Investigating the potential clinical significance of long non-coding RNA 00092 in patients with breast cancer

Jingquan Li[1], Fanghao Lu[2], Xin Shao[3], Bosen You[3]

[1]Department of Medical Oncology, Harbin Medical University Cancer Hospital, Harbin, China; [2]Department of Pathophysiology, Harbin Medical University, Harbin, China; [3]Department of Surgery, The Second Affiliated Hospital of Harbin Medical University, Harbin, China

Contributions: (I) Conception and design: J Li; (II) Administrative support: F Lu; (III) Provision of study materials or patients: J Li; (IV) Collection and assembly of data: X Shao; (V) Data analysis and interpretation: B You; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Jingquan Li. Department of Medical Oncology, Harbin Medical University Cancer Hospital, No. 150 Haping Road, Nangang District, Harbin 150081, China. Email: lijingquan@hrbmu.edu.cn.

Background: Aberrant promoter methylation and its resultant aberrant gene expression are important epigenetic mechanisms that promote the development of breast cancer (BC). However, the prognostic value of this type of methylation-driven gene in BC is unknown.

Methods: To identify DNA methylation-driven long non-coding RNAs (lncRNAs), a comprehensive analysis of RNA-sequencing and DNA methylation data of 1,200 clinical samples was performed. Differentially expressed lncRNAs (DELs) and survival-related lncRNAs in BC were identified using the R package. The function of the lncRNA was evaluated using Kaplan-Meier and receiver operating characteristic (ROC) curve analyses. The expression of the key lncRNA in tissues and cells was detected by quantitative real-time polymerase chain reaction (qRT-PCR). Biological functions of the key lncRNA were analyzed using Gene Ontology and Kyoto Encyclopedia of Genes and Genomes analyses. The Connectivity Map (CMap) was used to search for small-molecule targeted drugs for the key lncRNA. The functions of the key lncRNA in BC progression were investigated using cell proliferation and cell cycle assays.

Results: A total of 14 methylation-driven lncRNAs, 526 DELs, and 93 survival-associated lncRNAs were identified. The above data were intersected, and a unique lncRNA, *LINC00092*, was obtained. *LINC00092* was hypermethylated and hypoexpressed in both BC tissues and cell lines. *LINC00092* was found to be a diagnostic marker for BC, with its low expression being associated with poor prognosis (P=0.013). *LINC00092* overexpression inhibited the proliferation and cell cycle of BC cells *in vitro*. Nimesulide and sulpiride were screened out as potential targeted therapeutic drugs for *LINC00092* in BC, and sulpiride was observed to partially reverse the proliferative effect of (small interfer) si-*LINC00092* on BC cells.

Conclusions: *LINC00092* is a methylation-driven lncRNA in BC and could be a potential therapeutic target for this disease.

Keywords: Breast cancer (BC); methylation-driven lncRNA; long non-coding RNA 00092 (LINC00092); prognosis; cell cycle

Submitted Mar 21, 2022. Accepted for publication May 20, 2022.

doi: 10.21037/atm-22-1956

View this article at: https://dx.doi.org/10.21037/atm-22-1956

Introduction

Breast cancer (BC) affects more women around the world than any other cancer, and its incidence has been steadily increasing (1). According to GLOBOCAN 2018, there were approximately 2.1 million cases of BC worldwide in 2018 and about 630,000 deaths caused by the disease (2). Recent diagnostic and therapeutic advances have greatly improved the survival rates among patients with early-stage BC (3).

However, the mortality rate of BC is still not satisfactory, and the disease remains the second largest contributor to cancer-related death in women, with metastasis being the major factor (4).

BC features a sequence of key driver genomic events (5). Gene expression levels and DNA methylation landscapes in BC have both been extensively studied (6,7). Consensus molecular subtypes have been identified for clinical stratification and precision medicine (8). However, subtypes were generated in a diagnostic way, the prognostic differences in patients couldn't be reflected (9). In general, the prognosis of BC patients is often highly dependent on the individual. The heterogeneity of BC makes it challenging to predict the prognosis and determine treatment recommendations (10). Therefore, it is critical to develop effective biomarkers to improve the clinical outcomes of patients with BC.

Long non-coding RNAs (lncRNAs) are noncoding transcripts with a length of more than 200 nucleotides (nt) but no ability to code for proteins. Dysregulated expression of lncRNAs has been discovered to be involved in cancer advancement and progression. Some lncRNAs, such as *HOTAIR* (11), *SAMMSON* (12), *MALAT1* (13), and *H19* (14), have been found to serve as biomarkers and prognostic factors, and others have been discovered to have a bearing on regulation of the cell cycle, tumor growth, and medication response. Since lncRNAs play a vital function in cancer carcinogenesis, bioinformatics analysis is necessary to find more lncRNAs that could potentially be used as biomarkers and prognostic predictors.

Previous researches have shown that DNA methylation has a key role in the genesis and progression of BC (15,16). DNA methylation, a frequent epigenetic modification in eukaryotic genomes, is constantly involved in gene expression and histone modification control (17,18). Aberrant DNA methylation has been shown to be a key mechanism of oncogenic activation, and both hypo- and hypermethylation events have been observed in cancer (19). The revelation of methylation map could be the crux to understanding the epigenetic drivers of cancer (20). The DNA methylation was regulated by DNA methylation regulators such as methyltransferase and demethylase (21). Uncovering the molecular mechanisms of DNA methylation remains a great challenge.

In this study, LINC00092 was identified to be the key methylation-driven lncRNA in BC. LINC00092 was hypermethylated and lowly expressed in BC tissues and cells. LINC00092 was found to be a diagnostic biomarker for BC, with its low expression being associated with poor prognosis. *In vitro* experiments showed that overexpression of LINC00092 inhibited the proliferation and cell cycle progression of BC cells. The targeted small-molecule drug sulpiride demonstrated the ability to partially rescue the pro-proliferative effect caused by LINC00092 deletion. We present the following article in accordance with the MDAR reporting checklist (available at https://atm.amegroups.com/article/view/10.21037/atm-22-1956/rc).

Methods

Data extraction

The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov) was used to obtain 1,223 RNA-sequencing profiles (1,110 from BC samples and 113 from normal samples), 892 DNA methylation profiles (796 from BC samples and 96 from normal samples), and clinical information of patients with BC. BC gene expression data from TCGA were obtained with the Illumina HiSeq 2000 RNA sequencing platform, and DNA methylation data were acquired with Illumina Infinium Human Methylation 450 platform.

Patients and samples

Twenty paired tumor tissues and matching adjacent non-tumor tissues were collected from patients undergoing surgery for BC at the Harbin Medical University Cancer Hospital. All of the patients provided written informed consent prior to enrollment. All of the samples were immediately frozen in liquid nitrogen for later experiments. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Ethics Committee of the Harbin Medical University Cancer Hospital (No. 2019-185).

Identification of key methylation-driven genes

In this study, DNA methylation data and paired gene expression data of BC samples were analyzed using the MethylMix package in R software. DNA methylation driver lncRNAs that significantly affect the expression of the corresponding gene were identified according to the screening criteria ($|logFC| > 1$, $P<0.01$; $Cor < -0.3$, $P<0.05$). Differentially expressed lncRNAs (DELs) between patients with BC and paracancerous controls were then

Table 1 Primers used for qRT-PCR in the study

Primers	Primer sequences
LINC00092	Forward primer: 5′-CCCCGACACTGAGGACTTTT-3′
	Reverse primer: 5′-AGGTTCCTGGTTTGGGTTGG-3′
Unmethylation	Forward primer: 5′-GTTATTTAGGTTGGAGTGTAGTGGTG-3′
	Reverse primer: 5′-AACAAATCACAAAATCAAAAAATCA-3′
Methylation	Forward primer: 5′-TTATTTAGGTTGGAGTGTAGTGGC-3′
	Reverse primer: 5′-GAATCACAAAATCAAAAAATCGAA-3′

qRT-PCR, quantitative real-time polymerase chain reaction.

identified from the TCGA data using the edge R package. The $|\log_2 FC| > 0$ and adjusted false discovery rate (FDR) to a $P<1e^{-5}$ were used as cutoff values for DELs screening. Survival-associated lncRNAs were identified using Cox univariate survival analysis (P<1e-5). Finally, we intersected the DNA methylation-driven lncRNAs, DELs, and survival-associated lncRNAs by drawing a Venn diagram.

