Inferences of individual drug response-related long non-coding RNAs based on integrating multi-omics data in breast cancer

Hao Cui, Hanqing Kong, Fuhui Peng, Chunjing Wang, Dandan Zhang, Jiawei Tian, Lei Zhang

PII: S2162-2531(20)30081-0
DOI: https://doi.org/10.1016/j.omtn.2020.01.038
Reference: OMTN 831

To appear in: Molecular Therapy: Nucleic Acid

Received Date: 26 October 2019
Accepted Date: 31 January 2020

Please cite this article as: Cui H, Kong H, Peng F, Wang C, Zhang D, Tian J, Zhang L, Inferences of individual drug response-related long non-coding RNAs based on integrating multi-omics data in breast cancer, Molecular Therapy: Nucleic Acid (2020), doi: https://doi.org/10.1016/j.omtn.2020.01.038.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 The Author(s).
Inferences of individual drug response-related long non-coding RNAs based on integrating multi-omics data in breast cancer

Hao Cui¹, Hanqing Kong¹, Fuhui Peng¹, Chunjing Wang², Dandan Zhang³, Jiawei Tian¹*, Lei Zhang¹*

¹ Department of Ultrasound, the Second Affiliated Hospital of Harbin Medical University, Harbin 150081, China
² Department of General Surgery, the Second Affiliated Hospital of Harbin Medical University, Harbin 150081, China
³ Department of Ultrasound, Heilongjiang provincial hospital, Harbin 150081, China

Corresponding Author:
Prof Lei Zhang, Tel: +86 451 86605253; Fax: +86 451 86675845; Email: zhanglei6@hrbmu.edu.cn.

Prof Jiawei Tian, Tel: +86 451 86605811; Fax: +86 451 86675845; Email: jwtian2004@163.com.

Running title: Drug-related IncRNAs follow multi-omics in BRCA

Key words: Breast cancer, drug response, long non-coding RNAs, multi-omics integration, prognosis
ABSTRACT

Differences in individual drug responses are an obstacle in breast cancer (BRCA) treatment, so predicting responses would help to plan treatment strategies. The accumulation of cancer molecular profiling and drug response data provide opportunities and challenges to identify novel molecular signatures and mechanisms of tumor responsiveness to drugs in BRCA. This study evaluated drug responses with a multi-omics integrated system that depended on long non-coding RNAs (lncRNAs). We identified drug response-related lncRNAs (DRlnc) by combining expression data of lncRNA, microRNA, messenger RNA, methylation levels, and somatic mutations, and the survival data of cancer patients treated with drugs. We constructed an integrated and computational multi-omics approach to identify DRlncs for diverse chemotherapeutic drugs in BRCA. Some DRlncs were identified with adriamycin, cytoxan, tamoxifen, and all samples for BRCA patients. These DRlncs showed specific features regarding both expression and computational accuracies. The DR-gene co-expression networks were constructed and analyzed. Key DRlncs such as HOXA-AS2 in drug adriamycin was characterized. The experimental analysis also suggested that HOXA-AS2 was a key DRlnc in adriamycin drug resistance in BRCA patients. Some DRlncs were associated with survival and some specific functions. A possible mechanism of DRlnc HOXA-AS2 in the adriamycin drug response for BRCA resistance was inferred. In summary, this study provides a framework for
IncRNA-based evaluation of clinical drug responses in BRCA. Understanding the underlying molecular mechanisms of drug responses will facilitate improved responses to chemotherapy and outcomes of BRCA treatment.
INTRODUCTION

Breast cancer (BRCA), with a global incidence of nearly 1.7 million new cases each year and over 520,000 deaths\textsuperscript{1-3}, is still the focus of scientific research. Currently, chemotherapy remains one part of standardized treatment for BRCA. Several chemotherapeutic drugs including cytoxan, Adriamycin, fluorouracil, methotrexate, and tamoxifen are first-line medications. However, the development of drug resistance to these common chemotherapeutic drugs is the major obstacle for successful treatment\textsuperscript{4-6}. Drug resistance varies among different individuals\textsuperscript{7, 8} and molecular targeted combination therapy is an important approach to overcome drug resistance. Identifying the mechanisms that promote chemoresistance is therefore important in improving the efficacy of chemotherapy, which could treat BRCA by preventing tumor growth recurrence.

Long non-coding RNAs (lncRNAs) lack the ability to encode proteins, but play a key role in the occurrence and development of diseases\textsuperscript{9-11}. The lncRNAs have dual effects in BRCA metastasis by regulating invasion, migration, and distant metastasis of BRCA cells\textsuperscript{12}. For example, LINC00963 promotes tumorigenesis and radioresistance in BRCA and represents a potential target for the treatment of BRCA\textsuperscript{13}. The lncRNA termed HOTAIR shows increased expression in primary breast tumors and metastases, and the expression level in primary tumors is a good predictor of eventual metastasis and death\textsuperscript{14}. The transcript levels of GAS5, another lncRNA, were significantly reduced in BRCA,
by controlling apoptosis and downregulation\textsuperscript{15}. SNHG5 promoted breast cancer proliferation and cell cycle progression by upregulation of PCNA expression\textsuperscript{16}.

Recently, some studies have shown that IncRNAs play important roles in the tumor drug resistance regulatory network, which can lead to drug resistance of tumor cells through abnormal regulation of target genes. Fernando et al. reported that high expression of IncRNA BALA-2 was related to prednisolone resistance in acute lymphocytic leukemia (ALL), and knocking out BALA-2 improved the sensitivity of B-ALL to chemotherapy drugs\textsuperscript{17}. The functional loss of IncRNA TP53TG1 led to abnormal activity of the YBX1 protein, which prevented cancer cell death induced by anti-tumor drugs, and led to drug resistance of many common anti-tumor drugs such as 5-fluorouracil (chemotherapy drugs for gastric cancer and colon cancer) and irinotecan (chemotherapy drugs for colorectal cancer and BRCA)\textsuperscript{18}. Furthermore, some IncRNAs are related to chemoresistance in BRCA. LncRNA ARA was found in the BRCA adriamycin resistant cell lines, McF-7/ADR and McF-7, by using a LncRNA microarray\textsuperscript{19}. However, there are still many drug response-related Incs (Dlnc) in BRCA, which urgently need integration of data to systematically develop efficient and reliable algorithms to predict the DRlncs in BRCA patients.

