Cellular Physiology and Biochemistry

Cell Physiol Biochem 2017;44:2476-2490
DOI: 10.1159/000486170
Published online: December 18, 2017
Accepted: November 29, 2017

Original Paper

Long Noncoding RNA Linc00152 Functions as a Tumor Propellant in Pan-Cancer

Shouping Xu[a] Lin Wan[a] Huizi Yin[a] Hongbiao Xu[a] Wei Zheng[a] Meiying Shen[a] Zhongmin Zhang[a] Da Pang[a,b]

[a]Department of Breast Surgery, Harbin Medical University Cancer Hospital, Harbin, [b]Heilongjiang Academy of Medical Sciences, Harbin, China

Key Words
Linc00152 • Tumorigenesis • Pan-cancer • Breast cancer

Abstract
Background/Aims: The oncogenic role of linc00152 in pan-cancer is unclear. ***Methods:*** In this study, RNA-Seq of 33 breast specimens was performed, and the expression of linc00152 was validated by qPCR using 50 paired breast cancer tissues and adjacent normal tissues. This result combined with the expression of linc00152 in pan-cancer was revalidated by Gene Expression Omnibus and The Cancer Genome Atlas data. Next, the oncogenic roles of linc00152 in view of prognosis, chemoresistance, genomic and epigenetic regulation, including DNA methylation and histone modification, potential biological function enrichment, and basic molecular function in pan-cancer, were also evaluated *in vitro* and *in vivo*. ***Results:*** Linc00152 is upregulated in pan-cancer, especially in progressive cancer, and the high expression of linc00152 may lead to a worse prognosis and chemoresistance in pan-cancer patients. Amplification, DNA hypomethylation, promoter-like lncRNA characteristics and super-enhancer regulation are the drivers that lead to the upregulation of linc00152 in pan-cancer. Meanwhile, linc00152 was involved in cancer-related pathways, infection and immune response-associated pathways by enriched analysis using TCGA data. Finally, linc00152 was confirmed to promote the proliferation, migration and invasion in MDA-MB-231, SGC-7901 and 786-O. Moreover, RIP and RNA pull-down assays indicated that linc00152 can bind to EZH2 directly. ***Conclusion:*** All of the results indicated that linc00152 acted as an oncogenic propellant from various perspectives, and it may be an effective therapy target in pan-cancer.

Introduction

Breast cancer is the most commonly diagnosed malignant tumor in women worldwide. The incidence of breast cancer is 22.9% in all malignant tumors among women. Moreover, it causes 13.7% of cancer death among female tumor sufferers [1, 2]. As a well-acknowledged life-threatening disease, breast cancer also seriously dampens the life quality of sufferers,

Da Pang

Department of Breast Surgery, Harbin Medical University Cancer Hospital
150 Haping Road, Harbin, 150040, (China)
Tel. +86-0451-86298613, Fax +86-0451-86298613; E-Mail pangda@ems.hrbmu.edu.cn

including psychological problems. Of great importance is that the identification and validation of oncogenic genes that are closely associated with the development of breast cancer are necessary and obligatory.

As novel regulators of cancer development, long non-coding RNAs (LncRNAs) have been attracting increased attention by scientists. LncRNAs are characterized by a length of more than 200 nt but do not encode any protein [3]. LncRNAs have more diversified molecular mechanisms of expression and regulation, such as activating or inhibiting the expression of target genes by direct binding, or by participating in histone modification and the elevation of regulatory factors; lncRNAs may participate in the regulation of gene expression through two types or all signal molecules, decoy molecules, guidance molecules or scaffolding molecules separately or collectively [4-9]. The dysregulation of lncRNAs has been functionally associated with human cancers. The dysregulation of lncRNAs is correlated with promoted proliferation, chemoresistance, invasion and metastasis [10-14]. For example, HOTAIR accelerates such cancer biology in pan-cancers, including breast cancer, glioblastoma, gastric cancer, pancreatic cancer, thyroid cancer, ovarian cancer and lung cancer [4, 15]. Thus, the identification and validation of lncRNAs that act as HOTAIR in pan-cancer is essential and may remain a huge challenge for us to accomplish.

The aim of this study was to identify and validate the oncogenic role of linc00152 in breast cancer and other cancers. In this study, linc00152 expression was higher in breast cancer than in adjacent normal tissues. This result was obtained from the RNA-Seq analysis of 33 breast cancer tissues and adjacent normal tissues and the qPCR validation of 50 paired tissues in our cohort. Next, such results were reexamined and revalidated in the Gene Expression Omnibus (GEO) data and The Cancer Genome Atlas (TCGA) data, respectively. Moreover, the oncogenic roles of linc00152 in the prognosis, chemoresistance, upstream regulation, histone modification and basic molecular function in pan-cancer were assessed subsequently in this article. Overall, linc00152 functions as a tumor propellant in pan-cancer, and it may be an effective therapy target that regulates cancer progression.

Materials and Methods

Clinical samples and RNA-seq

Thirty-three samples (15 breast cancer tissues, 15 adjacent normal tissues and 3 normal breast tissues) were obtained from patients without chemotherapy or radiotherapy at the Department of Breast Surgery at Harbin Medical University Cancer Hospital. This study protocol conformed to the clinical research guidelines and was approved by the research ethics committee of the Harbin Medical University Cancer Hospital. Written informed consent was obtained from all of the patients who participated in this study. For RNA-seq, ribosomal RNA was removed using Ribo-Zero™ Gold kits (Epicentre, Wisconsin, USA). Sequencing libraries were generated according to the manufacturer's recommendations with varied index labeling using the NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina (NEB, Ipswich, USA). The libraries were then sequenced on an Illumina HiSeq 2500 platform, and 100-bp paired-end reads were generated.

GEO and TCGA data analyses

GEO datasets were downloaded from the NCBI website (http://www.ncbi.nlm.nih.gov/geo/). The genome-wide linc00152 expression profiles for pan-cancer were downloaded from TCGA (https://tcga-data.nci.nih.gov/). For the microarray in GEO, we adjusted the signal values for low-abundance genes. Moreover, the invariant genes and low-variation genes were filtered. Genes that were detected in less than 50% of the profiled samples were also filtered. The SAM method was applied, and we implemented a series of steps to estimate the significance of difference and false discovery rate for each filtered gene [16]. The hierarchical clustering for genes and samples and meta-analysis for survival was conducted as previously described [17, 18]. DNA methylation data were downloaded from the Illumina Infinium HumanMethylation450 Beadchip in the TCGA Data Portal. The DNA methylation level ranged from 0 (least methylated) to 1 (most methylated). The methylation level is given by beta = Methylated probe intensity (M) /(Unmethylated probe intensity (U) + Methylated probe intensity (M) + 100).

Guilt-by-association analysis

Data from the TCGA was evaluated to compute a pairwise Pearson correlation between the expression of linc00152 and all of the genes. Only associated genes with an absolute $r \geq 0.6$ and a significant correlation ($P < 0.05$) were retained. Gene ontology term enrichment and KEGG pathway analysis were analyzed with these genes using DAVID, as previously described [19, 20].

