Inferences of Individual Drug Response-Related Long Non-coding RNAs Based on Integrating Multi-omics Data in Breast Cancer

Hao Cui,[1] Hanqing Kong,[1] Fuhui Peng,[1] Chunjing Wang,[2] Dandan Zhang,[3] Jiawei Tian,[1] and Lei Zhang[1]

[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

Differences in individual drug responses are obstacles 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 (DRlncs) by combining expression data of lncRNA, microRNA, messenger RNA, methylation levels, 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 DRlnc-gene co-expression networks were constructed and analyzed. Key DRlncs, such as HOXA-AS2 (Ensembl: ENSG00000253552), in the drug Adriamycin were characterized. The experimental analysis also suggested that HOXA-AS2 (Ensembl: ENSG00000253552) 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 (Ensembl: ENSG00000253552) in the Adriamycin drug response for BRCA resistance was inferred. In summary, this study provides a framework for lncRNA-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,[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 ma-

jor obstacle for successful treatment.[4–6] Drug resistance varies among different individuals,[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.[9–11] The lncRNAs have dual effects in BRCA metastasis by regulating invasion, migration, and distant metastasis of BRCA cells.[12] For example, LINC00963 (Ensembl: ENSG00000204054) promotes tumorigenesis and radioresistance in BRCA, and represents a potential target for the treatment of BRCA.[13] The lncRNA termed HOTAIR (Ensembl: ENSG00000228630) 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.[14] The transcript levels of GAS5 (Ensembl: ENSG00000234741), another lncRNA, were significantly reduced in BRCA, by controlling apoptosis and downregulation.[15] SNHG5 (Ensembl: ENSG00000203875) promoted BRCA proliferation and cell-cycle progression by upregulation of PCNA (Ensembl: ENSG00000132646) expression.[16]

Recently, some studies have shown that lncRNAs 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.[17] reported that high expression of lncRNA BALR-2 (Ensembl: ENSG00000237819) was related to prednisolone resistance in acute lymphocytic leukemia (ALL), and knocking out BALR-2 (Ensembl: ENSG00000237819) improved the sensitivity of B-ALL to chemotherapy drugs. The functional loss of lncRNA TP53TG1 (Ensembl: ENSG00000182165) led to abnormal

activity of the YBX1 (Ensembl: ENSG00000065978) 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).[18] Furthermore, some lncRNAs 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.[19] However, there still urgently need to systematically develop efficient and reliable algorithms to predict the drug response-related lncRNAs (DRlncs) based on integrating data in BRCA patients.

lncRNAs cannot function independently, but rather interact with other molecules and coordinate with each other to play roles in physiological and pathological processes.[20] 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.[21] Some reports have shown that network analysis of aberrantly methylated lncRNAs in cancers indicates that lncRNAs with aberrant methylation patterns might be involved in cancer development and progression.[22] Ning et al.[23] developed a database to identify disease-associated single-nucleotide polymorphisms (SNPs) in human lncRNAs. Khurana et al.[24] 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, and 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 were also characterized. In addition, some experiments were performed to confirm that HOXA-AS2 (Ensembl: ENSG00000253552) 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 lncRNA discovery.

