Comprehensive analysis of the aberrantly expressed IncRNA-associated ceRNA network in breast cancer

TAYIER TUERSONG1*, LINLIN LI1*, ZUMURETI ABULAITI2 and SHUMEI FENG3

1Department of Pharmacology, Pharmacy College; 2Department of Internal Medicine, Traditional Medical college; 3Department of Histology and Embryology, Basic Medical Sciences College, Xinjiang Medical University, Ürümqi, Xinjiang Uyghur Autonomous Region 830001, P.R. China

Received July 20, 2018; Accepted April 1, 2019

DOI: 10.3892/mmr.2019.10165

Abstract. Previous studies have suggested that long non-coding RNAs (lncRNAs) are closely associated with human diseases, particularly cancer, including cancer of the lung, breast and stomach. A variety of lncRNAs are abnormally expressed in cancer and participate in several pathways including cell proliferation and apoptosis; these elements are closely associated with the development of cancer. The Cancer Genome Atlas (TCGA) is an important cancer database. It consists of clinical data, genomic variation, mRNA, microRNA (miRNA) and lncRNAs expression, methylation and other data for various types of human cancer. In the present study, differential expression of RNA was identified using the edgeR package. A total 1,222 RNA sequencing profiles from patients with breast cancer were downloaded from TCGA. A competing endogenous RNA (ceRNA) network was constructed for breast cancer based on mircode and miRTarBase. The top 10 lncRNAs were selected using cox regression analysis. Survival analysis was performed using Kaplan-Meier analysis. A total of 1,028 breast cancer-associated lncRNAs and 89 miRNAs (fold change >2; P<0.05) were identified; among these, 93 lncRNAs and 19 miRNAs were included in the ceRNA network. Subsequently, 10 basic lncRNAs were selected and their associations with overall survival were identified. In addition, 5 lncRNAs (ADAM metallopeptidase with thrombospondin type 1 motif 9-antisense RNA 1, AL513123.1, chromosome 10 open reading frame 126, long intergenic non-protein coding RNA 536 and Wilms tumor 1 antisense RNA) were identified to be significantly associated with overall survival (P<0.05, log rank test). These results suggested that mRNAs, lncRNAs and miRNAs were involved in pathological mechanisms of breast cancer. The newly-identified ceRNA network included 93 breast cancer-specific lncRNAs, 19 miRNAs and 27 mRNAs. The results of the present study highlight the potential of lncRNAs in understanding the development and pathogenesis of breast cancer, and suggest novel concepts and an experimental basis for the identification of prognostic biomarkers and therapeutic targets for breast cancer.

Introduction

Breast cancer is one of the most widespread types of cancer in females, and its incidence is increasing yearly. It is expected that by 2018, the United States of America will have ~266,120 incident cases of invasive breast cancer and 63,960 cases of in situ breast cancer. Furthermore, it is estimated that 40,920 women will succumb to breast cancer (1). Similarly, in China, the incidence of breast cancer is the primary cause of mortality in women <45 years, followed by colorectal, liver and esophageal cancer (2). Following systemic treatment, patients with breast cancer continue to experience recurrence and metastasis, and the majority of patients eventually succumb to metastases. Therefore, it is important to identify diagnostic and prognostic markers, and potential therapeutic targets (3,4). The present study explored the mechanisms through which long non-coding RNAs (lncRNAs) act as competing endogenous RNAs (ceRNAs) to regulate target genes and participate in the prognosis and pathogenesis of breast cancer.

Over previous decades, 70-90% of the transcribed human genome has been identified. Relevant data indicate that the proportion of genes encoding proteins in the genome is >2%, and non-coding RNA accounts for the majority of the human transcriptome (5). Non-coding RNAs are a large class of RNA molecules that do not encode proteins, but which serve regulatory roles. Non-coding RNAs may be divided into three classes, by length: <50 nucleotides (nt), including microRNA (miRNA), small interfering RNA and Piwi-interacting RNA; 50-500 nt, including ribosomal RNA, transfer RNA, small nuclear RNA, small nucleolar RNA and IncRNA; >500 nt, including long mRNA-like non-coding RNAs and lncRNAs without polyA tails (6). IncRNA is a generic term for a class of
RNA molecules with lengths of >200 nucleotides, and is one of the most active fields of study in molecular biology at present. IncRNAs regulate the gene expression of tumor cells through multiple modes of action, and therefore are widely involved in the occurrence and metastasis of tumors (7-9). IncRNAs are thought to serve an important role in the development of cancer; however, only a few have been well-characterized regarding their functional roles (10).

