

RESEARCH

Open Access



Oncogenic long noncoding RNA landscape in breast cancer

Shouping Xu¹, Dejia Kong¹, Qianlin Chen¹, Yanyan Ping² and Da Pang^{1,3*}

Abstract

Background: Few long noncoding RNAs (lncRNAs) that act as oncogenic genes in breast cancer have been identified.

Methods: Oncogenic lncRNAs associated with tumorigenesis and worse survival outcomes were examined and validated in Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA), respectively. Then, the potential biological functions and expression regulation of these lncRNAs were studied via bioinformatics and genome data analysis. Moreover, progressive breast cancer subtype-specific lncRNAs were investigated via high-throughput sequencing in our cohort and TCGA validation. To elucidate the mechanisms of the regulation of these lncRNAs, genomic alterations from the TCGA, Broad, Sanger and BCCRC data, as well as epigenetic modifications from GEO data, were then applied and examined to meet this objective. Finally, cell proliferation assays, flow cytometry analyses and TUNEL assays were applied to validate the oncogenic roles of these lncRNAs in vitro.

Results: A cluster of oncogenic lncRNAs that was upregulated in breast cancer tissue and was associated with worse survival outcomes was identified. These oncogenic lncRNAs are involved in regulating immune system activation and the TGF-beta and Jak-STAT signalling pathways. Moreover, TINCR, LINC00511, and PPP1R26-AS1 were identified as subtype-specific lncRNAs associated with HER-2, triple-negative and luminal B subtypes of breast cancer, respectively. The up-regulation of these oncogenic lncRNAs is mainly caused by gene amplification in the genome in breast cancer and other solid tumours. Finally, the knockdown of TINCR, DSCAM-AS1 or HOTAIR inhibited breast cancer cell proliferation, increased apoptosis and inhibited cell cycle progression in vitro.

Conclusions: These findings enhance the landscape of known oncogenic lncRNAs in breast cancer and provide insights into their roles. This understanding may potentially aid in the comprehensive management of breast cancer.

Keywords: lncRNAs, Prognosis, Tumorigenesis, Breast cancer, TINCR

Background

Although great advances have been made in the management of breast cancer over the past decade, patient outcomes still merit consideration due to the high rate of tumour-specific death [1, 2]. The molecular mechanisms of tumorigenesis are still unclear in breast cancer. Thus, identifying new genes related to tumorigenesis and patient prognosis, as well as elucidating the molecular mechanisms underlying these oncogenic processes, are urgently required.

It is certified that less than 2% of the genome encodes proteins, but at least 75% are transcribed into noncoding RNAs [3]. Generally, long noncoding RNAs (lncRNAs) are defined as transcripts that are longer than 200 nucleotides and have no protein-coding capacity [4]. There are more than 120,000 transcripts annotated as lncRNAs in the human genome as released by the Encyclopedia of DNA Elements (ENCODE) Project Consortium (release 28) [5]. lncRNAs share common features with mRNAs: many of them are transcribed by RNA polymerase II and undergo 5'-capping, polyadenylation and splicing [3]. In addition, the histone profiles of lncRNAs are similar to protein-coding genes for the active histone markers H3K4me2, H3K4me3, H3K9ac and H3K27ac [6]. On the other hand, lncRNAs have several distinct features that distinguish them from protein-coding mRNAs.

* Correspondence: pangda@ems.hrbmu.edu.cn

¹Department of Breast Surgery, Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin 150081, China

³Heilongjiang Academy of Medical Sciences, 157 Baojian Road, Harbin 150086, China

Full list of author information is available at the end of the article



Generally, lncRNAs lack coding potential due to fewer exons [7]. However, there are some exceptions with the development of lncRNA investigations. LINC00961 regulates mTORC1 and muscle regeneration by encoding SPAR polypeptide [8]. Functional polypeptides of myoregulin encoded by LINC00948 and Dworf encoded by NONMMUG026737 have been reported to modulate SERCA pump activity [9, 10]. lncRNA CRNDE encodes a nuclear peptide, CRNDEP, which is overexpressed in highly proliferating tissues and is involved in cell turnover [11]. Moreover, many investigations have shown that lncRNAs are of relatively lower expression levels than protein-coding genes; however, they exhibit a more cell type-specific pattern [12–14]. Unlike mRNAs, most lncRNAs are localized to the nucleus, while mRNA is in the cytoplasm [6]. Finally, lncRNAs belong to evolutionary conserved families evolving faster than mRNAs, where sequence similarity is likely to be preserved mainly in regions with secondary structure formation [6, 12].

lncRNAs play an important role in regulating gene expression at various levels, including alternative splicing, regulation of protein activity, and alteration of protein localization, as well as chromatin modification, transcription, and posttranscriptional processing [3, 15–21]. The mechanisms by which lncRNAs contribute to the regulatory networks that underpin cancer development are diverse [22, 23]. Unlike microRNA or piwi-interacting RNA (piRNA), lncRNAs drive many important cancer phenotypes through their interactions with other cellular macromolecules such as DNA, protein, and RNA [22–26]. Regarding their role in malignancy, lncRNAs are associated with immortal-associated characteristics, including cell cycle regulation, survival, immune response or pluripotency in cancer cells [27–32]. Several common lncRNAs have been investigated in cancers. For example, HOTAIR promotes cancer metastasis by reprogramming the chromatin state in a manner dependent on PRC2 [33]. SchLAP1 contributes to the development of lethal cancer by antagonizing the tumour-suppressive functions of the SWI/SNF complex [34]. Another lncRNA, SAMMSON, increases the pro-oncogenic function via interacting with p32 in melanoma [35]. Despite growing knowledge regarding the molecular mechanisms of lncRNA functions in malignancy, the modes of action of most lncRNAs in breast cancer remain unclear. Aberrant expression of oncogenic lncRNAs may confer capacities for tumour initiation, growth, and metastasis in breast cancer cells, thus leading to a worse prognosis for patients. Similarly, applying antisense oligonucleotides can inhibit the expression of the candidate therapeutic target Malat1 successfully in an MMTV-PyMT mouse model with breast cancer [36]. Thus, unravelling the landscape of these oncogenic lncRNAs in breast cancer is essential and urgent.

The aim of this study was to identify the predictive capability of oncogenic lncRNAs for the tumourigenesis and prognosis of breast cancer. In this study, 1088 breast cancer patients from GEO data were selected and analysed to investigate the lncRNAs associated with tumourigenesis and outcomes. The aberrant long noncoding RNAs were then validated using TCGA data and high-throughput sequencing in our cohort. Genome-wide in silico analysis and in vitro assays revealed the potential biological functions of these lncRNAs, including their mechanisms of regulation and roles in tumourigenesis. Moreover, identification and validation of breast cancer subtype-specific lncRNAs and the determination of their association with clinical outcomes were also investigated.

Methods

Public data access and analysis

GEO data (GSE21653, GSE31448, GSE10810, GSE29431, GSE23177, GSE42568, and GSE48391) were downloaded and processed (<http://www.ncbi.nlm.nih.gov/geo/>). The genome-wide lncRNA expression profiles for breast cancer, renal cancer, and lung cancer were downloaded from TCGA (<https://tcga-data.nci.nih.gov/>). For the microarray analysis, we adjusted the signal values for low-abundance genes. A signal value lower than log3 was set to log3. Moreover, the invariant genes (i.e., same expression value across all samples) and low-variation genes were filtered. Genes that were detected in less than 50% of the profiled samples were also filtered. The SAM method was applied, and we implemented a series of steps to estimate the significance of difference and false discovery rate for each filtered gene as previously described [37]. A method to estimate the significance of difference and false discovery rate for each filtered gene was implemented as follows:

- (1) Calculate the exchange factor s_0 . First, we calculated the standard deviation of gene expression for all genes s_i , denoted s^α , as the α percentile for s_i . The relative difference in gene expression (d Score) at the α percentile is calculated as $d_i^\alpha = r_i / (s_i + s^\alpha)$, where r_i is the fold change of gene expression for gene i between two conditions. Next, each interval of the percentile value $q_1 < q_2 < \dots < q_{100}$ of the s_i and the mean absolute deviation of $d_i^\alpha v_j = mad \left\{ d_i^\alpha = r_i / (s_i + s^\alpha) \mid s_i \in [q_j, q_{j+1}] \right\}$ were calculated. Finally, we selected the α (denote as $\hat{\alpha}$) to make the CV (coefficient of variation) of the v_j achieve a minimum, and set the exchange factor $s_0 = s^{\hat{\alpha}}$.

- (2) Calculate the statistical value (d Score) for every gene i :

$$d_i = r_i / (s_i + s_0),$$

(3) Calculate the order statistic according to d_i :

$$d_{(1)} \leq d_{(2)} \leq \dots \leq d_{(i)} \leq \dots \leq d_{(p)},$$

(4) Perform 1000 permutations to estimate the expected distribution of the d Score. We denote the estimated statistical values:

$$d_{(1)}^* \leq d_{(2)}^* \leq \dots \leq d_{(i)}^* \leq \dots \leq d_{(p)}^*,$$

(5) Obtain the order statistic value under the permutation:

$$\bar{d}_{(i)} = \frac{\sum_{i=1}^{1000} d_{(i)}^*}{1000},$$

(6) By calculating the maximum distance between the order statistic $d_{(i)}$ and the expected order statistic $\bar{d}_{(i)}$, we would construct a series of reject regions for the q-value.

