such as cell activity, gene regulation, and so on (Fukushima and Kida, 2019; De Troyer et al., 2020; Labonté et al., 2020). Therefore, dissecting how lncRNAs regulate tumor progress attracts great interest from researchers (Majidinia and Yousefi, 2020). In breast cancer (BRCA), high levels of BCRT1, SPRY4-IT1, RP11-19E11, and FAM83H-AS1 were associated with clinical stage and prognosis (Giro-Perafita et al., 2020; Han et al., 2020; Liang et al., 2020; Song et al., 2020). These studies indicated that lncRNAs could be regarded as valuable biomarkers and therapeutic targets for cancer therapy.

Breast cancer is one of the common malignant tumors among women and ranks first among malignant tumors in women (DeSantis et al., 2019). In China, the age of patients suffering from breast cancer is getting lower (Rossi et al., 2019). However, treatments may temporarily control tumor growth but very often lead to a chemotherapy resistance status (Tang et al., 2016). So, it is very important and meaningful to find new markers or therapeutic targets. LINC01235 is located on the chromosome 9 p23 and expressed in a variety of cancer tissues. It has been reported that the expression of LINC01235 was specific in a specific environment or tissue and had a characteristic expression pattern in tumors (Tang et al., 2019; Xu et al., 2019). In gastric cancer, LINC01235 can promote gastric cancer cell metastasis through EMT and be used as a biomarker of prognosis (Tan et al., 2020). In breast cancer, the expression of LINC01235 was analyzed (Vishnubalaji et al., 2019). The high expression of LINC01235 was associated with poor prognosis of breast cancer patients (Li et al., 2021; Zhang K. et al., 2021). These results are consistent with the data we analyzed. However, the regulation of LINC01235 on the progression of breast cancer has not been reported. In addition, it is strange that the low expression of LINC01235 in breast cancer has a good prognosis, while the high expression of LINC01235 has a poor prognosis. At present, there is no relevant experimental data to explain this problem.

In this study, we explored the differentially expressed lncRNAs in the breast cancer based on The Cancer Genome Atlas (TCGA) dataset. Combined with the Kaplan–Meier (K–M) plotter analysis, protein–protein interaction (PPI) network, expression analysis, and RNA interference experiment, we found the most relevant mRNA (ESR1 and ADRA2A) of LINC01235, and there was an obvious negative feedback between LINC01235 and two mRNAs. Moreover, a series of interference analyses and validation experiments in MCF-7 cells were used to analyze the effect of LINC01235 on the progression of breast cancer. All results may be used to provide data reference for the study of LINC01235 to regulate breast cancer deterioration.

## MATERIALS AND METHODS

### Patients' Enrollment From TCGA–BRCA

Our research excluded any samples that had missing or insufficient data on age, TNM stage, OS time, lymph node metastases, and ER status. We retained both RNA-Seq and clinical data, which we used for further investigation. The detailed clinical and pathological characteristics of the study population are shown in **Table 1**. A total of 1,028 patients were pathologically diagnosed as having breast cancer. The median age of all patients was 52 years (26–91 years). Majority of the patients were white (68.58%) with high histological grade (91.21%).

### Screening of Differentially Expressed RNAs

Breast cancer RNA-seq sequencing expression profile data, miRNA sequencing (miRNA-seq) expression profile data, and clinical data were downloaded from the open access data of the TCGA database.<sup>1</sup> We obtained 30 adjacent normal and 300 tumor sample dates for mRNA and lncRNA analysis and obtained 10 adjacent normal and 357 tumor sample dates for miRNA analysis. The clinical information of the cases was also collected, including age, stage, pathologic tumor (pathologic T), pathologic node (pathologic N), pathologic metastasis (pathologic M), lymph node metastases, and ER status. Differentially expressed mRNAs, differentially expressed miRNAs, and differentially expressed lncRNAs between the breast cancer tumor tissues and the normal samples were screened using the limma package in R, with the criterion of  $|\log_2| \text{Fold Change}| > 1.5$ , adjusted  $p$ -value  $< 0.01$ , and FDR (false discovery rate)  $< 0.05$ .

### Kaplan–Meier Survival Curve Analysis

To investigate the impact of the expression levels of differentially expressed RNAs on the prognostic survival of breast cancer patients, K–M survival curve analysis was performed using the online tool “GEPIA.”<sup>2</sup> The statistical significance was set at  $p < 0.05$ . Then, the data obtained from the analysis are verified by Starbase<sup>3</sup>.

### RNA Extraction and Real-Time Quantitative Reverse Transcription Polymerase Chain Reaction

A total of 35 breast cancer tissues and 35 adjacent normal tissues were collected. This study was approved by the Ethical Committee of Hospital and was conducted in accordance with the Declaration of Helsinki. In addition, each patient offered written informed consent. All tissues were immediately stored at  $-80^\circ\text{C}$ .

<sup>1</sup><https://tcga-data.nci.nih.gov/>

<sup>2</sup><http://gepia2.cancer-pku.cn/#index>

<sup>3</sup><http://starbase.sysu.edu.cn/>

**TABLE 1** | Clinical pathological characteristics of 1,028 patients with breast cancer.

Parameter	Subtype	Patientsn (%)
Age (Years)	26 to ≤52	344 (33.46%)
	52 to ≤91	684 (66.54%)
Race	White	705 (68.58%)
	Black or African American	173 (16.83%)
	Asian	56 (5.45%)
	Not reported	94 (9.14%)
Pathologic stage	Stage i	170 (16.54%)
	Stage ii	584 (56.82%)
	Stage iii	231 (22.47%)
	Stage iv	19 (1.85%)
	Missing	24 (2.33%)
Pathologic T	T1	267 (25.97%)
	T2	597 (58.08%)
	T3	125 (12.16%)
	T4	36 (3.5%)
	TX	3 (0.29%)
Pathologic M	M0	855 (83.17%)
	M1	21 (2.04%)
	MX	152 (14.79%)
Pathologic N	N0	474 (46.11%)
	N1	346 (33.65%)
	N2	117 (11.38%)
	N3	71 (6.91%)
	NX	20 (1.95%)
Vital Status	Alive	885 (86.09%)
	Dead	143 (13.91%)

Total RNA of tissue was extracted using the RNeasy Mini Kit (Qiagen, Germany) according to the manufacturer's protocol. Then, the extracted RNA was qualitatively controlled and quantified by Nanodrop. The equivalent amount of RNA was reverse transcribed into cDNA with Golden star TM RT6 cDNA synthesis KIT ver.2 (TsingKe, Beijing, China). lncRNA and mRNA expression was then examined by real-time quantitative reverse transcription polymerase chain reaction (qRT-PCR) with Master qPCR Mix (SYBR GREEN 1) (TsingKe, Beijing, China) and Biosystems 7,500 Fast Real-time PCR System. GAPDH expression was used as endogenous control. The primers were designed using Primer 5.0 software (Table 2). The  $2^{-\Delta\Delta Ct}$  method was used to calculate the levels of lncRNA expression and mRNA expression. Excel was used to analyze the qRT-PCR data, and each reaction was performed in triplicate. Two groups were performed using *t*-test ( $p < 0.05$ ).

## Prediction of lncRNA Subcellular Localization

Long non-coding RNA sequences were obtained from LNCipedia.<sup>4</sup> lncRNA cellular localization was predicted by its sequence using lncLocator.<sup>5</sup> Localization is determined by the scores of different subcellular locations.

<sup>4</sup><https://lncipedia.org/>

<sup>5</sup><http://www.csbio.sjtu.edu.cn/bioinf/lncLocator/>

**TABLE 2** | The sequences of primers.

Gene symbol	Primer	Sequences (5'-3')
LINC01235	Forward primer	GCCTACCTTACCTGTGGCTC
	Reverse primer	CGTTGACCTGTGAGAGACCC
FGF14-AS2	Forward primer	AGTTCAGTTACCATCTTCA
	Reverse primer	AGGTTCCATAGTTGCCAGAC
AC055854.1	Forward primer	CAGTGTCTCAGGTCACGCTT
	Reverse primer	GTAGGGTGAGGGGAGGGAAT
AC124798.1	Forward primer	CACGTGGGCCAGAAAAAGTG
	Reverse primer	GAGGGAAGGTGGAAGTCAGC
ESR1	Forward primer	TGAACCATCACTGAGGTGG
	Reverse primer	TGGATCTGATGCAGTAGGG
ADRA2A	Forward primer	AGAAGTGGTAGCTCATCTCG
	Reverse primer	ATCATGATGAGGCAGGGAG

## Protein-Protein Interaction Network Construction

Three databases (Starbase,<sup>6</sup> DIANA tools,<sup>7</sup> and miRTarbas<sup>8</sup>) were used to verify the miRNA-protein coding gene mRNA and miRNA-lncRNA regulation correlation. Only the overlapping target genes were identified to enhance the reliability of the bioinformatics analysis. Then, overlapping target genes were used to construct the PPI network. The PPIs were analyzed using the STRING database (STRING<sup>9</sup>) and a combined score  $> 0.5$  was used as the cutoff criterion. Additionally, Cytoscape was used to construct and visualize the PPI network.

## Cox Proportional Regression Model Based on Differentially Expressed lncRNAs

In order to analyze the independent effect of individual lncRNA on the overall survival (OS) of breast cancer patients, we used online tools (SangerBox tools<sup>10</sup>) to carry out univariate and multivariate Cox proportional regression analysis. We constructed a Cox proportional hazard regression model to calculate the value at risk of each patient. Risk score = expression of lncRNA1  $\times$   $\beta_{lncRNA1}$  + expression of lncRNA2  $\times$   $\beta_{lncRNA2}$  + ..... + expression of lncRNAn  $\times$   $\beta_{lncRNAn}$ .  $\beta$  represents the multivariate Cox regression coefficient. Then, we calculated the survival rate of the high-risk group and the low-risk group, and drew 1-, 3-, and 5-year ROC (survival receiver operating characteristic) curves to test the feasibility of the predictive ability of the model.

## Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Analysis

Functional enrichment analysis of lncRNA-target genes was performed using Metascape. All statistically enriched terms

<sup>6</sup><http://starbase.sysu.edu.cn/>

<sup>7</sup>[http://carolina.imis.athena-innovation.gr/diana\\_tools/web/](http://carolina.imis.athena-innovation.gr/diana_tools/web/)

<sup>8</sup><http://miRTarBase.cuhk.edu.cn/>

<sup>9</sup><https://string-db.org/>

<sup>10</sup><http://sangerbox.com/Tool>

[Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG)] were identified based on accumulative hypergeometric *p* values.