Differentially expressed ceRNAs have been identified in a number of diseases, particularly in cancer. Previous studies have demonstrated that ceRNAs affect the proliferation, growth, differentiation, apoptosis and other biological behaviors of cancer cells (11-13). miRNA response elements (MREs) are partially complementary sequences located on mRNAs that bind to miRNAs, and may inhibit the expression of miRNA target genes. miRNAs regulate hundreds of mRNAs, and one miRNA may be regulated by hundreds of mRNAs. Numerous types of RNA molecules constitute a ceRNA regulatory network, and the more MREs that are shared between them, the greater the potential for communication or co-regulation of a target gene (14). Theoretically, almost every RNA molecule possesses at least one MRE binding site that binds to an miRNA, thereby forming a ceRNA association with the corresponding miRNA (15). Among these, the interaction mechanisms between IncRNA and miRNA include: miRNA binding to and degrading IncRNA; IncRNA acting as miRNA sponge by binding to adsorbed miRNA; and truncation of IncRNA to generate miRNA (16). There is evidence that IncRNAs interact with miRNAs and act as ceRNAs to regulate the expression of miRNAs. CeRNAs serve important roles together with cancer-associated genes, and they have demonstrated potential in clinical tumor diagnosis and treatment (17,18).

As the largest database of cancer gene information currently available, The Cancer Genome Atlas (TCGA) includes 33 types of cancers in situ, advancing the understanding of the molecular basis of the onset of these diseases and improving the ability to diagnose, treat and prevent cancer. TCGA comprises multiple levels of tumor data, including genomic, transcriptomic, proteomic and epigenetic data (19). In the present study, according to the analysis of RNA expression profiles among 1,109 breast cancer tumor and 113 non-tumor tissue samples, a IncRNA-miRNA-mRNA regulatory network was successfully constructed. Furthermore, Cox regression analysis was used to select the top 10 differentially expressed IncRNAs (DEIncRNAs), to understand the function of IncRNAs and their potential mechanisms in breast cancer. Survival analyses were also performed to identify prognostic genes.

Materials and methods

Study population. RNA sequencing data was downloaded from TCGA (https://www.cancergenome.nih.gov/). As of February 19, 2018, a total of 1,092 breast cancer cases were collected in the database. Univariate Cox analysis was performed using the library (survival) package in R software (version 3.4.3; R: A Language and Environment for Statistical Computing) (20) from the TCGA database. The present study adhered to the TCGA publication guidelines.

Screening differentially expressed genes. Breast cancer mRNA and miRNA sequencing data were derived from 1,222 samples, including 1,109 tumor samples (cohort Tumor) and 113 normal samples (cohort Normal). Normal sample and tumor sample data were then merged, and all expression values equal to zero were removed. Comparing the data from normal tissues with breast cancer samples, Perl language was used to extract the mRNA matrix file data from the RNA sequencing data (IncRNA, mRNA) downloaded by TCGA in the Windows Environment, running the program with cmd.exe, and then the data were converted into ensemble ID to obtain the gene symbol matrix file. The gene symbol matrix file contained the IncRNA and mRNA matrix, and gets-IncRNA symbol.pl and get-mRNA symbol.pl scripts were used to obtain the IncRNA symbol and mRNA symbol matrix data, respectively. DEIncRNAs were measured using IncRNA symbol matrix and ‘edgeR’ package in R software (version 3.4.3), with thresholds of log2 fold change (FC) >2.0 and adjusted P<0.01. DEmiRNAs were defined as those with log2FC >2.0 and adjusted P<0.01.

Constructing the ceRNA network in breast cancer. DEIncRNAs and DEmiRNAs were predicted using the miRcode (http://www.mircode.org/), starBase (http://starbase.sysu.edu.cn/) (21), using mirTarBase online software to perform target gene predictions on the screened differential miRNAs (22). TargetScan (http://www.targetscan.org/vert_72/) and miRDB (http://mirdb.org/) databases were used to predict the targeting relationship between DEmiRNA and DEsmRNA. The target DEmiRNA is entered into the database, which then shows all the target genes that may interact with the target microRNA, and then further screens the genes involved by the construction of a ceRNA network (23). To understand the functions of miRNA and IncRNA with the ceRNA network, gene co-expression network analysis was used, and the results were visualized with Cytoscape 3.6.0 (24).

Top 10 aberrantly expressed IncRNAs in breast cancer. The top 10 IncRNAs were selected by receiver operating characteristics (ROC) curve and area under the curve (AUC) analysis. The prognostic roles of IncRNAs were examined with Kaplan-Meier curve analysis, and the log-rank test was conducted to distinguish survival time. P<0.05 was considered to indicate a statistically significant difference.

Analysis the IncRNAs with gene expression profiling interactive analysis (GEPIA). Similarly, IncRNA expression levels in para-nocancerous tissues and cancer tissues were compared using the GEPIA database (http://gepia.cancer-pku.cn) (25), which was used to analyze the RNA sequencing data of 33 types of cancers from TCGA database.