(7) For a fixed delta value, we computed the difference $\Delta_{(i)} = d_{(i)} - \bar{d}_{(i)}$ and found the nearest $\Delta_{(i)}$ for gene i. The cut-up was marked as $\min\{\Delta_{(i)} \geq \text{delta}\}$ for the positive gene and the cut-down as $\max\{\Delta_{(i)} \leq -\text{delta}\}$ as for negative gene. Next, we called these genes significantly positive genes, whose difference was larger than the cut-up value, and significantly negative genes, whose difference was smaller than the cut-down value.

(8) Estimate the false discovery rate (FDR):

$$FDR = \frac{V_{(p)}}{R_{(p)}},$$

where $V_{(p)}$ is the number of positive genes called in the 1000 permutations, and $R_{(p)}$ is the median of the number of false-positive genes for the above permutations.

(9) Obtain the q-value for gene i by selecting the minimum of the FDR for the 50 delta values determined in step (7).

Welch's t-test (unequal variances) and analysis of variance were also applied for two-group and multiple-group analyses, respectively. For multiple-comparison analysis, the q-value was used to control the false discovery rate. Clustering heatmap analysis was performed according to previous study [38].

Guilt-by-association analysis

To identify a list of lncRNAs positively and negatively correlated with the target genes, data from the TCGA were evaluated to compute a pairwise Pearson correlation between the expression of the target lncRNA and all the genes. Only associated genes with an absolute $r \geq 0.4$ and a significant correlation ($P < 0.05$) were retained. Gene ontology term enrichment (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were analysed using DAVID as previously described [39, 40].

Patients and clinical samples

Thirty invasive breast cancer and adjacent non-cancerous tissues were obtained from patients who

had not received either chemotherapy or radiotherapy and who were treated at the Department of Breast Surgery at Harbin Medical University Cancer Hospital in 2014. This study was approved by the research ethics committee of the Harbin Medical University Cancer Hospital. Written informed consent was obtained from all the patients who participated in the study. The clinicopathological characteristics of the patients are presented in Additional file 1: Table S1.

Library preparation for lncRNA sequencing

A total of 3 μg of RNA per sample was used for downstream RNA sample preparations. Ribosomal RNA was removed using the Ribo-Zero™ Gold kit (Epicentre, Wisconsin, USA). Subsequently, sequencing libraries were generated according to the manufacturer's recommendations. The libraries were sequenced on an Illumina HiSeq 2500 platform, and 100-bp paired-end reads were generated. Raw sequencing and processed RNA-Seq data of this study have been deposited to the NCBI Gene Expression Omnibus database under accession number GSE71651 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=obcxosaur&xppwx&acc=GSE71651>).

Cell culture experiments

MDA-MB-453 and MCF-7 cells were cultured in DMEM (Invitrogen, Carlsbad, CA) containing 10% foetal bovine serum and 100 units/ml penicillin/streptomycin at 37 °C in an atmosphere containing 5% CO₂. UACC812 and T47D cells were cultured in 1640 medium (Invitrogen, Carlsbad, CA) containing 10% foetal bovine serum and 100 units/ml penicillin/streptomycin at 37 °C in an atmosphere containing 5% CO₂. All the cell lines were obtained from the Chinese Type Culture Collection, Chinese Academy of Sciences. Cells were used during their logarithmic growth phase.

Small interfering RNA (siRNA) and qRT-PCR

Breast cells were grown in complete medium before transfection with siRNAs using Lipofectamine 2000 (Life Technologies, 11,668–019) according to the manufacturer's protocol. Two siRNAs each were designed to target TINCR (siR_1: 5'-GCAUGAAGUAGCAGGUAUUUU-3' and siR_2: 5'-GAUCCCGAGUGAGUCAGAAUU-3'), HOTAIR (siR_1: 5'-CCACAUGAACGCCAGAGAUU-3' and siR_2: 5'-GAACGGGAGUACAGAGAUU-3') and DSCAM-AS1 (siR_1: 5'-ACUCAUCAUGUACCCAUUUCUUA-3' and siR_2: 5'-CCUCCUCCAACUGCCAUUUAUUUAU-3'). qRT-PCR analysis was performed using the SYBR-Green method, and the specific sequences of the primers used were as follows: 5'-TGTGGCCCAAACCTCAGGGATACAT-3' (forward) and 5'-AGATGACAGTGGCTGGAGTTGTCA-

3'(reverse) for TINCR, 5'-GTCCCTAATATCCCGGAGGT-3' (forward) and 5'-GCAGGCTTCTAAATCCGTC-3' (reverse) for HOTAIR; 5'-GATCCTTGTGGTCTCACTCC-3'(forward) and 5'-ATGCCTATGTGGGTGATTGG-3'(reverse) for DSCAM-AS1; and 5'-TTTGATGGTGACCTGGGAAT-3' (forward) and 5'-GAACATCTGCTGGTTCACA-3' (reverse) for ERBB2; 5'-ACCACAGTCCATGCCATCAC-3' (forward) and 5'-TCCACCCTGTTGCTGTA-3' (reverse) for GAPDH. Primers for MiR-125b and U6 were obtained as previously described [41]. Quantitative normalization of target cDNA was performed for each sample using GAPDH/U6 expression as an internal control. The relative levels of TINCR, HOTAIR, DSCAM-AS1, ERBB2, MiR-125b vs. GAPDH/U6 were determined by the comparative CT ($2^{-\Delta\Delta CT}$) method.

Cell proliferation assays

Cell proliferation assays were performed using the Cell Counting Kit-8 according to the manufacturer's instructions (Beyotime, Shanghai, China). Briefly, 2×10^3 cells were seeded in a 96-well plate. Cell proliferation was assessed at 24, 48, and 72 h. After the addition of 20 μ l of WST-1 reagents per well, cultures were incubated for 2 h, and the absorbance was measured at 450 nm using a microplate reader (BioTek, VT, United States).

Flow cytometry

An Annexin-PE Apoptosis detection kit (BD Biosciences, San Jose, CA) was used to examine cell apoptosis according to the manufacturer's instructions. Briefly, cells were washed twice in cold PBS and then harvested and resuspended in $1 \times$ binding buffer. Next, 100 μ l of the cell solution (1×10^5 cells) was transferred to a 5-ml culture tube, and 5 μ l of annexin V-PE and 5 μ l of 7-AAD were added into culture tube. The cells were gently vortexed and incubated for 15 min at RT (25 °C) in the dark. Next, 400 μ l of $1 \times$ binding buffer was added to each tube, and apoptosis analysis was performed using a FACScan instrument (Becton Dickinson, Mountain View, CA, USA). For cell cycle analysis, the CycleTEST™ PLUS DNA Reagent Kit (BD, Cat No.340242) was used, and the experiment was performed as previously described [42].

TUNEL assays

Apoptosis-induced DNA fragmentation was performed using the transfers-mediated deoxyuridine triphosphate (dUTP)-digoxigenin nick end-labelling (TUNEL) assay. UACC812 and MDA-MB-453 cells were plated in 24-well flat-bottomed plates at a density of 1×10^5 cells per well, and cells were fixed in 4% (*w/v*) paraformaldehyde at 4 °C for 25 min. TUNEL staining was examined using the in situ cell death detection kit (Roche), and the nuclei were stained with DAPI for 10 min according to the

manufacturer's instructions. The numbers of TUNEL-positive cells were captured with a fluorescence microscope (Olympus), and the ratio of apoptosis cells was determined with ImagePro Plus software.

Statistical analyses

The times of OS and RFS were calculated as the time from surgery until the occurrence of death or relapse, respectively. The expression of lncRNA was dichotomized using a study-specific median expression as the cut-off to define "high value" at or above the median versus "low value" below the median. The differences between the groups in our in vitro experiments were analysed using Student's t-test. Spearman correlation coefficients were calculated for correlation analysis. All the experiments were performed in triplicate, and SPSS 16.0 software (SPSS, Chicago, IL) was used for statistical analysis. All statistical tests were two-sided, and $P < 0.05$ was considered to be statistically significant.

Results

Aberrant gene expression landscape in breast cancer

Two cohorts containing seven datasets, including 1088 breast cancer patients, were selected from GEO data and were analysed in this study. In cohort I (GSE21653 and GSE31448), 137 lncRNAs and 8914 coding genes showed significantly aberrant expression between breast cancer and normal tissue ($P < 0.05$); however, in cohort II (GSE10810, GSE29431, GSE23177, GSE42568, and GSE48391), 164 lncRNAs and 9685 coding genes were differentially expressed in breast cancer and normal tissues ($P < 0.05$) (Fig. 1a, b and Additional file 2: Table S2 and Additional file 3: Table S3). Next, the overlapping genes were compared across these two cohorts. There were 30 and 25 lncRNAs upregulated and downregulated, respectively, in both of these cohorts ($P < 0.05$) (Fig. 1c and Additional file 2: Table S2). For the coding genes, there were 1712 and 1269 mRNAs upregulated and downregulated, respectively ($P < 0.05$) (Fig. 1d and Additional file 3: Table S3). To further confirm the result in a third independent cohort, the aberrant lncRNAs were validated using TCGA data. As expected, all except one lncRNA (WASIRS1) showed similar expression profiles as those from the in silico analysis (Fig. 1e, f).