## Cell Culture and Cell Transfection

MCF-7 was purchased from the Beijing Beina Chuanglian Biotechnology Research Institute (Beijing, China). MCF-7 was cultured in Roswell Park Memorial Institute 1,640 medium (Gibco) supplemented with 10% fetal bovine serum (Thermo Scientific), 100 U/ml penicillin, and 100 µg/ml streptomycin under a humidified atmosphere of 5% CO<sub>2</sub> at 37°C. The cell lines have been identified and the fourth passage cells of MCF-7 were used to perform experiments in this study. MCF-7 cells were transfected with 50 nM siRNA (si-LINC01235, si-ESR1, and siADRA2A; Ribobio, Smart Silencer), pCDH-LINC01235 (Ribobio, Smart Silencer), and their NC controls for 48 h using Lipofectamine 3,000 reagent (Invitrogen, Shanghai, China) according to the manufacturer's instructions.

## Cell Counting Kit-8 Assays

Cell counting Kit-8 (CCK8) (Sigma, United States) was used to detect the proliferation of MCF-7 cells in 96-well plates. A total of 3,000 cells/well were incubated for 24 and 48 h. All cells were incubated with CCK-8 reagent (10 µl per well) for 3 h, and the absorbance of each well was detected by enzyme labeling instrument made by Thermo Company at 450 nm wavelength. Each experiment was conducted three times.

## Wound-Healing Assay

MCF-7 cells were inoculated in a petri dish (96 wells, Corning, United States) with  $5 \times 10^5$  cells per well. The temperature of the culture medium was 37°C and the concentration of CO<sub>2</sub> was 5%. Then, the fusion cell monolayer was scratched with the tip of a 200-ml aseptic pipette, and Opti-MEM reduced serum medium (Gibco, United States) was added. The scratches were recorded by a microscope at 0, 24, and 48 h later. The closure rate was evaluated by ImageJ software.

## Transwell Assay

After 48 h of transfection, MCF-7 cells were collected to prepare single-cell suspension. The MCF-7 cell suspension ( $3 \times 10^3$  cell/well) was added to the upper chamber (Corning, United States), and the Roswell Park Memorial Institute 1,640 medium (RPMI-1640, 20% fetal bovine serum) was added to the 24-well plate in the lower chamber. The upper chamber was coated with matrix glue. After 24 h, the cells were fixed with 4% paraformaldehyde (Sigma, United States) and stained with 1% crystal violet (Sigma, United States). Cells were observed and counted under a light microscope (Olympus, Japan).

## Statistical Analysis

SPSS and GraphPad Prism were used to analyze the experimental data. Student's *t*-test or ANOVA variance test was used for comparison. All variables were expressed as mean  $\pm$  standard deviation (SD), which were obtained from three independent experiments. *p* < 0.05 was a significant difference.

## RESULTS

### Differentially Expressed lncRNAs, miRNAs, and mRNAs in Breast Cancer Patients

In order to study the potential role of RNAs (lncRNAs, miRNAs, and mRNAs) in the development of breast cancer, we identified the differentially expressed RNAs using the limma package in R. The RNA expression profile data of breast cancer patients were downloaded from the open access data of the TCGA database. By setting the criteria thresholds to  $\log_2$  |Fold Change| > 1.5, adjusted *p*-value < 0.01, and FDR < 0.05, a total of 59 differentially expressed lncRNAs (24 upregulated and 35 downregulated) (Table 3), 88 differentially expressed miRNAs (42 upregulated and 46 downregulated) (Table 4), and 1,464 differentially expressed mRNAs (520 upregulated and 944 downregulated) (Table 5) were identified. Volcano plots (Figure 1A) and heatmap (Figure 1B) showed the number of differentially expressed RNAs identified from the TCGA database.

### Long Non-coding RNAs Associated With Prognosis and Clinical Parameters in Breast Cancer Patients

The prognosis of breast cancer according to the expression of differentially expressed lncRNAs was performed based on the dates obtained from the TCGA database by GEPIA database. A total of 1,097 cases were included in analysis. Among the 59

**TABLE 3** | Top10 (up- and down-regulated) of differentially expressed lncRNAs in normal tissues and breast cancer tissues.

ID	Symbol	logFC	AveExpr	t	p-Value
<b>Up regulation</b>					
ENSG00000204049	AL391421.1	1.960	-0.039	6.401	2.21E-10
ENSG00000227036	LINC00511	2.068	1.255	10.253	1.03E-23
ENSG00000218416	PP14571	2.094	1.086	6.844	1.22E-11
ENSG00000253125	AC055854.1	2.095	0.459	6.058	1.84E-09
ENSG00000265415	AC099850.3	2.098	0.270	12.976	3.80E-36
ENSG00000267751	AC009005.1	2.127	0.225	12.458	1.31E-33
ENSG00000228630	HOTAIR	2.216	0.872	6.119	1.27E-09
ENSG00000254973	AC105219.4	2.257	0.163	16.541	1.42E-55
ENSG00000261039	LINC02544	3.297	0.393	16.893	1.18E-57
ENSG00000230838	LINC01614	5.693	1.161	26.773	1.96E-124
<b>Down regulation</b>					
ENSG00000249669	CARMN	-4.206	0.747	-41.050	1.31E-231
ENSG00000180139	ACTA2-AS1	-2.854	0.985	-25.018	1.03E-111
ENSG00000196167	COLCA1	-2.671	0.273	-17.435	6.43E-61
ENSG00000215386	MIR99AHG	-2.517	1.231	-24.477	7.65E-108
ENSG00000234899	SOX9-AS1	-2.491	0.948	-14.640	8.26E-45
ENSG00000272143	FGF14-AS2	-2.394	-0.071	-21.585	1.07E-87
ENSG00000229847	EMX2OS	-2.388	0.598	-19.807	7.22E-76
ENSG00000234456	MAGI2-AS3	-2.365	3.204	-31.014	5.89E-156
ENSG00000267272	LINC01140	-2.346	0.470	-23.226	5.17E-99
ENSG00000270547	LINC01235	-2.242	1.199	-16.445	5.20E-55

**TABLE 4** | Top10 (up- and down-regulated) of differentially expressed miRNAs in normal tissues and breast cancer tissues.

Symbol	logFC	AveExpr	t	PValue	FDR
<b>Up regulation</b>					
hsa-miR-200a-5p	2.5124	8.8659	18.1566	2.73E-65	4.85E-64
hsa-miR-184	2.5491	2.7064	7.4899	1.33E-13	3.15E-13
hsa-miR-200a-3p	2.6410	7.7297	17.8111	3.70E-63	6.10E-62
hsa-miR-183-5p	2.9236	13.5554	24.5103	6.47E-108	4.80E-106
hsa-miR-190b	2.9605	3.0982	10.7373	9.58E-26	3.77E-25
hsa-miR-429	2.9959	6.3372	19.0099	1.18E-70	2.50E-69
hsa-miR-141-3p	3.0892	9.6728	22.9897	3.17E-97	1.41E-95
hsa-miR-96-5p	3.1552	4.7125	21.9740	3.13E-90	1.27E-88
hsa-miR-1307-5p	3.1966	7.3873	21.6791	3.17E-88	1.09E-86
hsa-miR-592	3.4271	1.2954	19.9167	1.70E-76	4.71E-75
<b>Down regulation</b>					
hsa-miR-486-5p	-4.0994	5.2776	-25.5942	1.13E-115	1.01E-113
hsa-miR-139-3p	-3.6041	2.6593	-27.8961	1.89E-132	4.20E-130
hsa-miR-204-5p	-3.5308	0.9813	-23.6935	3.83E-102	2.44E-100
hsa-miR-139-5p	-3.1947	5.2511	-31.1418	1.39E-156	6.17E-154
hsa-miR-451a	-3.0756	7.0965	-19.9251	1.49E-76	4.43E-75
hsa-miR-5683	-2.9633	2.3067	-15.7227	8.50E-51	9.23E-50
hsa-miR-144-5p	-2.7977	5.1749	-19.1086	2.77E-71	6.17E-70
hsa-miR-1247-3p	-2.7420	3.8113	-14.7795	1.53E-45	1.36E-44
hsa-miR-452-5p	-2.4861	4.7375	-16.3033	3.88E-54	4.80E-53
hsa-miR-145-5p	-2.4557	10.3201	-26.6543	2.35E-123	2.62E-121

deregulated lncRNAs, we found that the FGF14-AS2, LINC01235, AC055854.1, and AC124798.1 were statistically related to the OS in breast cancer patients ( $p < 0.05$ ; **Figure 2A**). The high expression levels of LINC01235 predicted poor prognosis, while the other three lncRNAs FGF14-AS2, AC055854.1, and AC124798.1 were positively correlated with patient survival. These results suggested that the four lncRNAs might play important roles in the development of breast cancer. Meanwhile, the expression levels of four differentially expressed lncRNAs (FGF14-AS2, LINC01235, AC055854.1, and AC124798.1) were analyzed using a univariate Cox proportional hazards regression model. The analysis results are consistent with those of the K-M curve. FGF14-AS2, LINC01235, AC055854.1, and AC124798.1 were identified as prognostic factors of breast cancer patients ( $p < 0.05$ ; **Figure 2B**).

Moreover, we evaluated the relationship among four lncRNA (FGF14-AS2, LINC01235, AC055854.1, and AC124798.1) expression levels and various clinicopathological parameters of breast cancer patients. Expression data and clinical characteristics were obtained from the TCGA-BRCA database. The results showed that the expression of FGF14-AS2, LINC01235, AC055854.1, and AC124798.1 was significantly correlated with tumor stage, pathologic TNM, and ER status ( $p < 0.05$ ; **Figure 2C**). From the multivariate Cox regression analysis, LINC01235 and AC124798.1 were further screened from the above four lncRNAs with prognostic significance (**Table 6**). The hazard ratios of AC124798.1 (HR = 0.695; 95% CI, 0.848–1.064;  $p = 0.037$ ) were less than 1; however, the hazard ratios of LINC01235 (HR = 1.688, 95% CI, 0.788–1.884;  $p = 0.017$ )