Cell culture experiments

MDA-MB-231 cells were cultured in DMEM (Invitrogen, Carlsbad, CA) containing 10% fetal bovine serum and 100 units/ml penicillin/streptomycin at 37°C in an atmosphere without CO_2. SGC-7901 and 786-O cells were cultured in a 1640 medium (Invitrogen, Carlsbad, CA) containing 10% fetal bovine serum and 100 units/ml penicillin/streptomycin at 37°C in an atmosphere containing 5% CO_2. All of the cell lines were obtained from the Chinese Type Culture Collection, Chinese Academy of Sciences. Cells were used during their logarithmic growth phase.

Linc00152 knockdown and qRT-PCR

To silence linc00152 expression, an shRNA sequence targeting LINC00152 was produced and cloned into the lentivirus vector. Two shRNAs were designed to target linc00152 (shR_1: 5′-TGTGGACTCTGAGGCCTCTGCATTT -3′ and shR_2: 5′- TCTATGTGTCTTAATCCCTTGTCC T-3′). An shRNA without sequence homology to human genes was used as the negative control. Recombinant lentivirus was generated from 293T cells, and the stable cells infected with LINC00152 shRNA were selected by puromycin. The knockdown efficiency of the LINC00152 was examined by qRT-PCR. qRT-PCR was performed using the SYBR-Green method, and the specific sequences of the primers used were 5′- CTGGATGGTCGCTGCTTTTT-3′ (forward) and 5′- GATCTGAAGACAGGCACGGG-3′ (reverse) for linc00152; and 5′- GCACCGTCAAGGCTG AGAAC-3′ (forward) and 5′- GGATCTCGCTCCTGGAAGATG-3′ (reverse) for GAPDH. Quantitative normalization of linc00152 cDNA was performed for each sample using GAPDH expression as an internal control. The relative levels of linc00152 vs. GAPDH were determined by the comparative CT ($2^{-\Delta\Delta CT}$) method.

Cell proliferation, migration and invasion assay

Cell proliferation assays were performed using the Cell Counting Kit-8 according to the manufacturer's instructions (Beyotime, Shanghai, China). Briefly, 2×10^3 cells were seeded in a 96-well plate. Cell proliferation was assessed at 24, 48, and 72 h. After the addition of 20 µl of WST-1 reagent per well, the cultures were incubated for 2 h, and the absorbance was measured at 490 nm using a microplate reader (BioTek, VT, United States). For migration assay, MDA-MB-231 cells were seeded into the upper chambers of transwell culture plates, each with an 8-µm pore membrane insert (Corning, Shanghai, China). L-15 medium supplemented with 20% FBS was placed in the lower chambers. For invasion assay, the transwell chambers were firstly coated with matrigel solution. Then, MDA-MB-231 cells were seeded into the upper chambers of the transwells and L-15 medium with 20% FBS (600 µL) was added to the lower chambers. After incubation for 48 h in migration or invasion assays, cells that had penetrated through to the lower surface of the membrane were fixed with 4% paraformaldehyde for 30 min, stained with crystal violet for 30 min, and counted the cells in six randomly chosen fields.

Flow cytometry

An Annexin-PE Apoptosis detection kit (BD Biosciences, San Jose, CA) was used to examine cell apoptosis according to the manufacturer's instructions. Briefly, cells were washed twice in cold PBS, harvested and resuspended in 1 × binding buffer. Next, 100 µl of the cell solution (1×10^5 cells) was transferred into a 5-ml culture tube, and 5 µl of annexin V-PE and 5 µl of 7-AAD were added. The cells were gently vortexed and incubated for 15 min at RT (25°C) in the dark. Next, 400 µl of 1 × binding buffer was added to each tube, and apoptosis analysis was performed on a FACScan instrument (Becton Dickinson, Mountain View, CA, USA).

Xenografts in mice

Stable knockdown cells or control cells were subcutaneously injected into two groups of athymic female BALB/c mice (5 weeks old). After three weeks injection, all mice were sacrificed and the tumors were isolated and weighed. The study was carried out in accordance with the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The experimental protocols were finally approved by the committee on animal experimentation of Harbin Medical University.

RNA-binding protein immunoprecipitation assay (RIP) and RNA pull-down

RIP was performed according to the manufacturer's instructions (Millipore, Billerica, MA, USA). Briefly, cells were lysed in lysis buffer, and then were incubated with RIP buffer containing magnetic beads conjugated with human anti-EZH2 antibodies (Millipore, Billerica, MA, USA), and negative control IgG (Millipore, Billerica, MA, USA). Immunoprecipitated RNA was isolated and analyzed by qRT-PCR assays. RNA pull-down was performed using a Magnetic RNA-Protein Pull-Down Kit (Pierce Biotechnology, USA) in accordance with the manufacturer's instructions.

Statistical analyses

The expression of linc00152 was dichotomized using the median expression as the cut-off to define "high value" at or above the median versus "low value" below the median. Overall survival was calculated as the time from surgery until the occurrence of death. Any event-free survival was calculated as the time from surgery until the occurrence of any relapse or death. Metastasis-free survival was calculated as the time from surgery until the occurrence of distant relapse. The differences between the groups in our *in vitro* experiments were analyzed using Student's t-test. Spearman correlation coefficients were calculated for correlation analysis. All of the experiments were performed in triplicate, and SPSS 16.0 software (SPSS, Chicago, IL) was used for statistical analysis. All of the statistical tests were two-sided, and $P < 0.05$ was considered to be statistically significant.

Results

Linc00152 is upregulated in breast cancer

Thirty-three specimens, including breast cancer tissue (N=15) and adjacent normal tissue (N=15), as well as breast tissue from non-cancer patients (N=3), were obtained and subjected to RNA-Seq analysis performed by our cohort. Linc00152 showed higher expression in cancer tissues than in non-cancer tissues (Fig. 1A). Next, this result was validated by TCGA and GEO data. Linc00152 showed high expression in breast cancer tissues in 837 patients from TCGA and 10 datasets (GSE20685, GSE12763, GSE21422, GSE10810, GSE7904, GSE29431, GSE42568, GSE10780, GSE48390 and GSE21653) containing 1, 215 patients from GEO (Fig. 1B and (for all online suppl. material, see www.karger.com/doi/10.1159/000486170) Fig. S1). To further confirm the linc00152 expression model, 50 paired breast cancer and non-cancer tissues were applied to revalidate the result in our cohort. As expected, linc00152 showed higher expression in cancer tissues than in its non-cancer counterpart significantly (Fig. 1C).