Potential biological functions of aberrantly expressed lncRNAs in breast cancer

To explore the possible functions of these lncRNAs, guilt-by-association analyses were used, as described in the Methods section. GO and KEGG pathway analysis were also investigated for the identified lncRNAs. GOTERM_BP_DIRECT was selected to investigate the related biological process of enriched genes. The main related functions of the upregulated lncRNAs were

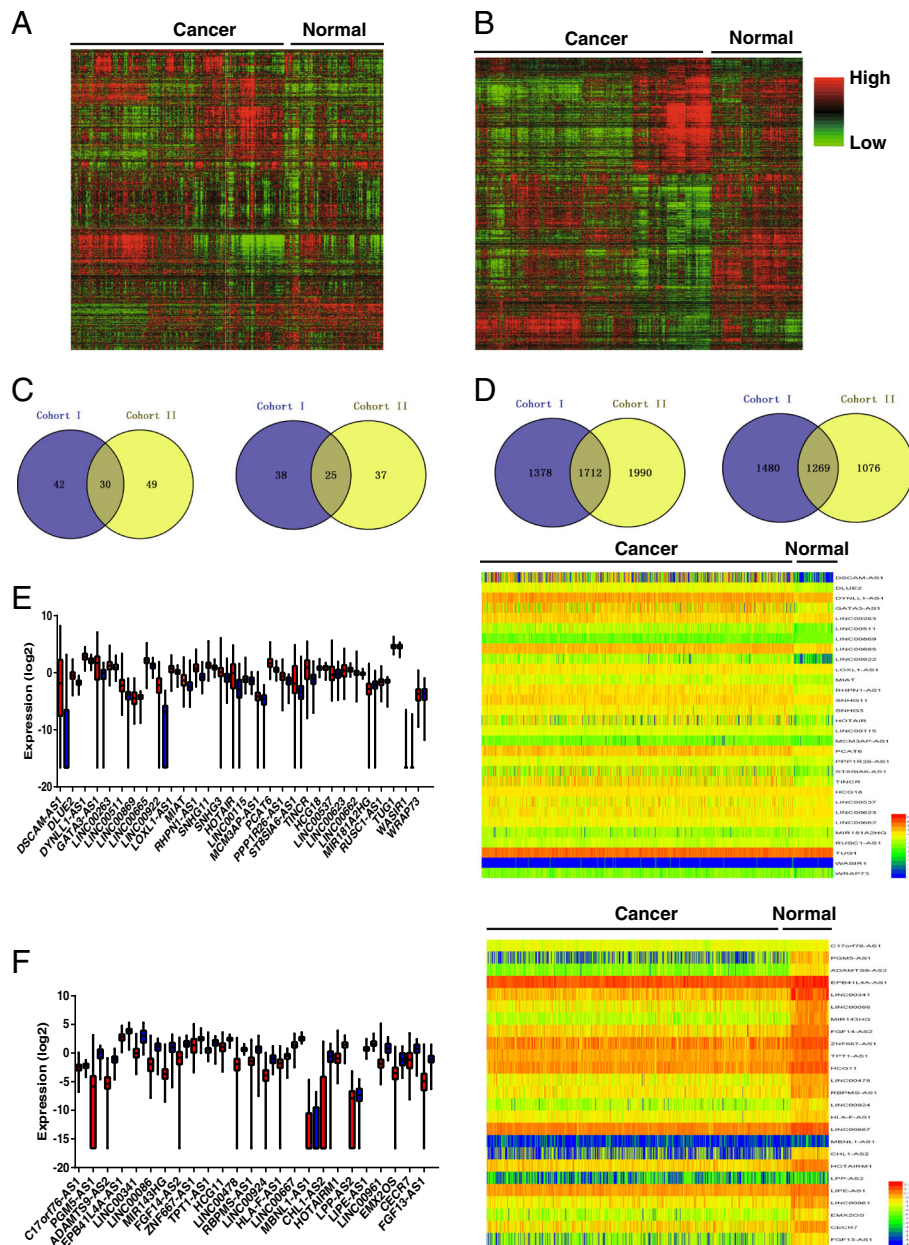


Fig. 1 Aberrant gene expression profile in breast cancer. **a** Hierarchical clustering of differentially expressed genes in breast cancer relative to normal tissue in cohort I ($N = 623$). **b** Hierarchical clustering of differentially expressed genes in breast cancer relative to normal tissue in cohort II ($N = 465$). Red through blue colour indicates high to low expression levels, respectively. **c** Upregulated (left) and downregulated (right) lncRNAs from cohort I and cohort II identified via Venny online software analysis (<http://bioinfogp.cnb.csic.es/tools/venny/>). **d** Upregulated (left) and downregulated coding genes (right) from cohort I and cohort II. **e** Expression levels of 30 upregulated lncRNAs in cancer relative to normal tissue (left) and hierarchical clustering of these genes (right) in the TCGA. **f** Expression levels of 25 downregulated lncRNAs in cancer relative to normal tissue (left) and hierarchical clustering of these genes (right) in the TCGA. The red columns indicate the expression levels of lncRNAs in cancer tissues, and the blue columns represent normal tissues

regulation of lymphocyte activation, extracellular matrix organization, and cell adhesion, and the main pathways or proteins that were involved included cell adhesion molecules, the T-cell receptor signalling pathway, ECM-receptor interaction and the Jak-STAT signalling pathway (Fig. 2a, b and Additional file 4: Table S4). The results of

GOTERM_CC_DIRECT and GOTERM_MF_DIRECT for upregulated lncRNAs were also performed (Additional file 5: Table S5). For the downregulated lncRNAs, cell adhesion, extracellular matrix organization, extracellular structure organization and angiogenesis were identified in the GO function analysis, and focal adhesion,

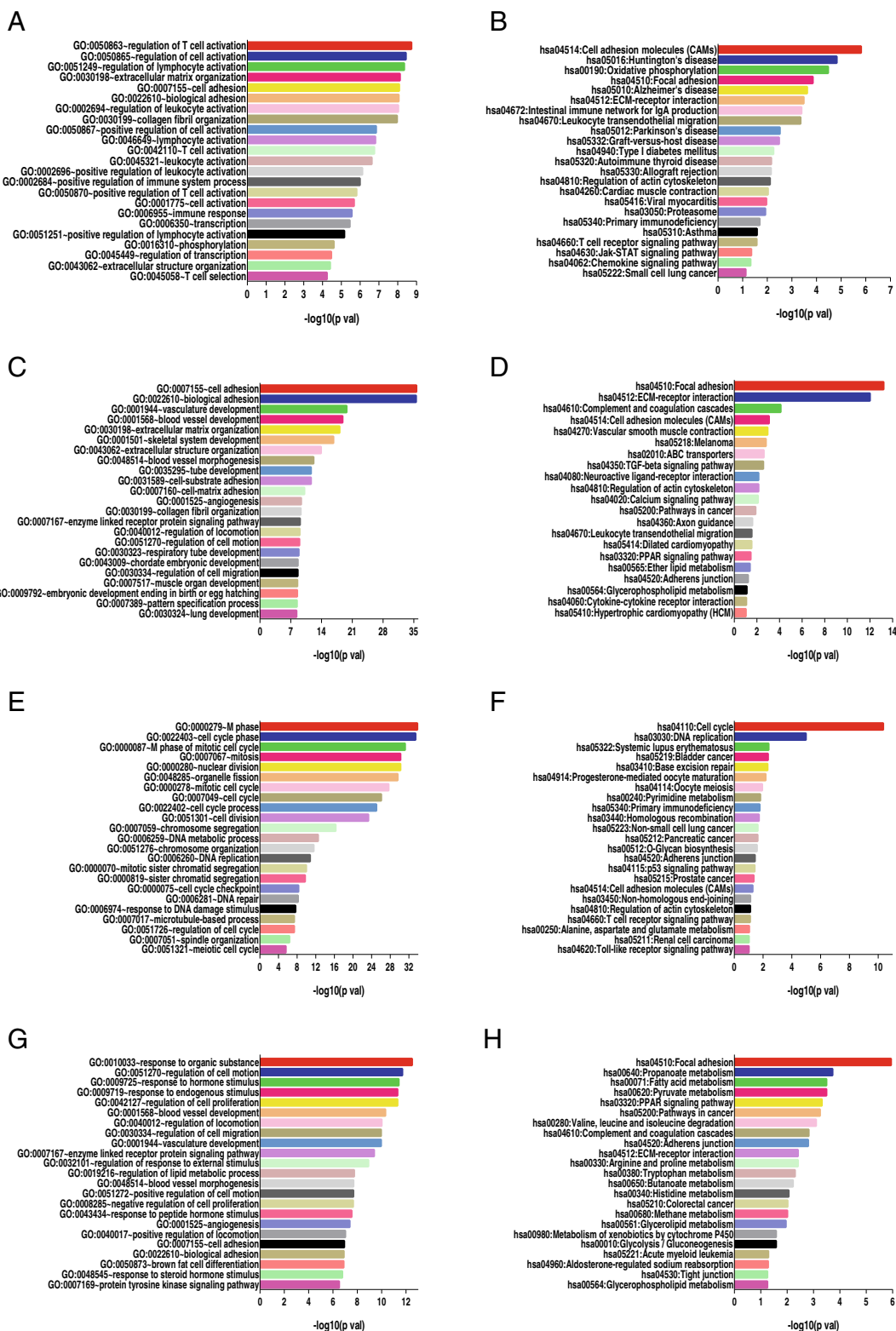


Fig. 2 (See legend on next page.)

(See figure on previous page.)