Receiver operating characteristic (ROC) curve and logistic regression analysis

To determine the specificity and sensitivity of *LINC00092* for diagnosing BC, we next conducted ROC curve analysis and calculated the area under the ROC curve (AUC) using the MedCalc software. To further evaluate the efficacy of *LINC00092* in diagnosing BC, we used the TCGA database to undertake five-fold cross-validation using a logistic regression model. The performance of the logistic regression model was assessed based on its precision, recall, accuracy, and F1-score. High precision reflected the high accuracy of the LINC00092 for predicting BC; recall referred to the number of true positives divided by the number of actual positives; accuracy was defined as the proportion of correctly forecasted results across all cases. The F1-score was selected as a comprehensive evaluation index because the precision and recall affected each other and their values could not be optimally large simultaneously. The F1-score was calculated by taking the harmonic mean of the precision and recall.

Cell culture and cell transfection

The American Type Culture Collection (ATCC) supplied the human BC cell lines (HCC38, DU4475, SKBR3, MCF7, and T47D) and human breast non-tumorigenic epithelial cell line (MCF-10A) used in this study. All cell lines were cultured according to ATCC's protocols and supplemented with 10% fetal bovine serum with 100 μg/mL penicillin and 100 mg/mL streptomycin. All cell lines were grown at 37 ℃, in a humid 5% CO_2/air environment. Small interfering RNA (siRNA) directly against human LINC00092 (si-LINC00092) and the vehicle control (si-NC) were purchased from GenePharma (Shanghai, China). Full-length complementary DNA (cDNA) of human LINC00092 was also synthesized and cloned into the expression vector pCDNA3.1 by GenePharma (Shanghai, China). Transfection of cells with pcDNA-LINC00092, si-LINC00092, and the negative controls was performed using Lipofectamine 2000 (Invitrogen) on the basis of the manufacturer's instruction. To determine the transfection efficiency, the cell media were replaced with fresh media after 6 hours of transfection and cultured for a further 48 hours. Cells were used for experiments after confirmation of successful transfection.

DNA isolation, bisulfite conversion, and methylation-specific PCR assays

To determine if *LINC00092* had a CpG island in its promoter, we searched by RefSeq transcripts from the UCSC Genome Browser database (http://genome.ucsc.edu/) and the default parameters in the CpG Island Searcher Software. The input sequence included 1000 bp upstream and 500 bp downstream of the transcription start site. Methylation-specific PCR (MSP) was used to evaluate the main gene harboring a promoter CpG-island in BC tissues and cell lines. Methyl Primer Express 1.0 (Applied Biosystems, USA) was used to construct the primers, and their sequences are listed in *Table 1*. Genomic DNA was extracted from tissue and cell samples on the basis of the manufacturer's instructions using a Gentra Puregene Tissue Kit (Qiagen, Germany). Universal methylated human DNA standard (ZYMO Research, Irvine, CA, USA) was utilized as the positive control. A Qubit 3.0 Fluorometer (Thermo-Fisher Scientific, Waltham, MA, USA) was used to identify the absolute DNA quantity of the patient tissue samples and cell samples, as well as the quantity of

universal methylated and non-methylated human DNA. Then, according to the supplier's recommended protocol, an EZ DNA Methylation-Gold Kit (ZYMO Research) was used for bisulfite modification of 0.5 μg of DNA. The bisulfite-modified DNA was eluted and diluted to a final concentration of 10 ng/L. The methylation level of the LINC00092 was quantified according to the MethyLight Taqman probe protocol.

RNA extraction and qRT-PCR

Total RNA was extracted from patient tissue samples and cell samples using TRIzol reagent (Life Technologies), and the SuperScript first-strand synthesis system (Tiangen Biotech, Beijing, China) was used to synthesize cDNA. SYBR Green qRT-PCR was used to evaluate the relative levels of LINC00092, which were standardized to GAPDH mRNA expression. For relative quantification and statistical analysis, the comparative $2^{-\Delta\Delta Ct}$ approach was applied. The primer sequences used were as follows: for LINC00092, 5'-CCTCTCAGCCTCCAGCGTTG-3' (forward) and 5'-TGCTCTTGCTCACTCACACTCC-3' (reverse); for GAPDH, 5'-ATTCAACGGCACAGTCAAGG-3' and 5'-GCAGAAGGGGCGGAGATGA-3'.

Functional enrichment analysis and Cmap analysis

Pearson's correlation coefficient (r) was used to screen and identify LINC00092-associated protein-coding genes (PCGs) in this study. Those PCGs that met the screening criteria of $P<0.05$ and $|r|>0.25$ were determined to be LINC00092 related. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses of lncRNA co-expression genes were performed using the Bioconductor clusterProfiler package (https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html). The Connectivity Map (CMap) was used to search for small-molecule targeted agents for *LINC00092* in BC. Small-molecule agents with an average connective score of <–0.2 and $P<0.05$ were identified as potential therapeutic drugs for BC.

Cell proliferation assays

Cell proliferation was evaluated with the Cell Counting Kit (CCK)-8 assays (Dojindo Laboratories, China) according to the manufacturer's instructions. Transfected HCC38 and T47D cells were seeded at a density of 1×10^3 cells/well in 96 wells plates. CCK-8 solution was added to each well at a final concentration of 10% and incubated for 1 hour at 37 ℃. The optical density (OD) values were measured at 450 nm using a microplate reader (BioTek, VT, USA) every 24 h for 4 days. To obtain the half maximal inhibitory concentration (IC50) values of the anti-breast cancer drugs nimesulide and sulpiride screened from the cMAP analysis, BC cells were treated with nimesulide (0, 20, 50, 80, 100, 120, and 160 μM) or sulpiride (0, 20, 50, 80, 100, 120, and 160 μM). After 72 hours of treatment, cell proliferation was detected by CCK-8.

Cell cycle analysis

Cells were collected, fixed in 70% ethanol at 4 ℃ overnight, and then stained with RNase A/propidium iodide (Sigma-Aldrich, USA). Flow cytometry (FACS, fluorescence-activated cell sorting) with ModFit software was used to determine the cell cycle distribution in the pcDNA-LINC00092, and negative control groups. FlowJo 10.6.2 (Treestar Inc., Ashland, OR, USA) software was used to analyze the cell percentage.

Statistical analysis

Statistical analysis was performed using R version 4.1.1 and package "pROC". Statistical analyses were performed by using one-way analysis of variance (ANOVA) and two-tailed Student's *t*-tests with GraphPad Prism 8. Pearson's correlation analysis was adopted to determine the linear relationship between two groups. Each experiment was repeated three times or more, and all data were expressed as mean ± standard deviation (SD). $P<0.05$ was considered to show a statistically significant difference.

Results

Identification of methylation-driven lncRNA in BC

With the MethylMix R package, we identified 14 methylation-driven lncRNAs (*Figure 1A*), which are listed in *Table 2*. As shown in the scatter plot (*Figure 1B*), analysis of gene expression profiles from the TCGA identified 526 DELs, including 354 which were upregulated and 172 which were downregulated in BC samples compared to normal control tissues. Univariate Cox survival analysis identified 93 lncRNAs with $P<0.05$ and hazard ratio (HR) <1 (*Figure 1C*). Grounded on the cut-off criteria,