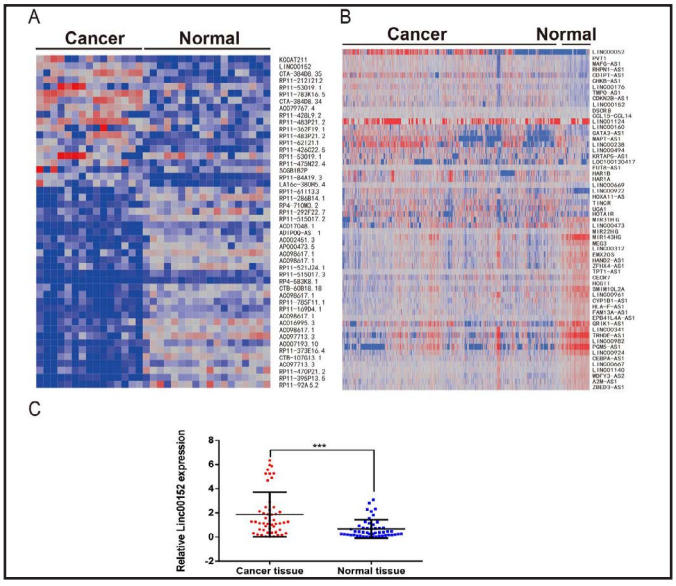

Fig. 1. Linc00152 is upregulated in breast cancer. (A) Hierarchical clustering of differentially expressed genes in breast cancer relative to normal tissue in our cohort (N=33). The blue through red color indicates low to high expression levels, respectively. (B) Hierarchical clustering of differentially expressed genes in breast cancer relative to normal tissue in the TCGA cohort (N=942). The blue through red color indicates low to high expression levels, respectively. (C) Linc00152 expression in 50 paired breast cancer and non-cancer tissues. The red columns indicate the expression levels of linc00152 in cancer tissues, and the blue columns represent normal tissues. *** P<0.001.

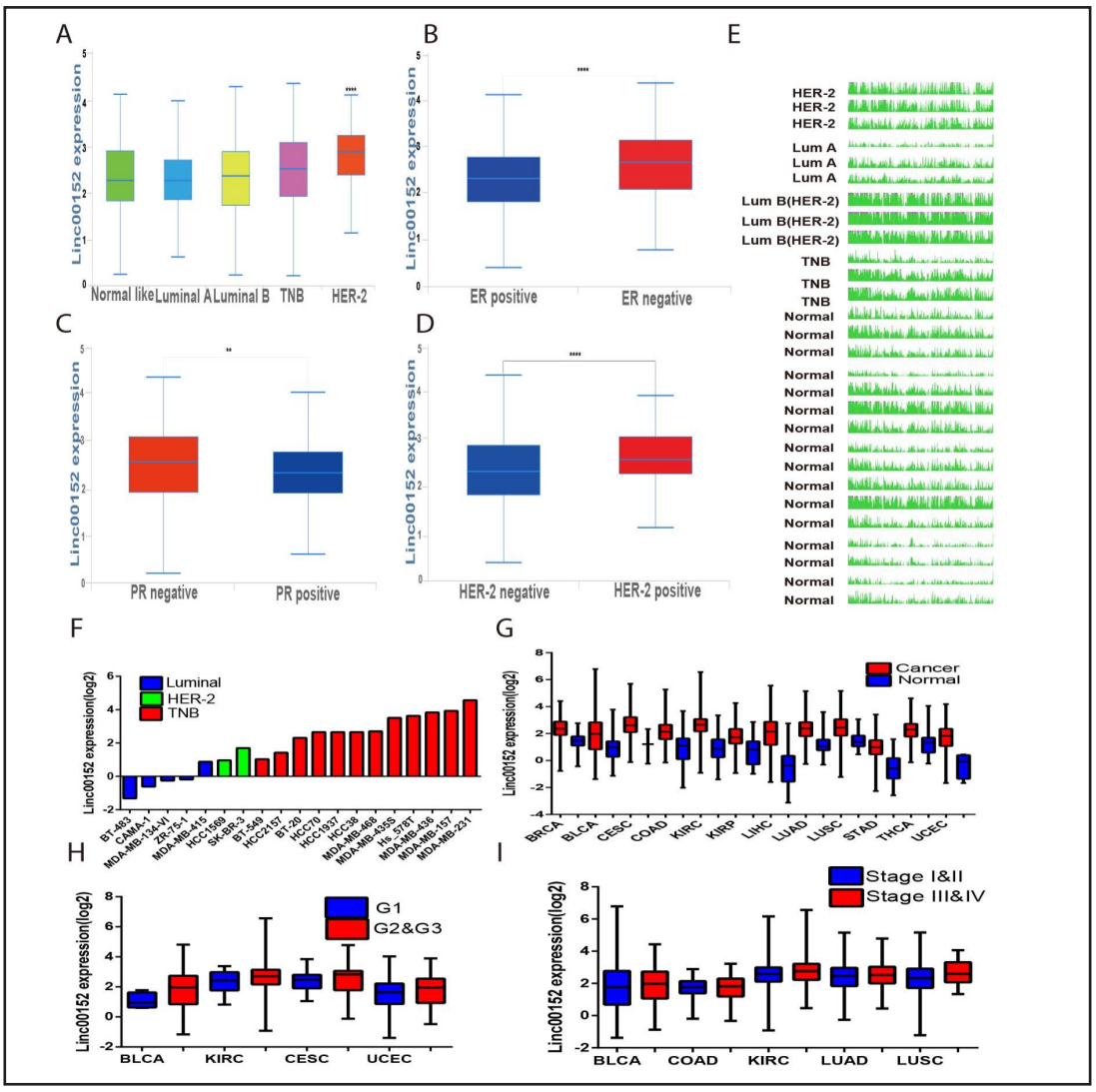

Fig. 2. Linc00152 is closely related to cancer with progressive characteristics. (A) Linc00152 expression in different molecular subtypes of breast cancer in TCGA. (B) Linc00152 expression in ER-positive and -negative breast cancer in TCGA. (C) Linc00152 expression in PR-positive and -negative breast cancer in TCGA. (D) Linc00152 expression in HER-2-positive and -negative breast cancer in TCGA. (E) RNA-Seq data visualization of linc00152 in 33 samples with different subtypes. (F) Linc00152 expression in different breast cancer cell lines in CCLE. (G) Linc00152 expression in different types of cancer in TCGA. (H) Linc00152 expression in different grades of bladder carcinoma, renal clear cell carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma and endometrioid carcinoma in TCGA. (I) Linc00152 expression in different TNM stages of bladder carcinoma, renal clear cell carcinoma, colon adenocarcinoma, lung adenocarcinoma and lung squamous cell carcinoma in TCGA. The expression level of linc00152 was measured by log2 FPKM. * P<0.05, ** P<0.01, *** P<0.001, **** P<0.0001.

Linc00152 is closely related to cancer with progressive characteristics

For breast cancer, linc00152 was expressed very highly in patients with progressive triple-negative and HER-2 subtypes in TCGA (Fig. 2A). Meanwhile, its expression was also high in patients with ER-negative, PR-negative and HER-2–positive cancers than in their counterparts (Fig. 2B, 2C and 2D). Next, linc00152 was expressed at a lower level in breast cancer with the luminal A subtype but showed higher expression in other progressive subtypes in our 33 RNA-seq specimens (Fig. 2E). Breast cancer cell lines were also