RESULTS

Identification of Individual DRlncs Based on Multi-omics 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 (Ensembl: ENSG00000241048), AL691403 (Ensembl: ENSG00000258740), BX324167 (Ensembl: ENSG00000235091), and HOX-AS2 (Ensembl: ENSG00000253552) all showed more significant results. DRlncs AC117386 (Ensembl: ENSG00000241048), AL691403 (Ensembl: ENSG00000258740), and BX324167 (Ensembl: ENSG00000235091) were influenced by a single factor; however, HOX-AS2 (Ensembl: ENSG00000253552) 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 was different for diverse drugs. For Adriamycin, DRlncs, which were influenced by two kinds of factor, 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 lncRNAs, including AC010680 (Ensembl: ENSG00000267784), AC073611 (Ensembl: ENSG00000257605), AP001107 (Ensembl: ENSG00000245156), and AC008443 (Ensembl: ENSG00000233937), 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 (Ensembl: ENSG00000233937) was influenced by 14 genes, 32 methylation sites, and 4 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 drug (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 drug (Figures 2B–2E). 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 (Ensembl: ENSG00000253552), an lncRNA located between the HOXA3 (Ensembl: ENSG00000105997) and HOXA4 (Ensembl: ENSG00000197576) genes in the HOXA cluster, has been characterized as an oncogene in many kinds of cancer, including BRCA.[25–27] Specifically, HOXA-AS2 (Ensembl: ENSG00000253552) plays an important role in the resistance of acute myeloid leukemia cells to Adriamycin.[28] In our analysis, HOXA-AS2 (Ensembl: ENSG00000253552) was identified as a DRlnc in Adriamycin (p < 0.001). HOXA-AS2 (Ensembl: ENSG00000253552) was upregulated in Adriamycin-resistant BRCA samples (Figure 2G). The HNRNPA2B1 (Ensembl: ENSG00000122566) gene, belonging to the A/B subfamily of ubiquitously expressed heterogeneous nuclear ribonucleoproteins, was downregulated in Adriamycin-resistant BRCA samples (Figure 2G). HOXA-AS2
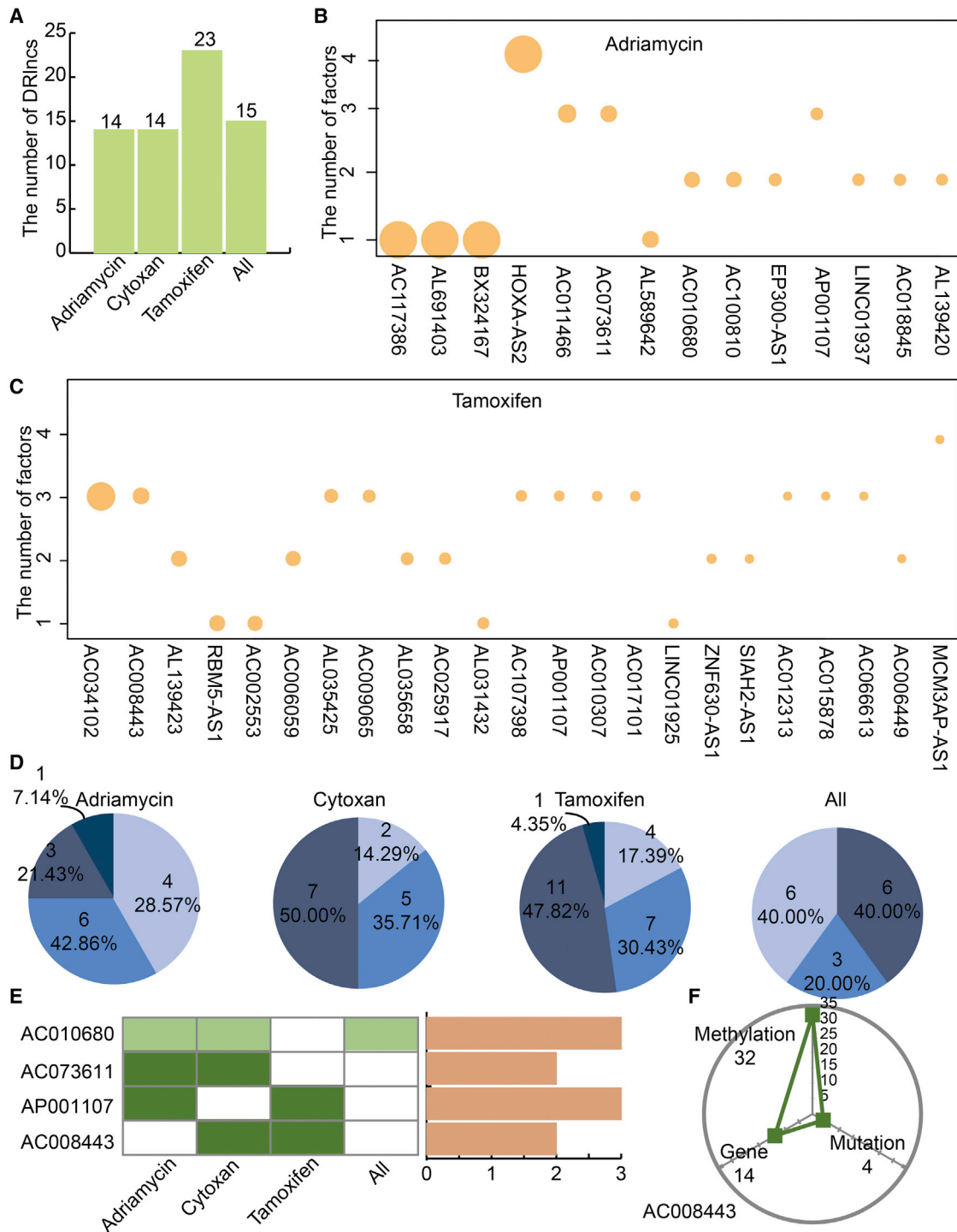
Figure 1. Identification of Individual DRlncs Based on Multi-omics Data Integration for Multiple Drugs in BRCA Patients

(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 axes 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 of factors for multiple drugs. (E) The heatmap 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 (Ensembl: ENSG00000233937).
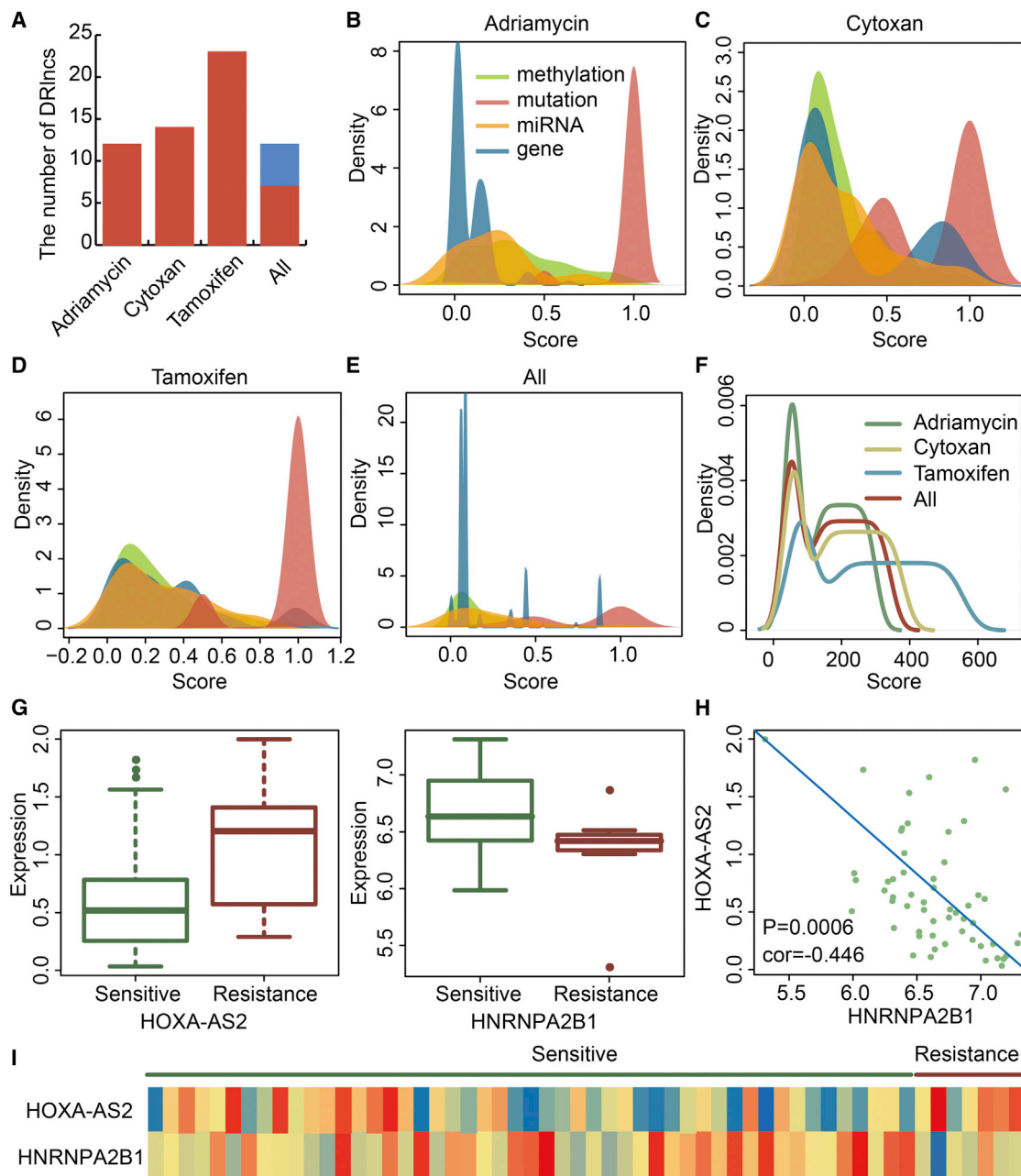
Figure 2. Individual DRlncs Showed Specific Features for Diverse Drugs in BRCA Patients