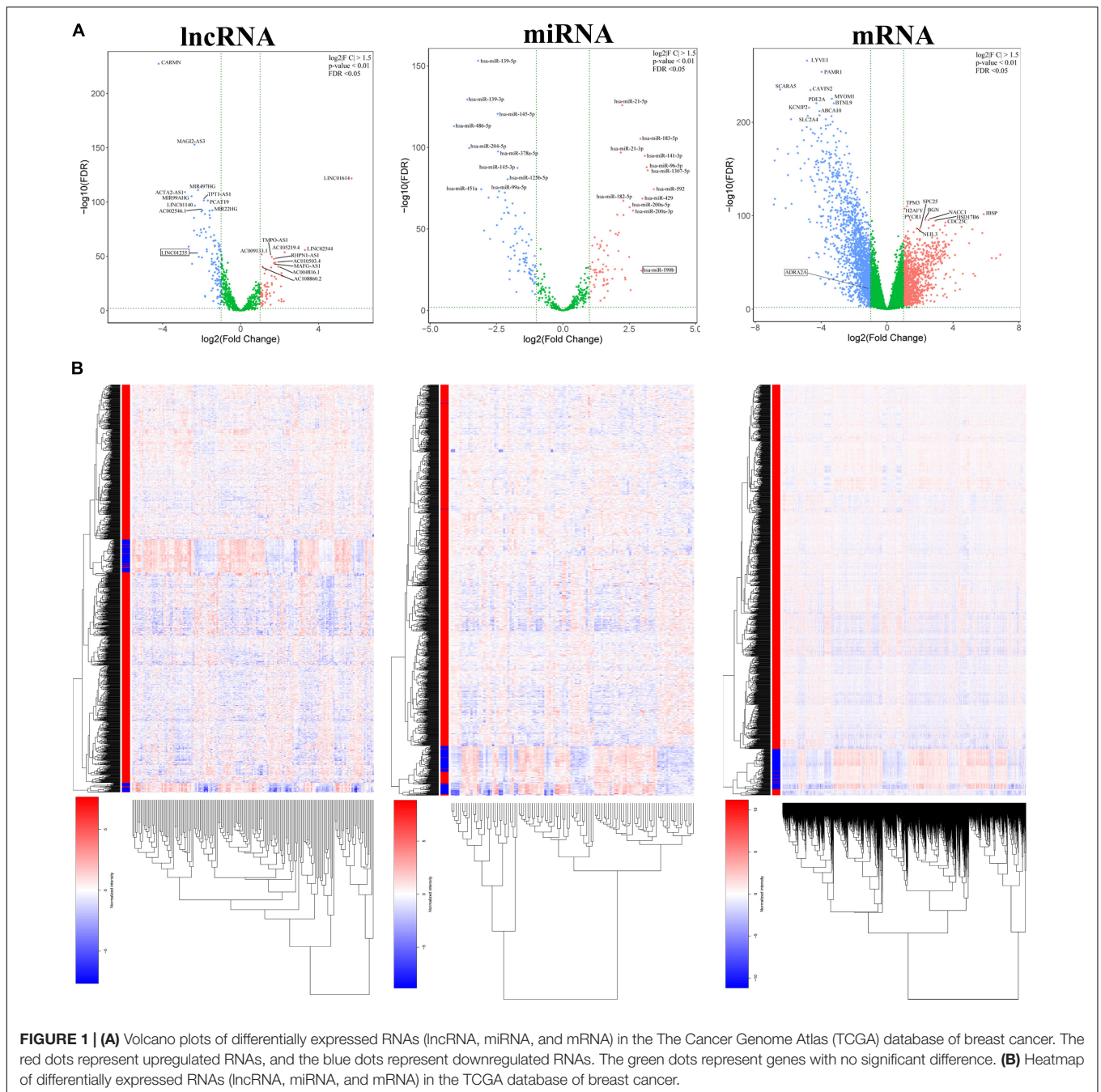
were greater than 1. Furthermore, LINC01235 and AC124798.1 were used to construct a risk score model. The risk scores for samples from TCGA were calculated using the formula: risk score =  $0.2036 \times$  expression of LINC01235 –  $0.0551 \times$  expression of AC124798.1. On the basis of the median risk score, breast cancer patients were divided into a high-risk and a low-risk group. The risk curve, scatterplot, and K-M analysis showed that the high-risk group displayed worse OS (**Figure 2D-1,2**). In order to confirm our findings, we calculated the risk score for the testing sets including GSE21653 ( $n = 252$ ) and GSE42568 ( $n = 104$ ). With the same cutoff, the patients from each testing set were separately classified into low-risk and high-risk groups. Similar to the findings obtained from the TCGA-BRCA set, patients in the high-risk group had shorter survival time than patients in the low-risk group (**Supplementary Figure 1A**). Besides, ROC curve analysis showed that the AUC of 1, 3, and 5 years was 0.62, 0.59, and 0.62, respectively (**Figure 2D-3**). All the results indicated that the risk score model has prognostic value. Therefore, these above results confirmed that LINC01235 and AC124798.1 have significant prognostic significance in breast cancer patients.

## Expression Validation of LINC01235 and AC124798.1 in Breast Cancer Tissues

As shown in **Figure 3A**, the bioinformatics analysis results showed that the expression level of LINC01235 was downregulated, and AC124798.1 was significantly increased compared to the normal tissues. To analyze the accuracy and reliability of present bioinformatics analysis, qPCR was used

**TABLE 5** | Top10 (up- and down-regulated) of differentially expressed mRNAs in normal tissues and breast cancer tissues.

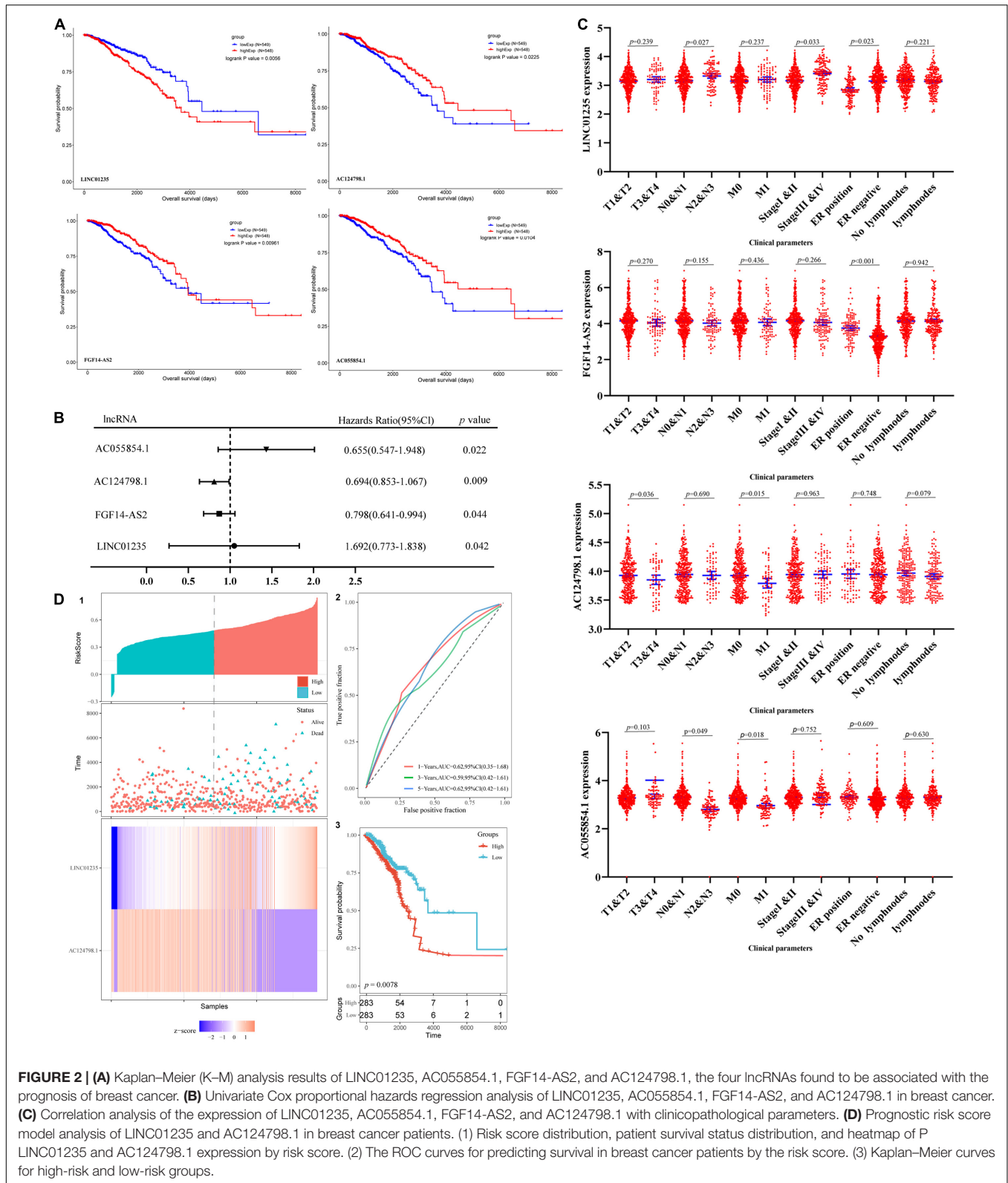
Symbol	logFC	AveExpr	t	p-Value	FDR
<b>Up regulation</b>					
COMP	4.5911	5.6252	13.6328	1.77E-39	1.38E-38
SLC24A2	4.9779	1.0255	19.3263	9.37E-73	2.27E-71
PLPP4	5.1782	1.7559	21.5819	1.12E-87	3.83E-86
MMP1	5.2342	1.9926	13.6401	1.62E-39	1.27E-38
COL11A1	5.3334	5.9023	13.1931	3.10E-37	2.24E-36
MMP11	5.7762	6.9196	20.9930	1.05E-83	3.30E-82
IBSP	5.8592	-0.4205	23.9461	4.48E-104	2.34E-102
MMP13	6.3473	2.6126	16.6949	1.76E-56	2.47E-55
CST1	6.6093	0.7129	16.6713	2.43E-56	3.37E-55
COL10A1	6.8648	5.663608	17.3076	3.81E-60	6.08E-59
<b>Down regulation</b>					
ADH1B	-6.8751	2.8739	-25.4161	1.41E-114	9.33E-113
TUSC5	-6.8274	0.5241	-31.8335	3.98E-162	9.37E-160
ADIPOQ	-6.8086	2.1216	-26.2480	1.32E-120	1.00E-118
CIDEC	-6.6436	0.7323	-33.4936	1.16E-174	3.95E-172
SCARA5	-6.4982	0.2681	-42.1088	1.71E-239	8.72E-236
FABP4	-6.1494	3.8951	-24.8773	1.05E-110	6.41E-109
PLIN1	-6.0173	3.2862	-29.7342	2.32E-146	3.18E-144
GPD1	-6.0055	2.6358	-32.8030	1.93E-169	5.56E-167
AQP7	-5.8304	0.5324	-37.7592	6.58E-207	6.70E-204
PLIN4	-5.8014	4.3026	-25.5145	2.75E-115	1.85E-113



to evaluate LINC01235 and AC124798.1 expression level in 35 breast cancer tissues and 35 adjacent tissues. We found that only LINC01235 was consistent with bioinformatics analysis results. The expression of LINC01235 in breast cancer tissues was significantly downregulated when compared to breast normal tissues ( $p < 0.001$ ; **Figure 3B**). Specifically, it was strange that LINC01235 is downregulated in breast cancer tissues when compared to normal specimens; however, its low expression correlates with a favorable prognosis. This may be the reverse regulation of LINC01235 in the regulation of breast cancer, which negatively affects the progression of breast cancer by

regulating another factor. The specific mechanism remains to be further studied.