LncRNAs cannot function independently, but interact with other molecules and coordinate with each other to play roles in physiological and pathological...
processes. Some studies have indicated that lncRNA specifically binds to microRNA (miRNA) and indirectly regulates the expression of miRNA target genes, thereby participating in biological processes related to malignant tumors. Some reports have shown that network analysis of aberrantly methylated lncRNAs in cancers indicate that lncRNAs with aberrant methylation patterns might be involved in cancer development and progression. Ning et al. developed a database to identify disease-associated single nucleotide polymorphisms (SNP)s in human lncRNAs. Khurana et al. showed that genetic mutations in lncRNA were closely related to the development of malignant tumors and chemotherapy resistance. However, the above reports of lncRNA regulation are single reports, and few studies have been conducted by multi-omics. Importantly, more studies are needed to apply multi-omics in the DRlncs in BRCA.

In the present study, an integrated approach was developed to identify DRlncs in BRCA based on the expression of lncRNA, miRNA, messenger RNA (mRNA), methylation levels and somatic mutations, and the survival data of cancer patients treated with drugs. Some DRlncs were identified for adriamycin, Cytoxan, Tamoxifen, and all samples for BRCA patients. The DR gene co-expression networks were constructed and analyzed. Key DRlncs such as HOXA-AS2 in drug resistance to adriamycin was also characterized. In addition, some experiments were performed to confirm that HOXA-AS2 was related to adriamycin resistance in BRCA cell lines, including T47D cells and...
MDA-MB-231 cells. Some DRlncs were associated with survival and other specific functions. Collectively, this study provided a feasible strategy for cancer drug repositioning as well as novel findings regarding cancer-associated IncRNA discovery.
RESULTS

Identification of individual DRlncs based on multi-omic data integration for multiple drugs in BRCA patients

An integrated computational algorithm was developed to identify DRlncs for multiple drugs in BRCA based on multiple factors including genes, miRNAs, methylations, and somatic mutations. A total of 14, 14, 23, and 15 DRlncs were identified for adriamycin, cytoxan, tamoxifen, and all BRCA samples, respectively (Figure 1A). These DRlncs could be influenced by single or multiple factors in their drug responses in BRCA patients. Single or multiple factors had no impact on the P-values of DRlncs. For adriamycin, DRlncs AC117386, AL691403, BX324167, and HOX-AS2 all showed more significant results. DRlncs AC117386, AL691403, and BX324167 were influenced by a single factor; however, HOX-AS2 was influenced by four factors (Figure 1B). A similar phenomenon was also discovered for Tamoxifen, Cytoxan, and all BRCA samples (Figure 1C, Figure S1). The proportion of factor numbers were different for diverse drugs. For adriamycin, DRlncs, which were influenced by two kinds of factors, were the largest contributor to all DRlncs (Figure 1D). However, DRlncs that were influenced by two kinds of factors involved the maximum proportion for cytoxan, tamoxifen, and all samples. Four IncRNAs, including AC010680, AC073611, AP001107, and AC008443, were identified as DRlncs in multiple drugs (Figure 1E). These DRlncs could be considered as common DRlncs and may have participated in multiple drug
responses during BRCA treatments. DRlnc AC008443 was influenced by 14 genes, 32 methylation sites, and four somatic mutations in BRCA samples with cytoxan and tamoxifen (Figure 1F).

**Individual DRlncs showed specific features for diverse drugs in BRCA patients.**

To depict the characteristics of DRlncs for diverse drugs in BRCA, we further characterized the DRlncs. Most DRlncs were upregulated in drug resistant samples for the four kinds of drugs (Figure 2A). Risk score profiles were created to characterize the compactness of DRlncs for each drug in BRCA patients. Risk scores of somatic mutations showed a bimodal distribution in all four kinds of drugs (Figure 2B, C, D, E). Risk scores of methylation and miRNA showed unimodal distributions. Although there was an obvious difference of risk scores in diverse drugs, the last risk scores showed similar distributions in all drugs (Figure 2F). The results indicated that the trends of the last risk scores were consistent for all drugs. In addition, we also analyzed the expressions of DRlncs. For example, DRlnc HOXA-AS2, an IncRNA located between the HOXA3 and HOXA4 genes in the HOXA cluster, has been characterized as an oncogene in many kinds of cancers including BRCA\textsuperscript{25-27}. Specifically, HOXA-AS2 plays an important role in the resistance of acute myeloid leukemia cells to adriamycin\textsuperscript{28}. In our analysis, HOXA-AS2 was identified as a DRlnc in adriamycin (P < 0.001). HOXA-AS2 was upregulated in
adriamycin resistant BRCA samples (Figure 2G). The *HNRNPA2B1* gene, belonging to the A/B subfamily of ubiquitously expressed heterogeneous nuclear ribonucleoproteins, was downregulated in adriamycin resistant BRCA samples (Figure 2G). HOXA-AS2 and HNRNPA2B1 were negatively correlated (P = 0.0006, correlation = −0.446) in BRCA samples treated with adriamycin (Figure 2H, I). All the results indicated that computational accuracy and expression features of DRlncs were informative in BRCA patients.