(A) The bar plot shows the number of upregulated and downregulated DRlncs for diverse drugs. (B–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 boxplots show the expressions of HOXA-AS2 (Ensembl: ENSG00000253552) and HNRNPA2B1 (Ensembl: ENSG00000122566) for Adriamycin-sensitive (green) and -resistant (red) BRCA patients. (H) The point plot shows the co-expression between HOXA-AS2 (Ensembl: ENSG00000253552) and HNRNPA2B1 (Ensembl: ENSG00000122566). (I) The heatmap shows the expressions of HOXA-AS2 (Ensembl: ENSG00000253552) and HNRNPA2B1 (Ensembl: ENSG00000122566). Red and blue represent high and low expression levels, respectively.

(Ensembl: ENSG00000253552) and HNRNPA2B1 (Ensembl: ENSG00000122566) were negatively correlated (p = 0.0006, correlation = −0.446) in BRCA samples treated with Adriamycin (Figures 2H and 2I). All of 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 (Ensembl: ENSG00000253552) and HNRNPA2B1 (Ensembl: ENSG00000122566) 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 (Ensembl: ENSG00000245667), HOXA-AS2 (Ensembl: ENSG00000253552), and AC008443 (Ensembl: ENSG00000233937), 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 DRlnc-gene co-expression network for Adriamycin was constructed (Figure 3D). DRlncs HOXA-AS2 (Ensembl: ENSG00000253552) and AC073611 (Ensembl: ENSG00000257605) could be considered as two key hubs in the Adriamycin-related DRlnc-gene co-expression network. DRlncs were negatively correlated with the CAPRIN1 (Ensembl: ENSG00000135387), UPF1 (Ensembl: ENSG00000005007), and CSTF2 (Ensembl: ENSG00000101811) genes (Figure 3E). In addition, DRlnc-gene co-expression networks for Cytoxan, Tamoxifen, and all samples were also constructed (Figures 3F–3H). All of the above results indicated that the DRlnc-gene co-expression networks may help to understand the mechanism of DRlnc for diverse drugs in BRCA.