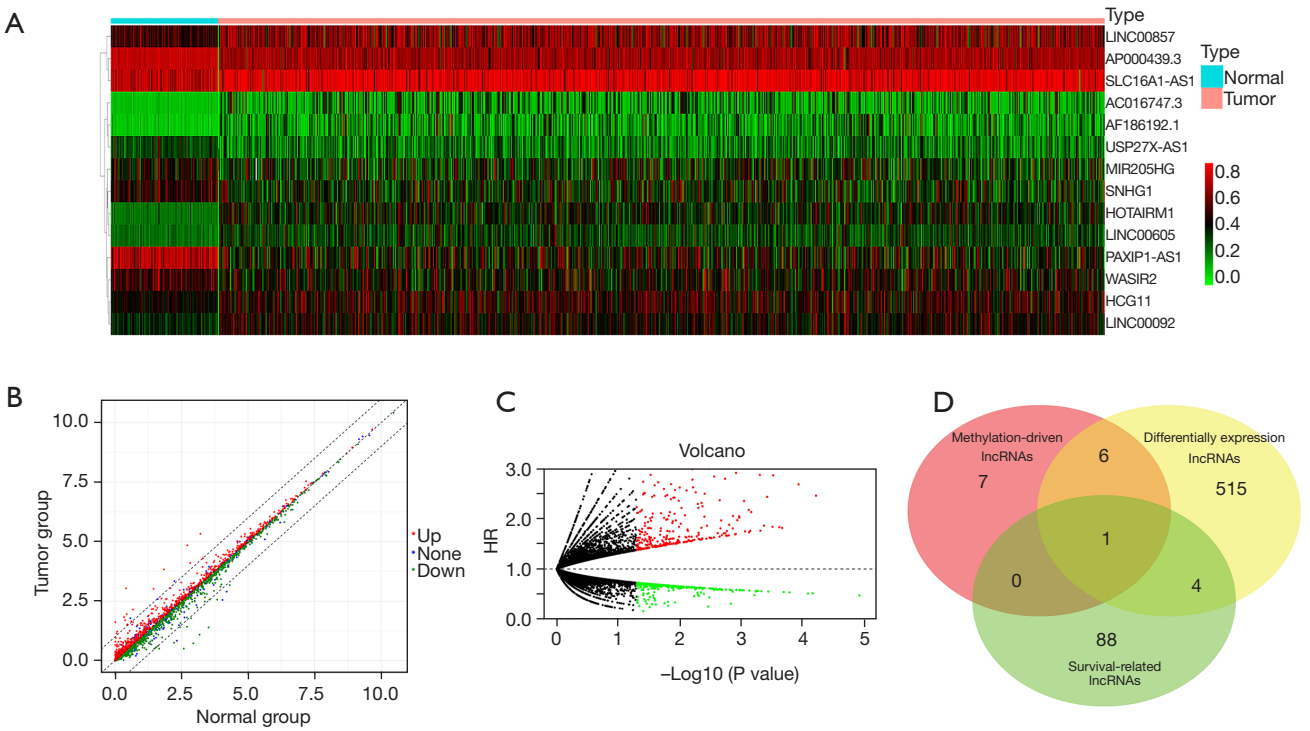


Figure 1 Identification of key methylation driver lncRNAs in BC after data integration. (A) Heat map of methylation-driven lncRNAs between BC and normal tissue; red represents hypermethylated lncRNAs and green represents hypomethylated lncRNAs. (B) DELs were identified from TCGA gene expression profile data. (C) Survival-related lncRNAs identified from BC survival data. (D) One overlapping upregulated differential lncRNA (LINC00092) was obtained through integration of the three datasets (methylation-driven lncRNAs, differentially expression lncRNAs, and survival-related lncRNAs). HR, hazard ratio; BC, breast cancer; TCGA, The Cancer Genome Atlas; DELs, differentially expressed lncRNAs.

Table 2 Fourteen methylation-driven lncRNAs in BC

Gene	Normal mean	Tumor mean	logFC	P value	Adjust P value	Cor	Cor P value	NaN_result
AC016747.3	0.129	0.259	1.003	3.90E-14	1.37E-11	–0.615	8.23E-83	lncRNA
AF186192.1	0.117	0.243	1.06	4.62E-24	1.62E-21	–0.444	3.18E-39	lncRNA
AP000439.3	0.83	0.77	–0.107	4.02E-27	1.41E-24	–0.475	2.29E-45	lncRNA
HCG11	0.488	0.555	0.186	6.47E-10	2.27E-07	–0.613	6.98E-82	lncRNA
HOTAIRM1	0.275	0.452	0.719	1.04E-48	3.66E-46	–0.317	1.08E-19	lncRNA
LINC00092	0.417	0.519	0.314	6.40E-27	2.25E-24	–0.327	6.48E-21	lncRNA
LINC00605	0.284	0.388	0.448	3.45E-28	1.21E-25	–0.35	4.89E-24	lncRNA
LINC00857	0.569	0.7	0.299	3.58E-25	1.26E-22	–0.554	3.11E-64	lncRNA
MIR205HG	0.492	0.422	–0.22	3.21E-06	0.0011	–0.375	1.64E-27	lncRNA
PAXIP1-AS1	0.796	0.49	–0.7	3.49E-48	1.23E-45	–0.335	5.58E-22	lncRNA
SLC16A1-AS1	0.807	0.893	0.147	1.42E-31	4.99E-29	–0.318	6.50E-20	lncRNA
SNHG1	0.539	0.432	–0.321	7.45E-16	2.61E-13	–0.433	3.79E-37	lncRNA
USP27X-AS1	0.375	0.274	–0.452	1.22E-20	4.27E-18	–0.356	8.26E-25	lncRNA
WASIR2	0.577	0.498	–0.213	6.02E-14	2.11E-11	–0.471	1.67E-44	lncRNA

BC, breast cancer.

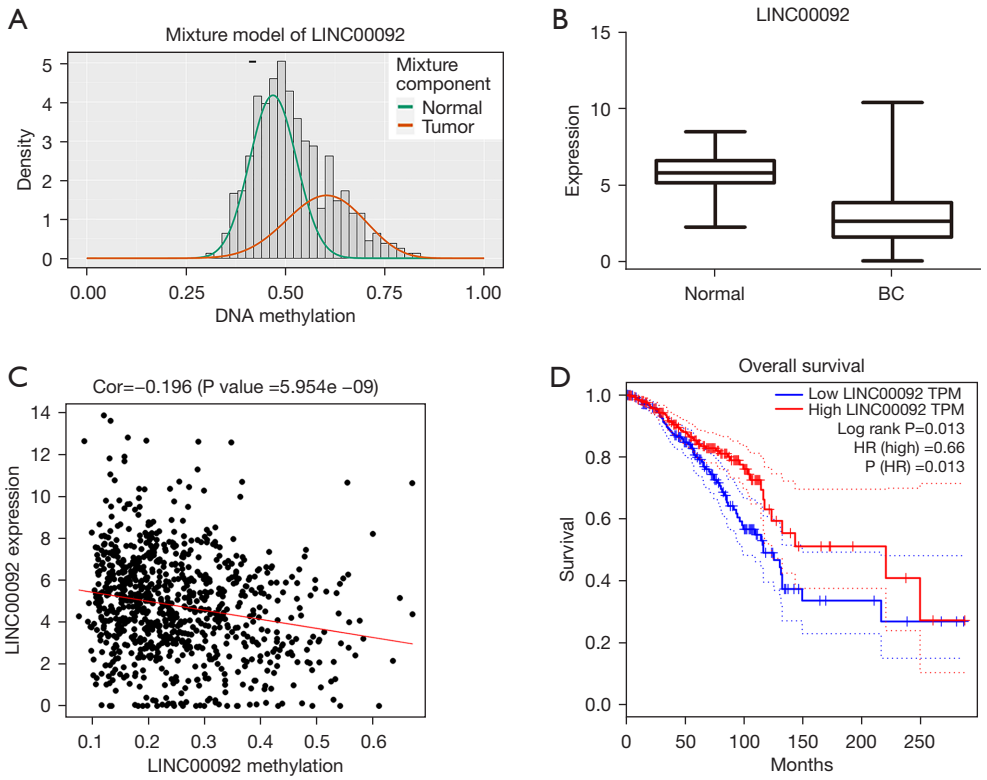


Figure 2 LINC00092 is hypermethylated and lowly expressed in BC. (A) LINC00092 is hypo-methylated in normal breast tissues and hyper-methylated in BC tissues. (B) DNA methylation levels of LINC00092 in the TCGA dataset in BC and normal controls. (C) LINC00092 expression is negatively correlated with DNA methylation in BC. (D) Kaplan-Meier analysis of the overall survival of patients with BC was analyzed by in the TCGA dataset. BC, breast cancer; TPM, transcripts per million; HR, hazard ratio; TCGA, The Cancer Genome Atlas.

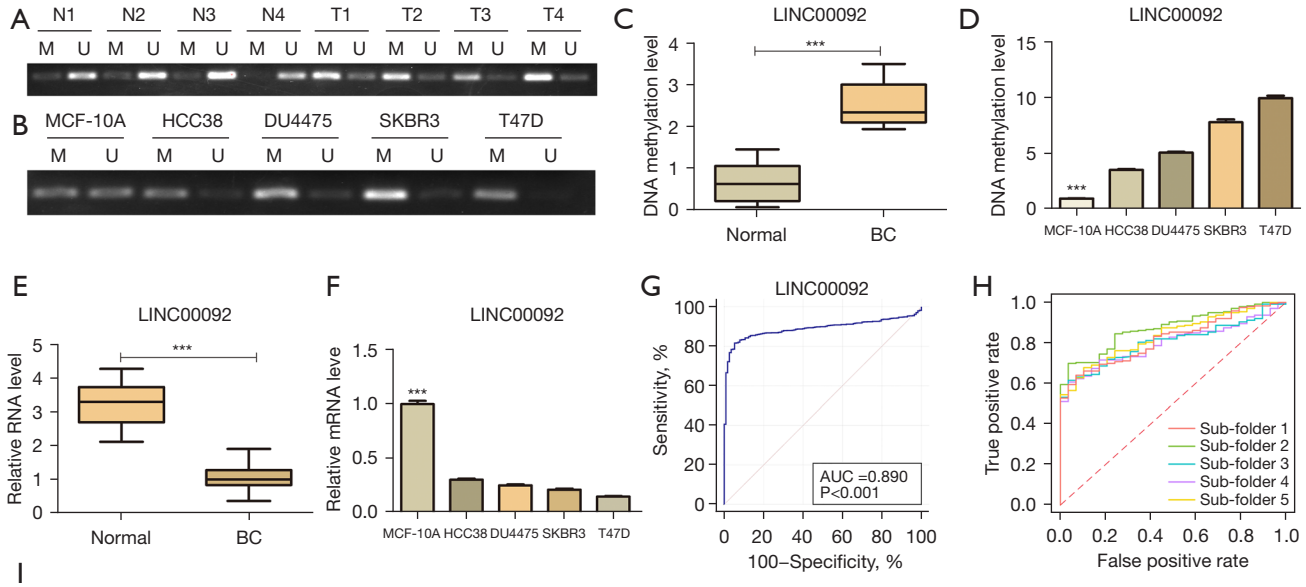
key methylation driver gene was selected by integrated analysis and we gained 1 overlapping upregulated DELs (LINC00092) (*Figure 1D*).