The DRlncs-gene co-expression networks related to drug response in BRCA

The strong co-expression between HOXA-AS2 and HNRNPA2B1 suggested that interactions between DRlncs-genes may be key ways of exploring the roles of DRlncs in drug response for BRCA patients. For each DRlnc that interacted with genes in diverse drugs, Pearson’s correlation coefficients were calculated and only significant ones with P < 0.05 were extracted as effective DRlnc-gene interactions to construct co-expression networks for diverse drugs in BRCA. There were both positive and negative co-expression interactions for cytoxan, tamoxifen, and all samples except adriamycin (Figure 3A). A total DRlnc-gene co-expression network including adriamycin, cytoxan, tamoxifen, and all drugs was constructed (Figure 3B). Some hub and key DRlncs including AC006064, HOXA-AS2, and AC008443 were discovered in this total co-expression network. The total DRlnc-gene network exhibited a scale-free
distribution ($R^2 = 0.827$), which was a specific topological feature of transcriptional regulatory networks (Figure 3C). The DRInc-gene co-expression network for adriamycin was constructed (Figure 3D). DRInc HOXA-AS2 and AC073611 could be considered as two key hubs in the adriamycin-related DRInc-gene co-expression network. DRInc were negatively correlated with the $CAPR1N1$, $UPF1$, and $CSTF2$ genes (Figure 3E). In addition, $DRInc$-gene co-expression networks for cytoxan, tamoxifen, and all samples were also constructed (Figure 3F, G, H). All the above results indicated that the $DRInc$-gene co-expression networks may help to understand the mechanism of DRInc for diverse drugs in BRCA.

Some DRInc were associated with survival of BRCA patients.

To evaluate the potential value of DRInc as prognostic biomarkers for BRCA, we determined whether the DRInc in diverse drugs were associated with survival. We used median values of DRInc expressions as the cut-off to test the survival of BRCA patients. DRInc BX322234 in cytoxan, AC006059 in tamoxifen, and AL158151 and UBXN10-AS1 in all samples were associated with survival (Figure 4A-D). In addition, BRCA patients in the high expression group exhibited a significantly shorter median overall survival than those in the low expression group in all DRInc except AC006059 for Tamoxifen. The results indicated that some DRInc were associated with survival and served as specific prognostic biomarkers.
The functions and mechanism of DRlncs for diverse drugs in BRCA patients

We performed Gene Ontology (GO) enrichment analyses for all DRlncs with diverse drugs. Although DRlncs for diverse drugs were enriched in some different GO terms, most of them were enrichment in COPII vesicle coating related with GO terms (Figure 5A, B, C, D). Coat protein complex II (COPII) is a type of vesicle coat protein that transports from the rough endoplasmic reticulum to the Golgi apparatus. COPII vesicle coating and Golgi vesicle-mediated transport are closely related with development and treatment of BRCA. We inferred that these DRlncs participated in the process of COPII vesicle coating and Golgi vesicle-mediated transport to influence the drug response in BRCA patients. In addition, we inferred the possible mechanism of DRlncs in drug responses in BRCA. For example, we found HOXA-AS2 was a key DRlnc in adriamycin for BRCA patients (Figure 5E). DRlnc HOXA-AS2 was regulated by three important genes including EIF4A3, FMR1, and HNRNPA2B1. These three genes were downregulated in adriamycin-resistant BRCA patients, to promote overexpression of DRlnc HOXA-AS2, thus resulting in drug resistance to adriamycin in BRCA patients. This series of changes also caused changes in other multi-omics. The methylation levels of cg09865454 and expression of mir-107 were dysregulated in drug-resistant BRCA patients. Overall, these results indicated the necessary of identifying DRlncs by
integrating multi-omics data in BRCA. In addition, these findings could provide novel insights for characterizing the role of DRlnc in drug resistant BRCA.

Validation of the role of DRlnc HOXA-AS2 in BRCA

We analyzed the expression of HOXA-AS2 in BRCA tissues from patients with or without adriamycin-based neoadjuvant chemotherapy before surgery. Real-time PCR assays showed that, compared with samples from patients without neoadjuvant chemotherapy, the HOXA-AS2 expressions in samples from patients with adriamycin-based neoadjuvant chemotherapy were significantly higher (Fig. 6A). Moreover, adriamycin-resistant T47D cells dramatically expressed higher levels of HOXA-AS2 than normal T47D cells. The same expression pattern was also confirmed in MDA-MB-231 cells and adriamycin-resistant MDA-MB-231 cells (Fig. 6B and C). To characterize the function of HOXA-AS2 in BRCA cells, the efficiency of HOXA-AS2 knockdown in BRCA cells was detected by real-time PCR assays (Fig. 6D). Fig. 6E shows that the adriamycin-resistant BRCA cells had higher colony formation ability after treatment with 5 mg/L adriamycin when compared with normal BRCA cells. However, HOXA-AS2 silencing reversed the colony formation ability of adriamycin-resistant BRCA cells (Fig. 6E). These results demonstrated that HOXA-AS2 might promote the adriamycin resistance of BRCA cells.
DISCUSSION

Chemotherapy is currently the primary treatment for cancer including BRCA, but its effectiveness is limited by individual differences in drug responses. Drug resistance is one of the causes of BRCA treatment failure. How to evaluate individual drug responses is therefore an urgent need for BRCA treatment. Most current studies have suggested the roles of some lncRNAs in drug responses. However, these studies focused on individual lncRNAs, which lacked a global view of lncRNAs for multiple drugs in BRCA. In the present study, a comprehensive and computational approach was developed to identify DRlncs for multiple drugs based on genes, miRNAs, methylations, and somatic mutations in BRCA. Thus, some DRlncs were identified and analyzed for adriamycin, cytoxan, tamoxifen, and all BRCA samples. The co-expressions of DRlnc-gene networks for multiple drugs were constructed and some key DRlncs including HOXA-AS2 were identified. Some DRlncs were associated with survival and specific functions. In our analysis, most DRlncs were only identified for one kind of drug, and showed drug specificity. Only four DRlncs were present in multiple drug groups. These DRlncs may have common functions with diverse drugs for BRCA resistance.