Survival analysis. Using the survival time data for breast cancer in TCGA, the ‘Survival’ package in R software (version 3.4.3) was used to analyze the specific IncRNA, miRNA and mRNA associated with survival. Kaplan-Meier survival analysis was performed to analyze the correlation between DEIncRNAs signature and breast cancer patient prognosis. P<0.05 was considered to indicate a statistically significant difference.
**Results**

**DElncRNAs, DEMRNAs and DEmiRNAs.** RNA expression levels in 1,109 breast cancer tumor and 113 normal tissue samples were investigated. Genes with |log2 Fc| >1.5 and adjusted P<0.01 were considered as differentially expressed. The significant DEmiRNAs and DEmiRNAs was identified in breast cancer samples. A total of 2,155 DEmRNAs, 89 DEmiRNAs and 1,028 IncRNAs were identified from TCGA data set using the ‘edgeR’ package in R. Of these, 1,434 mRNAs, 68 miRNAs and 777 IncRNAs were upregulated and 721 mRNAs, 21 miRNAs and 251 IncRNAs were downregulated in breast cancer tissues compared with normal tissues. The RNAs hierarchical clustering analyses are presented in Fig. 1A, and it was demonstrated that the expression levels of these three types of RNAs were significantly differentiated compared with the normal tissues. Finally, volcano plots were generated, and differences between the normal and tumor groups were identified (Fig. 1B).

**miRNA targeting IncRNAs predicted by miRcode and starBase.** A total of 89 breast cancer-associated miRNAs were identified to be expressed in breast cancer and normal tissues. A total of 19 miRNAs were selected from 89 breast cancer-associated miRNAs in TCGA (|log2 FC| >3; P<0.001; Table I). It was identified that 19 miRNAs interacted with 93 IncRNAs (Table II).

**miRNA targets predicted by miRTarBase.** Based on the miRNAs described in Table I, miRNA-targeted mRNA were identified by searching the TargetScan, miRDB and miRTarBase databases. A total of 27 mRNAs were identified including titin, chordin like 1, hydroxycarboxylic acid receptor 2, transforming growth factor β receptor 2, A-kinase anchoring
protein 12, sprout RTK signaling antagonist 2, kelch like family member 40, sterile α motif domain containing 5, transcription elongation factor A like 7, KIT proto-oncogene receptor tyrosine kinase, cyclin B1, cell adhesion molecule L1 like, WAS protein family member 3, fibroblast growth factor (FGF) 2, SHc binding and spindle associated 1, serine rich and transmembrane domain containing 1, karyopherin α2 (KPNA2), neurotrophic receptor tyrosine kinase 2 (nTrK2), cyclin e2, secreted frizzled related protein 1 (SFrP1), par-6 family cell polarity regulator β, ELAV like RNA binding protein 2, FGF receptor 3, cadherin 2, C-C motif chemokine ligand 20, solute carrier family 1 member 1 and homeobox B5 (Table III).

**Construction of a ceRNA network in breast cancer.** To improve the understanding of the role of IncRNAs in breast cancer, a ceRNA network was constructed based on co-expressed IncRNAs-miRNAs and miRNAs-mRNAs. As demonstrated in Fig. 2, the ceRNA network was composed of 93 IncRNAs, 27 mRNAs and 19 miRNAs.

**ROC analysis of breast cancer-specific IncRNAs.** The 93 miRNAs targeting IncRNAs in breast cancer predicted by miRcode and starBase were selected. Then, Perl script was used to obtain the survival analysis data, and the clinical data file (IncRNA name, patient survival time, survival state) was combined with the survival data for univariate Cox analysis. In the univariate Cox analysis data, when the hazard ratio \( HR >1 \), this indicated that the higher the gene expression, the higher the risk of mortality, that is, the expression of the gene is contrary to survival. \( HR <1 \), indicated that the higher the expression of gene, the lower the risk of mortality, that is, the higher the survival rate. Survival-associated genes were then listed in order of ascending P-value, and the top 13 IncRNAs with \( P<0.05 \) were selected. IncRNAs were selected subsequent to entering the screening gene code in R software (version 3.4.3). Finally, the top 10 IncRNAs associated with breast cancer survival analysis were screened (Table IV; Fig. 3). These included: ADAM metallopeptidase with thrombospondin type 1 motif 9-antisense RNA 1 (ADA MTS9-AS1), AC061992.1, Wilms tumor 1 antisense RNA (WT1-AS), long intergenic non-protein coding RNA 536 (LINC00536), AL391421.1, SLIT-ROBO Rho GTPase activating protein 3 antisense RNA 2 (SRGAP3-AS2), Prader-Willi region non-protein coding RNA 1 (PWRN1), family with sequence similarity 230 member G (AC007731.1), chromosome 10 open reading frame 126 (C10orf126) and AL513123.1. Of these, the levels of ADA MTS9-AS1 and PWRN1 were downregulated in breast cancer tissues. The other 8 IncRNAs were upregulated in cancer tissues compared with normal tissues. A total of 6 IncRNAs exhibited high prognostic values for distinguishing tumor tissues from normal tissues, with AUC values of >0.99.