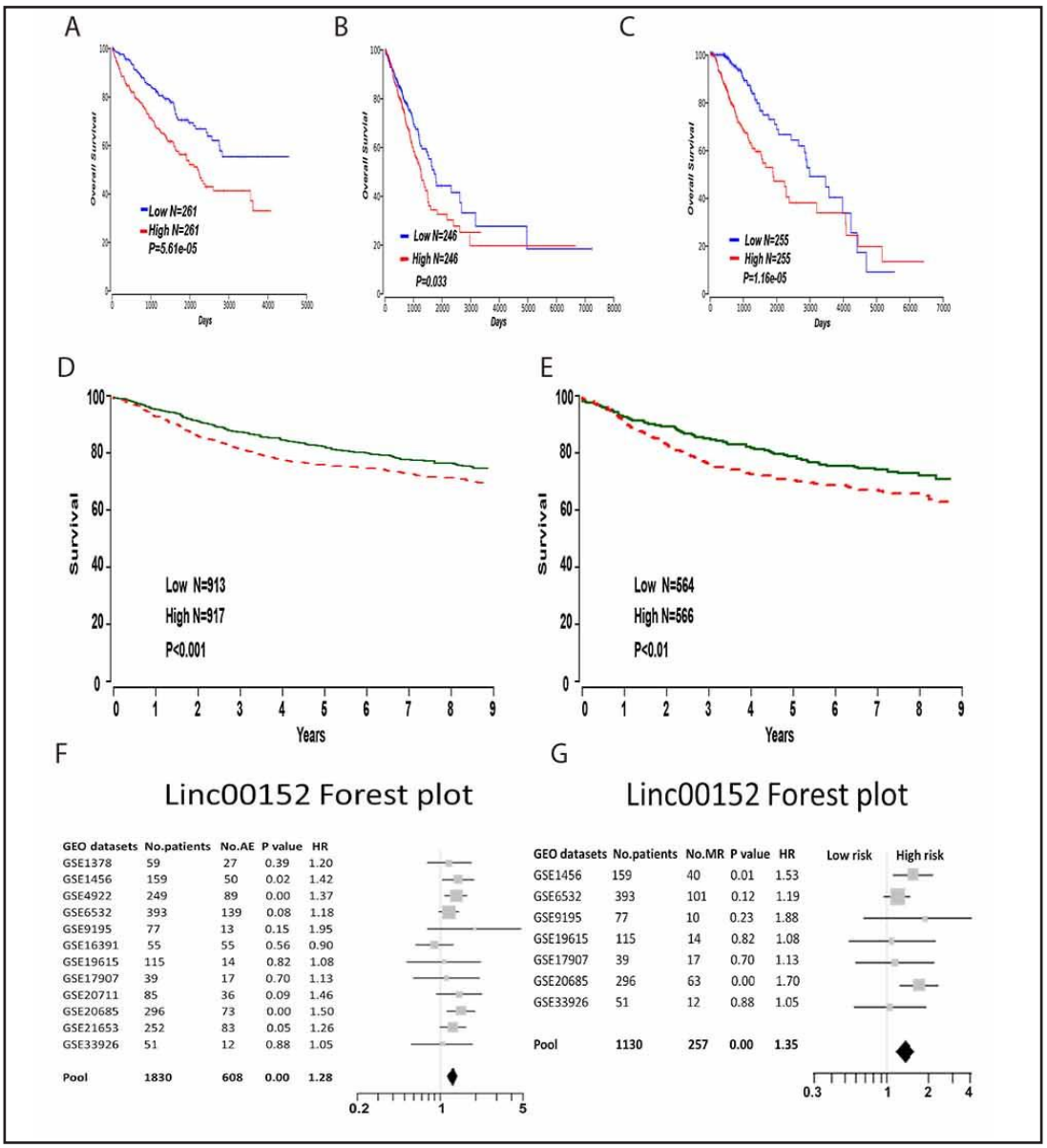

Fig. 3. High expression of linc00152 leads to a worse prognosis in pan-cancer patients. (A-C) High expression of linc00152 correlated with a worse overall prognosis in patients with renal clear cell carcinoma (A) or lung adenocarcinoma (B) or lower grade glioma (C) in TCGA. D&E. High expression of linc00152 correlated with a worse any event-free survival (D) and metastasis-free survival (E) in 12 datasets with 5,802 breast patients in GEO. G&H, Meta-analyses of the relationship between any event-free survival (F, N=1,830) and metastasis-free survival (G, N=1,130) and the expression of linc00152 in breast cancer patients in GEO.

investigated from CCLE (http://software.broadinstitute.org/software/cprg/?q=node/11) to validate this result. As expected, linc00152 showed higher expression in triple-negative or HER-2 overexpression cell lines than in luminal cancer cell lines (Fig. 2F). Additionally, other cancers were examined to investigate the linc00152 expression model. The expression levels of linc00152 were higher in bladder carcinoma, cervical squamous cell carcinoma and endocervical adenocarcinoma, colon adenocarcinoma, renal clear cell carcinoma, renal papillary cell carcinoma, hepatocellular carcinoma, lung adenocarcinoma, lung squamous cell carcinoma, stomach adenocarcinoma, thyroid carcinoma and endometrioid carcinoma than in their non-cancer counterparts in TCGA (Fig. 2G). Moreover, high-grade and high TNM

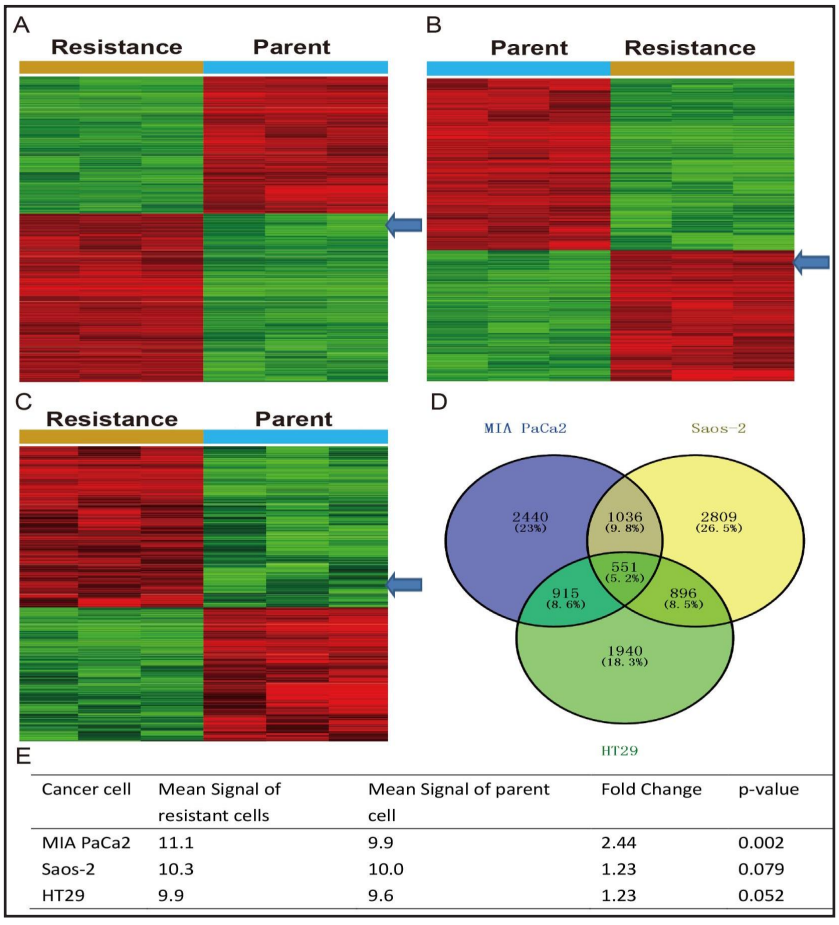

Fig. 4. Linc00152 induces chemoresistance in pan-cancer. (A-C) High expression of linc00152 was observed in MIA PaCa2 methotrexate-resistant cells (A), Saos-2 methotrexate- resistant cells (B), and HT29 methotrexate-resistant cells (C), as well as in their parental cells in the GSE16089 dataset. (D & E). Linc00152 is one of the upregulated genes among 551 overlapping genes in these three cell line cohorts via Venny online software analysis (D) and was upregulated significantly in resistant cells than in parental cells (E).