Studies have suggested that aberrant expression of lncRNAs are responsible for drug resistance in human cancers including BRCA\textsuperscript{30-32}. Multi-omics changes usually appear in the process of lncRNA participation in drug responses. For example, H3K27me3 induces multidrug resistance in small cell
lung cancer by affecting HOXA1 DNA methylation via regulation of the IncRNA HOTAIR\textsuperscript{33}. The IncRNA UCA1 is physically associated with the enhancer of zeste homolog 2, which suppresses the expression of p21 through histone methylation (H3K27me3) on the p21 promoter\textsuperscript{34}. Silencing IncRNA COMET markedly increases sensitivity to vemurafenib, a common inhibitor of mutated B-raf in BRAF- and RET-driven papillary thyroid carcinomas\textsuperscript{36}. The CT genotype of rs3200401 MALAT1 polymorphism could serve as a toxicity biomarker\textsuperscript{36}. LncRNA SNHG6 promotes chemoresistance through ULK1-induced autophagy by sponging miR-26a-5p in colorectal cancer cells\textsuperscript{37}. LncRNA H19 overexpression induces bortezomib resistance in multiple myeloma by targeting MCL-1 via miR-29b-3p\textsuperscript{33}. Multi-omic changes including those of the genome, transcriptome, and epigenome commonly contribute to drug responses and resistance in cancer. These results indicate that the multi-omics integration approach is an effective way to identify DRlncs in BRCA.

In the present study, a key DRlnc, HOXA-AS2, was identified in adriamycin groups. Furthermore, we confirmed this finding by using cell lines and specimens of BRCA, with real-time quantitative PCR and colony formation assays. HOXA-AS2 is a lncRNA located between the HOXA3 and HOXA4 genes in the HOXA cluster. Recent studies indicated that HOXA-AS2 represses apoptosis in NB4 promyelocytic leukemia cells treated with transretinoic acid\textsuperscript{38}, promotes the proliferation of gastric cancer via
P21/PLK3/DDIT3[39], and accelerates tumorigenesis of hepatocellular carcinomas[27]. The lncRNA, HOXA-AS2 promotes proliferation and invasion of BRCA by acting as a miR-520c-3p sponge[40]. Subsequent studies suggested that HOXA-AS2 was significantly overexpressed in a variety of tumors and was associated with poor prognoses of these tumors. The role of HOXA-AS2 in anticancer drug resistance has been confirmed in acute myeloid leukemia. A study showed that HOXA-AS2 was significantly upregulated in bone marrow samples from acute myeloid leukemia patients after treatment with adriamycin-based chemotherapy[28]. In our study, we also hypothesized that lncRNA HOXA-AS2 may contribute to the adriamycin response in BRCA patients.

This study opens a new avenue to leverage publicly available molecular data to evaluate clinical drug responses, and contributes to realizing personalized medicine. We propose that the significant DRlncs may be regarded as new markers for drug resistance, which indicates a possible novel mechanism in BRCA. In the process of clinical treatment, DRlncs could act as specific biomarkers for use in deciding chemotherapy and evaluating treatment effects. In future studies, diverse BRCA subtypes, stages, and other clinical characteristics should be distinguished to identify DRlncs. More DRlncs should be validated by studies of tumor tissues, model animals, and cell lines. Taken together, identification of significant DRlncs will provide more chances for development of BRCA therapeutics and clinical drug use, as well as a better
understanding of drug resistance mechanisms.

MATERIALS AND METHODS

Molecular expression and clinical information of BRCA patients

Large-scale mRNA, IncRNA, and miRNA expression profiles (Illumina HiSeq (Illumina, San Diego, CA, USA) 3), DNA methylation level profiles (Illumina Infinium HumanMethylation450 level 3), somatic mutation data (Illumina), clinical follow-up survival time, and clinical drug treatment records of cancer patients were obtained from the cancer genome atlas (TCGA) data portal (TCGA, Release: 07-18-2019, https://xenabrowser.net/datapages/?cohort=TCGA%20Breast%20Cancer%20(BRCA)&removeHub=https%3A%2F%2Fxena.treehouse.gi.ucsc.edu%3A443).

To filter genes, IncRNAs, miRNAs, and methylations not expressed across all BRCA samples, the items with expression values of 0 in all BRCA samples were excluded. To allow log transformation, any remaining expression values of 0 were set to the minimum value of all samples and all values were log₂ transformed. Log₂ transformations were performed for genes, IncRNAs, miRNAs, and methylations. The somatic mutations that were annotated as confirmed SNPs were screened out.

Distinguishing drug sensitive and resistant BRCA patients

All molecular information and clinical survival data of BRCA patients with drug
treatment were retained for subsequent analysis. BRCA patients with “complete response” were considered as the drug sensitive group. BRCA patients with “stable disease,” “clinical progressive disease,” and “partial response” were considered as the drug resistant group. According to sample numbers, adriamycin, cytoxan, and tamoxifen were extracted as candidate drugs to follow the analyses. In addition, all BRCA patients with indiscriminate use of drugs were also considered.