As shown in **Figure 2C** we have analyzed the relationship between the expression level of LINC01235 and various clinical parameters of breast cancer patients. The analysis results showed that increased expression of LINC01235 was significantly correlated with pathological N ( $p < 0.05$ ), tumor stage ( $p < 0.05$ ), and ER status ( $p < 0.05$ ). In addition, univariate analysis revealed that LINC01235 expression, pathologic TNM ( $p < 0.05$ ), tumor stage ( $p < 0.05$ ), and ER status ( $p < 0.05$ ) were significantly correlated with the OS of breast cancer patients (**Table 7**).



Multivariate analysis showed that LINC01235 expression might be an independent factor for the prognosis of breast cancer patients (Figure 3C and Table 8). Meanwhile, the ROC curve

AUC of LINC01235 expression for predicting survival was 0.616 (Figure 3D), which indicated that LINC01235 possessed the potential prognostic ability of breast cancer. In order to further

verify the clinical value of LINC01235, we combined the two datasets GSE21653 and GSE42568 as independent testing sets for ROC. The area under receiver operating characteristic (AUROC) was 0.8124 ( $p < 0.0001$ ; **Supplementary Figure 1B**). These results indicated that breast cancer patients with high levels of LINC01235 expression are more likely to promote breast cancer progression.

## Construction of LINC01235-Related mRNA PPI Network

The cellular localization of lncRNAs determines their molecular mechanism. We first explored the subcellular localization of those lncRNAs by using lncLocator. LINC01235 was identified to be located in the cytoplasm (**Figure 4A**; score = 0.60), which indicated that LINC01235 was more likely to exert their biological functions through ceRNA mechanism. The target miRNAs of LINC01235 were analyzed by DIANA tools (**Supplementary Table 1**). Then, we summarized the same miRNAs between target miRNAs of LINC01235 and differentially expressed miRNAs from TCGA. Only miR-190b was the overlap miRNA (**Figure 4B**), which was significantly expressed compared with non-tumor tissues in TCGA breast cancer (**Figure 4C**). Whereafter, the target mRNAs of miR-190b were also analyzed by DIANA tools (**Supplementary Table 2**). Following the same method, we obtained 107 overlap mRNAs between targeted

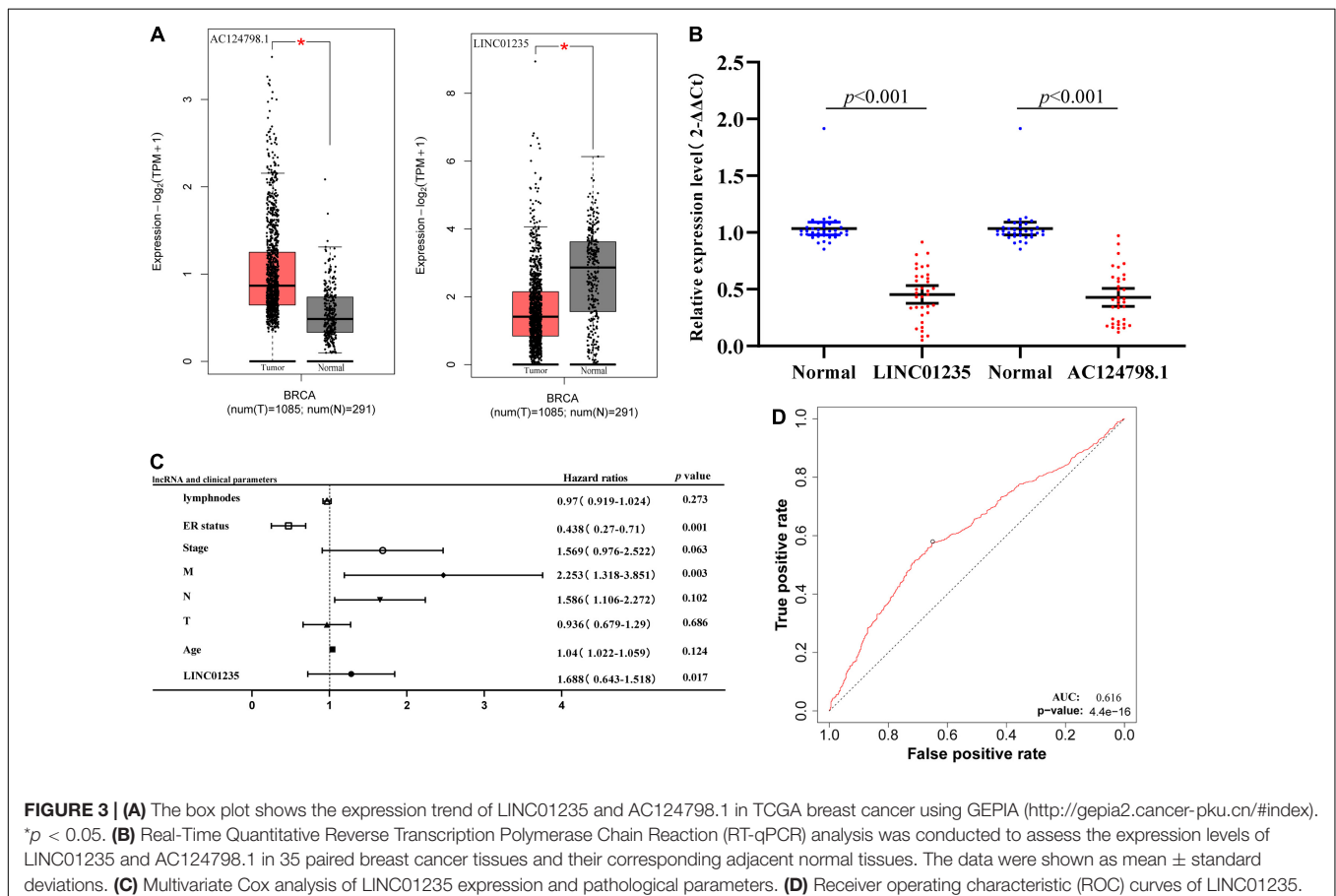
**TABLE 6** | Multivariate cox regression analysis of FGF14-AS2, LINC01235, AC055854.1, and AC124798.1 associated with survival in breast cancer patients.

	Coefficient	HR	SE	P-value
AC055854.1	0.055	1.056	0.328	0.867
AC124798.1	-0.055	0.695	0.058	0.037
FGF14-AS2	-0.221	0.802	0.112	0.148
LINC01235	0.204	1.688	0.222	0.017

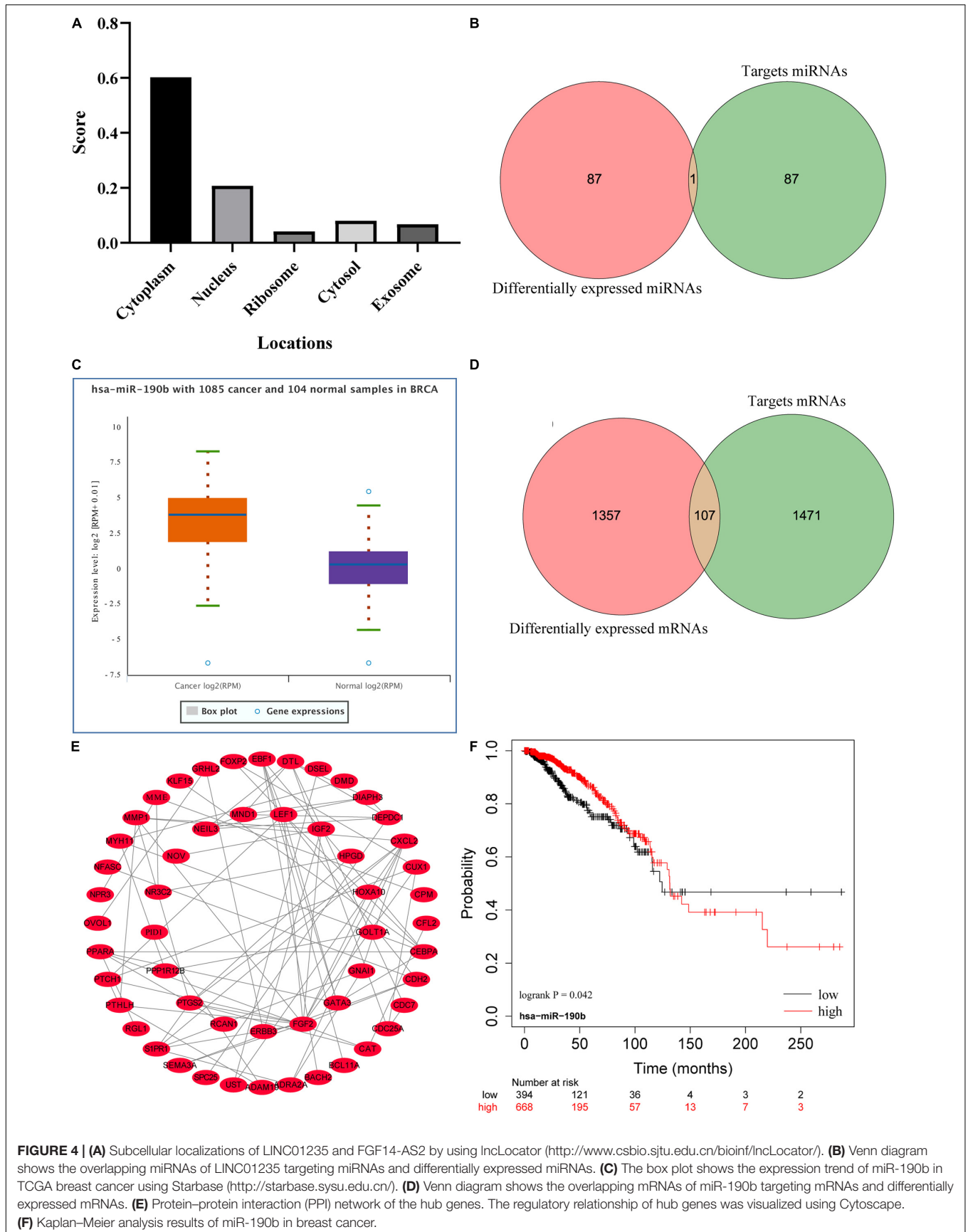
mRNAs and differentially expressed mRNAs (**Figure 4D** and **Supplementary Table 3**). Then, 53 hub mRNAs were obtained by constructing the PPI network of overlap mRNAs (**Figure 4E** and **Supplementary Table 4**). Besides, the prognosis of breast cancer according to the expression of miR-190b was determined. The results showed that miR-190b was statistically related to the OS in breast cancer patients ( $p < 0.05$ ; **Figure 4F**).