Fig. 2 Potential biological function of aberrantly expressed lncRNAs in breast cancer. **a & b** Gene ontology enrichment analysis (*left*) and KEGG analysis (*right*) for 30 upregulated lncRNAs from cohort I and cohort II. **c & d** Gene ontology enrichment analysis (*left*) and KEGG analysis (*right*) for 25 downregulated lncRNAs from cohort I and cohort II. **e & f** Gene ontology enrichment analysis for 1712 upregulated coding genes (*left*) and 1269 downregulated coding genes (*right*) from cohort I and cohort II. **g & h** KEGG analysis for 1712 upregulated coding genes (*left*) and 1269 downregulated coding genes (*right*) from cohort I and cohort II. *P*-values <0.05 were defined as statistically significant. The *vertical axis* represents the biological procession or pathway category, and the *horizontal axis* represents the $-\log_{10}(P)$ value of these significant biological processions or pathways

ECM-receptor interaction, ABC transporter activity and the TGF-beta signalling pathway were identified in the KEGG pathway analysis (Fig. 2c, d and Additional file 4: Table S4). Moreover, GO and KEGG analyses were performed for the coding genes. The results indicated that the upregulated genes were associated with cell cycle phase and mitosis, whereas the downregulated genes were associated with the regulation of cell motility, cell proliferation, cell adhesion, and angiogenesis (Fig. 2e, f and Additional file 6: Table S6). For KEGG analysis, cell cycle, DNA replication, the p53 signalling pathway, and the Toll-like receptor signalling pathway were identified for the upregulated genes, and focal adhesion, the PPAR signalling pathway, ECM-receptor interaction and glycolysis/gluconeogenesis were identified for the downregulated genes (Fig. 2g, h and Additional file 6: Table S6).

Oncogenic lncRNAs associated with progressive breast cancer

Breast cancer is highly heterogeneous and can be divided into the following five subclasses: luminal A, luminal B (HER-2 positive), luminal B (HER-2 negative), HER-2 subtype and triple negative [1]. Owing to the aggressive nature of the triple-negative and HER-2 subtypes, we wondered whether aggressive breast cancer subtype-specific lncRNAs (BCSPLs) might exist. Thus, 33 specimens, including each type of breast cancer tissue ($N = 15$) and adjacent normal tissues ($N = 15$), as well as breast tissue from non-cancer patients ($N = 3$), were obtained and subjected to RNA-Seq analysis performed by our research group (Fig. 3a and Additional file 7: Table S7). Among the 30 upregulated lncRNAs from the *in silico* analysis, three lncRNAs were confirmed to be subtype specific in our RNA-Seq results: TINCR, LINC00511, and PPP1R26-AS1 represented the HER-2, triple negative and luminal B subtypes, respectively (Fig. 3b–d). To avoid selection bias in the samples for BCSPL, TCGA data were used to validate these results. As expected, we obtained consistent results from the analysis of the public data (Fig. 3e–g).

Considering triple-negative breast cancer as one of the worst subtypes in breast cancer, LINC00511, as one of the triple-negative-subtype-specific lncRNAs, was selected for further analysis. By guilt-by-association

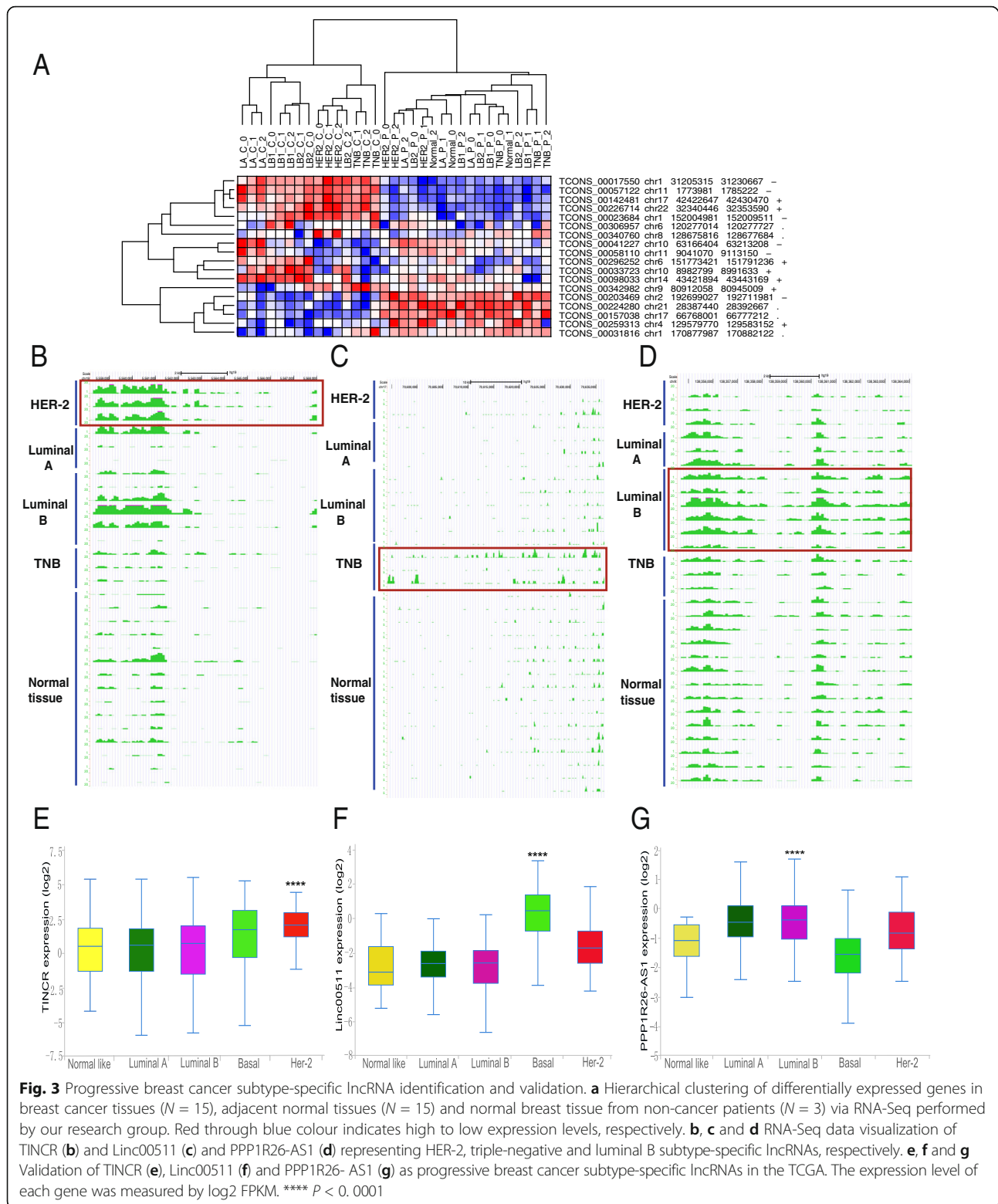
analysis in the TCGA set, LINC00511 was associated with cell adhesion, cell migration, ERK1 and ERK2 cascade regulation and so on by GO analysis (Additional file 8: Table S8). Moreover, it was involved in glycosphingolipid biosynthesis, cell adhesion molecules, proteoglycans and the p53 signalling pathway by KEGG analysis (Additional file 8: Table S8). Next, LINC00511 expression was higher in patients with BRCA1 mutation than in those with wild-type expression (Additional file 9: Figure S1a). Moreover, this phenomenon was observed in patients with RB1 mutation or TP53 mutation (Additional file 9: Figure S1b, c). Finally, the StarBase v2.0 database was applied to examine the relationship of LINC00511 and tumour suppressor microRNAs [43]. In this process, LINC00511 had binding sites of miR-29, miR-16, miR-195 and miR-497, respectively (Additional file 10: Table S9). Meanwhile, these microRNAs were reported to be tumour suppressors [44–47]. Thus, LINC00511 may be a *bona fide* therapy target in triple-negative breast cancer.

Oncogenic lncRNAs are correlated with worse survival outcomes

Next, we investigated the relationship between breast cancer prognosis and the 30 aberrant upregulated lncRNAs obtained via *in silico* analysis from cohorts I & II. GEO was rechecked to identify datasets that included all such genes and the patients' long-term follow-up information. Finally, 26 datasets with 4140 breast cancer patients met these criteria and were used to identify breast cancer prognosis-associated lncRNAs (BCPALs) (Additional file 11: Table S10). Among these lncRNAs, HOTAIR, LINC00115, MCM3AP-AS1, TINCR, PPP1R26-AS1, and DSCAM-AS1 were confirmed to be BCPALs via log-rank overall survival analysis (Fig. 4a–f). In addition to overall survival, increased expression of HOTAIR, MCM3AP-AS1, and PCAT6 indicated worse relapse-free survival in these patients (Fig. 4g–i). Among the six BCPALs, only the increased expression of HOTAIR was associated with a worse prognosis; the remaining five BCPALs were first reported in this study.