**Additional analysis for the selected IncRNAs expression.** GEPIA results (Fig. 4A) revealed upregulation of ADAMTS9-AS1 in kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP) and prostate adenocarcinoma (PRAD). As demonstrated in Fig. 4B, the results also identified that WT1-AS levels were significantly increased in acute myeloid leukemia, ovarian cancer (OV), uterine corpus endometrial carcinoma (UCEC) and uterine carcinosarcomas compared with in normal tissues. Notably,
Table II. DElncRNAs interacting with the 19 DEMiRNAs retrieved from the miCode database.

| lncRNA     | miRNAs                                      |
|------------|---------------------------------------------|
| AGAP11     | hsa-miR-141, hsa-miR-200a, hsa-miR-145, hsa-miR-182, hsa-miR-206, hsa-miR-204, hsa-miR-21, hsa-miR-375 |
| C2orf48    | hsa-miR-183, hsa-miR-204, hsa-miR-122       |
| SHANK2-AS3 | hsa-miR-96, hsa-miR-145, hsa-miR-187, hsa-miR-204, hsa-miR-122 |
| C20orf166-AS1 | hsa-miR-301b, hsa-miR-183, hsa-miR-429, hsa-miR-375 |
| C15orf54   | hsa-miR-301b, hsa-miR-182, hsa-miR-206, hsa-miR-429, hsa-miR-375 |
| AC127496.1 | hsa-miR-96, hsa-miR-182, hsa-miR-183, hsa-miR-204, hsa-miR-210, hsa-miR-122 |
| MIR7-3HG   | hsa-miR-145, hsa-miR-204                    |
| LINC00305  | hsa-miR-144, hsa-miR-204                    |
| C10orf91   | hsa-miR-429, hsa-miR-204, hsa-miR-122       |
| WT1-AS     | hsa-miR-96, hsa-miR-141, hsa-miR-200a, hsa-miR-145, hsa-miR-182, hsa-miR-206, hsa-miR-429, hsa-miR-375 |
| LINC00518  | hsa-miR-141, hsa-miR-200a, hsa-miR-145, hsa-miR-206, hsa-miR-204, hsa-miR-375 |
| LINC00221  | hsa-miR-301b, hsa-miR-96, hsa-miR-182, hsa-miR-204, hsa-miR-21 |
| TCL6       | hsa-miR-301b, hsa-miR-96, hsa-miR-137, hsa-miR-144, hsa-miR-145, hsa-miR-182, hsa-miR-183, hsa-miR-187, hsa-miR-206, hsa-miR-204, hsa-miR-210, hsa-miR-122, hsa-miR-375 |
| AC009093.1 | hsa-miR-183, hsa-miR-122                     |
| C10orf102  | hsa-miR-96, hsa-miR-145, hsa-miR-206, hsa-miR-429, hsa-miR-21 |
| AC135178.1 | hsa-miR-122                                   |
| AF241725.1 | hsa-miR-145                                   |
| MUC2       | hsa-miR-145, hsa-miR-182, hsa-miR-183, hsa-miR-184, hsa-miR-429, hsa-miR-210, hsa-miR-122 |
| RMRP       | hsa-miR-206, hsa-miR-122                      |
| C1orf137   | hsa-miR-204                                   |
| AL391421.1 | hsa-miR-137, hsa-miR-144                      |
| TDRG1      | hsa-miR-122                                   |
| C10orf126  | hsa-miR-141, hsa-miR-200a, hsa-miR-206, hsa-miR-375 |
| MUC19      | hsa-miR-301b, hsa-miR-96, hsa-miR-137, hsa-miR-144, hsa-miR-145, hsa-miR-182, hsa-miR-184, hsa-miR-187, hsa-miR-206, hsa-miR-429, hsa-miR-204, hsa-miR-122, hsa-miR-375 |
| UCA1       | hsa-miR-96, hsa-miR-182, hsa-miR-184, hsa-miR-206, hsa-miR-122 |
| LINC00488  | hsa-miR-96, hsa-miR-144, hsa-miR-206, hsa-miR-21, hsa-miR-122 |
| LINC00243  | hsa-miR-96, hsa-miR-145, hsa-miR-182, hsa-miR-206, hsa-miR-122, hsa-miR-375 |
| AL356479.1 | hsa-miR-429                                   |
| AL356310.1 | hsa-miR-206                                   |
| AC080037.1 | hsa-miR-429                                   |
| AL513123.1 | hsa-miR-141, hsa-miR-200a, hsa-miR-183, hsa-miR-183, hsa-miR-122 |
| LGALS8-AS1 | hsa-miR-122                                   |
| SMCRC2     | hsa-miR-145, hsa-miR-183, hsa-miR-204          |
| PHEX-AS1   | hsa-miR-301b, hsa-miR-96, hsa-miR-145, hsa-miR-182, hsa-miR-184, hsa-miR-187, hsa-miR-206, hsa-miR-429, hsa-miR-204, hsa-miR-122, hsa-miR-375 |
| LINC00466  | hsa-miR-96, hsa-miR-137, hsa-miR-141, hsa-miR-200a, hsa-miR-144, hsa-miR-183, hsa-miR-206, hsa-miR-429, hsa-miR-204, hsa-miR-21 |
| CHL1-AS2   | hsa-miR-183                                   |
| TSSC1-IT1  | hsa-miR-137, hsa-miR-141, hsa-miR-200a, hsa-miR-183 |
| LINC00337  | hsa-miR-145, hsa-miR-182, hsa-miR-375          |
| LINC00113  | hsa-miR-145                                   |
| LINC00351  | hsa-miR-21, hsa-miR-375                        |
| ADIPOQ-AS1 | hsa-miR-144, hsa-miR-145, hsa-miR-182, hsa-miR-183, hsa-miR-184, hsa-miR-122, hsa-miR-375 |
| NAALAD2-AS2 | hsa-miR-183, hsa-miR-206                      |
| LINC00355  | hsa-miR-141, hsa-miR-200a, hsa-miR-122         |
| LINC00392  | hsa-miR-183                                   |
Table II. Continued.