E

Cancer cell	Mean Signal of resistant cells	Mean Signal of parent cell	Fold Change	p-value
MIA PaCa2	11.1	9.9	2.44	0.002
Saos-2	10.3	10.0	1.23	0.079
HT29	9.9	9.6	1.23	0.052

stage were also accompanied with the high expression of linc00152 in bladder carcinoma and renal clear cell carcinoma (Fig. 2H and 2I).

High expression of linc00152 leads to a worse prognosis in pan-cancer patients

Next, the relationship between linc00152 expression and prognosis in pan-cancer was investigated via TCGA and GEO data. High expression of linc00152 indicates a worse overall survival in patients with renal clear cell carcinoma or lung adenocarcinoma or lower grade glioma (P<0.001; P=0.033; P<0.001, respectively) (Fig. 3A, 3B and 3C). Meanwhile, high expression of linc00152 in patients with breast cancer also indicates a trend of worse overall survival without significance (data not supplied). Thus, all datasets from GEO containing the data of linc00152 expression and complete follow-up information were collected to perform the survival analysis. As expected, high expression of linc00152 in breast cancer patients showed a worse any event-free survival (AES) and metastasis-free survival (MFS) in 12 datasets with 5, 802 patients in GEO (P<0.001; P<0.01, respectively) (Fig. 3D, 3E and see online suppl. material, Table S1). Moreover, meta-analyses were also performed to investigate the relationship between linc00152 expression and any event-free survival and metastasis-free survival. Patients with high expression of linc00152 have a higher risk of AES and MFS than those with low expression of linc00152 (P<0.001, HR=1.28; P<0.001, HR=1.36, respectively) (Fig. 3F and 3G).

Linc00152 induces chemoresistance in pan-cancer

The relationship between linc00152 and chemoresistance was also investigated subsequently. GSE16089 and GSE16446 were selected to meet the study need. In the GSE16089 dataset, seven cell lines representative of different types of cancers, including

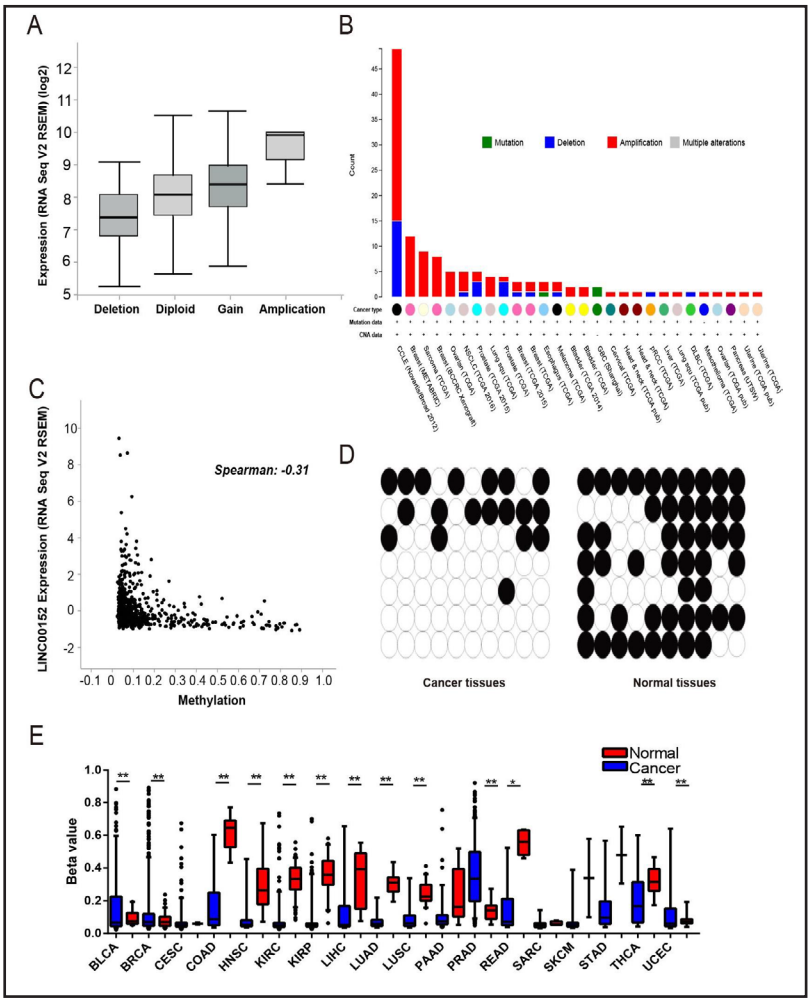

Fig. 5. Regulation of Linc00152 expression. (A) The relationship between the copy number variation and upregulation of linc00152 in breast cancer in TCGA. (B) Relationship between the copy number variation and upregulation of linc00152 in pan-cancer in TCGA. (C) Linc00152 expression was inversely correlated with DNA methylation of the linc00152 locus in breast cancer in TCGA. (D) Linc00152 was hypomethylated in hepatic carcinoma tissues compared with para-cancer tissues via bisulfite sequencing polymerase chain reaction [22]. (E) DNA methylation in the linc00152 locus in pan-cancer tissues and that in normal tissues in 450K data from TCGA. The blue columns indicate the β value in cancer tissues, and the red columns represent that in normal tissues. * P<0.05, ** P<0.01.

colon cancer (HT29 and Caco2), breast cancer (MCF7 and MDA-MB-468), pancreatic cancer (MIA PaCa-2), erythroblastic leukemia (K562) and osteosarcoma (Saos-2), were applied to investigate the differential expression pattern between sensitive and methotrexate-resistant cells by microarrays. Linc00152 indicated a higher expression in MIA PaCa2 methotrexate-resistant cells, Saos-2 methotrexate-resistant cells, and HT29 methotrexate-resistant cells and their parental cells (Fig. 4A, 4B and 4C). Linc00152 is one of the upregulated genes of the 551 overlapping genes in these cell line cohorts (Fig. 4D and 4E). In the GSE16446 dataset, anthracycline-resistance genes were examined by microarray analysis of 149 breast cancer patients with neoadjuvant chemotherapy. Linc00152 was upregulated in patients without a pathological complete response (data not supplied). In addition to these results, a previous study identified that the linc00152/miR-193a-3p/erbB-4/AKT signaling axis confers oxaliplatin resistance in colon cancer [21]. Overall, linc00152 may lead to chemoresistance in pan-cancer.