Some DRlncs Were Associated with Survival of BRCA Patients

To evaluate the potential value of DRlncs as prognostic biomarkers for BRCA, we determined whether the DRlncs in diverse drugs were associated with survival. We used median values of DRlnc expressions as the cutoff to test the survival of BRCA patients. DRlnc BX322234 (Ensembl: ENSG00000226445) in Cytoxan, AC006059 (Ensembl: ENSG00000230084) in Tamoxifen, and AL158151 (Ensembl: ENSG00000204055) and UBXN10-AS1 (Ensembl: ENSG00000225986) in all samples were associated with survival (Figures 4A–4D). 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 DRlncs except AC006059 (Ensembl: ENSG00000230084) for Tamoxifen. The results indicated that some DRlncs 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 coat protein complex II (COPII) vesicle coating related with GO terms (Figures 5A–5D). 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.[29] 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 (Ensembl: ENSG00000253552) was a key DRlnc in Adriamycin for BRCA patients (Figure 5E). DRlnc HOXA-AS2 (Ensembl: ENSG00000253552) was regulated by three important genes, including EIF4A3 (Ensembl: ENSG00000141543), FMR1 (Ensembl: ENSG00000102081), and HNRNPA2B1 (Ensembl: ENSG00000122566). These three genes were downregulated in Adriamycin-resistant BRCA patients, to promote overexpression of DRlnc HOXA-AS2 (Ensembl: ENSG00000253552), 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 necessity 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 (Ensembl: ENSG00000253552) 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 (Ensembl: ENSG00000253552) expressions in samples from patients with Adriamycin-based neoadjuvant chemotherapy were significantly higher (Figure 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 (Figures 6B and 6C). To characterize the function of HOXA-AS2 (Ensembl: ENSG00000253552) in BRCA cells, we detected the efficiency of HOXA-AS2 (Ensembl: ENSG00000253552) knockdown in BRCA cells by real-time PCR assays (Figure 6D). Figure 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 (Ensembl: ENSG00000253552) silencing reversed the colony formation ability of Adriamycin-resistant BRCA cells (Figure 6E). These results demonstrated that HOXA-AS2 (Ensembl: ENSG00000253552) 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
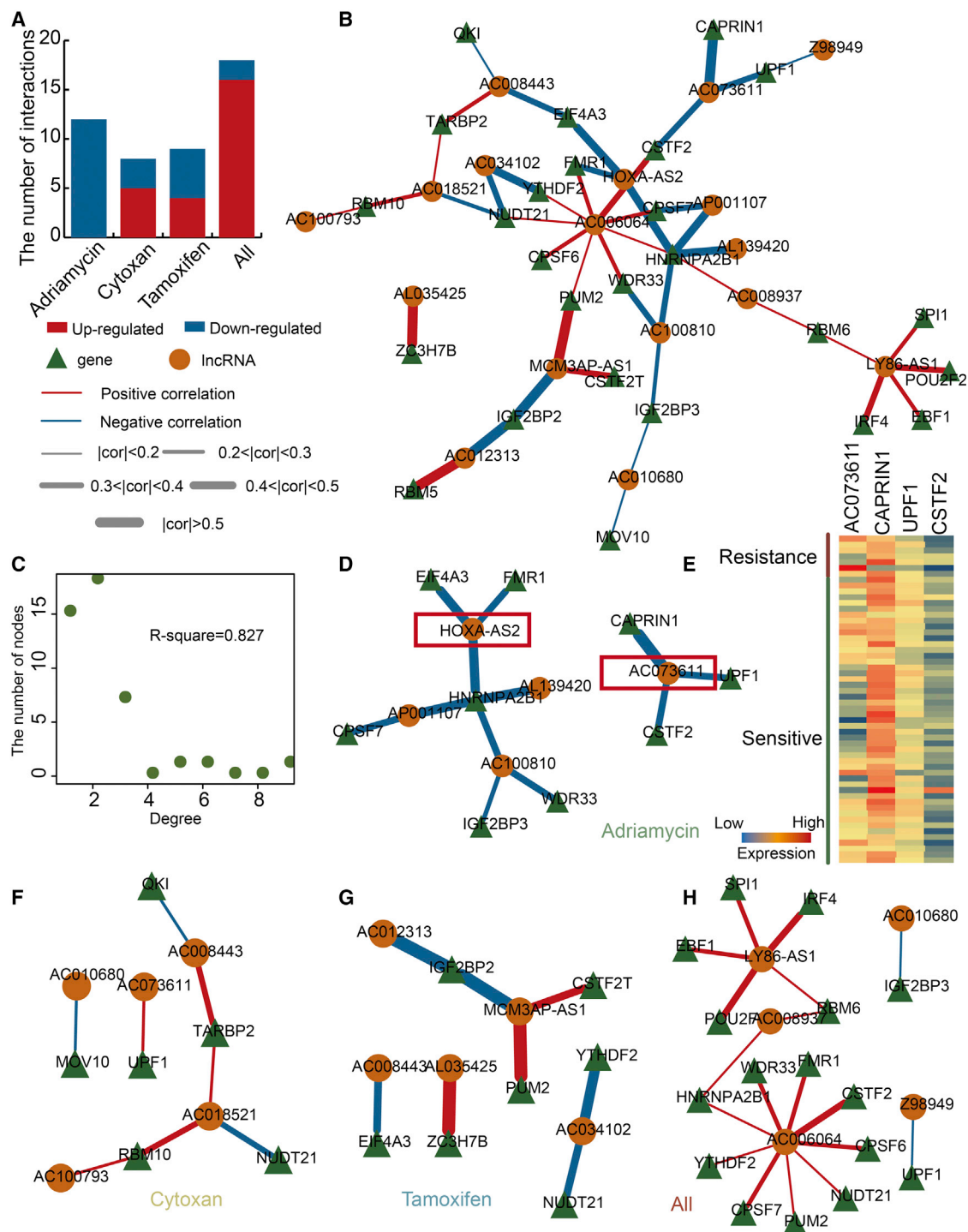
Figure 3. The DRlnc-Gene Co-expression Networks Related to Drug Responses in Breast Cancer

(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 represents correlation levels between DRlnc and genes. (C) The plots show the degree of distribution of the total DRlnc-gene co-expression network. (D) The DRlnc-gene co-expression network for Adriamycin. (E) The heatmap shows the expression level of DRlnc AC073611 (Ensembl: ENSG00000257605) and its interacting genes, including *CAPRIN1*, *UPF1* (Ensembl: ENSG00000005007), and *CSTF2* (Ensembl: ENSG00000101811). (F–H) The DRlnc-gene co-expression network for (F) Cytoxan, (G) Tamoxifen, and (H) all samples.
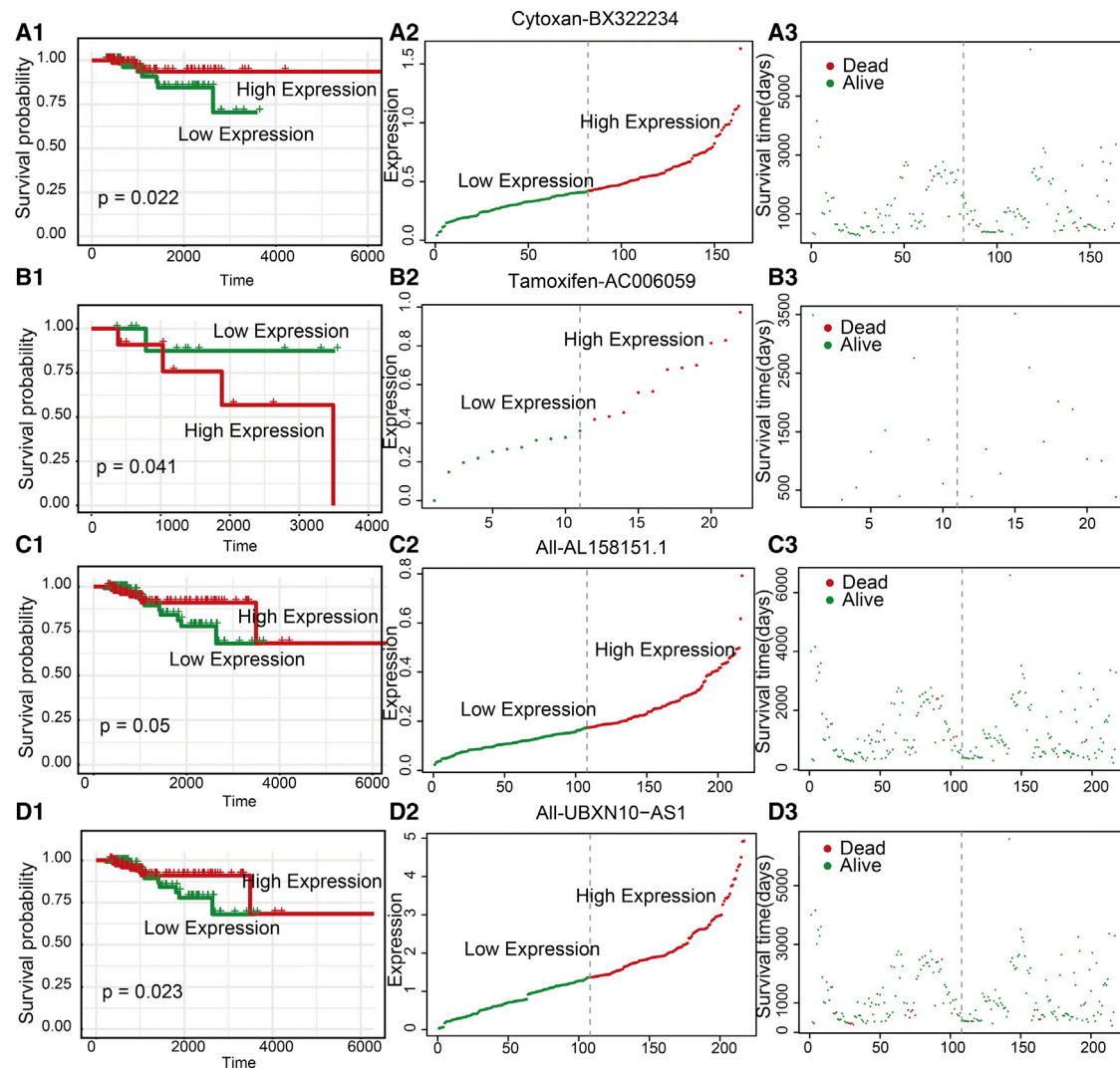
Figure 4. Some DRlncs Were Associated with Survival of Breast Cancer Patients