**Method workflow overview**

The workflow was divided into three phases. First, the *t*-test was used to identify the differential expression levels of lncRNAs, genes, miRNAs, and methylations between sensitive and resistant BRCA patients. Differentially expressed lncRNAs ($P < 0.05$) were considered candidate lncRNAs for follow-up analysis. The numbers of somatic mutations in drug sensitive and resistant samples were counted. BEDTools was used to map all somatic mutations and methylation sites to lncRNAs. The experimentally validated gene-lncRNA and miRNA-lncRNA interactions were all obtained from the RAID 2.0 database (www.rna-society.org/raid/). Only gene-lncRNA and miRNA-lncRNA interactions supported by strong experiments were extracted. Second, DRlncs were identified by integrating multi-omics scores including scores of genes, miRNAs, methylations, and somatic mutations based on a multi-dimensional rank method. Finally, 1,000 randomly permuting multi-omics
scores were performed to compare the final multi-dimensional rank risk scores with permutation risk scores to obtain significant results. \( P < 0.05 \) of the permutation results was selected as a threshold value to generate significantly DRlncs for BRCA.

**Identification of DRlncs for diverse chemotherapeutic drugs in BRCA**

An integrated computational workflow was designed to identify DRlncs for diverse chemotherapeutic drugs in BRCA. First, corresponding genes, miRNAs, methylation sites, and somatic mutations for all candidate differential expressed IncRNAs of diverse chemotherapeutic were extracted. The individual scores were represented by their P-values. The scores were represented by minimum values if there were a number of corresponding genes, miRNAs, or methylation sites for the same IncRNA. Second, somatic mutation scores were calculated as follows:

\[
S_{\text{mut}} = \frac{1}{|N_{\text{sensitive}} - N_{\text{resistant}}|}
\]

Where \( S_{\text{mut}} \) represented the score of somatic mutation. \( N_{\text{sensitive}} \) and \( N_{\text{resistant}} \) represented the number of somatic mutations for a IncRNA in sensitive and resistant BRCA patients. Third, an equally-weighted multiple ranking approach was conducted based on four risk scores including \( S_{\text{gene}}, S_{\text{miRNA}}, S_{\text{meth}}, \) and \( S_{\text{mut}} \) for each IncRNA to generate the last risk score. After this step, each IncRNA was assigned a final risk score and ranked by these final risk scores. The IncRNAs were considered as DRlncs if their permutation P-values were
smaller than 0.05. Thus, some DRlnCs were discovered for diverse chemotherapeutic drugs in BRCA.

**Construction of a DRInc-gene co-expression network for diverse chemotherapeutic drugs in BRCA**

Pearson correlation coefficients were calculated for each interacting DRInc-gene pair based on gene and IncRNA expression profiles for diverse chemotherapeutic drugs in BRCA. The significantly co-expressed (P < 0.05) DRInc-genes were extracted for constructing networks. The co-expressed networks for diverse chemotherapeutic drugs in BRCA were constructed using Cytoscape 3.3.0 (http://www.cytoscape.org/). In addition, the degree analysis and R-square score was also performed by Cytoscape.

**Survival analysis of DRlnCs in BRCA patients**

For each DRInc of chemotherapeutic drugs in BRCA, we classified samples into two groups based on median values of IncRNA expression values. Kaplan-Meier survival analysis was performed for the two clustered groups, and statistical significance assessed using the log-rank test. P< 0.05 was regarded as statistically significant. All analyses were performed within the R 3.2.3 framework (https://www.r-project.org).

**Functional enrichment analysis for DRlnCs for diverse chemotherapeutic**
drugs in BRCA

DRlncs in diverse chemotherapeutic drugs were selected for GO functional enrichment analysis. Online Enrichr tool (http://amp.pharm.mssm.edu/Enrichr/) was used with default parameters for functional enrichment\(^4^4\). We identified enriched GO terms (P < 0.01) for diverse chemotherapeutic drugs in BRCA.

Cell lines and BRCA specimens

The BRCA cell lines (T47D and MDA-MB-231) were obtained from the Chinese Type Culture Collection, Chinese Academy of Science (Beijing, China). T47D cells were maintained in Dulbecco’s Minimal Eagle Medium supplemented with 10% fetal bovine serum. MDA-MB-231 cells were cultured in Leibovitz’s L-15 Medium containing 10% fetal bovine serum. All cells were cultured in a humidified atmosphere containing 5% CO\(_2\) and at 37°C.

Fresh BRCA tissues were obtained from 60 patients who received no anticancer treatment before undergoing surgery in the Department of General Surgery of the Second Affiliated Hospital of Harbin Medical University, Harbin, China between January 2012 and January 2014. Specimens from another 60 patients who received neoadjuvant chemotherapy before surgery were also collected. All patients provided consent according to the ethical standards of the Declaration of Helsinki. The study was approved by the Ethics and Scientific committee of Harbin Medical University. All specimens were frozen at −80°C after surgical resection.
Real-time quantitative PCR

Total RNA was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA, USA). The cDNA was synthesized under standard conditions. Real-time PCR was performed by using the SYBR Green kit (Thermo Fisher Scientific, Waltham, MA, USA) following the manufacturer’s instructions. The HOXA-AS2 expression was normalized to 18S tRNA for each sample. The primers for real-time quantitative PCR detection are listed as follows: HOXA-AS2, 5’-CCCGTAGGAAGAACCGATGA-3’ (forward) and 5’-TTTAGGCTTTCGCAGCAGC-3’ (reverse); 18S rRNA, 5’-GTAACCCGTTGAACCCATT-3’ (forward) and 5’-CCATCCAATCGGTAGGCG-3’ (reverse). The real-time quantitative PCR reactions were repeated three times. Relative RNA expression was calculated using the $\Delta\Delta Ct$ method.

Colony formation assay

Indicated BRCA cells were seeded in 6-well plates at the density of 1,000 cells/well and cultured in medium for 2 weeks. The cells were then washed with phosphate-buffered saline and fixed in methanol for 20 min. Crystal Violet was used to stain the cells for 15 min, after which the cells were photographed using a digital camera.