LINC00092 is hypermethylated and lowly expressed in BC

To verify the level of LINC00092 DNA methylation in BC tissues and cell lines, we examined this and found that LINC00092 was low DNA methylation in normal breast tissues and high DNA methylation in BC tissues (*Figure 2A*). In contrast, the RNA expression levels of LINC00092 were considerably higher in normal breast tissues than in BC tissues (*Figure 2B*). The DNA methylation levels of LINC00092 were negatively correlated with the RNA expression levels of LINC00092 (*Figure 2C*). In the TCGA dataset, Kaplan-Meier analysis of the overall survival of patients with BC revealed a significant connection between low LINC00092 expression and poor survival in BC patients (P=0.013, *Figure 2D*). To assess the tumor resistance of LINC00092 in BC, we analyzed its methylation and expression in chemotherapy patients. The methylation levels of LINC00092 were not significant between complete remission/partial remission (CR/PR) patients and stable disease/progressive disease (SD/PD) patients (P=0.2, Figure S1A), but the expression levels of LINC00092 was significantly increased in CR/PR patients compared to SD/PD patients (P=0.026, Figure S1B).

LINC00092 is downregulated in BC and is a potential diagnostic biomarker

To investigate the role of LINC00092 in BC progression, we analyzed LINC00092 methylation and expression levels in BC tissues (N=20) and adjacent non-tumor tissues (N=20) by performing qRT-PCR. The levels of LINC00092 methylation and expression were also detected in BC



I

	Accuracy	Precision	Recall	F measure	AUC
Sub-folder 1	0.896	0.895	1.000	0.945	0.890
Sub-folder 2	0.899	0.902	0.996	0.946	0.873
Sub-folder 3	0.907	0.908	0.996	0.950	0.881
Sub-folder 4	0.907	0.905	1.000	0.950	0.909
Sub-folder 5	0.903	0.902	1.000	0.948	0.899
Average	0.902	0.902	0.998	0.948	0.891

Figure 3 Identification of *LINC00092* genes for the diagnosis of BC. (A) LINC00092 promoter methylation in tumor samples from patients with BC; U and M are amplicons generated from unmethylated and methylated LINC00092 promoter allele-specific primers, respectively. (B) LINC00092 promoter methylation in BC cell line samples. (C) Quantification of LINC00092 methylation levels in BC tissues versus adjacent non-tumor tissues. (D) Quantification of LINC00092 methylation levels in BC cell lines and the MCF-10A cell line. (E) The level of LINC00092 in BC tissues (N=20) compared with that in adjacent non-tumor tissues was assessed by qRT-PCR. (F) The relative level of LINC00092 was assessed by qRT-PCR in BC cell lines and the MCF-10A cell line. (G) The sensitivity and specificity of LINC00092 expression as a diagnostic biomarker for BC were assessed by ROC curve analysis using the TCGA database. (H) ROC curve of LINC00092 in five-fold cross-validation. (I) ROC evaluation parameter table of five-fold validation; all data are expressed as average values. ***P<0.001. BC, breast cancer; qRT-PCR, quantitative real-time polymerase chain reaction; ROC, receiver operating characteristic; AUC, area under the curve; TCGA, The Cancer Genome Atlas.

cell lines. As shown in the chart, the methylation levels of LINC00092 were significantly higher in BC tissues compared with adjacent non-tumor tissues (*Figure 3A,3B*). Compared to the normal human breast epithelial cell line MCF-10A, BC cell lines (HCC38, DU4475, SKBR3, MCF7, and T47D) had higher LINC00092 methylation levels (*Figure 3C,3D*). The levels of LINC00092 RNA expression were then evaluated in BC patient tissue samples and BC cell lines. Compared to normal tissue adjacent to cancer and normal breast epithelial cell lines, LINC00092 was lowly expressed at the RNA level in all patient samples (*Figure 3E*) and in all BC cell lines (*Figure 3F*). Among the BC cell lines, the T47D cell line had the lowest LINC00092 expression and was therefore selected for all subsequent experiments. These findings suggested that LINC00092 may play a role in BC carcinogenesis as a suppressor.

Next, ROC curve analysis was conducted using TCGA data to determine the value of LINC00092 in diagnosing BC. The diagnostic model showed an AUC of 0.890 for discriminating breast cancer from normal samples, indicating that LINC00092 has high sensitivity and specificity in the diagnosis of breast cancer (*Figure 3G*). Subsequently, the efficiency of LINC00092 in diagnosing BC was evaluated using logistic regression. The mean

AUC (0.890±0.05) was calculated from the ROC curve of LINC00092 (*Figure 3H*). The accuracy, precision, recall, and F1-score averaged 0.902, 0.902, 0.998, and 0.948, respectively (*Figure 3I*). These findings showed that LINC00092 performed well in distinguishing BC from normal control samples.

The expression levels of LINC00092 were significantly up-regulated in TNBC patients from non-TNBC patients P<2.22e-16, Figure S2A). And Receiver Operating Characteristic (ROC) analysis displayed that LINC00092 could sensitively distinguish Basal patients from non-basal patients (AUC =0.838, Figure S2B). In addition, the expression levels of LINC00092 were significantly up-regulated in TNBC patients from non-TNBC patients P<2.22e-16, Figure S2C). And ROC analysis displayed that LINC00092 could sensitively distinguish Basal patients from non-basal patients (AUC =0.877, Figure S2D). However, the expression levels of LINC00092 were not significant between M0 and M1 stage (P=0.36, Figure S2E), and the AUC was 0.557(Figure S2F).

LINC00092 inhibits the proliferation and cell cycle of BC cells in vitro

To investigate the potential mechanism and biological functions of LINC00092, we screened 354 relevant PCGs by performing a genome-wide co-expression analysis of LINC00092 with cut-off values of P<0.05 and |r|>0.25. Then, we used GO functional enrichment analysis to assess the functionality of these genes and discovered that the LINC00092-related PCGs were strongly linked to DNA replication (GO:0006260), double-strand break repair via break-induced replication (GO:0000727), G1/S transition of mitotic cell cycle (GO:0000082), DNA-dependent DNA replication (GO:0006261), mitotic metaphase plate congression (GO:0007080), nuclear DNA replication (GO:0033260), negative regulation of nuclear division (GO:0051784), metaphase plate congression (GO:0051310), DNA replication initiation (GO:0006270), chromosomal region (GO:0098687), condensed nuclear chromosome (GO:0000794), condensed chromosome (GO:0000793), centromeric region (GO:0000779), filopodium (GO:0030175), and DNA replication origin binding (GO:0003688) (*Figure 4A*, *Table 3*). KEGG pathway analysis further demonstrated that the PCGs were strongly linked to the cell cycle (hsa04110), alanine, aspartate, and glutamate metabolism (hsa00250), glycosphingolipid biosynthesis-lacto and neolacto series (hsa00601), human T-cell leukemia virus 1 infection (hsa05166), Fanconi anemia pathway (hsa03460), insulin secretion (hsa04911), Th1 and Th2 cell differentiation (hsa04658), DNA replication (hsa03030), apoptosis (hsa04210), cytokine-cytokine receptor interaction (hsa04060), the p53 signaling pathway (hsa04115), and biosynthesis of amino acids (hsa01230) (*Figure 4B*, *Table 4*).

To investigate the role of LINC00092 in BC oncogenesis, HCC38 and T47D cells were transfected with pcDNA-LINC00092 plasmid to upregulate the expression of LINC00092. The results revealed that LINC00092 expression was increased in the pcDNA-LINC00092 group as compared to the pcDNA-NC group (P<0.001, Figure S3A). The CCK-8 assay revealed that LINC00092 overexpression inhibited HCC38 and T47D cell proliferation (*Figure 4C*,*4D*). Cell cycle analysis was then performed using flow cytometry. In HCC38 and T47D cells, overexpression of LINC00092 increased the proportion of S-phase cells and decreased the proportion of G1-phase cells compared to pcDNA-NC (P<0.01, *Figure 4E*,*4F*). The results showed that LINC00092 overexpression caused BC cell cycle block in S phase.