## Functional Enrichment Analysis of LINC01235-Related Hub Target Genes

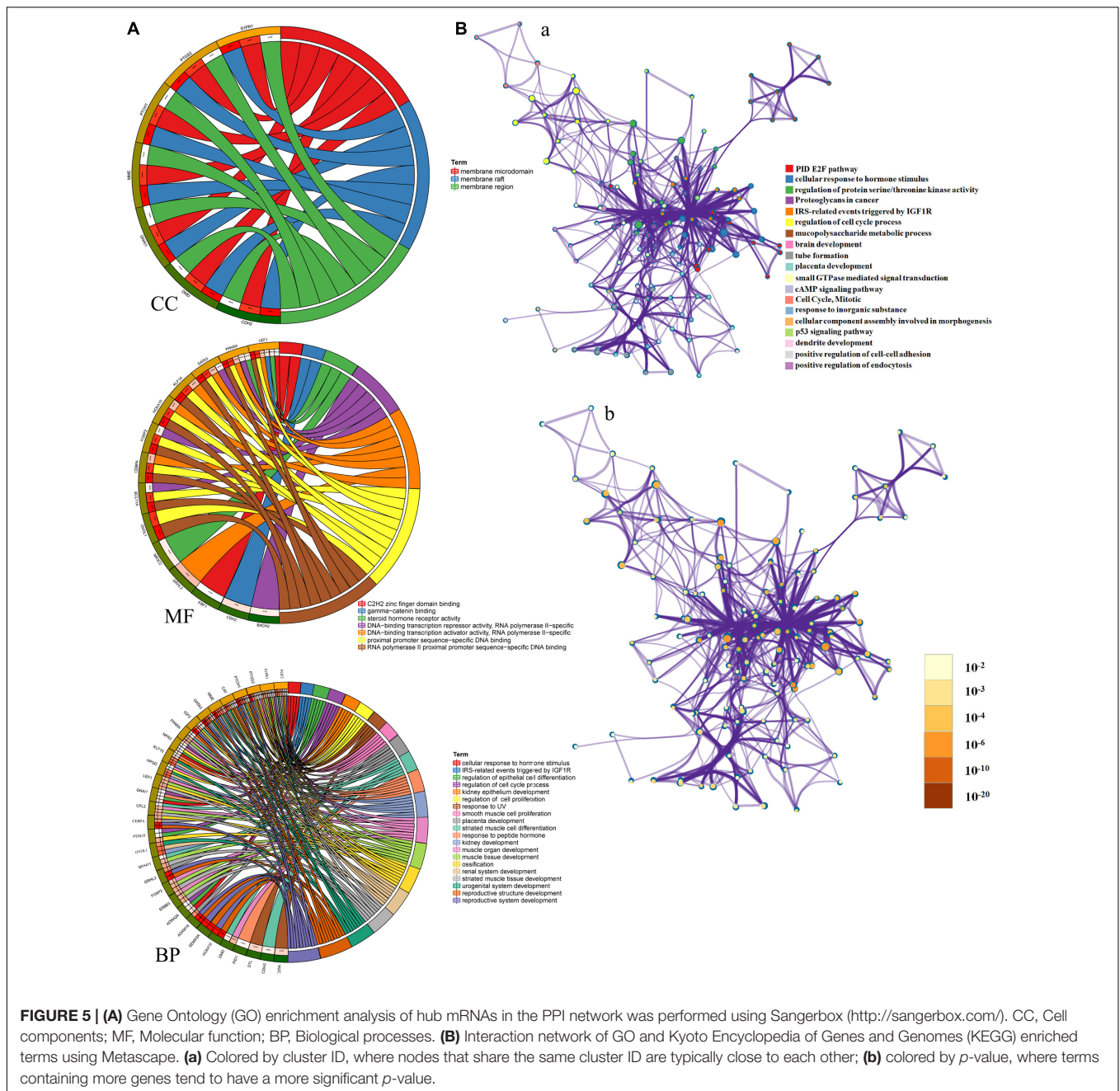
To better understand these LINC01235-related hub target mRNAs, we first performed GO functional annotation and KEGG pathway enrichment analysis by Enricher database and Metascape. The results showed that LINC01235 was also obviously associated with cellular response to hormone stimulus,







**FIGURE 4 | (A)** Subcellular localizations of LINC01235 and FGF14-AS2 by using lncLocator (<http://www.csbio.sjtu.edu.cn/bioinf/lncLocator/>). **(B)** Venn diagram shows the overlapping miRNAs of LINC01235 targeting miRNAs and differentially expressed miRNAs. **(C)** The box plot shows the expression trend of miR-190b in TCGA breast cancer using Starbase (<http://starbase.sysu.edu.cn/>). **(D)** Venn diagram shows the overlapping mRNAs of miR-190b targeting mRNAs and differentially expressed mRNAs. **(E)** Protein-protein interaction (PPI) network of the hub genes. The regulatory relationship of hub genes was visualized using Cytoscape. **(F)** Kaplan-Meier analysis results of miR-190b in breast cancer.



**FIGURE 5 | (A)** Gene Ontology (GO) enrichment analysis of hub mRNAs in the PPI network was performed using Sangerbox (<http://sangerbox.com/>). CC, Cell components; MF, Molecular function; BP, Biological processes. **(B)** Interaction network of GO and Kyoto Encyclopedia of Genes and Genomes (KEGG) enriched terms using Metascape. **(a)** Colored by cluster ID, where nodes that share the same cluster ID are typically close to each other; **(b)** colored by  $p$ -value, where terms containing more genes tend to have a more significant  $p$ -value.

regulation of protein serine/threonine kinase activity, IRS-related events triggered by IGF1R, regulation of epithelial cell differentiation, regulation of cell cycle process, regulation of epithelial cell proliferation, and regulation of epithelial cell proliferation. For KEGG pathway enrichment analysis, these LINC01235-related hub target mRNAs were significantly enriched in cAMP signaling pathway, PID E2F pathway, p53 signaling pathway, and proteoglycans in cancer (Figure 5A, CC, MF, and BP). To further capture the relationships between the terms, a subset of enriched terms has been selected and rendered as a network plot. Each node in the network represents an enriched term and is colored first by its cluster ID (Figure 5B-a)

and then by its  $p$ -value (Figure 5B-b). By the color and size of each point, we can clearly identify biological functions or pathways that are significantly enriched, as well as associations with other biological functions or pathways.

## LINC01235 Contributes to MCF-7 Cell Proliferation, Migration, and Invasion

In order to further analyze the biological function of LINC01235, we used siRNA and overexpression plasmids to alter the expression level of LINC01235 in MCF-7 cells and analyzed the effect of LINC01235 on MCF-7 cell proliferation, migration, and

**TABLE 7** | Univariate analysis of LINC01235 in breast cancer patients.

Clinical characteristic	HR	LOWER	UPER	P
Age	1.033	1.015	1.051	0.079
T	1.418	1.086	1.852	0.010
N	1.712	1.361	2.155	<0.001
M	2.675	1.609	4.449	<0.001
Stage	1.941	1.421	2.651	<0.001
ERstatus	0.556	0.346	0.896	0.016
lymph nodes	1.047	1.012	1.083	0.308
LINC01235	1.792	0.773	1.838	0.042

**TABLE 8** | Multivariate survival analysis of LINC01235 in breast cancer patients.

Clinical characteristic	HR	LOWER	UPER	P
LINC01235	1.688	0.643	1.518	0.017
Age	1.04	1.022	1.059	0.124
T	0.936	0.679	1.29	0.686
N	1.586	1.106	2.272	0.102
M	2.253	1.318	3.851	0.003
Stage	1.569	0.976	2.522	0.063
ER status	0.438	0.27	0.71	0.001
lymph nodes	0.97	0.919	1.024	0.273

invasion. qPCR results revealed successful transfection of the overexpression plasmid (pCDH-LINC01235) and si-LINC01235 (Figure 6A). The effect of LINC01235 on cell proliferation was detected by the CCK-8 assay. Compared with the negative control (NC) group, pCDH-LINC01235 significantly promoted MCF-7 proliferation, whereas MCF-7 proliferation was significantly inhibited by si-LINC01235 ( $p < 0.001$ ; Figure 6B). Meanwhile, the wound-healing assay revealed that pCDH-LINC01235 significantly enhanced MCF-7 migration, and si-LINC01235 showed a notably slower scratch closure rate than control cells (Figures 6C,D), which revealed that silencing LINC01235 inhibited MCF-7 migration ( $p < 0.001$ ). Furthermore, the transwell assay demonstrated that LINC01235 knockdown MCF-7s displayed significantly lower invasion potential than the control cells ( $p < 0.001$ ; Figures 6E,F). Collectively, these results suggest that the expression level of LINC01235 affected the proliferation, migration, and invasion of breast cancer cells.

## LINC01235 Negatively Regulates ER Breast Cancer-Related Gene ESR1

In a previous study, Gonzalez-Malerva found 227 resistant candidate mRNAs when analyzing high-throughput ectopic expression profiles of tamoxifen resistance in breast cancer (Gonzalez-Malerva et al., 2011). As shown in Figure 7D, the Venn diagram indicated that 2/53 (3.77%) of hub mRNAs (ADRA2A and DTL) were overlapped with breast cancer drug resistance genes. In the previous data analysis results, we found that the expression level of LINC01235 was related to the ER status of breast cancer patients. High expression of LINC01235 was significantly correlated with ER positivity ( $p < 0.05$ , Figure 2C). ESR1 protein is an estrogen receptor 1, and its dysregulation would lead to treatment resistance

and metastasis in patients with ER-positive breast cancer. So, is there a correlation between LINC01235 and ESR1? What kind of relationship exists? First, we analyzed the correlation of LINC01235, ESR1, ADRA2A, and DTL by online software (Figure 7C). The results showed that LINC01235 was negatively linked to ESR1 ( $R = -0.43$ ,  $p = 3.3e-31$ ), DTL ( $R = -0.14$ ,  $p = 1.8e-06$ ), and ADRA2A ( $R = -0.427$ ,  $p = 3.32e-03$ ).