(A1-D1) 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. (A2-D2) The expression distribution of the DRlnc in each drug. (A3-D3) The patient survival status of the DRlnc for diverse drugs in BRCA patients.

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 (Ensembl: ENSG00000253552), were identified. Some DRlncs were associated with survival and specific functions. In our analysis, most DRlncs were identified for only 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 is responsible for drug resistance in humancancers, including BRCA.[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 (Ensembl: ENSG00000105991) DNA methylation via regulation of the lncRNA HOTAIR (Ensembl: ENSG00000228630).[33] The lncRNA UCA1 (Ensembl: ENSG00000214049) is physically associated with the enhancer of zeste homolog 2, which suppresses the expression of p21 (Ensembl: ENSG00000124762) through histone methylation (H3K27me3) on the p21 (Ensembl: ENSG00000124762)
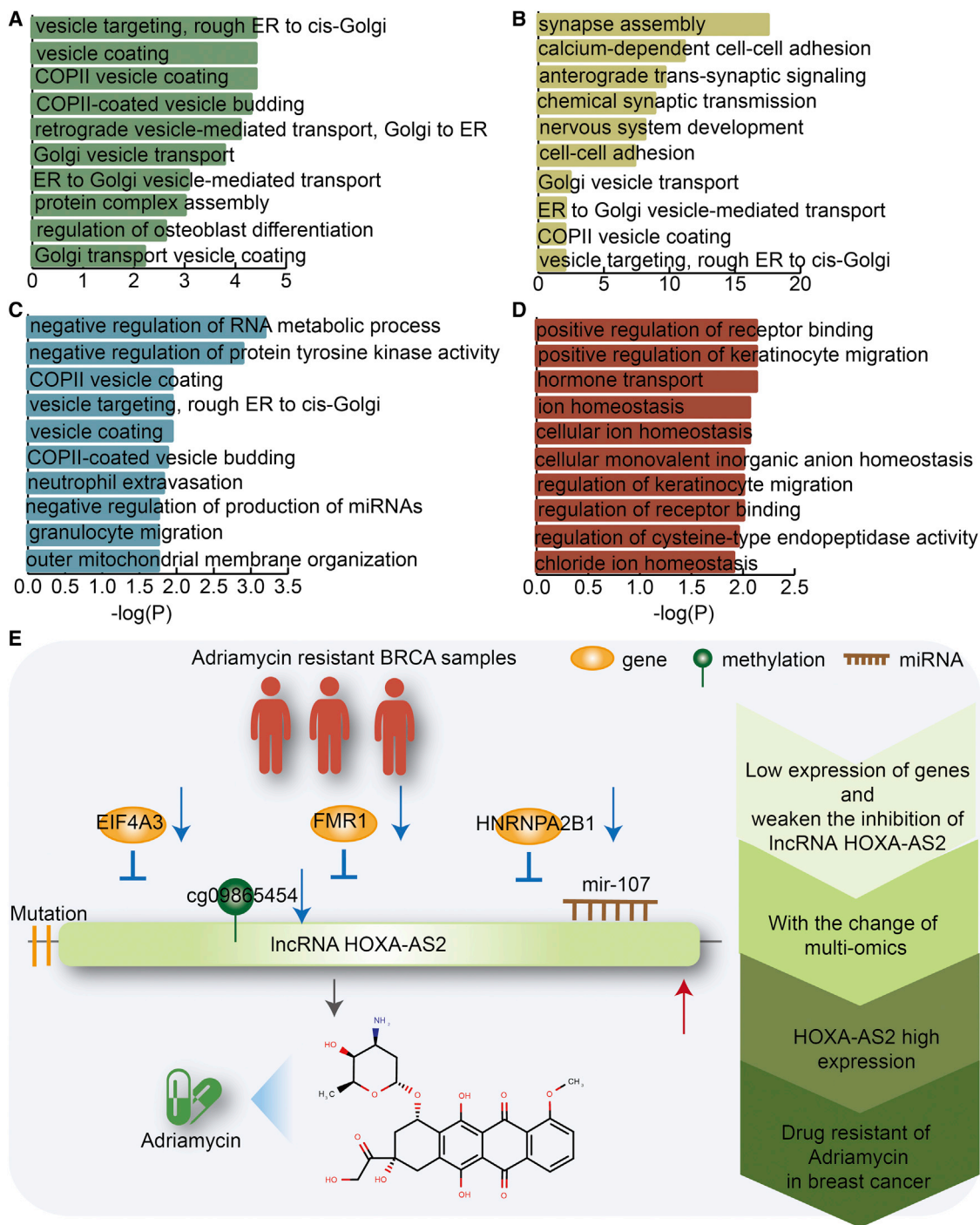
Figure 5. The Functions and Mechanism of DRlncs for Diverse Drugs in BRCA Patients