Regulation of Linc00152 expression

The relationship between copy number variations and the upregulation of linc00152 was performed in breast cancer by TCGA data. Obvious amplification was observed in breast cancer samples with high linc00152 expression (Fig. 5A). Moreover, the result can be obtained in pan-cancer (Fig. 5B). Linc00152 expression was inversely correlated with the DNA methylation of the linc00152 locus in breast cancer in TCGA (Fig. 5C). A previous study showed that the promoter region of linc00152 was hypomethylated in hepatic carcinoma tissues than in para-cancer tissues via bisulfite sequencing polymerase chain reaction [22]

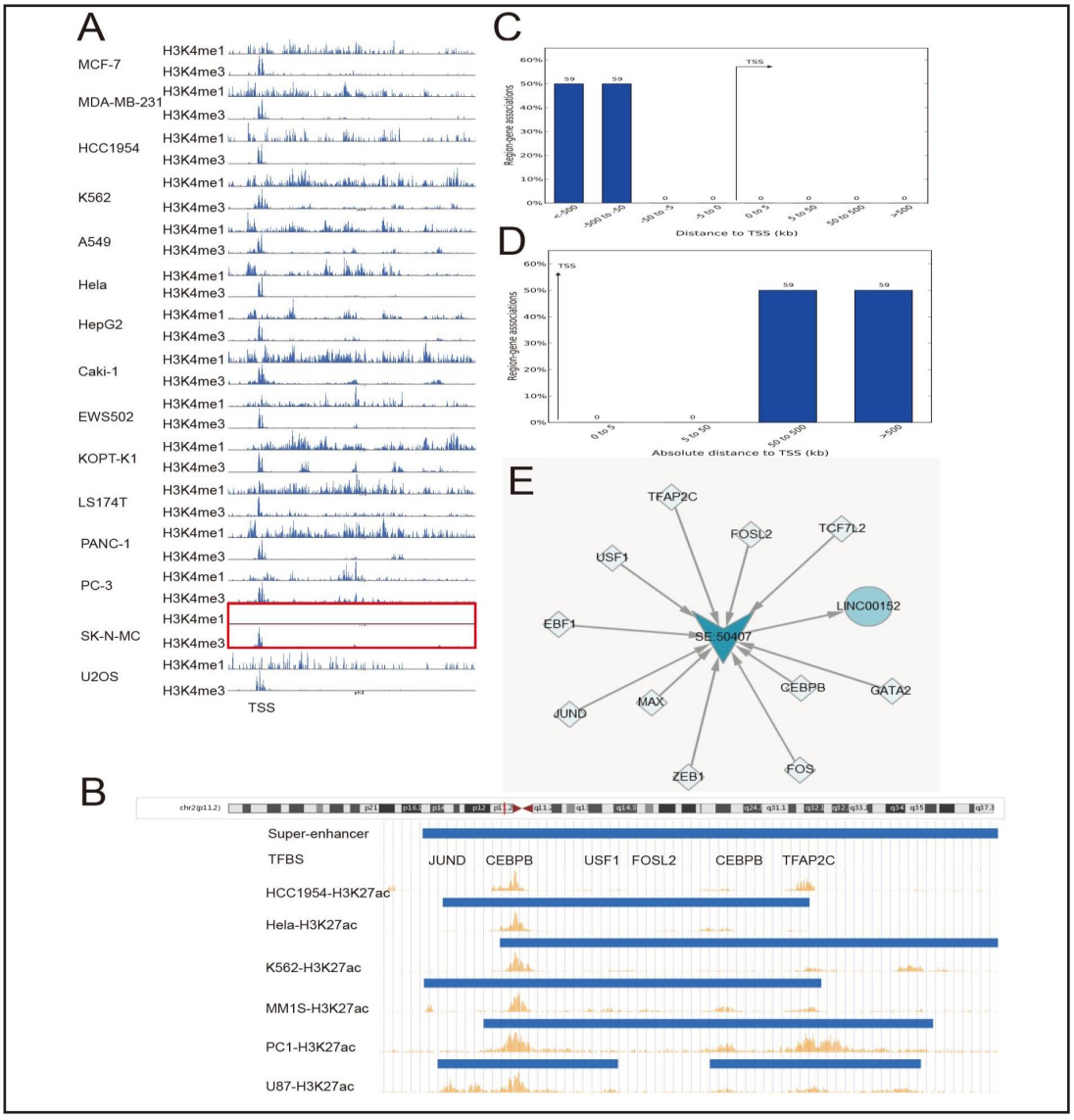

Fig. 6. Linc00152 and promoter/super-enhancer regulation. (A) Linc00152 as a promoter-like lncRNA via analysis of the ratio of H3K4me3/H3K4me1>1.2 in transcription start site of 15 cancer cell lines in the ENCODE data, especially in SK-N-MC cells marked with red. (B) Linc00152 is regulated by super-enhancers in six cancer cell lines in the SEA database. (C & D) The distance of 50 to 500 kb to the target gene transcription start sites was observed in 59 super-enhancers. (E) SE50407, a super-enhancer that regulates the expression of linc00152, was bound by oncogenic transcription factors.

(Fig. 5D). For the upregulation of linc00152 in pan-cancer, we hypothesized that the DNA methylation level would be lower in the linc00152 locus in cancer tissues than that in normal tissues. Thus, the 450K data from TCGA was downloaded and analyzed. The results showed that low linc00152 locus methylation was observed in 17 types of cancers than in their non-cancer counterparts (Fig. 5E).

Linc00152 and promoter/super-enhancer regulation

Next, the relationship between linc00152 expression and histone modification was analyzed using ENCODE data. Based on the ratio of H3K4me3/H3K4me1, lncRNA can be divided into promoter-like lncRNA and enhancer-like lncRNA groups [23]. Linc00152 with the ratio of H3K4me3/H3K4me1>1.2 at the transcription start site is regarded as a