Sulpride mediates the stimulative effect of LINC00092 on BC cell proliferation

CMap analysis was conducted to search for small-molecule drugs that could potentially be employed as latent targeted therapeutic drugs for LINC00092 in BC. Two such drugs were identified, nimesulide (mean connective score =–0.521; P=0.02286) and sulpiride (mean connective score =–0.537; P=0.00274, *Figure 5A*), and their chemical structures are shown in *Figures 5B*,*5C*.

The effects of nimesulide and sulpiride on cell growth were assessed in the T47D cell line (*Figure 5D*). Nimesulide and sulpiride were found to suppress the growth of T47D cells in a dose-dependent manner, with the IC50 values being 77.04 μM and 40.6 μM, respectively. Due to having a much lower IC50 value, sulpiride was selected for the follow-up rescue trial.

Considering that sulpiride was negatively correlated with LINC00092 (*Figure 5A*), we hypothesized that it may affect BC cell proliferation via LINC00092. A set of si-RNAs were transfected into T47D cells to knock down the expression of LINC00092. LINC00092 expression was significantly reduced in the si-LINC00092 group compared to in the si-NC group (Figure S3B). Therefore, we next investigated whether sulpiride could influence the si-LINC00092-mediated proliferation effect in BC cells.

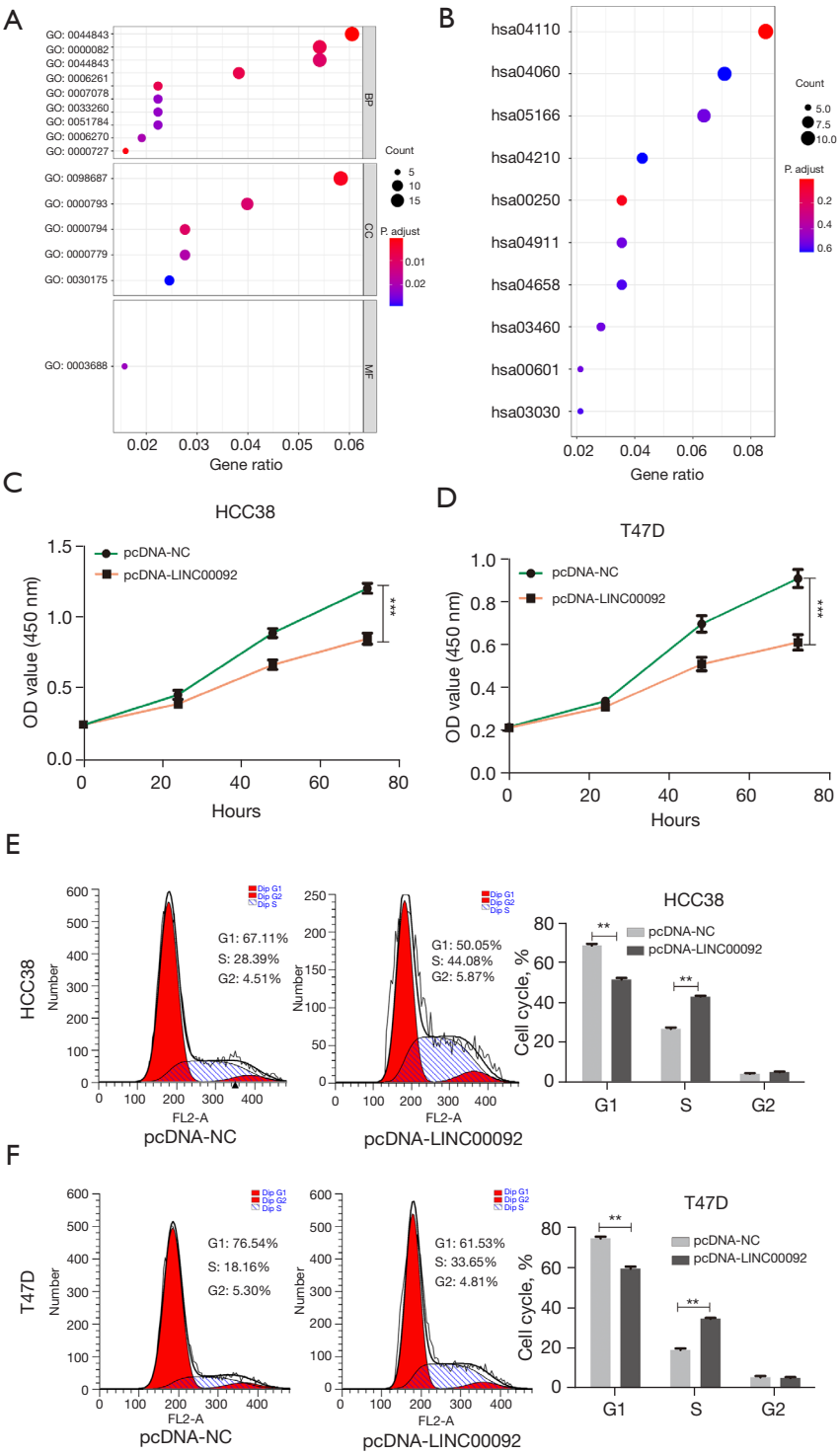


Figure 4 Pathway analysis of LINC00092 co-expressed genes. (A) Bubble diagram of significant enrichment pathway in GO analysis. (B) Bubble plot of significant enrichment pathway in KEGG analysis. (C,D) The proliferation curves of HCC38 and T47D cells after transfection with pcDNA-LINC00092 or pcDNA-NC were determined by CCK-8 assay. (E,F) The cell cycle distributions of HCC38 and T47D cells were analyzed by FACS at 48 hours following transfection (n=3). Representative plots are shown. **P<0.01, and ***P<0.001. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; CCK-8, Cell Counting Kit-8; OD, optical density; FACS, fluorescence-activated cell sorting.

Table 3 Significant GO functional enrichment pathways for the LINC00092-related PCGs

Ontology	ID	Description	P value
BP	GO:0006260	DNA replication	1.72E-07
BP	GO:0000727	Double-strand break repair via break-induced replication	8.90E-07
BP	GO:0000082	G1/S transition of mitotic cell cycle	6.90E-06
BP	GO:0006261	DNA-dependent DNA replication	8.72E-06
BP	GO:0007080	Mitotic metaphase plate congression	1.18E-05
BP	GO:0044843	Cell cycle G1/S phase transition	1.88E-05
BP	GO:0006270	DNA replication initiation	4.16E-05
BP	GO:0033260	Nuclear DNA replication	6.00E-05
BP	GO:0051784	Negative regulation of nuclear division	6.00E-05
BP	GO:0051310	Metaphase plate congression	6.68E-05
BP	GO:0000070	Mitotic sister chromatid segregation	8.29E-05
BP	GO:0051315	Attachment of mitotic spindle microtubules to kinetochore	8.89E-05
BP	GO:0044786	Cell cycle DNA replication	0.000100776
BP	GO:0000280	Nuclear division	0.000109593
BP	GO:0000819	Sister chromatid segregation	0.000114045
BP	GO:0006268	DNA unwinding involved in DNA replication	0.000116977
BP	GO:0030010	Establishment of cell polarity	0.000127679
BP	GO:0045839	Negative regulation of mitotic nuclear division	0.000216659
BP	GO:0007094	Mitotic spindle assembly checkpoint	0.00023162
BP	GO:0031577	Spindle checkpoint	0.00023162
BP	GO:0071173	Spindle assembly checkpoint	0.00023162
BP	GO:0071174	Mitotic spindle checkpoint	0.00023162
BP	GO:0022409	Positive regulation of cell-cell adhesion	0.000257495
BP	GO:0045841	Negative regulation of mitotic metaphase/anaphase transition	0.000305389
BP	GO:0050000	Chromosome localization	0.000344166
BP	GO:0051303	Establishment of chromosome localization	0.000344166
BP	GO:1902100	Negative regulation of metaphase/anaphase transition of cell cycle	0.00034834
BP	GO:0048285	Organelle fission	0.000419863
BP	GO:0022407	Regulation of cell-cell adhesion	0.000432523
BP	GO:0051383	Kinetochore organization	0.000434528
BP	GO:0071459	Protein localization to chromosome, centromeric region	0.000434528
BP	GO:2000816	Negative regulation of mitotic sister chromatid separation	0.000447778
BP	GO:1905819	Negative regulation of chromosome separation	0.000504848
BP	GO:0007059	Chromosome segregation	0.000509115
BP	GO:0006310	DNA recombination	0.000516322

Table 3 (*continued*)

Table 3 (*continued*)

Ontology	ID	Description	P value
BP	GO:0006541	Glutamine metabolic process	0.000519157
BP	GO:0051220	Cytoplasmic sequestering of protein	0.000519157
CC	GO:0098687	Chromosomal region	7.25E-06
CC	GO:0000794	Condensed nuclear chromosome	5.65E-05
CC	GO:0000793	Condensed chromosome	9.54E-05
CC	GO:0000779	Condensed chromosome, centromeric region	0.000208369
CC	GO:0030175	Filopodium	0.000400143
MF	GO:0003688	DNA replication origin binding	3.94E-05

GO, Gene Ontology; PCGs, protein-coding genes; BP, biological process; CC, cellular component; MF, molecular function.