(A–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 (Ensembl: ENSG00000253552) in drug responses for BRCA patients.

promoter.[34] Silencing lncRNA COMET (Ensembl: ENSG00000231210) markedly increases sensitivity to vemurafenib, a common inhibitor of mutated B-raf in BRAF- and RET-driven papillary thyroid carci-nomas.[35] The CT genotype (rs3200410 genotypes include CC, CT, TT) of rs3200401 MALAT1 (Ensembl: ENSG00000251562) polymorphism could serve as a toxicity biomarker.[36] lncRNA
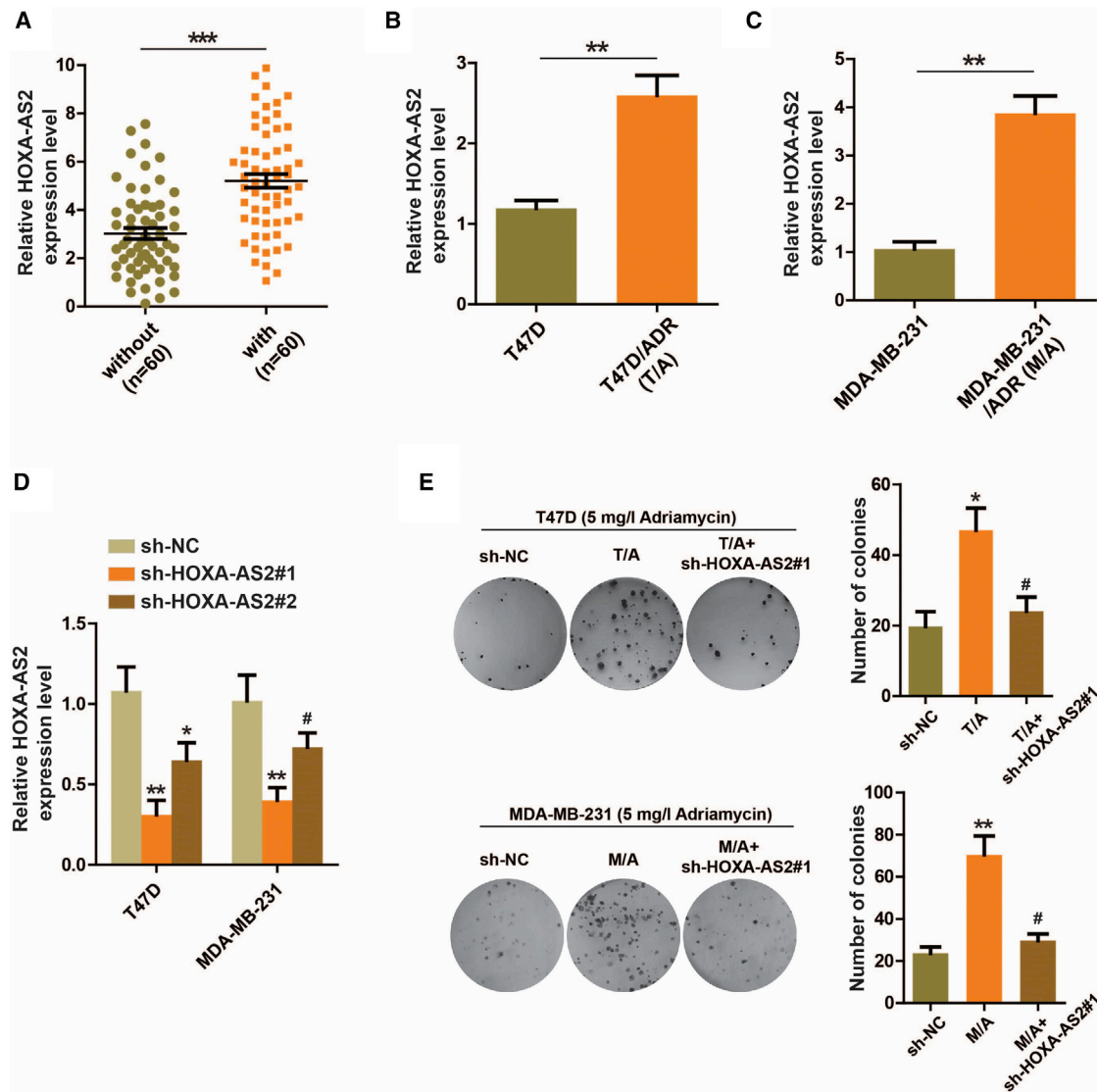
Figure 6. HOXA-AS2 Promotes the Adriamycin Resistance of BRCA Cells

(A) Quantitative real-time PCR was used to detect the expression of HOXA-AS2 in BRCA tissues from patients with or without Adriamycin-based neoadjuvant chemotherapy before surgery. The expression of HOXA-AS2 was examined by real-time PCR in normal or adriamycin-resistant (B) T47D and (C) MDA-MB-231 cells. (D) The efficiency of HOXA-AS2 (Ensembl: ENSG00000253552) knockdown by shRNAs was confirmed in BRCA cells. (E) With 5 mg/L Adriamycin treatment, colony formation assays were performed to analyze the proliferative ability of normal, Adriamycin-resistant, or HOXA-AS2 (Ensembl: ENSG00000253552)-knockdown Adriamycin-resistant BRCA cells.