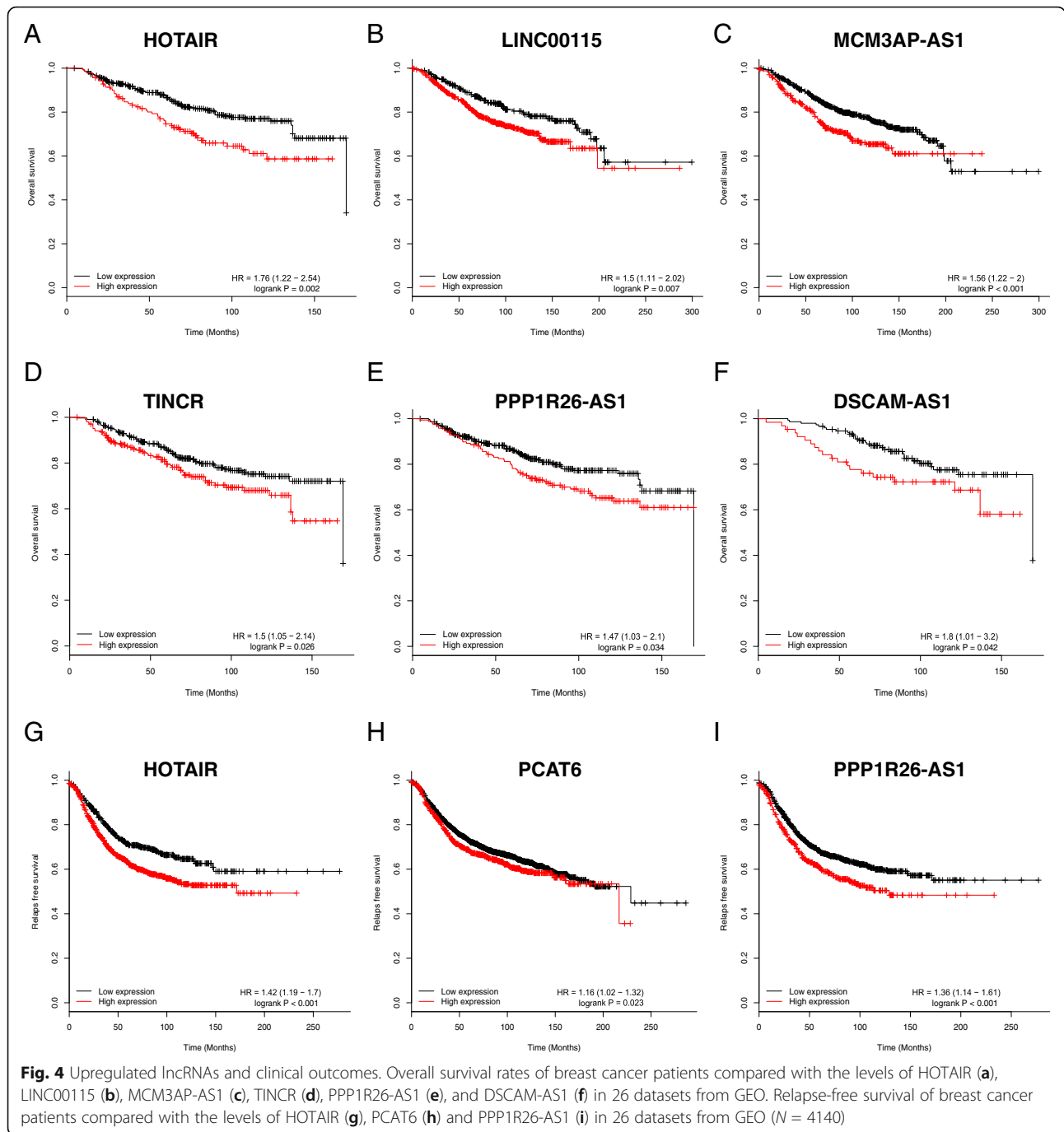
Regulation of the oncogenic lncRNAs

The mechanisms of the upregulation of 30 lncRNAs were next explored via a genomic approach. Mutation, deletion, amplification, and multiple alterations were



investigated for these lncRNAs in the TCGA, Broad and Sanger data. The results showed that the upregulation of these 30 lncRNAs is mainly caused by gene amplification in breast cancer (Fig. 5a). Moreover, gene

amplification may lead to the upregulation of these lncRNAs in pan-cancer, particularly ovarian cancer, bladder cancer and uterine cancer (Fig. 5b). Next, the potential upstream regulation of these lncRNAs via



transcription factor and epigenetic modifications was next identified in breast cancer (<http://www.ncbi.nlm.nih.gov/gds/>). The results indicated that DYNLL1-AS1, GATA3-AS1, LINC00263, LINC00922, MIAT, RHPN1-AS1, SNHG11, HOTAIR, MCM3AP-AS1, PCAT6, PPP1R26-AS1 and TINCR are positively regulated by oestrogen, whereas other lncRNAs, including DLEU2, LINC00511, LOXL1-AS1, SNHG3, LINC00115, MIR181-A2HG, and WRAP73, are negatively regulated by

oestrogen (Fig. 5c). Moreover, WRAP73, LINC00115, and PCAT6 may be negatively regulated by the Kruppel-like zinc finger protein ZNF217, whereas LOXL1-AS1 and MIAT may be positively regulated by this protein (Additional file 12: Table S11). DNA methylation and other transcription factors, such as GATA3 and LIM-only protein 4, were found to be involved in the regulation of similar lncRNAs (Additional file 12: Table S11).

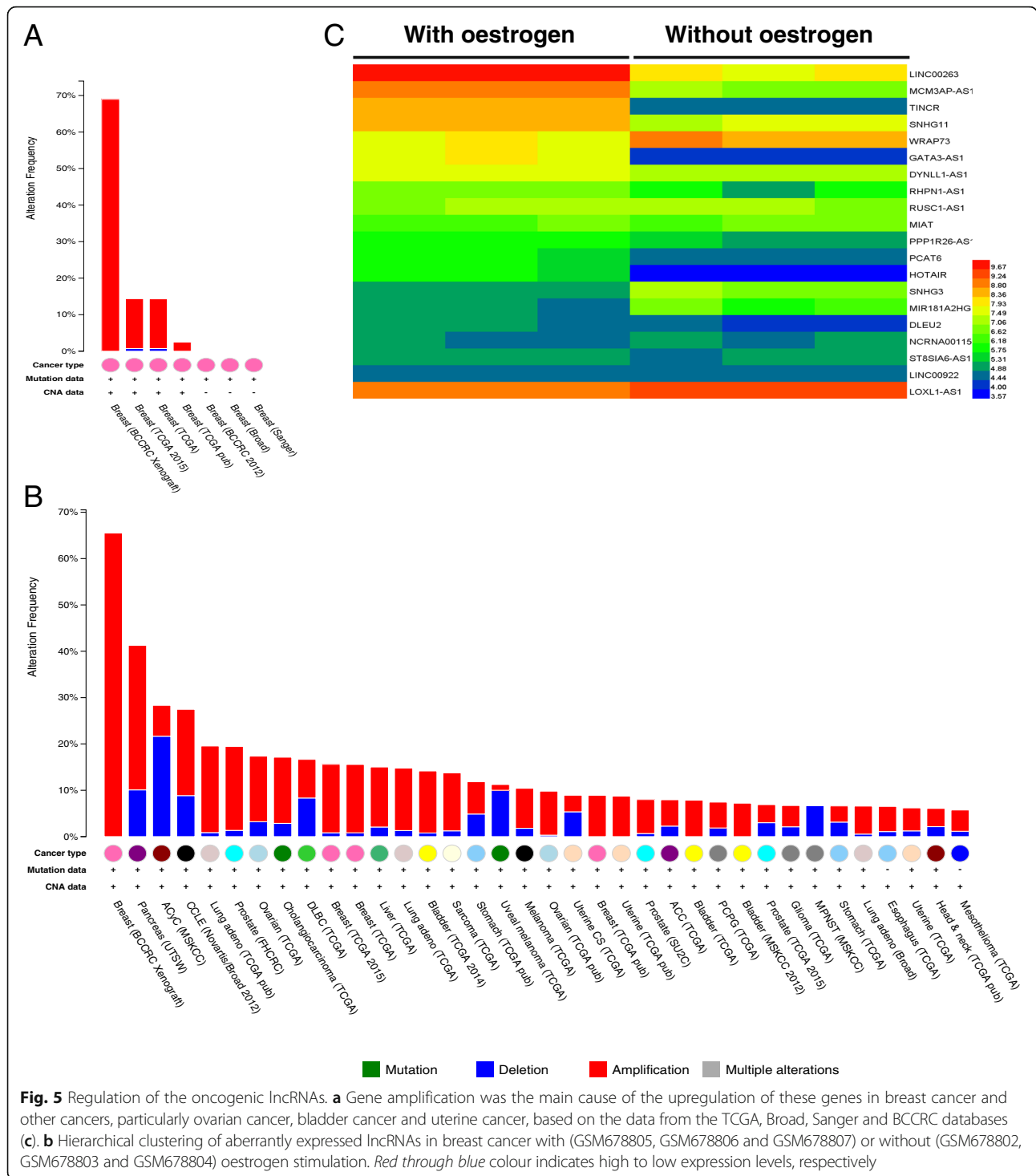


Fig. 5 Regulation of the oncogenic lncRNAs. **a** Gene amplification was the main cause of the upregulation of these genes in breast cancer and other cancers, particularly ovarian cancer, bladder cancer and uterine cancer, based on the data from the TCGA, Broad, Sanger and BCCRC databases (**c**). **b** Hierarchical clustering of aberrantly expressed lncRNAs in breast cancer with (GSM678805, GSM678806 and GSM678807) or without (GSM678802, GSM678803 and GSM678804) oestrogen stimulation. Red through blue colour indicates high to low expression levels, respectively

Validation of oncogenic roles of lncRNAs

To interpret the potential biological functions of these lncRNAs in tumourigenesis, we chose TINCR, HOTAIR and DSCAM-AS1 for further biological function research because these lncRNAs showed the most significant differential expression in cohort I and cohort II. Due to the increased expression of such oncogenic

lncRNAs being associated with worse outcomes, we sought to determine whether these lncRNAs might be involved in breast cancer tumourigenesis. The expression level of TINCR was examined in breast cancer lines (Fig. 6a). UACC812 and MDA-MB-453 cells were selected for the biological function study of TINCR. Two different siRNAs were designed and used to knock down