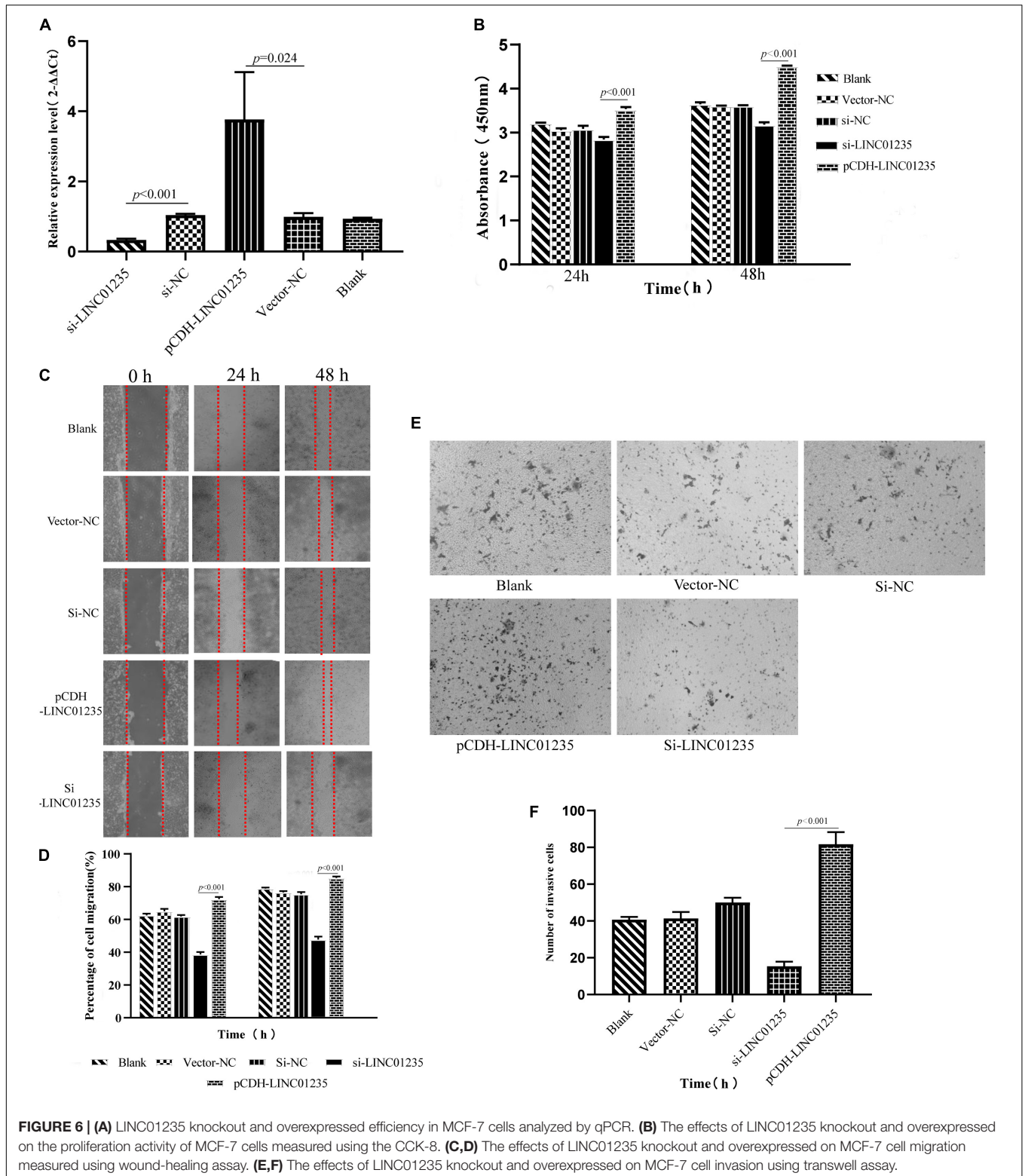
The above data indicated a strong correlation among LINC01235, ESR1, and ADRA2A. In order to further verify the correctness of the analysis results, we performed qPCR and correlation analysis on 35 breast cancer clinical samples (Figures 7A,B). Expectedly, the analysis results were consistent with the online prediction results. In the next step, we further analyzed the expression relationships among LINC01235, ESR1, and ADRA2A through cell experiments *in vitro*. The expression of LINC01235 was knocked out in MCF-7 cells, and the qPCR results showed that the expression of ESR1 and ADRA2A was significantly increased (Figure 7E). These further verified the above correlation analysis results. On the other hand, MCF-7 was transfected with si-ESR1 and si-ADRA2A, respectively, and then LINC01235 was silenced. CCK8 results showed that the cell proliferation was increased in the si-ESR1 and si-ADRA2A group and was significantly reduced in the si-LINC01235 group. However, there was no significant change in the cell proliferation rate in the co-transfection group (Figure 7F). These indicated that silencing ESR1 or ADRA2A could noticeably reverse the effect of LINC01235 knockdown on the proliferation.

Based on the above experimental analysis, there may be a reverse regulation mode among ESR1, ADRA2A, and LINC01235. As shown in Figure 8, low expression of LINC01235 regulates the biological stability of ESR1 and maintains the stability of ADRA2A, leading to a good prognosis of patients. The high expression of LINC01235 leads not only to ESR1 imbalance and drug resistance, but also to abnormal expression of ADRA2A, leading to the deterioration of breast cancer. For breast cancer caused by low expression of LINC01235, this needs to be further explored by more experiments and analysis.

## DISCUSSION

At present, some progress has been made in the diagnosis and treatment of breast cancer, but the prognosis of patients with breast cancer is still not optimistic (Ruddy and Ganz, 2019; Waks and Winer, 2019). In recent years, with the development of high-throughput technology, more and more attention has been paid to the function and regulation of lncRNA in diseases. Especially in tumors, lncRNAs not only act as oncogenes, but also act as tumor suppressors (Choudhari et al., 2020; Statello et al., 2021; Xing et al., 2021). However, there are still a lot of lncRNAs to be studied in breast cancer. Comprehensive analysis of lncRNAs in breast cancer can not only fully understand the new mechanism regulating the progression of breast cancer, but also identify new prognostic factors.

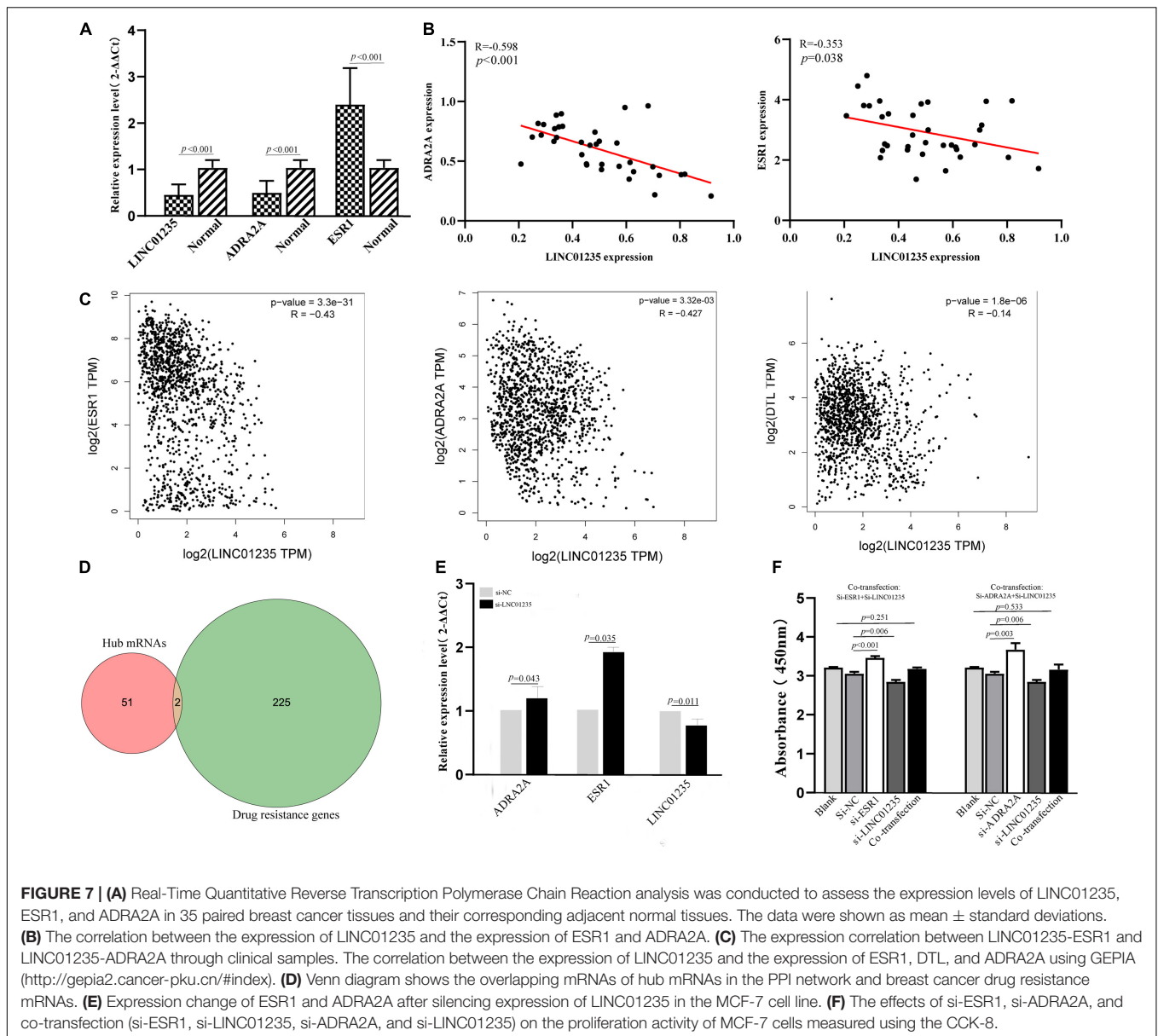
With the continuous maturity of high-throughput technologies, a large number of disease databases have been established. Miecznikowski obtained gene expression profiles



**FIGURE 6 | (A)** LINC01235 knockdown and overexpressed efficiency in MCF-7 cells analyzed by qPCR. **(B)** The effects of LINC01235 knockdown and overexpressed on the proliferation activity of MCF-7 cells measured using the CCK-8. **(C,D)** The effects of LINC01235 knockdown and overexpressed on MCF-7 cell migration measured using wound-healing assay. **(E,F)** The effects of LINC01235 knockdown and overexpressed on MCF-7 cell invasion using transwell assay.

of five breast cancers through Gene Expression Omnibus (GEO). Through the comparative analysis of the Cox risk ratio model, many genetic molecular markers were found and the survival rate of patients after receiving treatment was estimated

(Miecznikowski et al., 2010; Barrett et al., 2013). Goh used high-throughput data to analyze the interaction network of human genes and diseases and found that the protein products of disease genes showed a tendency to correlate in the interaction network