| IncRNA       | miRNAs                                      |
|--------------|---------------------------------------------|
| HOTAIR       | hsa-miR-301b, hsa-miR-206, hsa-miR-204, hsa-miR-21, hsa-miR-375 |
| SRGAP3-AS2   | hsa-miR-145, hsa-miR-206                    |
| HCG23        | hsa-miR-145                                  |
| LINC00200    | hsa-miR-183, hsa-miR-204                    |
| LINC00404    | hsa-miR-141, hsa-miR-200a                   |
| SACS-AS1     | hsa-miR-144, hsa-miR-187                    |
| EMX2OS       | hsa-miR-182, hsa-miR-183, hsa-miR-184, hsa-miR-210 |
| ATXN8OS      | hsa-miR-145, hsa-miR-183, hsa-miR-204, hsa-miR-210, hsa-miR-122 |
| LINC00210    | hsa-miR-96, hsa-miR-206, hsa-miR-204         |
| TLR8-AS1     | hsa-miR-182, hsa-miR-187, hsa-miR-206, hsa-miR-204 |
| LINC00460    | hsa-miR-206, hsa-miR-429                    |
| POU6F2-AS2   | hsa-miR-137, hsa-miR-187, hsa-miR-375        |
| BOK-AS1      | hsa-miR-184                                 |
| MAGI2-AS3    | hsa-miR-137, hsa-miR-141, hsa-miR-200a, hsa-miR-144, hsa-miR-145, hsa-miR-429, hsa-miR-204, hsa-miR-210, hsa-miR-122 |
| CHL1-AS1     | hsa-miR-137, hsa-miR-187                    |
| FNDC1-IT1    | hsa-miR-144                                 |
| DSCAM-AS1    | hsa-miR-137, hsa-miR-141, hsa-miR-200a, hsa-miR-122 |
| LINC00484    | hsa-miR-141, hsa-miR-200a, hsa-miR-187, hsa-miR-206, hsa-miR-122 |
| ARHGEF7-AS2  | hsa-miR-187, hsa-miR-210, hsa-miR-122, hsa-miR-375 |
| RBMS3-AS3    | hsa-miR-96, hsa-miR-182, hsa-miR-204         |
| LINC00445    | hsa-miR-375                                 |
| NDP-AS1      | hsa-miR-145, hsa-miR-206, hsa-miR-122       |
| AC009121.1   | hsa-miR-141, hsa-miR-200a                   |
| CLRN1-AS1    | hsa-miR-137, hsa-miR-206, hsa-miR-429, hsa-miR-204 |
| PDZRN3-AS1   | hsa-miR-141, hsa-miR-200a                   |
| AC080129.1   | hsa-miR-122                                 |
| AL109754.1   | hsa-miR-122                                 |
| MME-AS1      | hsa-miR-100, hsa-miR-182, hsa-miR-429       |
| LSAMP-AS1    | hsa-miR-183, hsa-miR-375                    |
| ADAMTS9-AS1  | hsa-miR-301b, hsa-miR-96, hsa-miR-144, hsa-miR-145, hsa-miR-182, hsa-miR-21 |
| SYNR-AS1     | hsa-miR-96, hsa-miR-182, hsa-miR-375         |
| ADAMTS9-AS2  | hsa-miR-301b, hsa-miR-96, hsa-miR-137, hsa-miR-141, hsa-miR-200a, hsa-miR-144, hsa-miR-145 hsa-miR-182, hsa-miR-183, hsa-miR-184, hsa-miR-204, hsa-miR-210, hsa-miR-122, hsa-miR-375 |
| AC007731.1   | hsa-miR-183                                 |
| AL139002.1   | hsa-miR-301b, hsa-miR-210                   |
| PEX5L-AS1    | hsa-miR-141, hsa-miR-200a                   |
| AC061992.1   | hsa-miR-301b, hsa-miR-204, hsa-miR-122      |
| AP000553.1   | hsa-miR-122                                 |
| LINC00461    | hsa-miR-96, hsa-miR-137, hsa-miR-141, hsa-miR-200a, hsa-miR-144, hsa-miR-145, hsa-miR-204, hsa-miR-210, hsa-miR-122 |
| ALDH1L1-AS2  | hsa-miR-301b, hsa-miR-145, hsa-miR-210      |
| MAST4-IT1    | hsa-miR-204                                 |
| LINC00536    | hsa-miR-96, hsa-miR-137, hsa-miR-182, hsa-miR-204 |
| OPCML-IT1    | hsa-miR-184, hsa-miR-375                    |
| C8orf49      | hsa-miR-301b, hsa-miR-100, hsa-miR-184, hsa-miR-429, hsa-miR-122, hsa-miR-375 |
| AL589642.1   | hsa-miR-141, hsa-miR-200a, hsa-miR-145, hsa-miR-182, hsa-miR-183, hsa-miR-429, hsa-miR-204, hsa-miR-210, hsa-miR-122 |
| AC040173.1   | hsa-miR-96, hsa-miR-144, hsa-miR-183, hsa-miR-429 |
| LINC00524    | hsa-miR-204                                 |
the increased expression of LINC00536 was demonstrated in breast invasive carcinoma, while a decreased expression was demonstrated in testicular germ cell tumors (TGCT) (Fig. 4C). Furthermore, increased SRGAP3-AS2 expression was identified in adenoid cystic carcinoma, cholangiocarcinoma, OV and UCEC, while significant downregulation of SRGAP3-AS2 was observed in lung adenocarcinoma and lung squamous cell carcinoma (LUSC) (Fig. 4D). PWRN1 was observed to be downregulated in thyroid carcinoma and upregulated in TGCT, pheochromocytoma and paraganglioma (Fig. 4E). A decreased expression of C10orf126 was identified in KIRC and liver hepatocellular carcinoma (Fig. 4F).