SNHG6 (Ensembl: ENSG00000245910) promotes chemoresistance through ULK1 (Ensembl: ENSG00000177169)-induced autophagy by sponging miR-26a-5p in colorectal cancer cells.[37] lncRNA H19 (Ensembl: ENSG00000130600) overexpression induces bortezomib resistance in multiple myeloma by targeting MCL-1 (Ensembl: ENSG00000143384) via miR-29b-3p.[33] Multiomics 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 (Ensembl: ENSG00000253552), 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 (Ensembl: ENSG00000253552) is a lncRNA located between the *HOXA3* (Ensembl: ENSG00000105997) and *HOXA4* (Ensembl: ENSG00000197576) genes in the HOXA cluster. Recent studies indicated that HOXA-AS2 (Ensembl: ENSG00000253552) represses apoptosis in NB4 promyelocytic leukemia cells treated with transretinoic acid,[38] promotes the proliferation of gastric cancer via P21/PLK3 (Ensembl: ENSG00000173846)/DDIT3

(Ensembl: ENSG00000175197),[39] and accelerates tumorigenesis of hepatocellular carcinomas.[27] The lncRNA HOXA-AS2 (Ensembl: ENSG00000253552) promotes proliferation and invasion of BRCA by acting as a miR-520c-3p sponge.[40] Subsequent studies suggested that HOXA-AS2 (Ensembl: ENSG00000253552) was significantly overexpressed in a variety of tumors and was associated with poor prognoses of these tumors. The role of HOXA-AS2 (Ensembl: ENSG00000253552) in anticancer drug resistance has been confirmed in acute myeloid leukemia. A study showed that HOXA-AS2 (Ensembl: ENSG00000253552) 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 (Ensembl: ENSG00000253552) 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, lncRNA, and miRNA expression profiles (Illumina HiSeq), 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, lncRNAs, miRNAs, and methylations not expressed across all BRCA samples, we excluded the items with expression values of 0 in all BRCA samples. To allow log transformation, any remaining expression values of 0 were set to the minimum value of all samples, and all values were $\log_2$ transformed. $\log_2$ transformations were performed for genes, lncRNAs, 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.[41] The experimentally validated gene-lncRNA and miRNA-lncRNA interactions were all obtained from the RAID 2.0 database[42] (http://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 significant 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 differentially expressed lncRNAs 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 lncRNA. Second, somatic mutation scores were calculated as follows:

$$s_{mut} = \frac{1}{|N_{sensitive} - N_{resis\,tan\,t}|}$$

where S_{mut} represented the score of somatic mutation, and $N_{sensitive}$ and $N_{resistant}$ represented the number of somatic mutations for a lncRNA in sensitive and resistant BRCA patients. Third, an equally weighted multiple ranking approach was conducted based on four risk scores, including S_{gene}, S_{miRNA}, S_{meth}, and S_{mut}, for each lncRNA to generate the last risk score.[43] After this step, each lncRNA was assigned a final risk score and ranked by these final risk scores. The lncRNAs 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 DRlnc-Gene Co-expression Network for Diverse Chemotherapeutic Drugs in BRCA

Pearson's correlation coefficients were calculated for each interacting DRlnc-gene pair based on gene and lncRNA expression profiles for

diverse chemotherapeutic drugs in BRCA. The significantly co-expressed ($p < 0.05$) DRlnc-genes were extracted for constructing networks. The co-expressed networks for diverse chemotherapeutic drugs in BRCA were constructed using Cytoscape 3.3.0 (https://cytoscape.org/). In addition, the degree analysis and R^2 score were also performed by Cytoscape.

Survival Analysis of DRlncs in BRCA Patients

For each DRlnc of chemotherapeutic drugs in BRCA, we classified samples into two groups based on median values of lncRNA expression values. Kaplan-Meier survival analysis was performed for the two clustered groups, and statistical significance was 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.[44] 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's 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.

Quantitative Real-Time 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 (Ensembl: ENSG00000253552) expression was normalized to 18S tRNA for each sample. The primers for quantitative real-time PCR detection are listed as follows: HOXA-AS2 (Ensembl: ENSG00000253552), 5′-CCCGTAGGAA GAACCGATGA-3′ (forward) and 5′-TTTAGGCCTTCGCAGA CA

GC-3′ (reverse); 18S rRNA, 5′-GTAACCCGTTGAACCCCATT-3′ (forward) and 5′-CCATCCAATCGGTAGTAGCG-3′ (reverse). The quantitative real-time PCRs were repeated three times. Relative RNA expression was calculated using the $^{-\Delta\Delta}$Ct method.

Colony Formation Assay

Indicated BRCA cells were seeded in six-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.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtn.2020.01.038.

AUTHOR CONTRIBUTIONS

LZ and JT conceived and designed the experiments. HC, HK and FP collected and analyzed data. CW and DZ validated the method and data. HC and LZ wrote this manuscript. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no competing interests.

REFERENCES

1. Siegel, R.L., Miller, K.D., and Jemal, A. (2018). Cancer statistics, 2018. CA Cancer J. Clin. 68, 7–30.

2. DeSantis, C.E., Ma, J., Goding Sauer, A., Newman, L.A., and Jemal, A. (2017). Breast cancer statistics, 2017, racial disparity in mortality by state. CA Cancer J. Clin. 67, 439–448.

3. DeSantis, C.E., Ma, J., Gaudet, M.M., Newman, L.A., Miller, K.D., Goding Sauer, A., Jemal, A., and Siegel, R.L. (2019). Breast cancer statistics, 2019. CA Cancer J. Clin. 69, 438–451.

4. Rodrigues-Ferreira, S., Nehlig, A., Moindjie, H., Monchecourt, C., Seiler, C., Marangoni, E., Chateau-Joubert, S., Dujaric, M.E., Servant, N., Asselain, B., et al. (2019). Improving breast cancer sensitivity to paclitaxel by increasing aneuploidy. Proc. Natl. Acad. Sci. USA 116, 23691–23697.