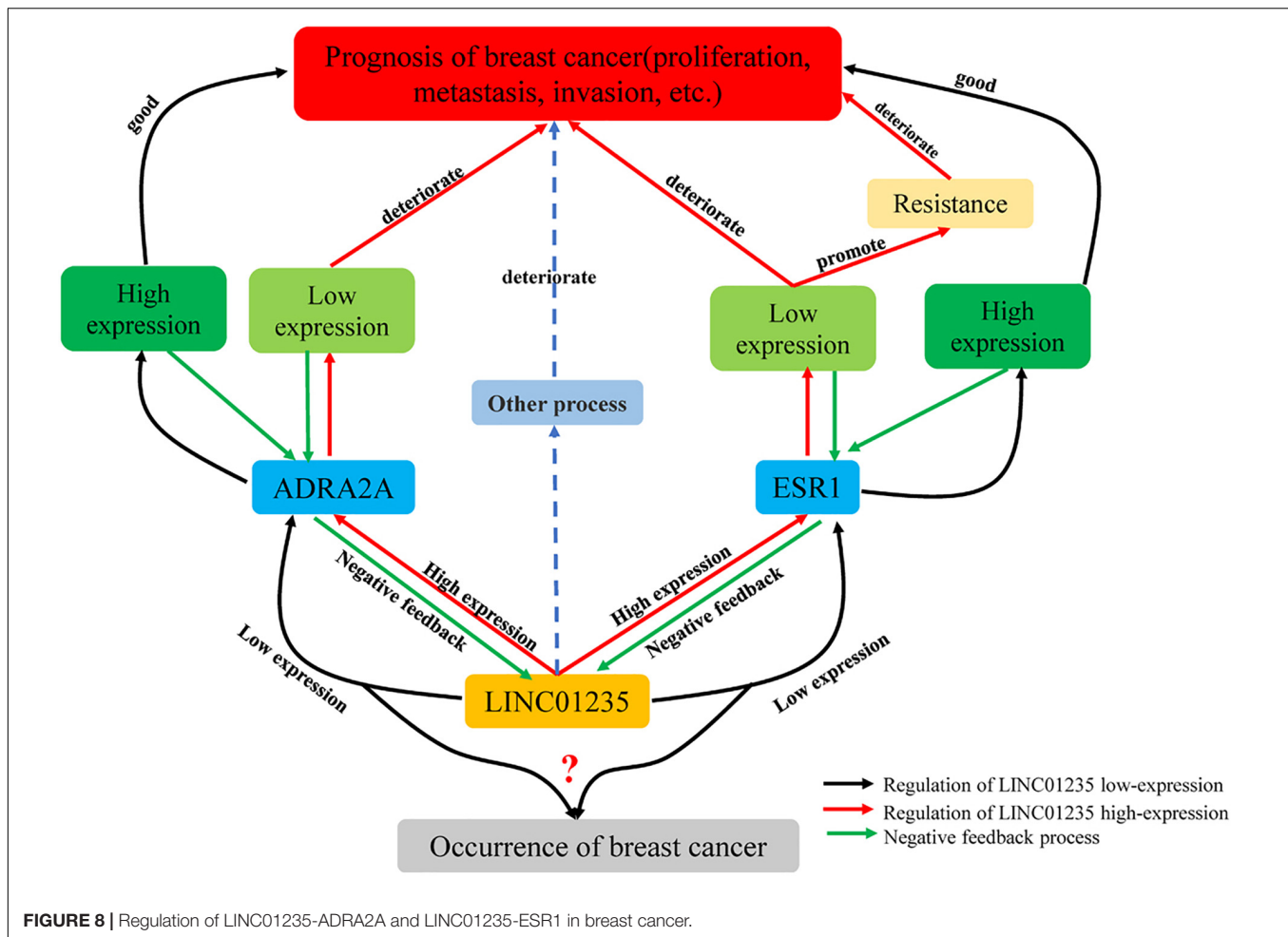


(Goh et al., 2004). This new research method makes research no longer limited to a specific experiment or a series of specific experiments, reduces heterogeneity, improves scientific research efficiency, and reveals the potential biological significance of the data and the value of follow-up research. The clinical treatment process provides suitable diagnostic and therapeutic targets.

In this study, we identified differentially expressed lncRNAs from TCGA-BRCA data and found that FGF14-AS2, LINC01235, AC055854.1, and AC124798.1 were significantly associated with the OS of breast cancer patients. Through univariate and multivariate Cox analysis, LINC01235 and AC124798.1 were determined to be the most significant prognostic factors for breast cancer. QPCR detection confirmed that LINC01235 was identical to data analysis results in breast cancer tissue. Previous studies suggest that LINC01235 is a

long-stranded non-coding RNA, located on chromosome 9 p23, which has a characteristic expression pattern because of its inconsistent expression in different tumor tissues (Tang et al., 2019; Xu et al., 2019). Studies have shown that LINC01235 was highly expressed in gastric cancer, and silencing LINC01235 could inhibit the EMT characteristics of cells *in vitro* (Zhang C. et al., 2021). In breast cancer, the researchers found that the expression of LINC01235 was downregulated, and it was speculated that the upregulation of LINC01235 might represent a more invasive phenotype, but this study was not carried out in depth (Vishnubalaji et al., 2019).

According to differentially expressed miRNAs from TCGA-BRCA and LINC01235 targeted miRNA, miR-190b was screened out. According to the report, miR-190b was increased in breast cancer and was involved in proliferation, migration, and invasion



of breast cancer cells by targeting MYLIP. miR-190b could be used as a biological indicator of ER-positive breast cancer (Cizeron-Clairac et al., 2015; Zhao et al., 2020). The above studies indicate that miR-190b regulates tumor growth and may be involved in breast cancer progression. Meanwhile, two candidate genes (ADRA2A and DTL) obtained by comparison screening have been reported in cancer. In breast cancer, for patients with luminal tumors, ADRA2A was an independent predictor of DMFS and the only factor that preserves its importance (Rivero et al., 2019). DTL depletion resulted in the disruption of the mitotic proteins cyclin B, Cdt2, CDK1, and aurora as well as the upregulation of the cell cycle arrest gene p21 (Ishii et al., 2010; Slenn et al., 2014).

Based on the prognostic roles of potential lncRNAs using TCGA-breast cancer data, LINC01235 was found to be a favorable prognostic biomarker for patients with breast cancer. In this study, we found that the uncharacteristic expression of LINC01235 in breast cancer tissues was significantly lower than that in normal tissues, and the high expression of LINC01235 was associated with poor prognosis of breast cancer patients. Functional studies have shown that overexpression of LINC01235 (pCDH-LINC01235) could promote the proliferation, migration, and invasion of MCF-7 cells *in vitro*. However, LINC01235

played a tumor suppressor role in breast cancer. Besides, our results demonstrate that expression of LINC01235 was significantly positively correlated with pathological N ( $p < 0.05$ ), tumor stage ( $p < 0.05$ ), and ER status ( $p < 0.05$ ). ER-positive clinical stage samples had elevated expression of LINC01235, so what is the regulatory relationship between the expression of LINC01235 and ER coding genes? ER signal transduction has been widely recognized as an important event in the growth and progression of breast cancer. ER is a transcription factor encoded by ESR1 (Lei et al., 2019; K uchler et al., 2020). Through knockout experiment and bioinformatics analysis, we found that there was a negative correlation between LINC01235 and ESR1. Also, we found that there was a negative correlation between LINC01235 and ADRA2A, the hub gene (ESR1) of drug resistance. Furthermore, we co-transfected si-ESR1 or si-ADRA2A with si-LINC01235 in MCF-7, and then CCK8 results showed that ESR1 or ADRA2A silencing could partially reverse the proliferation effect mediated by LINC01235 knockout. These results indicate that there is a significant feedback loop of mutual inhibition between ADRA2A, ESR1, and LINC01235 in breast cancer cells. Together, our results revealed that LINC01235-ESR1 and LINC01235-ADRA2A could serve as important co-expression pairs in the progression of

breast cancer. In the future, more experiments need to be performed to further confirm the roles of LINC01235-ESR1 and LINC01235-ADRA2A in breast cancer.

Based on the current analytical data, low expression of LINC01235 will lead to the occurrence of breast cancer, and high expression of LINC01235 could lead to poor prognosis in patients with breast cancer. All these indicate that there is no correlation between the expression pattern of LINC01235 in breast cancer and survival data. It has been reported that this situation may be caused by the great heterogeneity of TCGA-BRCA samples (Vishnubalaji et al., 2019). Aznar pointed out that the DAPLE gene was low expressed in patients with early stage of colorectal tissues and normal patients, acting as a tumor suppressor, but begins to be high expressed in the late stage of cancer, indicating poor prognosis (Aznar et al., 2015). Additionally, previous studies have found that LINC01235 may be related to immune regulation in cancer (George et al., 2017). Therefore, it is speculated that the bidirectional regulation of LINC01235 may be an immune escape mechanism of breast cancer cells. After tumorigenesis, breast cancer cells may hijack LINC01235 and induce the inactivation of its promoter, methylation sequence, or regulator, thereby evading the immune system. Galanos found that p53 was a regulator of p21. In normal tissue samples, p21 acted as a tumor suppressor. However, when p53 was missing or inactivated, p21 will sharply promote tumor growth and spread throughout the body (Galanos et al., 2016). Moreover, according to our research data, we speculate that it may be caused by the negative feedback mechanism of LINC01235-ESR1 and LINC01235-ADRA2A. The high expression of LINC01235 caused the imbalance of ESR1 and led to drug resistance in patients with breast cancer, resulting in a poor prognosis. Similarly, the high expression of LINC01235 reduced the expression of ADRA2A. As a tumor suppressor gene, the low expression of ADRA2A directly made the prognosis of breast cancer worse. The overall effect may be the same as the data we analyzed; LINC01235 overexpression leads to poor prognosis in breast cancer patients. More experiments are needed to verify the regulation mechanism.

## CONCLUSION

LINC01235 plays a critical role in facilitating the proliferation, migration, and invasion of breast cancer, and our results demonstrated that the effect is achieved by LINC01235-ESR1 and LINC01235-ADRA2A negative feedback pairs. LINC01235 may act as an effective biomarker for the diagnosis and treatment of breast cancer in the future.

## REFERENCES

- Aznar, N., Midde, K. K., Dunkel, Y., Lopez-Sanchez, I., Pavlova, Y., Marivin, A., et al. (2015). Daple is a novel non-receptor GEF required for trimeric G protein activation in Wnt signaling. *Elife* 4:e07091. doi: 10.7554/eLife.07091
- Barrett, T., Wilhite, S. E., Ledoux, P., Evangelista, C., Kim, I. F., Tomashevsky, M., et al. (2013). NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 41, D991–D995. doi: 10.1093/nar/gks1193

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding authors.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by The Ethical Committee of Shuguang Hospital. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

ZL and ZQ conducted the data analysis and wrote the manuscripts. FC, LM, and SJ carried out data analysis and manuscript revision. JC designed the study. All authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

We thank all TCGA data builders and data contributors, as well as the team that built the GEPIA and Starbase online analysis page. We also thank Shanghai Ordovician Biotechnology Co., LTD., for providing the platform for biological information analysis.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.709514/full#supplementary-material>

**Supplementary Figure 1 | (A)** Kaplan–Meier estimates of the survival of GSE21653 testing-set patients ( $n = 252$ ) and GSE42568 testing-set patients ( $n = 104$ ) using the 2-lncRNA (LINC01235 and AC124798.1) signature. **(B)** ROC of LINC01235 based on GSE21653 testing-set and GSE42568 testing-set.