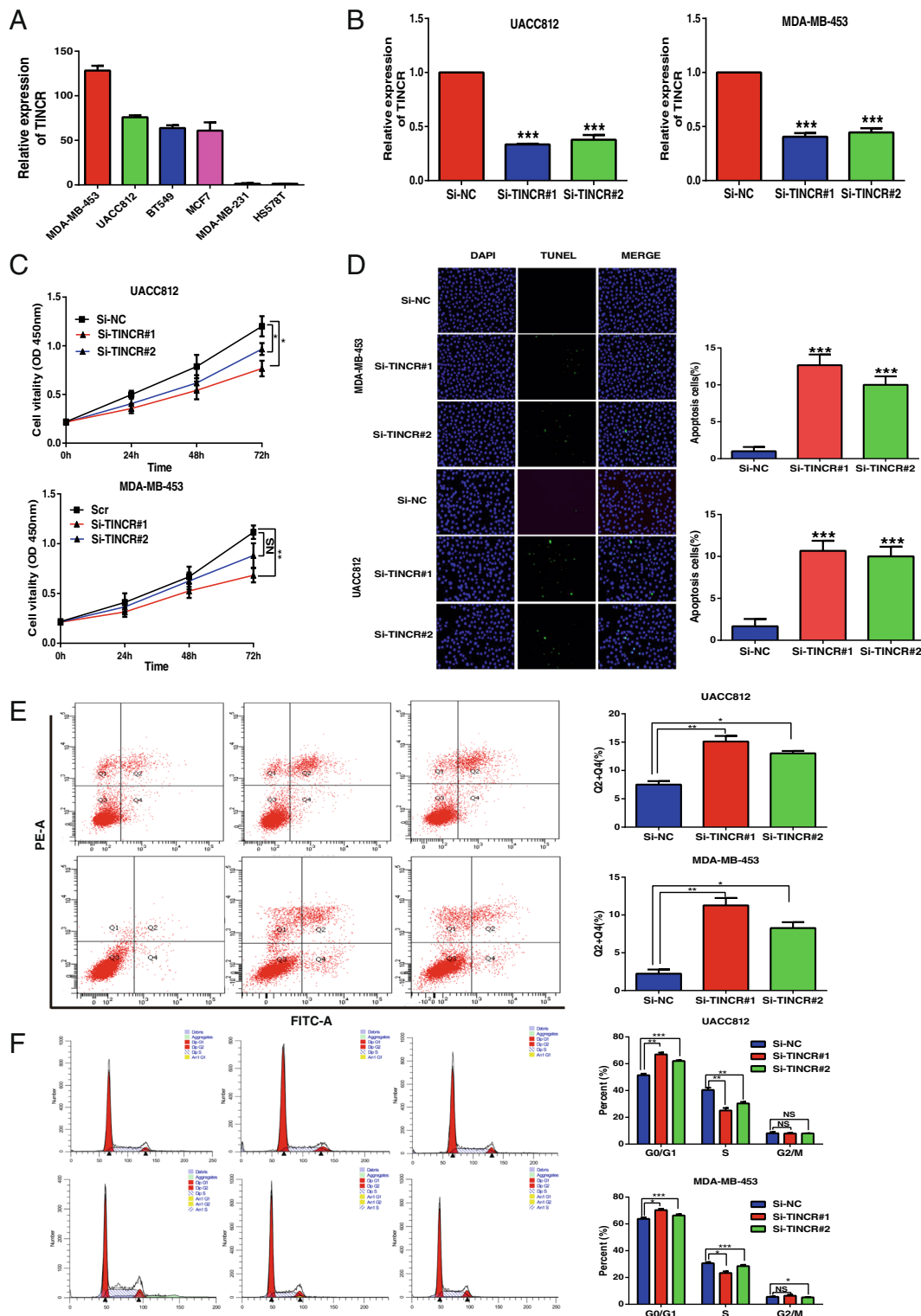


Fig. 6 (See legend on next page.)

(See figure on previous page.)

Fig. 6 Oncogenic functions of TINCR. **a** The expression level of TINCR was examined in breast cancer lines. **b** Knockdown efficiency of siRNAs targeting TINCR in UACC812 and MDA-MB-453 cells, as determined by qRT-PCR. Quantitative normalization of TINCR was performed in each sample using GAPDH expression as an internal control. The relative levels of TINCR vs. GAPDH were determined by the comparative CT ($2^{-\Delta\Delta CT}$) method. **c** The CCK-8 assay was conducted to measure cell proliferation in UACC812 and MDA-MB-453 cells after the knockdown of TINCR. **d** TUNEL apoptosis analysis of UACC812 and MDA-MB-453 cells after the knockdown of TINCR. **e** Flow cytometry apoptosis analysis of UACC812 and MDA-MB-453 cells after the knockdown of TINCR. **f** Determination of the cell cycle distribution in UACC812 and MDA-MB-453 cells after the knockdown of TINCR via flow cytometry. *** $P < 0.001$; ** $P < 0.01$; * $P < 0.05$. Each assay was performed in triplicate

TINCR. The knockdown efficiency was validated via qRT-PCR in vitro (Fig. 6b), and the CCK-8 assay results indicated that knockdown of TINCR inhibited UACC-812 or MDA-MB-453 cell proliferation, respectively (Fig. 6c). Moreover, increased apoptosis was observed after the downregulation of TINCR in UACC-812 and MDA-MB-453 cells via flow cytometry apoptosis detection and TUNEL assay, respectively (Fig. 6d, e). Finally, knockdown of TINCR, compared with controls, increased the percentage of the G0/G1 stage and decreased the percentage of S stage in UACC-812 and MDA-MB-453 cells (Fig. 6f).

Next, molecular interaction and potential oncogenic mechanism of TINCR were explored subsequently. Firstly, knockdown of TINCR increases the expression of Bax and decreases the expression of Bcl-2 in UACC-812 cell (Additional file 13: Figure S2a, b). This result was next validated using breast cancer data in TCGA (Additional file 13: Figure S2c, d). Secondly, co-expressed genes with TINCR were investigated (Additional file 14: Table S12) and cell proliferation function of these genes were detected via GO analysis (Additional file 13: Figure S2e). Finally, TargetScan database (release 3.1) was applied to predict microRNAs regulated by TINCR. MiR-125b was selected and validated in vitro using UACC-812 cell (Additional file 13: Figure S2f, g). It was reported that ERBB2 was a target of MiR-125b in solid cancers [41, 48, 49]. Thus, expression of MiR-125b and ERBB2 in breast cancer was investigated. As expected, MiR-125b was upregulated and ERBB2 was downregulated with TINCR knockdown (Additional file 13: Figure S2g, h). Taken together, TINCR promotes tumorigenesis partly via TINCR-MiR-125b-ERBB2 axis and its negative regulation of apoptosis pathway in breast cancer. For HOTAIR and DSCAM-AS1, such experiments to examine their oncogenic roles were carried out in breast cancer cell lines. As expected, the knockdown of DSCAM-AS1 or HOTAIR inhibits cell proliferation, promotes cell apoptosis, and increases the ratio of G0/G1 stage in MCF7 and T47D cells (Additional file 15: Figure S3a, b, c, d, e). Together, these results indicate that these lncRNAs may play oncogenic roles in the tumorigenesis of breast cancer.

Discussion

Oncogenic lncRNAs in breast cancer were identified and validated comprehensively via genome-wide in silico analysis and biological experiments in this study. To our knowledge, only a few lncRNAs that regulate the process of tumorigenesis and that are associated with the prognosis in breast cancer have been identified. In this study, the identified lncRNAs were found to be associated with essential biological functions in cancer, including the regulation of immune system activation, cell adhesion, angiogenesis, ABC transporter activity, and TGF-beta and Jak-STAT signalling. Moreover, TINCR, LINC00511, and PPP1R26-AS1 were identified as HER-2, triple-negative and luminal B subtype-specific lncRNAs, respectively. In addition, BCPALs, including HOTAIR, LINC00115, MCM3AP-AS1, TINCR, PPP1R26-AS1, and DSCAM-AS1, were identified and confirmed via log-rank analysis. Next, gene amplification in the genome appeared to be the main underlying mechanisms for the upregulation of these oncogenic lncRNAs. Finally, the oncogenic roles of TINCR, HOTAIR and DSCAM-AS1, which showed the most significant differential expression in cohort I and cohort II, were selected and performed in vitro. Knockdown of each of the above lncRNA inhibited proliferation in breast cancer cell lines, increased apoptosis and inhibited cell cycle progression. Together, these results indicated that the lncRNAs identified and validated in this study play oncogenic roles in breast cancer.