Disclosure statement

The authors declare no competing financial interests
Funding

This work was supported by the National Natural Science foundation of China under Grant No. 81701705, 81630048, 81974265 and 81901761. Postdoctoral Funding of Heilongjiang province under Grant No. LBH-Z17174. Scientific research project of Heilongjiang Health Committee No. 2019-050.
REFERENCES

1. Siegel, RL, Miller, KD, and Jemal, A (2018). Cancer statistics, 2018. CA Cancer J Clin 68: 7-30.
2. DeSantis, CE, Ma, J, Goding Sauer, A, Newman, LA, and Jemal, A (2017). Breast cancer statistics, 2017, racial disparity in mortality by state. CA Cancer J Clin 67: 439-448.
3. DeSantis, CE, Ma, J, Gaudet, MM, Newman, LA, Miller, KD, Goding Sauer, A, et al. (2019). Breast cancer statistics, 2019. CA Cancer J Clin 69: 438-451.
4. Rodrigues-Ferreira, S, Nehlig, A, Moinджie, H, Monchecourt, C, Seiler, C, Marangoni, E, et al. (2019). Improving breast cancer sensitivity to paclitaxel by increasing aneuploidy. Proc Natl Acad Sci U S A 116: 23691-23697.
5. Nounou, MI, ElAmrawy, F, Ahmed, N, Abdelraouf, K, Goda, S, and Syed-Sha-Qhattal, H (2015). Breast Cancer: Conventional Diagnosis and Treatment Modalities and Recent Patents and Technologies. Breast Cancer (Auckl) 9: 17-34.
6. Jayaraj, R, Nayagam, SG, Kar, A, Sathyakumar, S, Mohammed, H, Smiti, M, et al. (2019). Clinical Theragnostic Relationship between Drug-Resistance Specific miRNA Expressions, Chemotherapeutic Resistance, and Sensitivity in Breast Cancer: A Systematic Review and Meta-Analysis. Cells 8.
7. Vulsteke, C, Lambrechts, D, Dieudonne, A, Hatse, S, Brouwers, B, van Brussel, T, et al. (2013). Genetic variability in the multidrug resistance associated protein-1 (ABCC1/MRP1) predicts hematological toxicity in breast cancer patients receiving (neo-)adjuvant chemotherapy with 5-fluorouracil, epirubicin and cyclophosphamide (FEC). Ann Oncol 24: 1513-1525.
8. Liu, Q, Borcherding, NC, Shao, P, Maina, PK, Zhang, W, and Qi, HH (2020). Contribution of synergism between PHF8 and HER2 signalling to breast cancer development and drug resistance. EBioMedicine 51: 102612.
9. Cabili, MN, Trapnell, C, Goff, L, Koziol, M, Tazon-Vega, B, Regev, A, et al. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25: 1915-1927.
10. Guttmann, M, Garber, M, Levin, JZ, Donaghey, J, Robinson, J, Adiconis, X, et al. (2010). Ab initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. Nat Biotechnol 28: 503-510.
11. Gao, Y, Wang, P, Wang, Y, Ma, X, Zhi, H, Zhou, D, et al. (2019). Lnc2Cancer v2.0: updated database of experimentally supported long non-coding RNAs in human cancers. Nucleic Acids Res 47: D1028-D1033.
12. Huang, QY, Liu, GF, Qian, XL, Tang, LB, Huang, QY, and Xiong, LX (2019). Long Non-Coding RNA: Dual Effects on Breast Cancer Metastasis and Clinical Applications. Cancers (Basel) 11.
13. Zhang, N, Zeng, X, Sun, C, Guo, H, Wang, T, Wei, L, et al. (2019). LncRNA LINC00963 Promotes Tumorigenesis and Radioresistance in Breast Cancer by Sponging miR-324-3p and Inducing ACK1 Expression. Mol Ther Nucleic Acids 18: 871-881.
14. Gupta, RA, Shah, N, Wang, KC, Kim, J, Horlings, HM, Wong, DJ, et al. (2010). Long non-coding
RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**: 1071-1076.

15. Mourtada-Maarabouni, M, Pickard, MR, Hedge, VL, Farzaneh, F, and Williams, GT (2009). GASS, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **28**: 195-208.

16. Chi, JR, Yu, ZH, Liu, BW, Zhang, D, Ge, J, Yu, Y, et al. (2019). SNHG5 Promotes Breast Cancer Proliferation by Sponging the miR-154-5p/PCNA Axis. *Mol Ther Nucleic Acids* **17**: 138-149.

17. Fernando, TR, Rodriguez-Malave, NI, Waters, EV, Yan, W, Casero, D, Basso, G, et al. (2015). LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia. *Mol Cancer Res* **13**: 839-851.

18. Diaz-Lagares, A, Crujeiras, AB, Lopez-Serra, P, Soler, M, Setien, F, Goyal, A, et al. (2016). Epigenetic inactivation of the p53-induced long noncoding RNA TP53 target 1 in human cancer. *Proc Natl Acad Sci U S A* **113**: E7535-E7544.

19. Jiang, M, Huang, Q, Xie, Z, Wu, S, Zhang, X, Shen, A, et al. (2014). A novel long non-coding RNA-ARA: adriamycin resistance-associated. *Biochem Pharmacol* **87**: 254-283.

20. Guttman, M, and Rinn, JL (2012). Modular regulatory principles of large non-coding RNAs. *Nature* **482**: 339-346.

21. Sumazin, P, Yang, X, Chiu, HS, Chung, WJ, Iyer, A, Llobet-Navas, D, et al. (2011). An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. *Cell* **147**: 370-381.

22. Zhi, H, Ning, S, Li, X, Li, Y, Wu, W, and Li, X (2014). A novel reannotation strategy for dissecting DNA methylation patterns of human long intergenic non-coding RNAs in cancers. *Nucleic Acids Res* **42**: 8258-8270.