Survival analysis. The top 10 IncRNAs were selected using ROC analysis. As demonstrated in Fig. 5, the expression levels of 5 DElncRNAs, including ADAMTS9-AS1, AL513123.1, C10orf126, LINC00536 and WT1-AS, were associated with
Table IV. Top 10 aberrantly expressed lncRNAs and AUC values in breast cancer.

| lncRNA       | Gene ID           | Expression change | log2 FC (T/N) | AUC    | P-value |
|--------------|-------------------|-------------------|--------------|--------|---------|
| ADAMTS9-AS1  | ENSG00000241158   | Downregulated     | -2.746216755| 0.888  | <0.05   |
| AC061992.1   | ENSG00000266970   | Upregulated       | 2.140133288 | 0.867  | <0.05   |
| WT1-AS       | ENSG00000183242   | Upregulated       | 4.901360082 | 1.102  | <0.05   |
| LINC00536    | ENSG00000249917   | Upregulated       | 2.721025285 | 1.110  | <0.05   |
| AL391421.1   | ENSG00000204049   | Upregulated       | 3.149701497 | 0.940  | <0.05   |
| SRGAP3-AS2   | ENSG00000228723   | Upregulated       | 4.608297544 | 0.889  | <0.05   |
| PWRN1        | ENSG00000259905   | Downregulated     | -2.274243394| 1.243  | <0.05   |
| AC007751.1   | ENSG00000188280   | Upregulated       | 2.840014701 | 1.182  | <0.05   |
| C10orf126    | ENSG000002043655   | Upregulated      | 3.688082405 | 1.183  | <0.05   |
| AL513123.1   | ENSG00000236347   | Upregulated       | 2.905568621 | 1.194  | <0.05   |

lncRNA; long-non-coding RNA; FC, fold change; T/N, tumor/normal; AUC, area under the curve.

Figure 2. Competing endogenous RNA network in breast cancer. The red nodes represent increased level of expression, while the green nodes represent decreased level of expression. Rectangles represent miRNAs, diamonds represent lncRNAs and ellipses represent mRNAs. The gray edges denote lncRNA-miRNA-mRNA interactions. lncRNA, long non-coding RNA; miR, microRNA.
overall survival (P<0.05). A total of 2 miRNAs (hsa-miR-204 and hsa-miR-301b) were associated with overall survival (P<0.05; Fig. 6). In addition, 3 mRNA (KPNA2, NTRK2 and SFRP1) were associated with overall survival (P<0.05; Fig. 7).

Discussion

In previous years, much attention has been paid to breast cancer pathogenesis. Nevertheless, clinical outcomes remain highly heterogeneous. Nik-Zainal et al (26) analyzed genome-wide data from 560 patients with breast cancer and identified that 93 oncogenes involved in the encoded proteins may cause breast cancer. Prognosis is an important indicator of disease treatment. Bioinformatics methods have been used to investigate the prognostic significance of lncRNAs in breast cancer. Several studies has also indicated that the aberrant expression of lncRNAs contributes to the development of different types of cancer: Xiao et al (27) revealed that MALAT1 may be used as a ceRNA to promote the expression of ZEB2 by expanding miR-200s, and may be a therapeutic target for kidney cancer; Wang et al (28) demonstrated that UCA1 and AATBC are not only included in the ceRNA network, but are also associated with overall survival in muscle-invasive bladder cancer; Chen et al (25) suggested that the top 10 aberrantly expressed lncRNAs identified in their study served important roles in LUSC through an lncRNA-mRNA network; and Li et al (29) described a novel gastric cancer-specific ceRNA network including 11 lncRNAs, 9 miRNAs and 41 mRNAs.