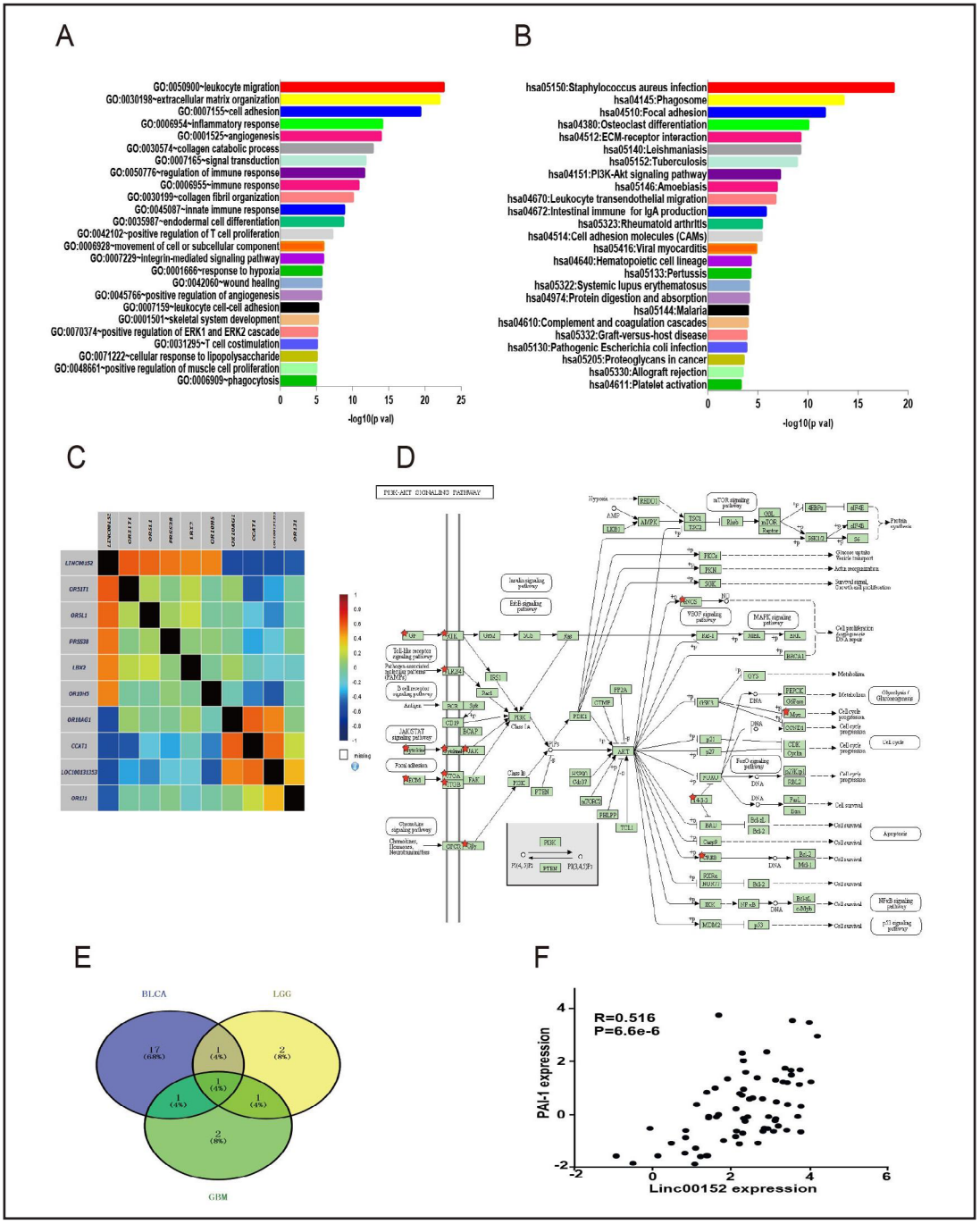

Fig. 7. Potential biological functions of linc00152. (A) Gene ontology enrichment analysis for linc00152 in TCGA. P-values < 0.05 were defined as statistically significant. The vertical axis represents the biological procession category, and the horizontal axis represents the –log10 (P value) of these significant biological processes. (B) KEGG analysis for linc00152 in TCGA. P-values < 0.05 were defined as statistically significant. The vertical axis represents the pathway category, and the horizontal axis represents the –log10 (P value) of these significant pathways. (C) Representative co-expressed genes obtained by guilt-by-association analysis in TCGA. (D) Co-expressed genes of linc00152 that participate in the PI3K-Akt signaling pathway. (E) Overlapping proteins that expressed positively or negative with linc00152 in bladder carcinoma, lower grade glioma and glioblastoma multiforme in TCGA RPPA data. (F) PAI-1 was positively expressed with linc00152 in glioblastoma multiforme (P<0.001).

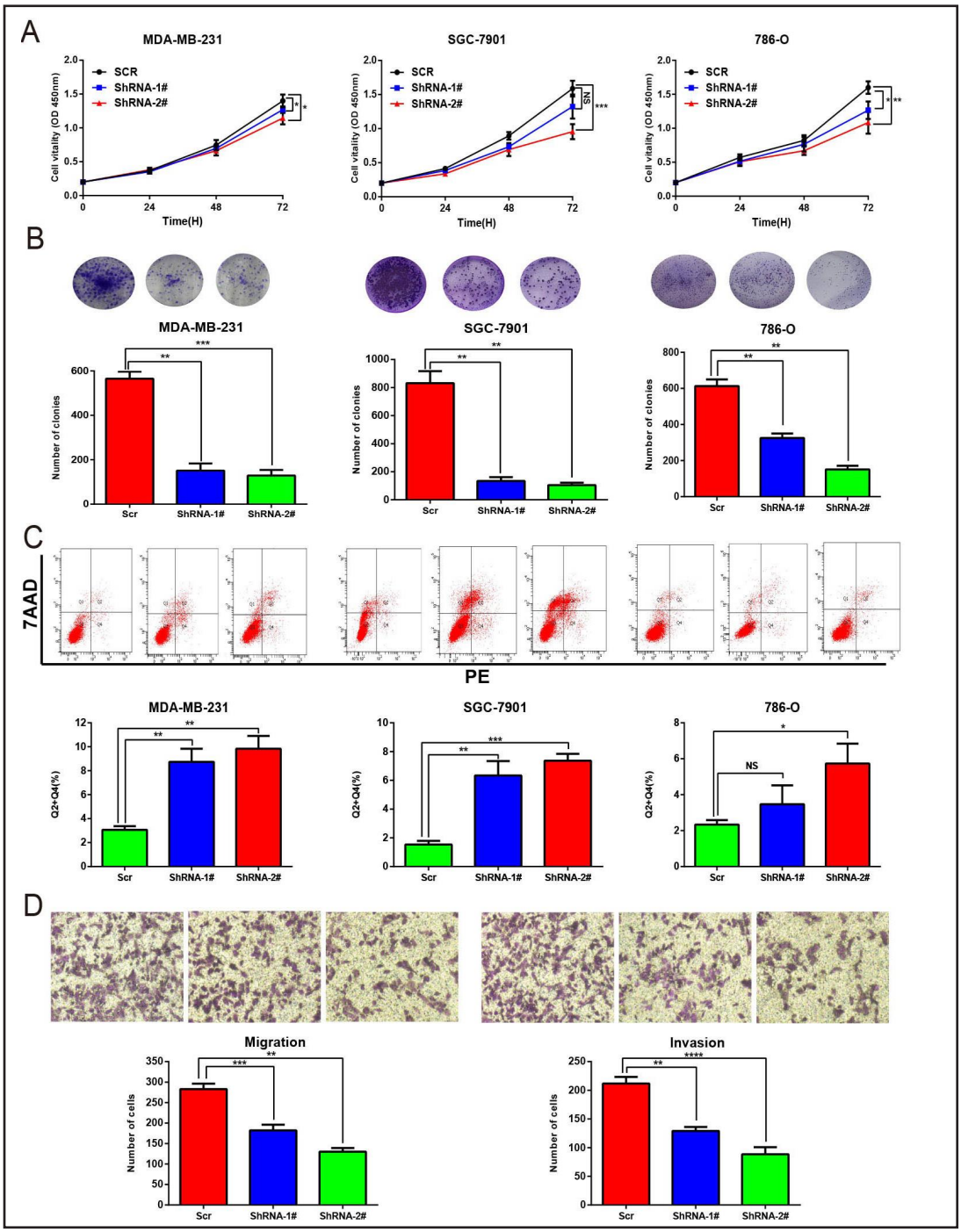

Fig. 8. Linc00152 promotes tumor cell proliferation, migration and invasion. (A) The CCK-8 assay indicated that the knockdown of linc00152 inhibited the proliferation of MDA-MB-231, SGC7901, and 786-O cancer cells, respectively. (B) The clonogenic assay indicated that the knockdown of linc00152 inhibited the clone formation of MDA-MB-231, SGC7901, and 786-O cancer cells, respectively. (C) Knockdown of linc00152 increased the apoptosis of MDA-MB-231, SGC7901, and 786-O cancer cells by flow cytometry, respectively. (D) Knockdown of linc00152 decreased migration and invasion in MDA-MB-231. * P<0.05, ** P<0.01, *** P<0.001. Each assay was performed in triplicate.

promoter-like lncRNA via the analysis of 15 cancer cell lines (Fig. 6A). Super-enhancers activate transcription, and they are assembled in a cell type-specific manner, which is

closely related to the binding of transcription factors and enrichment of H3K27ac [24]. The relationship between linc00152 upregulation and super-enhancer regulation was next investigated using the SEA database [25]. The results indicated that linc00152 is localized in the super-enhancer region in six cancer cell lines (Fig. 6B). The distance of 50 to 500 kb to the target gene transcription start sites is observed in 59 super-enhancers (Fig. 6C and 6D). SE50407, a super-enhancer that regulates the expression of linc00152, was bound by several oncogenic transcription factors, including FOS, ZEB-1, and MAX (Fig. 6E). Thus, Linc00152 as a promoter-like lncRNA and regulated by a super-enhancer may explain its overexpression in pan-cancer.