23. Ning, S, Yue, M, Wang, P, Liu, Y, Zhi, H, Zhang, Y, et al. (2017). LincSNP 2.0: an updated database for linking disease-associated SNPs to human long non-coding RNAs and their TFBSs. *Nucleic Acids Res* **45**: D74-D78.

24. Khurana, E, Fu, Y, Chakravarty, D, Demichelis, F, Rubin, MA, and Gerstein, M (2016). Role of non-coding sequence variants in cancer. *Nat Rev Genet* **17**: 93-108.

25. Wang, F, Wu, D, Chen, J, Chen, S, He, F, Fu, H, et al. (2019). Long non-coding RNA HOXA-AS2 promotes the migration, invasion and stemness of bladder cancer via regulating miR-125b/Smad2 axis. *Exp Cell Res* **375**: 1-10.

26. Wu, J, Li, M, and Zhang, Y (2019). Long noncoding RNA HOXA-AS2 regulates the expression of SCN3A by sponging miR-106a in breast cancer. *J Cell Biochem* **120**: 14465-14475.

27. Wang, F, Yang, H, Deng, Z, Su, Y, Fang, Q, and Yin, Z (2016). HOX Antisense lincRNA HOXA-AS2 Promotes Tumorigenesis of Hepatocellular Carcinoma. *Cell Physiol Biochem* **40**: 287-296.

28. Dong, X, Fang, Z, Yu, M, Zhang, L, Xiao, R, Li, X, et al. (2018). Knockdown of Long Noncoding RNA HOXA-AS2 Suppresses Chemoresistance of Acute Myeloid Leukemia via the miR-520c-3p/S100A4 Axis. *Cell Physiol Biochem* **51**: 886-896.

29. Ikeda, K, Horie-Inoue, K, Ueno, T, Suzuki, T, Sato, W, Shigekawa, T, et al. (2015). miR-378a-3p modulates tamoxifen sensitivity in breast cancer MCF-7 cells through targeting GOLT1A. *Sci Rep* **5**: 13170.

30. Majidinia, M, and Yousefi, B (2016). Long non-coding RNAs in cancer drug resistance development. *DNA Repair (Amst)* **45**: 25-33.

31. Nath, A, Lau, EYT, Lee, AM, Geeleher, P, Cho, WCS, and Huang, RS (2019). Discovering long
noncoding RNA predictors of anticancer drug sensitivity beyond protein-coding genes. *Proc Natl Acad Sci U S A.*

32. Yao, N, Fu, Y, Chen, L, Liu, Z, He, J, Zhu, Y, et al. (2019). Long non-coding RNA NONHSAT101069 promotes epirubicin resistance, migration, and invasion of breast cancer cells through NONHSAT101069/miR-129-5p/Twist1 axis. *Oncoogene.*

33. Fang, S, Shen, Y, Chen, B, Wu, Y, Jia, L, Li, Y, et al. (2018). H3K27me3 induces multidrug resistance in small cell lung cancer by affecting HOXA1 DNA methylation via regulation of the IncRNA HOTAIR. *Ann Transl Med* 6: 440.

34. Li, Z, Yu, D, Li, H, Lv, Y, and Li, S (2019). Long noncoding RNA UCA1 confers tamoxifen resistance in breast cancer endocrinotherapy through regulation of the EZH2/p21 axis and the PI3K/AKT signaling pathway. *Int J Oncol* 54: 1033-1042.

35. Esposito, R, Esposito, D, Pallante, P, Fusco, A, Ciccodicola, A, and Costa, V (2019). Oncogenic Properties of the Antisense IncRNA COMET in BRAF- and RET-Driven Papillary Thyroid Carcinomas. *Cancer Res* 79: 2124-2135.

36. Lampropoulou, DI, Aravantinos, G, Katifelis, H, Lazaris, F, Laschos, K, Theodosopoulos, T, et al. (2019). Long non-coding RNA polymorphisms and prediction of response to chemotherapy based on irinotecan in patients with metastatic colorectal cancer. *Cancer Biomark* 25: 213-221.

37. Wang, X, Lan, Z, He, J, Lai, Q, Yao, X, Li, Q, et al. (2019). LncRNA SNHG6 promotes chemoresistance through ULK1-induced autophagy by sponging miR-26a-5p in colorectal cancer cells. *Cancer Cell Int* 19: 234.

38. Zhao, H, Zhang, X, Frazao, JB, Condino-Neto, A, and Newburger, PE (2013). HOX antisense lincRNA HOXA-AS2 is an apoptosis repressor in all trans retinoic acid treated NB4 promyelocytic leukemia cells. *J Cell Biochem* 114: 2375-2383.

39. Xie, M, Sun, M, Zhu, YN, Xia, R, Liu, YW, Ding, J, et al. (2015). Long noncoding RNA HOXA-AS2 promotes gastric cancer proliferation by epigenetically silencing P21/PLK3/DDIT3 expression. *Oncotarget* 6: 33587-33601.

40. Fang, Y, Wang, J, Wu, F, Song, Y, Zhao, S, and Zhang, Q (2017). Long non-coding RNA HOXA-AS2 promotes proliferation and invasion of breast cancer by acting as a miR-520c-3p sponge. *Oncotarget* 8: 46090-46103.

41. Quinlan, AR, and Hall, IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26: 841-842.

42. Yi, Y, Zhao, Y, Li, C, Zhang, L, Huang, H, Li, Y, et al. (2017). RAID v2.0: an updated resource of RNA-associated interactions across organisms. *Nucleic Acids Res* 45: D115-D118.

43. Aerts, S, Lambrechts, D, Maity, S, Van Loo, P, Coessens, B, De Smet, F, et al. (2006). Gene prioritization through genomic data fusion. *Nat Biotechnol* 24: 537-544.