Among the 30 aberrantly upregulated lncRNAs identified in this study, only 7 have previously been reported to be associated with malignancy, and the potential functions of the other 23 lncRNAs remain a mystery. Increased SNHG3 expression is associated with malignant status and worse overall survival, recurrence-free survival and disease-free survival in hepatocellular carcinoma patients [50]. PCAT6 may play an oncogenic role in lung cancer progression, and it negatively correlates with the overall survival of lung cancer patients [51]. MIAT may promote the development of lung adenocarcinoma via the MIAT-miR-106-MAPK signalling pathway loop [52]. Oestrogen increases HOTAIR levels via GPER-mediated miR-148a inhibition and is an independent

prognostic marker of metastasis in breast cancer [53, 54]. Moreover, increased expression of DSCAM-AS1 was associated with a worse overall survival in breast cancer patients according to our *in silico* analysis, and knockdown of DSCAM-AS1 inhibited breast cancer cell proliferation, increased apoptosis and inhibited cell cycle progression in this study. We obtained similar results in a previous report [55]. Antisense lncRNAs can regulate their corresponding sense mRNAs at different levels, including transcriptional interference, imprinting, alternative splicing, translation or RNA editing [56]. However, the sense mRNA DSCAM was not regulated by DSCAM-AS1 [55]. Thus, trans-regulation may be associated with the role of DSCAM-AS1 in breast cancer. Regarding DLEU2 and TINCR, it appears to be controversial that DLEU2 and TINCR would act differently in various malignancies. DLEU2 inhibits cell proliferation and tumour progression through the regulation of miR-15a/miR-16-1 in chronic lymphocytic leukaemia [57, 58]. However, our *in silico* analysis suggest that, in breast cancer, it may be oncogenic, in contrast to its reported role as a tumour suppressor in chronic lymphocytic leukaemia. For TINCR, this lncRNA is induced by SP1 and promotes cell proliferation in gastric cancer and oesophageal squamous cell carcinoma, a finding that is in accordance with our results in breast cancer [59, 60]. However, in colorectal cancer, the loss of TINCR expression promotes proliferation and metastasis by activating EpCAM cleavage [61]. The functional diversity of DLEU2 and TINCR in various malignancies may explain this discrepancy, and subsequent research is necessary to reveal the potential roles of these two lncRNAs.

The identification of these oncogenic lncRNAs in this study may add vital significance for clinical practice. The study addresses three issues that are not present in other studies. First, a panel of oncogenic lncRNAs was identified via genome-wide *in silico* analysis. It provides a more efficient step in research methodology and expands the scope of studied lncRNAs that are associated with patient outcome. Hence, the drawbacks of a single lncRNA study model were avoided. Second, the breast cancer prognosis-associated lncRNAs identified via data integration and analysis are definite in this study, and they can be applied for the evaluation of cancer characteristics and patient survival potential. In this way, clinical examination of HOTAIR, LINC00115, MCM3AP-AS1, TINCR, PPP1R26-AS1, and DSCAM-AS1 may be beneficial for comprehensive management. Third, breast cancer subtype-specific lncRNAs were also obtained via high-throughput sequencing. Such oncogenic lncRNAs should be not underestimated for their clinical significance. From a clinical viewpoint, owing to the varying objectivity of pathologists and different qualities of antibodies to ER, PR, HER-2 and Ki-67 used in immunohistochemical methods, determining the molecular subtype

for an individual's breast cancer may be influenced by subjective and objective factors. Many obstacles, including economic and technical factors, must be overcome to achieve the standardization of immunohistochemical methods among laboratories and greater accessibility of quality assurance programmes. From our sequencing results and TCGA data, TINCR, LINC00511, and PPP1R26-AS1 were identified and were found to represent HER-2, triple-negative and luminal B subtypes of breast cancer, respectively. Thus, examination of these BCSPs may complement the use of the aforementioned parameters, and this approach may decrease the misdiagnosis rate and optimize costs in clinical practice.

We acknowledge several limitations of our study. First, not all the datasets that include lncRNA expression and follow-up information in the GEO could be accessed. Thus, the oncogenic lncRNA landscape of breast cancer indicated in this study would not be sufficiently comprehensive. Second, the biological functions and detailed mechanisms of these oncogenic lncRNAs were not elucidated. Additionally, further in-depth efforts will be needed to investigate other oncogenic lncRNAs besides TINCR, HOTAIR and DSCAM-AS1 with large sample sizes and a focus on molecular mechanisms.

Conclusions

Taken together, these findings broaden the oncogenic lncRNA landscape of breast cancer and provide insights into the roles of these oncogenic lncRNAs. The results may aid in the comprehensive management of breast cancer.

Additional files

Additional file 1: Table S1. The clinicopathological characteristics of the patients. (XLS 19 kb)

Additional file 2: Table S2. Differentially expressed genes in Cohort I. (XLS 96 kb)

Additional file 3: Table S3. Differentially expressed genes in Cohort II. (XLS 8892 kb)

Additional file 4: Table S4. Gene ontology term enrichment (GO) and KEGG pathway analysis for upregulated lncRNAs. (XLS 326 kb)

Additional file 5: Table S5. GOTERM_CC_DIRECT and GOTERM_MF_DIRECT for upregulated lncRNAs. (XLSX 41 kb)

Additional file 6: Table S6. Gene ontology term enrichment (GO) and KEGG pathway analysis for differentially expressed mRNAs. (XLS 909 kb)

Additional file 7: Table S7. Differentially expressed genes of RNA-seq in our Cohort. (XLS 2231 kb)

Additional file 8: Table S8. Gene ontology term enrichment (GO) and KEGG pathway analysis for linc00511 in TCGA. (XLSX 21 kb)

Additional file 9: Figure S1. LINC00511 expression was high in patients with BRCA1 or RB1 or TP53 mutation than those with wild type. (PDF 116 kb)

Additional file 10: Table S9. StarBase v2.0 database prediction for the relationship between LINC00511 and tumour suppressor microRNAs. (XLSX 9 kb)

Additional file 11: Table S10. 26 datasets with 4140 breast cancer patients for identifying prognosis-associated lncRNAs. (XLS 24 kb)

Additional file 12: Table S11. DNA methylation and transcription factors regulation for lncRNAs. (XLS 27 kb)

Additional file 13: Figure S2. TINCR mediates the tumorigenesis of breast cancer partly via regulation of apoptosis pathway and TINCR-MiR-125-ERBB2 axis in UACC-812 cell line. (PDF 1571 kb)

Additional file 14: Table S12. TINCR is negatively expressed with several genes involved in apoptosis pathway via analysis of TCGA breast cancer data. (XLSX 657 kb)

Additional file 15: Figure S3. Oncogenic functions of DSCAM-AS1 and HOTAIR. (PDF 462 kb)

Abbreviations

BCPALS: Breast cancer prognosis-associated lncRNAs; BCSPs: Breast cancer subtype-specific lncRNAs; CV: Coefficient of variation; ENCODE: Encyclopedia of DNA Elements; FDR: False discovery rate; GEO: Gene expression omnibus; GO: Gene ontology term enrichment; KEGG: Kyoto encyclopedia of genes and genomes; lncRNA: Long noncoding RNAs; piRNA: Piwi-interacting RNA; siRNA: Small interfering RNA; TCGA: The cancer genome atlas; TUNEL: Transferase mediated deoxyuridine triphosphate-digoxigenin nick end-labelling assay

Acknowledgements

The authors thank the patients, study investigators, and staff who participated in this study.

Funding

This work was supported by funding from the specific research fund for public service sector, National Health and Family Planning Commission of the People's Republic of China (Grant Number 201402003), the National Key Technology Support Programme (Grant Number 2014BAI09B08), the Project of Heilongjiang Province Applied Technology Research and Development (Grant Number GA13C201), and the National Natural Science Foundation of China (Grant Number 81602323), China Postdoctoral Science Foundation Grant (Grant Number 2016 M600262), Heilong Jiang postdoctoral Fund (Grant Number LBH-Z16163), Heilong Jiang province Health and Family Planning Commission Foundation Grant (Grant Number 2016-087), Graduate Innovation Fund of China and Russia (Grant Number YJSCX2015-65HYD).

Availability of data and materials

The authors declare that the data supporting the findings of this study are available within the article and its supplementary information files.

Authors' contributions

SX and YP contributed to the study design, RNA sequencing data and public data interpretation, and manuscript draft. DP contributed to the study design and supervision. DK, QC and contributed to the molecular biology experiments. All authors contributed to the review and revision of the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

This study protocol conformed to the clinical research guidelines and was approved by the research ethics committee of the Harbin Medical University Cancer Hospital. Written informed consent was obtained from all the patients who participated in this study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Breast Surgery, Harbin Medical University Cancer Hospital, 150 Haping Road, Harbin 150081, China. ²College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China. ³Heilongjiang Academy of Medical Sciences, 157 Baojian Road, Harbin 150086, China.