Table 4 The KEGG pathways are significantly enriched for the LINC00092-related PCGs

ID	Description	P value
hsa04110	Cell cycle	1.49E-06
hsa00250	Alanine, aspartate and glutamate metabolism	0.000419139
hsa00601	Glycosphingolipid biosynthesis—lacto and neolacto series	0.011199921
hsa05166	Human T-cell leukemia virus 1 infection	0.01427901
hsa03460	Fanconi anemia pathway	0.014319203
hsa04911	Insulin secretion	0.016978947
hsa04658	Th1 and Th2 cell differentiation	0.022082807
hsa03030	DNA replication	0.024407474
hsa04210	Apoptosis	0.031470418
hsa04060	Cytokine-cytokine receptor interaction	0.03332476
hsa04115	p53 signaling pathway	0.03832561
hsa01230	Biosynthesis of amino acids	0.041687559
hsa03008	Ribosome biogenesis in eukaryotes	0.044423777
hsa05412	Arrhythmogenic right ventricular cardiomyopathy	0.045211714

KEGG, Kyoto Encyclopedia of Genes and Genomes; PCGs, protein-coding genes.

We found that the proliferation-promoting effect of si-LINC00092 was partially reversed after treatment with sulpiride (*Figure 5E*).

Discussion

Epigenetic changes and modifications are a necessary part of the initiation and progression of tumorigenesis (22). Critical driver genomic events that contribute to oncogenesis are important mechanisms of the developmental processes of BC. Emerging data have revealed that lncRNAs play important roles in the occurrence and development of tumours via epigenetic regulation such as DNA methylation, histone modification, and chromatin remodelling, all of which have been linked to numerous biological markers. A growing number of studies have revealed a large number of lncRNAs that are involved in multiple diseases, including tumor progression, by regulating gene expression at the

A

Cmap name	Mean	N	Enrichment	P	Specificity	%Non-null
Nimesulide	−0.521	4	−0.68	0.02286	0.0129	75
Sulpiride	−0.537	5	−0.734	0.00274	0.0102	80

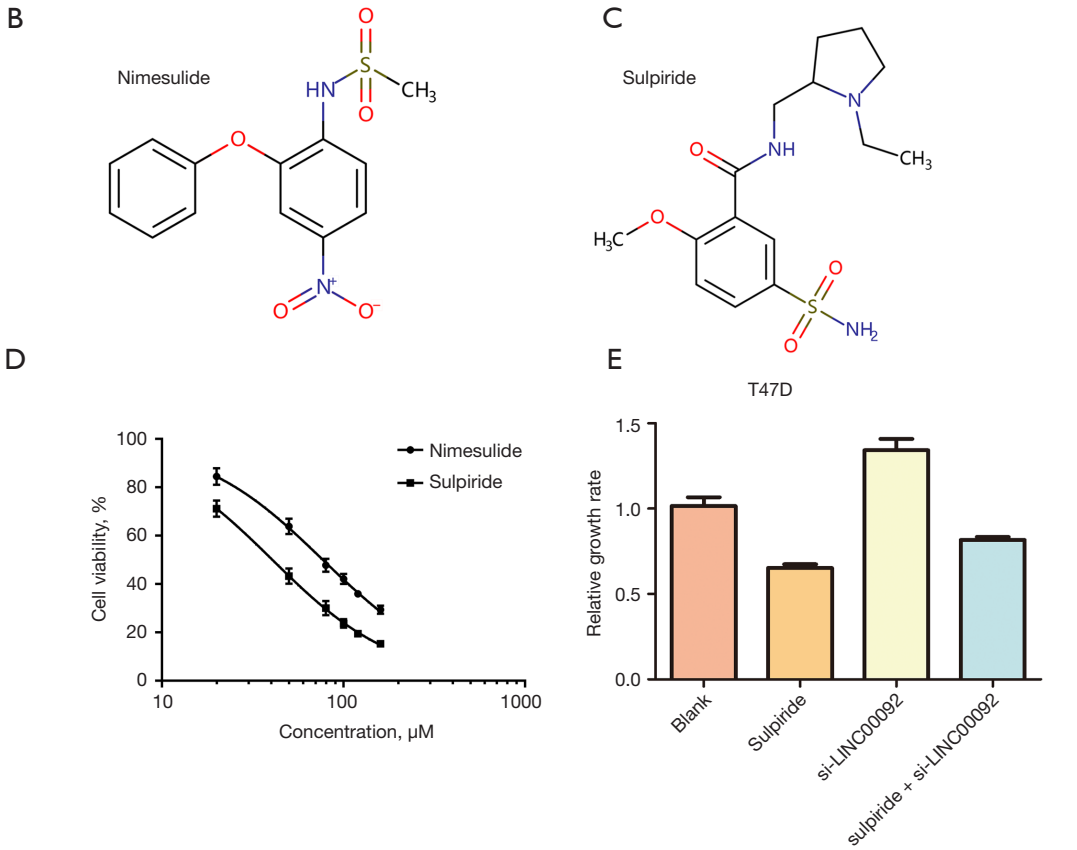


Figure 5 Sulpiride reverses si-LINC00092 mediated promotion of cell proliferation. (A) Results of CMap analysis. (B) The chemical structure of nimesulide. (C) The chemical structure of sulpiride. (D) BC cell lines were cultured for 72 hours with nimesulide and sulpiride at the concentrations specified, and CCK-8 assay was used to ascertain cell growth. (E) Sulpiride reversed si-LINC00092-mediated promotion of BC cell proliferation. CMap, Connectivity Map; BC, breast cancer; CCK-8, Cell Counting Kit-8.

epigenetic, transcriptional and post-transcriptional levels. In breast cancer studies, some lncRNA were identified as tumor-driven oncogenic lncRNA and a few were identified as tumor suppressor lncRNA. they are involved in cell growth, apoptosis, cell migration, invasion, and stemness of cancer cells. Therefore, such novel lncRNAs may serve as diagnostic and prognostic biomarkers and as potential therapeutic targets for breast cancer. For instance, oncogenes such as *PCAT-1* (23), *MALAT1* (24), and *LINC01614* (25), which are linked with pan-cancer development, have also been determined to be prognostic biological markers in BC.

In this study, 14 BC-specific methylation-driven LncRNAs, 526 DELs, and 93 survival-related lncRNAs were identified in BC. Among these lncRNAs, the only overlapping lncRNA (LINC00092) was identified as a methylation-driven biomarker for predicting the prognosis of BC. In previous research, low expression of LINC00092 has been identified as a prognostic biomarker for lung adenocarcinoma (26). Chen *et al.* reported that the *LINC00092* gene was involved in the activation of human cardiac fibroblasts and regulated glycolysis (27). LINC00092 was also reported to play a role in driving glycolysis and promoting ovarian cancer progression in ovarian cancer-associated fibroblasts (28). In recent years, DNA methylation is widely studied, and hypermethylation is linked to the suppression effect of oncogenes expression. In the current study, we detected the methylation and

expression levels of LINC00092 in clinical tumor samples. We found that LINC00092 was hypermethylated in BC tissues and cells as compared to non-tumor tissues and normal breast cells, respectively. Our findings also showed that LINC00092 expression was decreased considerably in BC tissues and cells as compared to non-tumor tissues and normal breast cells, respectively. Further, we discovered that patients with lower LINC00092 expression had poorer overall survival than did those with higher LINC00092 expression, which suggested that DNA methylation was crucial to the progression of BC and linked to poor clinical outcomes. Further suggesting that low expression of LINC00092 can be used to predict poor prognosis in breast cancer patients. ROC curve analysis validated the excellent sensitivity and specificity of LINC00092 as a potential marker for the diagnosis of BC. Overall, these results suggest that LINC00092 plays an essential role in breast cancer tumorigenesis and is a good biomarker for breast cancer diagnosis and prediction of prognosis.