ceRNA transcripts may be used to competitively bind the same MRE with miRNAs, thereby relieving or mitigating the inhibitory effects of miRNAs on target genes. MREs are the basis for miRNA function, and other non-coding genes also interact with miRNAs through MREs. Recently, Zhou et al (30) used a Pearson's correlation analysis of miRNA-gene pairs to construct breast cancer-specific ceRNA networks. Regulation of phosphatase and tensin homolog expression in a 3' untranslated region-and miRNA-dependent manner in breast cancer has been identified in another study (31). Furthermore, ceRNA crosstalk may be associated with regulatory mechanisms in breast cancer. Shen et al (32) and Yang et al (33) demonstrated that FAM83H-AS1 appeared to be a novel prognostic biomarker in breast cancer.
Cheang et al (34) revealed that patients with breast cancer with high HER-2 protein expression levels and Ki67 index exhibited poor recurrence and disease-specific survival rates. Yuan et al (35) identified that the ceRNA crosstalk network in triple negative breast cancer included the differential expression of 22 hub mRNAs, 11 miRNAs and 14 lncRNAs. The present study obtained lncRNA, mRNA and miRNA data from TCGA to construct breast cancer-associated ceRNA networks, studied the regulatory mechanisms of lncRNA as ceRNAs in the progression of cancer, and also investigated their potential as prognostic biomarkers and therapeutic targets in breast cancer. A ceRNA network with differential expression of 27 mRNAs, 19 miRNAs and 93 lncRNAs was constructed.

The present study investigated the lncRNA expression profiles of a large cohort of patients in TCGA. A total of 93 DElncRNAs were identified in breast cancer samples compared with the normal samples. Among these, the top 10 lncRNA were significantly associated with overall survival. Furthermore, it was observed that the lncRNAs ADAMTS9-AS1, AL513123.1, C10orf126, LINC00536 and WT1-AS were included in the ceRNA network. Therefore, we hypothesized that these lncRNAs may serve significant roles in the pathogenesis and prognosis of breast cancer. Wang et al (36) revealed that antisense lncRNA ADAMTS9-AS1 was associated with the nearby coding gene ADAMTS9, which was involved in ovarian cancer progression. Li et al (37) identified that the lncRNA...
Figure 5. Kaplan-Meier survival curves for 5 differentially expressed long non-coding RNAs associated with overall survival in breast cancer. (A) ADAMTS9-AS1 ($p=5.52\times10^{-103}$), (B) AL513123.1 ($p=1.17\times10^{-08}$), (C) C10orf126 ($p=4.35\times10^{-11}$), (D) LINC00536 ($p=5.03\times10^{-32}$) and (E) WT1-AS ($p=2.11\times10^{-37}$). The x-axis represents overall survival time (years) and the y-axis represents survival function. ADAMTS9-AS1, ADAM metallopeptidase with thrombospondin type 1 motif 9-antisense RNA 1; C10orf126, chromosome 10 open reading frame 126; LINC00536, long intergenic non-protein coding RNA 536; WT1-AS, Wilms tumor 1 antisense RNA.
Figure 6. Kaplan-Meier survival curves for two differentially expressed miRNAs associated with overall survival in breast cancer. (A) hsa-miR-204 (P=2.02x10^{-61}) and (B) hsa-miR-301b (P=7.76x10^{-40}). The x-axis represents overall survival time (years) and the y-axis represents survival function. hsa, *Homo sapiens*; miRNA, microRNA.

Figure 7. Kaplan-Meier survival curves for 3 mRNA associated with overall survival in breast cancer. (A) KPNA2 (P=8.08x10^{-81}), (B) NTRK2 (P=3.89x10^{-70}) and (C) SFRP1 (P=1.76x10^{-46}). The x-axis represents overall survival time (years) and the y-axis represents survival function. KPNA2, karyopherin α 2; NTRK2, neurotrophic receptor tyrosine kinase 2; SFRP1, secreted frizzled related protein 1.
ADAMTS9-AS1 served as a novel prognostic biomarker for clinical application in esophageal squamous cell carcinoma. Crippa et al. (38) hypothesized that LINC00536 disruption may have contributed to the onset of the clinical trichorhinophalangeal syndrome-like phenotype. Wang et al. (39) suggested that WT1 was overexpressed in human hepatocellular carcinoma tissues, and was negatively correlated with overall survival in patients with hepatocellular carcinoma.