5. Nounou, M.I., 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 (Suppl 2), 17–34.

6. Jayaraj, R., Nayagam, S.G., Kar, A., Sathyakumar, S., Mohammed, H., Smiti, M., Sabarimurugan, S., Kumarasamy, C., Priyadharshini, T., Gothandam, K.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, e1250.

7. Vulsteke, C., Lambrechts, D., Dieudonné, A., Hatse, S., Brouwers, B., van Brussel, T., Neven, P., Belmans, A., Schöffski, P., Paridaens, R., and Wildiers, H. (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, N.C., Shao, P., Maina, P.K., Zhang, W., and Qi, H.H. (2020). Contribution of synergism between PHF8 and HER2 signalling to breast cancer development and drug resistance. EBioMedicine 51, 102612.

9. Cabili, M.N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., and Rinn, J.L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev. 25, 1915–1927.

10. Guttman, M., Garber, M., Levin, J.Z., Donaghey, J., Robinson, J., Adiconis, X., Fan, L., Koziol, M.J., Gnirke, A., Nusbaum, C., 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., Li, X., Fang, Y., Shen, W., Xu, Y., et al. (2019). Lnc2Cancer v2.0: updated database of experimentally supported long non-coding RNAs in human cancers. Nucleic Acids Res. 47 (D1), D1028–D1033.

12. Huang, Q.Y., Liu, G.F., Qian, X.L., Tang, L.B., Huang, Q.Y., and Xiong, L.X. (2019). Long Non-Coding RNA: Dual Effects on Breast Cancer Metastasis and Clinical Applications. Cancers (Basel) 11, e1802.

13. Zhang, N., Zeng, X., Sun, C., Guo, H., Wang, T., Wei, L., Zhang, Y., Zhao, J., and Ma, X. (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, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 464, 1071–1076.

15. Mourtada-Maarabouni, M., Pickard, M.R., Hedge, V.L., Farzaneh, F., and Williams, G.T. (2009). GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene 28, 195–208.

16. Chi, J.R., Yu, Z.H., Liu, B.W., Zhang, D., Ge, J., Yu, Y., and Cao, X.C. (2019). SNHG5 Promotes Breast Cancer Proliferation by Sponging the miR-154-5p/PCNA Axis. Mol. Ther. Nucleic Acids 17, 138–149.

17. Fernando, T.R., Rodriguez-Malave, N.I., Waters, E.V., Yan, W., Casero, D., Basso, G., Pigazzi, M., and Rao, D.S. (2015). LncRNA Expression Discriminates Karyotype and Predicts Survival in B-Lymphoblastic Leukemia. Mol. Cancer Res. 13, 839–851.

18. Diaz-Lagares, A., Crujeiras, A.B., Lopez-Serra, P., Soler, M., Setien, F., Goyal, A., Sandoval, J., Hashimoto, Y., Martinez-Cardús, A., Gomez, A., et al. (2016). Epigenetic inactivation of the p53-induced long noncoding RNA TP53 target 1 in human cancer. Proc. Natl. Acad. Sci. USA 113, E7535–E7544.

19. Jiang, M., Huang, O., Xie, Z., Wu, S., Zhang, X., Shen, A., Liu, H., Chen, X., Wu, J., Lou, Y., et al. (2014). A novel long non-coding RNA-ARA: adriamycin resistance-associated. Biochem. Pharmacol. 87, 254–283.

20. Guttman, M., and Rinn, J.L. (2012). Modular regulatory principles of large non-coding RNAs. Nature 482, 339–346.

21. Sumazin, P., Yang, X., Chiu, H.S., Chung, W.J., Iyer, A., Llobet-Navas, D., Rajbhandari, P., Bansal, M., Guarnieri, P., Silva, J., and Califano, A. (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., Zhang, J., Gao, Y., Guo, M., Zhou, D., 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 (D1), D74–D78.

24. Khurana, E., Fu, Y., Chakravarty, D., Demichelis, F., Rubin, M.A., 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., Wu, Q., Liu, S., Li, X., and Wang, W. (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., Pan, G., and Liu, J. (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., Osaki, A., Saeki, T., Berezikov, E., Mano, H., and Inoue, S. (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, E.Y.T., Lee, A.M., Geeleher, P., Cho, W.C.S., and Huang, R.S. (2019). Discovering long noncoding RNA predictors of anticancer drug sensitivity beyond protein-coding genes. Proc. Natl. Acad. Sci. USA 116, 22020–22029.

32. Yao, N., Fu, Y., Chen, L., Liu, Z., He, J., Zhu, Y., Xia, T., and Wang, S. (2019). Long non-coding RNA NONHSAT101069 promotes epirubicin resistance, migration, and invasion of breast cancer cells through NONHSAT101069/miR-129-5p/Twist1 axis. Oncogene 38, 7216–7233.

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

34. Li, Z., Yu, D., Li, H., Lv, Y., and Li, S. (2019). Long non-coding 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 lncRNA COMET in BRAF- and RET-Driven Papillary Thyroid Carcinomas. Cancer Res. 79, 2124–2135.

36. Lampropoulou, D.I., Aravantinos, G., Katifelis, H., Lazaris, F., Laschos, K., Theodosopoulos, T., Papadimitriou, C., and Gazouli, M. (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., Liu, Y., Lai, H., Gu, C., Yan, 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., Frazão, J.B., Condino-Neto, A., and Newburger, P.E. (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, Y.N., Xia, R., Liu, Y.W., Ding, J., Ma, H.W., He, X.Z., Zhang, Z.H., Liu, Z.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, A.R., and Hall, I.M. (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., Liu, L., Hou, P., Cui, T., Tan, P., et al. (2017). RAID v2.0: an updated resource of RNA-associated interactions across organisms. Nucleic Acids Res. 45 (D1), D115–D118.

43. Aerts, S., Lambrechts, D., Maity, S., Van Loo, P., Coessens, B., De Smet, F., Tranchevent, L.C., De Moor, B., Marynen, P., Hassan, B., et al. (2006). Gene prioritization through genomic data fusion. Nat. Biotechnol. 24, 537–544.

44. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Res. 44 (W1), W90–W97.