There is currently no research into the molecular mechanism of LINC0092 in BC. As a result, the second stage in this study was to probe into the potential mechanism of LINC00092 in BC. GO and KEGG pathway analyses indicated that LINC00092-related genes were involved in DNA replication, double-strand break repair via break-induced replication, G1/S transition of mitotic cell cycle, DNA-dependent DNA replication, mitotic metaphase plate congression, nuclear DNA replication, negative regulation of nuclear division, DNA replication origin binding and cell cycle, DNA replication, apoptosis, cytokine-cytokine receptor interaction, and the p53 signaling pathway in BC. *In vitro* cell experiments in BC cell lines were then used to further understand the biological function of LINC00092 in BC progression. LINC00092 was overexpressed in BC cell lines via transfection of pcDNA-LINC00092, and we observed that LINC00092 overexpression decreased the proliferation of HCC38 and T47D cells. Therefore, we considered LINC00092 as a potential target for the treatment of breast cancer. Furthermore, we found that HCC38 and T47D cell cycles were also altered after overexpression of LINC00092, with the increased expression of LINC00092 causing more cells to block in the S phase. Thus, we speculate that the negative effect of LINC00092 overexpression on cell proliferation may be due to the cell cycle of BC cells being blocked in the S phase.

Targeted therapies are the focus of much research in oncology. CMap has been employed in a growing number of studies to find promising small molecules to treat a variety of diseases (29-31). For instance, Huang *et al.* predicted drugs that could overcome 5-FU resistance in colorectal cancer through a pharmacogenomic analysis (32). In the present study, we used CMap analysis to identify two small-molecule drugs (nimesulide and sulpiride) that can be utilized as latent targeted drugs for LINC00092 in BC. Nimesulide has been reported to inhibit pancreatic cancer cell proliferation and cause apoptosis by increasing the expression of PTEN (33). Previous research has also shown that sulpiride, a dopamine d2-like receptor antagonist, enhances the antitumor effects of dexamethasone in tumor stem cell-like cells resistant BC (34). Our study explored the effects of nimesulide and sulpiride on cell growth in the T47D cell line. The IC50 value of sulpiride acting on BC cells was lower than that of nimesulide, which indicated that sulpiride had greater potential for BC cell growth inhibition. Furthermore, our data showed that siRNA-mediated silencing of LINC00092 resulted in an obvious increase in the proliferation capacity of T47D cells, whereas the addition of sulpiride significantly inhibited the growth of LINC00092-silenced cells. Therefore, we hypothesize that LINC00092 could be a potential target for treating BC.

This study aimed to identify a promising methylation-driven lncRNA that could serve as a diagnostic and predictive tool for BC prognosis. Our findings suggest that LINC00092 could be a potential biological marker for independent diagnosis and prognostic prediction for patients with BC, as well as a promising candidate target for targeted BC therapy.

Acknowledgments

Funding: This study was supported by the Heilongjiang Provincial Natural Science Foundation of China (No. H2017048).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://atm.amegroups.com/article/view/10.21037/atm-22-1956/rc

Data Sharing Statement: Available at https://atm.amegroups.com/article/view/10.21037/atm-22-1956/dss

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://atm.

amegroups.com/article/view/10.21037/atm-22-1956/coif). All authors report that this study was supported by the Heilongjiang Provincial Natural Science Foundation of China (No. H2017048). The authors have no other conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Ethics Committee of the Harbin Medical University Cancer Hospital (No. 2019-185). All of the patients provided written informed consent prior to enrollment.

References

1. DeSantis C, Ma J, Bryan L, et al. Breast cancer statistics, 2013. CA Cancer J Clin 2014;64:52-62.
2. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
3. Waks AG, Winer EP. Breast Cancer Treatment: A Review. JAMA 2019;321:288-300.
4. Filipova A, Seifrtova M, Mokry J, et al. Breast cancer and cancer stem cells: a mini-review. Tumori 2014;100:363-9.
5. King CR, Swain SM, Porter L, et al. Heterogeneous expression of erbB-2 messenger RNA in human breast cancer. Cancer Res 1989;49:4185-91.
6. Vishnubalaji R, Alajez NM. Epigenetic regulation of triple negative breast cancer (TNBC) by TGF-β signaling. Sci Rep 2021;11:15410.
7. Luo C, Huang J, Guo Z, et al. Methylated biomarkers for breast cancer identified through public database analysis and plasma target capture sequencing. Ann Transl Med 2021;9:683.
8. Howlader N, Altekruse SF, Li CI, et al. US incidence of breast cancer subtypes defined by joint hormone receptor and HER2 status. J Natl Cancer Inst 2014;106:dju055.
9. Fragomeni SM, Sciallis A, Jeruss JS. Molecular Subtypes and Local-Regional Control of Breast Cancer. Surg Oncol Clin N Am 2018;27:95-120.
10. Burguin A, Diorio C, Durocher F. Breast Cancer Treatments: Updates and New Challenges. J Pers Med 2021;11:808.
11. Qu X, Alsager S, Zhuo Y, et al. HOX transcript antisense RNA (HOTAIR) in cancer. Cancer Lett 2019;454:90-7.
12. Yang S, Cai H, Hu B, et al. LncRNA SAMMSON negatively regulates miR-9-3p in hepatocellular carcinoma cells and has prognostic values. Biosci Rep 2019;39:BSR20190615.
13. Li ZX, Zhu QN, Zhang HB, et al. MALAT1: a potential biomarker in cancer. Cancer Manag Res 2018;10:6757-68.
14. Ghafouri-Fard S, Esmaeili M, Taheri M. H19 lncRNA: Roles in tumorigenesis. Biomed Pharmacother 2020;123:109774.
15. Noruzinia M, Tavakkoly-Bazzaz J, Ahmadvand M, et al. Young Breast Cancer: Novel Gene Methylation in WBC. Asian Pac J Cancer Prev 2021;22:2371-5.
16. Zhang D, Wang Y, Yang Q. A High Epigenetic Risk Score Shapes the Non-Inflamed Tumor Microenvironment in Breast Cancer. Front Mol Biosci 2021;8:675198.
17. Suzuki MM, Bird A. DNA methylation landscapes: provocative insights from epigenomics. Nat Rev Genet 2008;9:465-76.
18. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat Rev Genet 2013;14:204-20.
19. Dai X, Ren T, Zhang Y, et al. Methylation multiplicity and its clinical values in cancer. Expert Rev Mol Med 2021;23:e2.
20. Pang JB, Savas P, Fellowes AP, et al. Breast ductal carcinoma in situ carry mutational driver events representative of invasive breast cancer. Mod Pathol 2017;30:952-63.
21. Jin B, Robertson KD. DNA methyltransferases, DNA damage repair, and cancer. Adv Exp Med Biol 2013;754:3-29.
22. Dumitrescu RG. Interplay Between Genetic and Epigenetic Changes in Breast Cancer Subtypes. Methods Mol Biol 2018;1856:19-34.
23. Wang J, Chen X, Hu H, et al. PCAT-1 facilitates breast cancer progression via binding to RACK1 and enhancing oxygen-independent stability of HIF-1α. Mol Ther Nucleic Acids 2021;24:310-24.
24. Sun Z, Liu J, Liu J. The expression of lncRNA-MALAT1

in breast cancer patients and its influences on prognosis. Cell Mol Biol (Noisy-le-grand) 2020;66:72-8.
25. Vishnubalaji R, Shaath H, Elkord E, et al. Long non-coding RNA (lncRNA) transcriptional landscape in breast cancer identifies LINC01614 as non-favorable prognostic biomarker regulated by TGFβ and focal adhesion kinase (FAK) signaling. Cell Death Discov 2019;5:109.
26. Wang X, Wang X, Huang G, et al. Lower expression of LINC00092 in lung adenocarcinoma might mean poorer prognosis: A study based on data mining and bioinformatics. Medicine (Baltimore) 2020;99:e23012.
27. Chen ZT, Zhang HF, Wang M, et al. Long non-coding RNA Linc00092 inhibits cardiac fibroblast activation by altering glycolysis in an ERK-dependent manner. Cell Signal 2020;74:109708.
28. Zhao L, Ji G, Le X, et al. Long Noncoding RNA LINC00092 Acts in Cancer-Associated Fibroblasts to Drive Glycolysis and Progression of Ovarian Cancer. Cancer Res 2017;77:1369-82.
29. Lim N, Pavlidis P. Evaluation of connectivity map shows limited reproducibility in drug repositioning. Sci Rep 2021;11:17624.
30. He J, Yan H, Cai H, et al. Statistically controlled identification of differentially expressed genes in one-to-one cell line comparisons of the CMAP database for drug repositioning. J Transl Med 2017;15:198.
31. Liang P, Li J, Chen J, et al. Immunoprognostic model of lung adenocarcinoma and screening of sensitive drugs. Sci Rep 2022;12:7162.
32. Huang TC, Peng KC, Kuo TT, et al. Predicting Agents That Can Overcome 5-FU Resistance in Colorectal Cancers via Pharmacogenomic Analysis. Biomedicines 2021;9:882.
33. Chu M, Wang T, Sun A, et al. Nimesulide inhibits proliferation and induces apoptosis of pancreatic cancer cells by enhancing expression of PTEN. Exp Ther Med 2018;16:370-6.
34. Li J, Yao QY, Xue JS, et al. Dopamine D2 receptor antagonist sulpiride enhances dexamethasone responses in the treatment of drug-resistant and metastatic breast cancer. Acta Pharmacol Sin 2017;38:1282-96.