**Supplementary Table 1 |** The targets miRNAs of LINC01235 were analyzed by DIANA tools.

**Supplementary Table 2 |** The targets mRNAs of miR-190b were analyzed by DIANA tools.

**Supplementary Table 3 |** The same mRNAs between targeted mRNAs and differentially expressed mRNAs.

**Supplementary Table 4 |** 53 mRNAs in the protein–protein interaction network.

- Choudhari, R., Sedano, M. J., Harrison, A. L., Subramani, R., Lin, K. Y., Ramos, E. I., et al. (2020). Long noncoding RNAs in cancer: from discovery to therapeutic targets. *Adv. Clin. Chem.* 95, 105–147. doi: 10.1016/bs.acc.2019.08.003
- Cizeron-Clairac, G., Lallemand, F., Vacher, S., Lidereau, R., Bieche, I., and Callens, C. (2015). MiR-190b, the highest up-regulated miRNA in ER $\alpha$ -positive compared to ER $\alpha$ -negative breast tumors, a new biomarker in breast cancers? *BMC Cancer* 15:499. doi: 10.1186/s12885-015-1505-5

- De Troyer, L., Zhao, P., Pastor, T., Baietti, M. F., Barra, J., Vendramin, R., et al. (2020). Stress-induced lncRNA LASTR fosters cancer cell fitness by regulating the activity of the U4/U6 recycling factor SART3. *Nucleic Acids Res.* 48, 2502–2517. doi: 10.1093/nar/gkz1237
- DeSantis, C. E., Ma, J., Gaudet, M. M., Newman, L. A., Miller, K. D., Goding Sauer, A., et al. (2019). Breast cancer statistics, 2019. *CA Cancer J. Clin.* 69, 438–451. doi: 10.3322/caac.21583
- Fukushima, K., and Kida, H. (2019). A host lncRNA regulates the innate immune response to an RNA virus. *Cell. Mol. Immunol.* 16, 841–842. doi: 10.1038/s41423-019-0280-7
- Galanos, P., Vougas, K., Walter, D., Polyzos, A., Maya-Mendoza, A., Haagensen, E. J., et al. (2016). Chronic p53-independent p21 expression causes genomic instability by deregulating replication licensing. *Nat. Cell Biol.* 18, 777–789. doi: 10.1038/ncb3378
- George, J., Saito, M., Tsuta, K., Iwakawa, R., Shiraiishi, K., Scheel, A. H., et al. (2017). Genomic amplification of CD274 (PD-L1) in small-cell lung cancer. *Clin. Cancer Res.* 23, 1220–1226. doi: 10.1158/1078-0432.CCR-16-1069
- Giro-Perafita, A., Luo, L., Khodadadi-Jamayran, A., Thompson, M., Akgol Oksuz, B., Tsirigos, A., et al. (2020). LncRNA RP11-19E11 is an E2F1 target required for proliferation and survival of basal breast cancer. *NPJ Breast Cancer* 6:1. doi: 10.1038/s41523-019-0144-4
- Goh, C. S., Lan, N., Douglas, S. M., Wu, B., Echols, N., Smith, A., et al. (2004). Mining the structural genomics pipeline: identification of protein properties that affect high-throughput experimental analysis. *J. Mol. Biol.* 336, 115–130. doi: 10.1016/j.jmb.2003.11.053
- Gonzalez-Malerva, L., Park, J., Zou, L., Hu, Y., Moradpour, Z., Pearlberg, J., et al. (2011). High-throughput ectopic expression screen for tamoxifen resistance identifies an atypical kinase that blocks autophagy. *Proc. Natl. Acad. Sci. U.S.A.* 108, 2058–2063. doi: 10.1073/pnas.1018157108
- Han, C., Fu, Y., Zeng, N., Yin, J., and Li, Q. (2020). LncRNA FAM83H-AS1 promotes triple-negative breast cancer progression by regulating the miR-136-5p/metadherin axis. *Aging* 12, 3594–3616. doi: 10.18632/aging.102832
- Ishii, T., Shiohira, Y., Takami, T., Murakami, Y., Ohnishi, N., and Nishitani, H. (2010). Proliferating cell nuclear antigen-dependent rapid recruitment of Cdt1 and CRL4Cdt2 at DNA-damaged sites after UV irradiation in HeLa cells. *J. Biol. Chem.* 285, 41993–42000. doi: 10.1074/jbc.M110.161661
- Küchler, E. C., Meger, M. N., Ayumi Omori, M., Gerber, J. T., Carneiro Martins Neto, E., Silva Machado, N. C. D., et al. (2020). Association between oestrogen receptors and female temporomandibular disorders. *Acta Odontol. Scand.* 78, 181–188. doi: 10.1080/00016357.2019.1675904
- Labonté, B., Abdallah, K., Maussion, G., Yerko, V., Yang, J., Bittar, T., et al. (2020). Regulation of impulsive and aggressive behaviours by a novel lncRNA. *Mol. Psychiatry* 5, 1–14. doi: 10.1038/s41380-019-0637-4
- Lei, J. T., Gou, X., Seker, S., and Ellis, M. J. (2019). ESR1 alterations and metastasis in estrogen receptor positive breast cancer. *J. Cancer Metastasis Treat.* 5:38. doi: 10.20517/2394-4722.2019.12
- Li, X., Chen, J., Yu, Q., Huang, H., Liu, Z., Wang, C., et al. (2021). A signature of autophagy-related long non-coding RNA to predict the prognosis of breast cancer. *Front. Genet.* 12:569318. doi: 10.3389/fgene.2021.569318
- Liang, Y., Song, X., Li, Y., Chen, B., Zhao, W., Wang, L., et al. (2020). LncRNA BCRT1 promotes breast cancer progression by targeting miR-1303/PTBP3 axis. *Mol. Cancer* 19:85. doi: 10.1186/s12943-020-01206-5
- Majidinia, M., and Yousefi, B. (2020). Long non-coding RNAs in drug resistance of breast cancer. *Onco Targets Ther.* 13, 7075–7087. doi: 10.2147/OTT.S255226
- Miecznikowski, J. C., Wang, D., Liu, S., Sucheston, L., and Gold, D. (2010). Comparative survival analysis of breast cancer microarray studies identifies important prognostic genetic pathways. *BMC Cancer* 10:573. doi: 10.1186/1471-2407-10-573
- Peng, W. X., Koirala, P., and Mo, Y. (2017). LncRNA-mediated regulation of cell signaling in cancer. *Oncogene* 36, 5661–5667. doi: 10.1038/onc.2017.184
- Rivero, E. M., Martinez, L. M., Bruque, C. D., Gargiulo, L., Bruzzone, A., and Lüthy, I. A. (2019). Prognostic significance of  $\alpha$ - and  $\beta$ -adrenoceptor gene expression in breast cancer patients. *Br. J. Clin. Pharmacol.* 85, 2143–2154. doi: 10.1111/bcp.14030
- Rossi, L., Mazzara, C., and Pagani, O. (2019). Diagnosis and treatment of breast cancer in young women. *Curr. Treat. Opt. Oncol.* 20:86. doi: 10.1007/s11864-019-0685-7
- Ruddy, K. J., and Ganz, P. A. (2019). Treatment of nonmetastatic breast cancer. *JAMA* 321, 1716–1717. doi: 10.1001/jama.2019.3927
- Slenn, T. J., Morris, B., Havens, C. G., Freeman, R. M. Jr., Takahashi, T. S., and Walter, J. C. (2014). Thymine DNA glycosylase is a CRL4Cdt2 substrate. *J. Biol. Chem.* 289, 23043–23055. doi: 10.1074/jbc.M114.574194
- Song, X., Zhang, X., Wang, X., Chen, L., Jiang, L., Zheng, A., et al. (2020). LncRNA SPRY4-IT1 regulates breast cancer cell stemness through competitively binding miR-6882-3p with TCF7L2. *J. Cell. Mol. Med.* 24, 772–784. doi: 10.1111/jcmm.14786
- Statello, L., Guo, C. J., Chen, L. L., and Huarte, M. (2021). Gene regulation by long non-coding RNAs and its biological functions. *Nat. Rev. Mol. Cell Biol.* 22, 96–118. doi: 10.1038/s41580-020-00315-9
- Tan, Y. E., Xing, Y., Ran, B. L., Zhang, C., Pan, S. W., An, W., et al. (2020). LINC01235-TWIST2 feedback loop facilitates epithelial-mesenchymal transition in gastric cancer by inhibiting THBS2. *Aging* 12, 25060–25075. doi: 10.18632/aging.103979
- Tang, T., Guo, C., Xia, T., Zhang, R., Zen, K., Pan, Y., et al. (2019). LncCCAT1 promotes breast cancer stem cell function through activating WNT/ $\beta$ -catenin signaling. *Theranostics* 9, 7384–7402. doi: 10.7150/thno.37892
- Tang, Y., Wang, Y., Kiani, M. F., and Wang, B. (2016). Classification, treatment strategy, and associated drug resistance in breast cancer. *Clin. Breast Cancer* 16, 335–343. doi: 10.1016/j.clbc.2016.05.012
- Vishnubalaji, R., Shaath, H., Elkord, E., and Alajez, N. M. (2019). Long non-coding RNA (lncRNA) transcriptional landscape in breast cancer identifies LINC01614 as non-favorable prognostic biomarker regulated by TGF $\beta$  and focal adhesion kinase (FAK) signaling. *Cell Death Discov.* 5:109. doi: 10.1038/s41420-019-0190-6
- Waks, A. G., and Winer, E. P. (2019). Breast cancer treatment: a review. *JAMA* 321, 288–300. doi: 10.1001/jama.2018.19323
- Xing, C., Sun, S. G., Yue, Z. Q., and Bai, F. (2021). Role of lncRNA LUCAT1 in cancer. *Biomed. Pharmacother.* 134:111158. doi: 10.1016/j.biopha.2020.111158
- Xu, H., Jiang, Y., Xu, X., Su, X., Liu, Y., Ma, Y., et al. (2019). Inducible degradation of lncRNA Sros1 promotes IFN- $\gamma$ -mediated activation of innate immune responses by stabilizing Stat1 mRNA. *Nat. Immunol.* 20, 1621–1630. doi: 10.1038/s41590-019-0542-7
- Zhang, C., Liang, Y., Zhang, C. D., Pei, J. P., Wu, K. Z., Li, Y. Z., et al. (2021). The novel role and function of LINC01235 in metastasis of gastric cancer cells by inducing epithelial-mesenchymal transition. *Genomics* 113, 1504–1513. doi: 10.1016/j.ygeno.2021.03.027
- Zhang, K., Ping, L., Du, T., Liang, G., Huang, Y., Li, Z., et al. (2021). A ferroptosis-related lncRNAs signature predicts prognosis and immune microenvironment for breast cancer. *Front. Mol. Biosci.* 8:678877. doi: 10.3389/fmolb.2021.678877
- Zhao, L., Zhou, Y., Zhao, Y., Li, Q., Zhou, J., and Mao, Y. (2020). Long non-coding RNA TUSC8 inhibits breast cancer growth and metastasis via miR-190b-5p/MYLP axis. *Aging* 12, 2974–2991. doi: 10.18632/aging.102791

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Li, Qian, Chen, Jiang, Meng and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.