In the present study, a novel ceRNA network was constructed to identify the associations between miRNAs and IncRNAs. miRNAs also serve important roles in cell differentiation, biological development and disease progression. Several previous data have demonstrated that the mutual regulation between IncRNAs and miRNAs and their downstream target genes are closely associated with the occurrence and development of cancer (40-42). Zhang et al. (43) used reverse transcription quantitative polymerase chain reaction and in situ hybridization for breast cancer tissue samples and cell lines. Cell phenotype experiments were performed to verify that IncRNA-GAS5 exon 4 bound miR-21 and inhibited the occurrence and development of breast cancer cells (43). The present study demonstrated that hsa-miR-204 was down-regulated and hsa-miR-301b was upregulated in patients with breast cancer compared with healthy controls, and was associated with overall survival. Yuan et al. (44) revealed that miR-204 suppressed the proliferation and metastasis of gastric cancer cells. Todorova et al. (45) identified that miR-204 promoted prostate cancer-associated androgen-responsive genes and androgen receptor (AR) target genes and AR co-regulated molecules. Abmutalib et al. (46) hypothesized that hsa-miR-301b may be involved in regulating lymph node metastasis in papillary thyroid carcinoma via interactions with hepatic leukemia factor, hypoxia-inducible factor and REL/nuclear factor kappa-light-chain-enhancer of activated B cells. Geng et al. (47) demonstrated that overexpression of hsa-miR-301b significantly affected the cell cycle of human lung adenocarcinoma A549 cells. ceRNAs have been implicated in several biological processes, and abnormalities in the ceRNA network may lead to tumorigenesis.

In summary, the present study investigated IncRNA-mediated ceRNA interactions using miRNA, IncRNA and mRNA expression profiles in cancer and normal tissues. The results suggested that cancer-specific IncRNAs in breast cancer may be involved in the regulation of a complex ceRNA network. These data may provide novel insights into the clinical significance and regulatory mechanisms of IncRNA-mediated ceRNA networks, and identify novel IncRNAs as potential prognostic biomarkers and therapeutic targets for the diagnosis and treatment of breast cancer.

Acknowledgements
Not applicable.

Funding
The present study was supported by grants from National Natural Science Foundation of China (grant no. 81660324) and the Key discipline Construction of the 13th Five-Year Plan in Xinjiang, China - Plateau Discipline Project.

Availability of data and materials
All data analyzed during this study are included in this published article.

Authors’ contributions
SMF and LLL conceived and designed the manuscript. TT and ZA analyzed the data. TT wrote the paper. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References
1. Breastcancer.org: US Breast Cancer Statistics: http://www.breastcancer.org/symptoms/understand_bc/statistics. breast cancer org. Accessed February 18, 2019.
2. Linos E, Spanos D, Rosner BA, Linos K, Hesketh T, Qu JD, Gao YT, Zheng W and Colditz GA: Effects of reproductive and demographic changes on breast cancer incidence in China: A modeling analysis. J Natl Cancer Inst 100: 1352-1360, 2008.
3. Higgins MJ and Baselga J: Targeted therapies for breast cancer. J Clin Invest 121: 3797-3803, 2011.
4. Gyorffy B, Hatzis C, Sanft T, Hofstatter E, Aktas B and PusztaI: Multigene prognostic tests in breast cancer: Past, present, future. Breast Cancer Res 17: 11, 2015.
5. Esteller M: Non-coding RNAs in human disease. Nat Rev Genet 12: 861-874, 2011.
6. Ma L, Bajic VB and Zhang Z: On the classification of long non-coding RNAs. RNA Biol 10: 925-933, 2013.
7. Yuan SX, Yang F, Yang Y, Tao QF, Zhang J, Huang G, Yang Y, Wang RY, Yang S, Huo XS, et al. Long noncoding RNA associated with microvascular invasion in hepatocellular carcinoma promotes angiogenesis and serves as a predictor for hepatocellular carcinoma patients’ poor recurrence-free survival after hepatectomy. Hepatology 56: 2231-2241, 2012.
8. Hu X, Feng Y, Zhang D, Zhao SD, Hu Z, Greshock J, Zhang Y, Yang L, Zhong X, Wang LP, et al: A functional genomic approach identifies FAL1 as an oncogenic long noncoding RNA that associates with BMMI and represses p21 expression in cancer. Cancer Cell 26: 344-357, 2014.
9. Feng S, Zhang J, Su W, Bai S, Xiao L, Chen X, Lin J, Reddy RM, Chang AC, Beer DG and Chen Q: Overexpression of LINC00152 correlates with poor patient survival and knobdown impairs cell proliferation in lung cancer. Sci Rep 7: 2982, 2017.
10. Sang H, Liu H, Xiong P and Zhu M: Long non-coding RNA functions in lung cancer. Tumor Biol 36: 4027-4037, 2015.
11. Zhong Y, Du Y, Yang X, Mo Y, Fan C, Xiong F, Ren D, Ye X, Li C, Wang Y, et al: Circular RNAs function as ceRNAs to regulate and control human cancer progression. Mol Cancer 17: 79, 2018.
12. Wang Y, Lu T, Wang Q, Liu J and Jiao W: Circular RNAs: Crucial regulators in the human body (Review). Oncol Rep 40: 3119-3135, 2018.
13. Wang Y, Mo Y, Gong Z, Yang X, Yang M, Zhang S, Xiong F, Xiang B, Zhou M, Liao Q, et al: Circular RNAs in human cancer. Mol Cancer 16: 25, 2017.
14. Salmena L, Poliseno L, Tay Y, Kats L and Pandolfo M: A ceRNA hypothesis: The Rosetta Stone of a hidden RNA language? Cell 146: 353-358, 2011.