(English Language Editor: J. Reylonds)

Cite this article as: Li J, Lu F, Shao X, You B. Investigating the potential clinical significance of long non-coding RNA 00092 in patients with breast cancer. Ann Transl Med 2022;10(10):602. doi: 10.21037/atm-22-1956

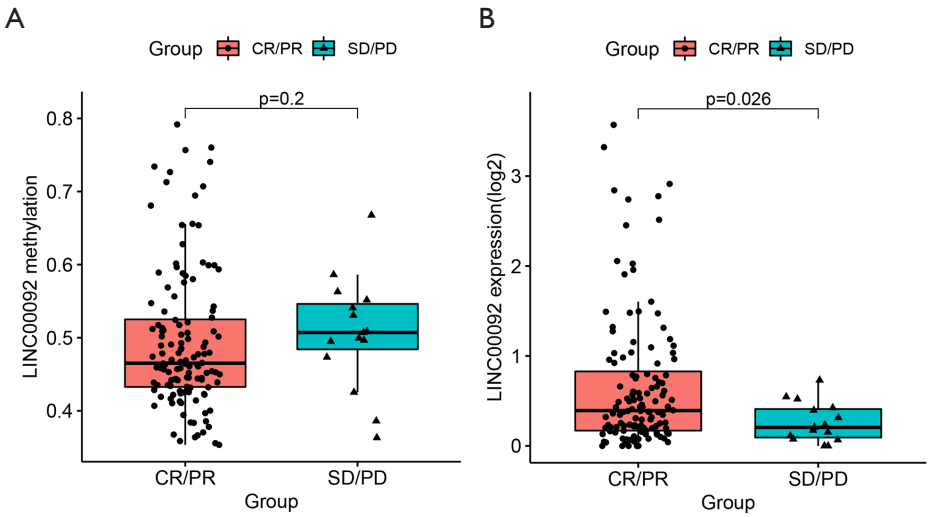


Figure S1 The methylation and expression levels of LINC00092 in chemotherapy patients. (A) Boxplot of LINC00092 methylation and (B) Boxplot of LINC00092 expression comparison between CR/PR and SD/PD. CR, complete remission; PD, progressive disease; PR, partial remission; SD, stable disease.

https://dx.doi.org/10.21037/atm-22-1956

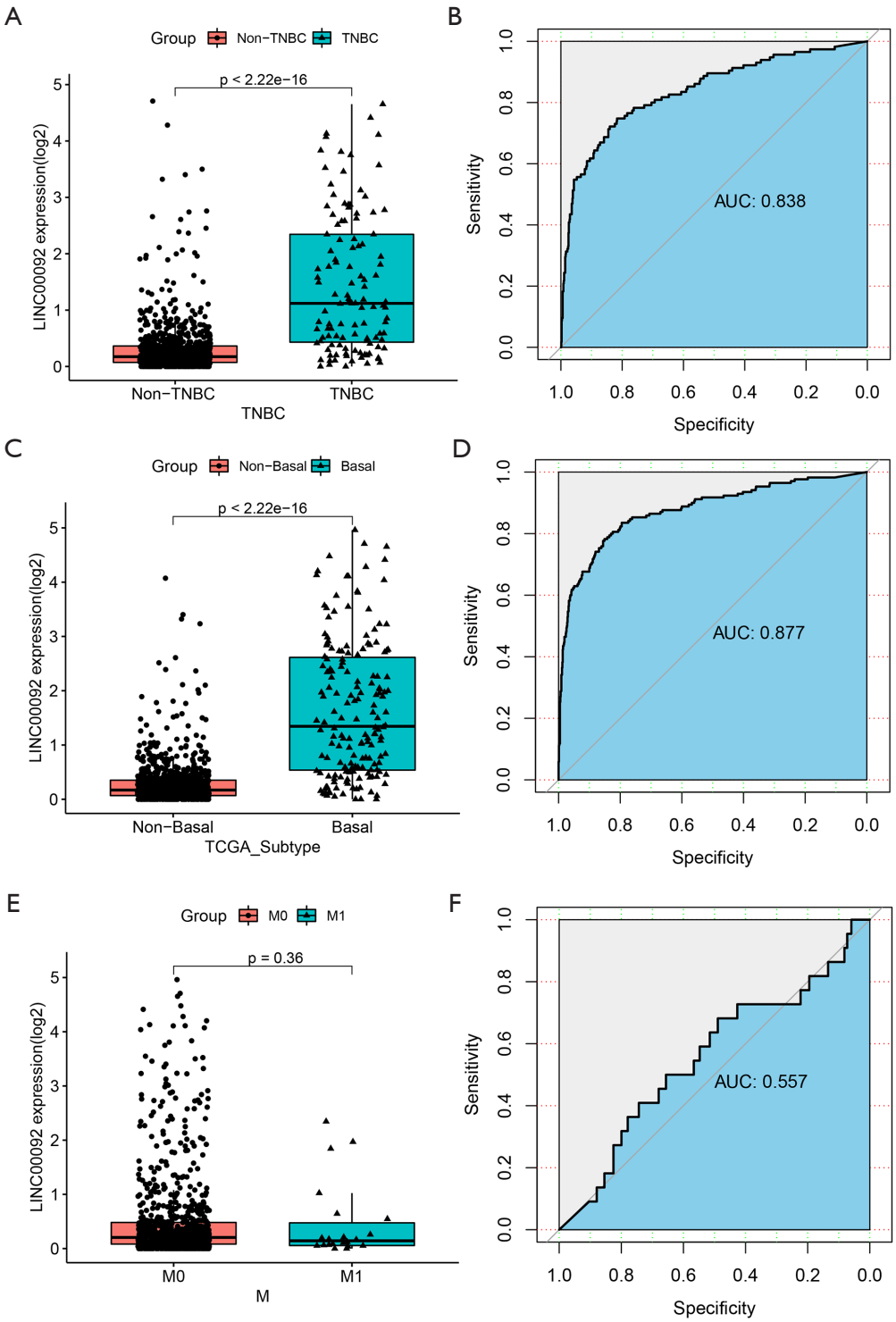


Figure S2 Determine the expression and clinical value of LINC00092. (A) Boxplot comparison between Non-TNBC and TNBC and (B) ROC curve was applied to evaluate the diagnostic value of LINC00092 expression for TNBC. (C) Boxplot comparison between Non-Basal and Basal and (D) ROC curve was applied to evaluate the diagnostic value of LINC00092 expression for Basal. (E) Boxplot comparison between M0 and M1 and (F) ROC curve was applied to evaluate the diagnostic value of LINC00092 expression for M1. TNBC, triple negative breast cancer; ROC, receiver operating characteristic; AUC, area under the curve.

https://dx.doi.org/10.21037/atm-22-1956

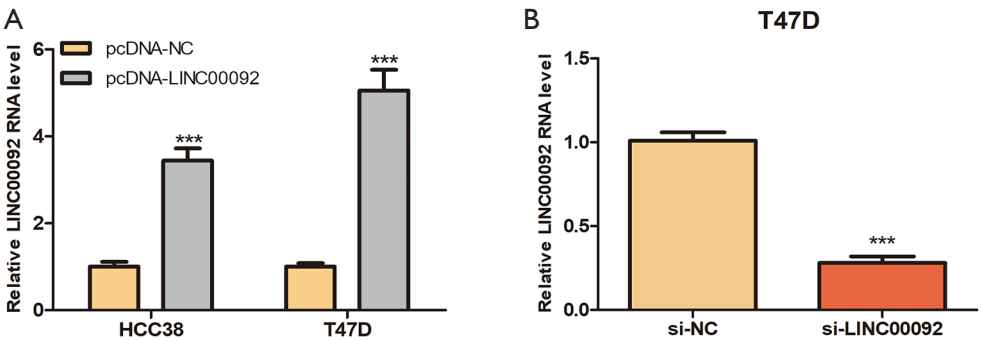


Figure S3 Detection of LINC00092 expression by qRT-PCR. (A) validation of LINC00092 overexpression; (B) validation of LINC00092 knockdown. qRT-PCR, quantitative real-time polymerase chain reaction. ***P<0.001.

https://dx.doi.org/10.21037/atm-22-1956