Received: 7 April 2017 Accepted: 10 July 2017

Published online: 24 July 2017

References

- Santana-Davila R, Perez EA. Treatment options for patients with triple-negative breast cancer. *J Hematol Oncol.* 2010;3:42.
- Smith AD, Roda D, Yap TA. Strategies for modern biomarker and drug development in oncology. *J Hematol Oncol.* 2014;7:70.
- Huarte M. The emerging role of lncRNAs in cancer. *Nat Med.* 2015;21(11):1253–61.
- Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell.* 2009;136(4):629–41.
- Jalali S, Gandhi S, Scaria V. Navigating the dynamic landscape of long noncoding RNA and protein-coding gene annotations in GENCODE. *Hum Genomics.* 2016;10(1):35.
- Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, et al. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 2012;22(9):1775–89.
- Batista PJ, Chang HY. Long noncoding RNAs: cellular address codes in development and disease. *Cell.* 2013;152(6):1298–307.
- Matsumoto A, Pasut A, Matsumoto M, Yamashita R, Fung J, Monteleone E, et al. mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature.* 2017;541(7636):228–32.
- Anderson DM, Anderson KM, Chang CL, Makarewicz CA, Nelson BR, McAnally JR, et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell.* 2015;160(4):595–606.
- Nelson BR, Makarewicz CA, Anderson DM, Winders BR, Troupes CD, Wu F, et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science.* 2016;351(6270):271–5.
- Szafron LM, Balcerak A, Grzybowska EA, Pienkowska-Grela B, Felisiak-Golabek A, Podgorska A, et al. The novel gene CRNDE encodes a nuclear peptide (CRNDEP) which is overexpressed in highly proliferating tissues. *PLoS One.* 2015;10(5):e0127475.
- Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 2011;25(18):1915–27.
- Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. *Nature.* 2012;489(7414):101–8.
- Ravasi T, Suzuki H, Pang KC, Katayama S, Furuno M, Okunishi R, et al. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res.* 2006;16(11):11–9.
- Beltran M, Puig I, Pena C, Garcia JM, Alvarez AB, Pena R, et al. A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev.* 2008;22(6):756–69.
- Hah N, Murakami S, Nagari A, Danko CG, Kraus WL. Enhancer transcripts mark active estrogen receptor binding sites. *Genome Res.* 2013;23(8):1210–23.
- Heo JB, Sung S. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science.* 2011;331(6013):76–9.
- Kwek KY, Murphy S, Furger A, Thomas B, O'Gorman W, Kimura H, et al. U1 snRNA associates with TFIID and regulates transcriptional initiation. *Nat Struct Biol.* 2002;9(11):800–5.
- Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, et al. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science.* 2008;322(5908):1717–20.
- Penny GD, Kay GF, Sheardown SA, Rastan S, Brockdorff N. Requirement for Xist in X chromosome inactivation. *Nature.* 1996;379(6561):131–7.
- Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Bruggmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell.* 2007;129(7):1311–23.
- Huang X, Xiao R, Pan S, Yang X, Yuan W, Tu Z, et al. Uncovering the roles of long non-coding RNAs in cancer stem cells. *J Hematol Oncol.* 2017;10(1):62.
- Zhang H, Chen Z, Wang X, Huang Z, He Z, Chen Y. Long non-coding RNA: a new player in cancer. *J Hematol Oncol.* 2013;6:37.
- Moyano M, Stefani G. piRNA involvement in genome stability and human cancer. *J Hematol Oncol.* 2015;8:38.

25. Naidu S, Magee P, Garofalo M. MiRNA-based therapeutic intervention of cancer. *J Hematol Oncol*. 2015;8:68.
26. Sanchez-Mejias A, Tay Y. Competing endogenous RNA networks: tying the essential knots for cancer biology and therapeutics. *J Hematol Oncol*. 2015;8:30.
27. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*. 2009;458(7235):223–7.
28. Kawasaki Y, Komiya M, Matsumura K, Negishi L, Suda S, Okuno M, et al. MYU, a target lncRNA for Wnt/c-Myc signaling, mediates induction of CDK6 to promote cell cycle progression. *Cell Rep*. 2016;16(10):2554–64.
29. Li Z, Gao B, Hao S, Tian W, Chen Y, Wang L, et al. Knockdown of lncRNA-PANDAR suppresses the proliferation, cell cycle and promotes apoptosis in thyroid cancer cells. *EXCLI J*. 2017;16:354–62.
30. Li Z, Rana TM. Decoding the noncoding: prospective of lncRNA-mediated innate immune regulation. *RNA Biol*. 2014;11(8):979–85.
31. Suravajhala P, Kogelman LJ, Mazzoni G, Kadarmideen HN. Potential role of lncRNA cyp2c91-protein interactions on diseases of the immune system. *Front Genet*. 2015;6:255.
32. Zhou M, Hou Y, Yang G, Zhang H, Tu G, Du YE, et al. lncRNA-Hh strengthen cancer stem cells generation in twist-positive breast cancer via activation of hedgehog signaling pathway. *Stem Cells*. 2016;34(1):55–66.
33. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*. 2010;464(7291):1071–6.
34. Prensner JR, Iyer MK, Sahu A, Asangani IA, Cao Q, Patel L, et al. The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. *Nat Genet*. 2013;45(11):1392–8.
35. Leucci E, Vendramin R, Spinazzi M, Laurette P, Fiers M, Wouters J, et al. Melanoma addiction to the long non-coding RNA SAMMSON. *Nature*. 2016;531(7595):518–22.
36. Mendell JT. Targeting a long noncoding RNA in breast cancer. *N Engl J Med*. 2016;374(23):2287–9.
37. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A*. 2001;98(9):5116–21.
38. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A*. 1998;95(25):14863–8.
39. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44–57.
40. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol*. 2005;4:Article17.
41. Tang XY, Zheng W, Ding M, Guo KJ, Yuan F, Feng H, et al. miR-125b acts as a tumor suppressor in chondrosarcoma cells by the sensitization to doxorubicin through direct targeting the ErbB2-regulated glucose metabolism. *Drug Des Devel Ther*. 2016;10:571–83.
42. Garcia-Castro B, Alvarez-Zavala M, Riveros-Magana AR, Ortiz-Lazareno PC, Ratkovich-Gonzalez S, Hernandez-Flores G, et al. Restoration of WNT4 inhibits cell growth in leukemia-derived cell lines. *BMC Cancer*. 2013;13:557.
43. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res*. 2014;42(Database issue):D92–7.
44. Rivas MA, Venturutti L, Huang YW, Schillaci R, Huang TH, Elizalde PV. Downregulation of the tumor-suppressor miR-16 via progesterin-mediated oncogenic signaling contributes to breast cancer development. *Breast Cancer Res*. 2012;14(3):R77.
45. Rostas JW 3rd, Pruitt HC, Metge BJ, Mitra A, Bailey SK, Bae S, et al. microRNA-29 negatively regulates EMT regulator N-myc interactor in breast cancer. *Mol Cancer*. 2014;13:200.
46. Wang Y, Zhang X, Zou C, Kung HF, Lin MC, Dress A, et al. miR-195 inhibits tumor growth and angiogenesis through modulating IRS1 in breast cancer. *Biomed Pharmacother*. 2016;80:95–101.
47. Wu Z, Li X, Cai X, Huang C, Zheng M. miR-497 inhibits epithelial mesenchymal transition in breast carcinoma by targeting Slug. *Tumour Biol*. 2016;37(6):7939–50.
48. Fassan M, Pizzi M, Realdon S, Balistreri M, Guzzardo V, Zagonel V, et al. The HER2-miR125a5p/miR125b loop in gastric and esophageal carcinogenesis. *Hum Pathol*. 2013;44(9):1804–10.
49. Yagishita S, Fujita Y, Kitazono S, Ko R, Nakadate Y, Sawada T, et al. Chemotherapy-regulated microRNA-125-HER2 pathway as a novel therapeutic target for Trastuzumab-mediated cellular cytotoxicity in small cell lung cancer. *Mol Cancer Ther*. 2015;14(6):1414–23.
50. Zhang T, Cao C, Wu D, Liu L. SNHG3 correlates with malignant status and poor prognosis in hepatocellular carcinoma. *Tumour Biol*. 2016;37(2):2379–85.
51. Wan L, Zhang L, Fan K, Cheng ZX, Sun QC, Wang JJ. Knockdown of Long Noncoding RNA PCAT6 Inhibits Proliferation and Invasion in Lung Cancer Cells. *Oncol Res*. 2016;24(3):161–70.
52. Li DS, Ainiwaer JL, Sheyhiding I, Zhang Z, Zhang LW. Identification of key long non-coding RNAs as competing endogenous RNAs for miRNA-mRNA in lung adenocarcinoma. *Eur Rev Med Pharmacol Sci*. 2016;20(11):2285–95.
53. Tao S, He H, Chen Q. Estradiol induces HOTAIR levels via GPER-mediated miR-148a inhibition in breast cancer. *J Transl Med*. 2015;13:131.
54. Yu X, Li Z. Long non-coding RNA HOTAIR: A novel oncogene (Review). *Mol Med Rep*. 2015;12(4):5611–8.
55. Niknafs YS, Han S, Ma T, Speers C, Zhang C, Wilder-Romans K, et al. The lncRNA landscape of breast cancer reveals a role for DSCAM-AS1 in breast cancer progression. *Nat Commun*. 2016;7:12791.
56. Magistri M, Faghihi MA, St Laurent G 3rd, Wahlestedt C. Regulation of chromatin structure by long noncoding RNAs: focus on natural antisense transcripts. *Trends Genet*. 2012;28(8):389–96.
57. Kasar S, Underbayev C, Yuan Y, Hanlon M, Aly S, Khan H, et al. Therapeutic implications of activation of the host gene (Dleu2) promoter for miR-15a/16-1 in chronic lymphocytic leukemia. *Oncogene*. 2014;33(25):3307–15.
58. Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, et al. The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. *Cancer Cell*. 2010;17(1):28–40.
59. Xu TP, Liu XX, Xia R, Yin L, Kong R, Chen WM, et al. SP1-induced upregulation of the long noncoding RNA TINCR regulates cell proliferation and apoptosis by affecting KLF2 mRNA stability in gastric cancer. *Oncogene*. 2015;34(45):5648–61.
60. Xu Y, Qiu M, Chen Y, Wang J, Xia W, Mao Q, et al. Long noncoding RNA, tissue differentiation-inducing nonprotein coding RNA is upregulated and promotes development of esophageal squamous cell carcinoma. *Dis Esophagus*. 2016;29(8):950–8.
61. Zhang ZY, Lu YX, Zhang ZY, Chang YY, Zheng L, Yuan L, et al. Loss of TINCR expression promotes proliferation, metastasis through activating EpCAM cleavage in colorectal cancer. *Oncotarget*. 2016;7(16):22639–49.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