44. Kuleshov, MV, Jones, MR, Rouillard, AD, Fernandez, NF, Duan, Q, Wang, Z, et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 44: W90-97.
**FIGURE LEGENDS:**

**Figure 1. Identification of individual DRlncs based on multi-omic data integration for multiple drugs in BRCA patients.** (A) The bar plot shows the number of DRlncs for multiple drugs in BRCA patients. (B) The point plot shows the DRlncs for adriamycin. The larger circles represent more significant P values of DRlncs. The x- and y-axis represent DRlncs and the number of factors. (C) The point plot shows the DRlncs for Tamoxifen. (D) The pie charts show the percent of diverse number factors for multiple drugs. (E) The heat map shows common DRlncs. Dark and light green represent DRlncs, which were influenced by three or two factors for multiple drugs, respectively. (F) The radar chart shows the numbers of methylation sites, genes, and mutations of DRlnc AC008443.

**Figure 2. Individual DRlncs showed specific features for diverse drugs in BRCA patients.** (A) The bar plot shows the number of up- and downregulated DRlncs for diverse drugs. (B, C, D, E) The density distribution curves show the distribution of scores of DRlncs for methylations (green), mutations (red), miRNAs (yellow), and genes (blue) with (B) adriamycin, (C) cytoxan, (D) tamoxifen, and (E) all samples. (F) The density distribution curves show the distribution of the last scores of DRlncs for adriamycin (green), cytoxan (yellow), tamoxifen (blue), and all samples (red). (G) The box plots show the expressions of HOXA-AS2 and HNRNPA2BA for adriamycin sensitive (green) and resistant (red) BRCA patients. (H) The point plot shows the co-expression between HOXA-AS2 and HNRNPA2BA. (I) The heat map shows the expressions of HOXA-AS2
and HNRNPA2BA. Red and blue represent high and low expression levels, respectively.

**Figure 3.** The DRlnc-gene co-expression networks related to drug responses in breast cancer. (A) The bar plot shows the number of positive and negative correlated interactions for diverse drugs. (B) The co-expression DRlnc-gene network for all drugs including adriamycin, cytoxan, tamoxifen, and all samples. The yellow circle and green triangle represent DRlnc and genes, respectively. The red and blue edges represent positive and negative interactions between DRlnc and genes, respectively. The thickness of edges represent correlation levels between DRlnc and genes. (C) The plots show the degree distribution of the total DRlnc-gene co-expression network. (D) The DRlnc-gene co-expression network for adriamycin. (E) The heat map shows the expression level of DRlnc AC073611 and its interacting genes including CAPR1N1, UPF1, and CSTF2 (F, G, H). The DRlnc-gene co-expression network for cytoxan, tamoxifen, and all samples.

**Figure 4.** Some DRlncs were associated with survival of breast cancer patients. (A) The Kaplan-Meier curves for the overall survival of two DRlnc groups with high and low risk expressions. The difference between the two curves was evaluated by a two-sided log-rank test. (B) The expression distribution of the DRlnc in each drug. (C) The patient survival status of the DRlnc for diverse drugs in BRCA patients.

**Figure 5.** The functions and mechanism of DRlncs for diverse drugs in BRCA patients. (A, B, C, D) Bar plots show the P-values of Gene Ontology enrichment terms for DRlncs for adriamycin (A), cytoxan (B), tamoxifen (C), and all samples (D). (E) The possible mechanism of DRlnc HOXA-AS2 in drug responses for BRCA patients.

**Figure 6.** HOXA-AS2 promotes the adriamycin resistance of BRCA cells. (A)
real-time quantitative PCR was used to detect the expression of HOXA-AS2 in BRCA tissues from patients with or without adriamycin-based neoadjuvant chemotherapy before surgery. (B) The expression of HOXA-AS2 was examined by real-time PCR in normal or adriamycin-resistant BRCA cells. (C) The efficiency of HOXA-AS2 knockdown by shRNAs was confirmed in BRCA cells. (D) With 5 mg/L adriamycin treatment, colony formation assays were performed to analyze the proliferative ability of normal, adriamycin-resistant, or HOXA-AS2-knockdown adriamycin-resistant BRCA cells.
A. vesicle targeting, rough ER to cis-Golgi
   vesicle coating
   COPII vesicle coating
   COPII-coated vesicle budding
   retrograde vesicle-mediated transport, Golgi to ER
   Golgi vesicle transport
   ER to Golgi vesicle-mediated transport
   protein complex assembly
   regulation of osteoblast differentiation
   Golgi transport vesicle coating

B. synapse assembly
   calcium-dependent cell-cell adhesion
   anterograde trans-synaptic signaling
   chemical synaptic transmission
   nervous system development
   cell-cell adhesion
   Golgi vesicle transport
   ER to Golgi vesicle-mediated transport
   COPII vesicle coating
   vesicle targeting, rough ER to cis-Golgi

C. negative regulation of RNA metabolic process
   negative regulation of protein tyrosine kinase activity
   COPII vesicle coating
   vesicle targeting, rough ER to cis-Golgi
   vesicle coating
   COPII-coated vesicle budding
   neutrophil extravasation
   negative regulation of production of miRNAs
   granulocyte migration
   outer mitochondrial membrane organization

D. positive regulation of receptor binding
   positive regulation of keratinocyte migration
   hormone transport
   ion homeostasis
   cellular ion homeostasis
   cellular monovalent inorganic anion homeostasis
   regulation of keratinocyte migration
   regulation of receptor binding
   regulation of cysteine-type endopeptidase activity
   chloride ion homeostasis

E. Adriamycin resistant BRCA samples
   ELF4A3
   FMR1
   HNRNPA2B1
   cg00805454
   mir-107
   IncRNA HOXA-AS2
   Low expression of genes and weaken the inhibition of IncRNA HOXA-AS2
   With the change of multi-omics
   HOXA-AS2 high expression
   Drug resistant of Adriamycin in breast cancer