Potential biological functions of linc00152

To explore the possible functions of linc00152, guilt-by-association analyses were applied using TCGA data, as described in the Methods section. Gene ontology term enrichment (GO) and KEGG pathway analysis were performed for the potential functions and pathways that linc00152 may be involved in using enriched genes. The main functions of linc00152 were leukocyte migration, extracellular matrix organization, immune response, and cell adhesion, and the main pathways that linc00152 were involved in included the *Staphylococcus aureus* infection, phagosome, focal adhesion and PI3K-Akt signaling pathways (Fig. 7A, 7B and see online suppl. material, Table S2). The representative co-expressed genes of linc00152 and genes that participate in PI3K-Akt signaling pathway were examined subsequently (Fig. 7C and 7D). The proteins that were expressed positively or negatively with linc00152 were explored in TCGA RPPA data. Bladder cancer, brain lower grade glioma and glioblastoma multiforme contained such data, and the results indicated that PAI-1 was the overlapping protein in three cohorts (Fig. 7E and see online suppl. material, Table S3). Moreover, PAI-1 was positively expressed with linc00152 in glioblastoma multiforme ($P<0.001$) (Fig. 7F). Additionally, previous studies have indicated that PAI-1 was involved in breast cancer, ovarian cancer, prostate cancer, colorectal cancer, and cervical cancer [26-30]. Thus, linc00152 may function as an oncogene in pan-cancer.

Linc00152 promotes tumor cell proliferation, migration and invasion

To elucidate the potential biological functions of linc00152 in tumorigenesis, a breast cancer cell line (MDA-MB-231), gastric cancer cell line (SGC-7901) and renal cancer cell line (786-O) were selected for functional study. Stable knockdown of linc00152 was performed using purified lentivirus with two different shRNAs (see online suppl. material, Fig. S2). The CCK-8 assay and clonogenic assay indicated that the knockdown of linc00152 inhibited the proliferation and clone formation of these three cancer cell lines (all $P<0.05$) (Fig. 8A and 8B). Moreover, increased apoptosis was observed after the downregulation of linc00152 in the three cancer cell lines by flow cytometry (all $P<0.05$) (Fig. 8C). Decreased migration and invasion were also observed in MDA-MB-231 cell line with knockdown of linc00152 (all $P<0.05$) (Fig. 8D). Next, SGC-7901 cells transfected with either scramble or shRNAs were applied to examine linc00152 role *in vivo*. Tumor growth in the shRNAs group was slower than that in the scramble group ($P< 0.001$) (see online suppl. material, Fig. S3A &B). Finally, RIP and RNA pull-down assays were applied to interpret the molecular mechanism of linc00152 in cancer. The results shown that endogenous linc00152 was enriched in the anti-EZH2 compared to the IgG in SGC-7901 cells ($P< 0.01$) (see online suppl. material, Fig. S3C). Furthermore, RNA pull-down assays confirmed the interaction (see online suppl. material, Fig. S3D). It was reported that EZH2 could bind to promoter of tumor suppressor genes and inhibited their expression [31-35]. Together, these results demonstrate linc00152 promotes cancer development via EZH2 directly.

Discussion

Linc00152, a 828-bp lncRNA that maps to chromosome 2p11.2, was initially defined as differentially hypomethylated during hepatocarcinogenesis [36]. In this study, the expression

of linc00152 in pan-cancer was performed, and the relationship between linc00152 expression and prognosis, chemoresistance, genomic and epigenetic regulation, including DNA methylation and histone modification, was also investigated. Additionally, we elucidated the potential biological functions of linc00152 in tumorigenesis and progression using three types of cancer *in vitro* (MDA-MB-231, SGC-7901 and 786-O). All of the results indicated that linc00152 acted as an oncogenic propellant in pan-cancer from various perspectives.

Previously published studies have revealed that the knockdown of linc00152 could inhibit cell proliferation and suppress cell cycle progression [22, 37, 38]. In our study, we obtained similar results to such studies. Moreover, the pathways downstream of linc00152 were not fully understood. Linc00152 was upregulated in hepatocellular carcinoma tissues and could induce cell proliferation *in vitro* and tumor growth *in vivo* by activating the mTOR signaling pathway [22]. The PI3K-Akt signaling pathway was discovered in our study via KEGG analysis. Moreover, linc00152 induces chemoresistance in pan-cancer via analysis of the public GEO datasets and published reports. Thus, linc00152 would be a *bona fide* oncogenic gene. Linc00152 may be associated with the regulation of the infection process which were annotated by GO and KEGG analyses, respectively (see online suppl. material, Table S2). Microbial infections may disrupt the physiological homeostasis; thus, linc00152 would favor tumorigenesis and cancer progression. Interestingly, linc00152 was closely related to the regulation of immunity annotated by GO and KEGG analyses, respectively (see online suppl. material, Table S2). Cancer immunotherapy is complicated in solid malignancies due to multiple factors, such as the patient's gut microbiome, their diet, and an underlying infection, which can affect or contribute to the success or failure of immune therapy. Yet, as complicated and incompletely understood is the immunology and the relationship between linc00152 expression and immune regulation observed in this study, targeting linc00152 would be an alternative strategy to improve the immunotherapy effect in solid cancer patients.

The mechanism of the upregulation of linc00152 in pan-cancer remains unclear. Transcription factor specificity protein 1 may induce its overexpression in gallbladder cancer [38]. Additionally, there is no other report regarding the exploration of such a mechanism. In this study, genomic and epigenetic regulation was applied to solve this issue. The results indicated that gene amplification was observed in almost all cancers, and low linc00152 locus methylation was observed in 17 types of cancer compared with their counterparts via analysis of Illumina Infinium HumanMethylation450 Beadchip in TCGA. Moreover, linc00152 is regarded as a promoter-like lncRNA via the analysis of H3K4me3/H3K4me1 enrichment in 15 cancer cell lines, and linc00152 was regulated by a super-enhancer. Genomic and epigenetic regulation of linc00152 explained the mechanism of its overexpression in pan-cancer. Moreover, other regulators such as transcription factors, microRNAs, RNA-binding protein and so on would be investigated in future work.

Overall, the comprehensive study of linc00152 in this work broadens the oncogenic lncRNA landscape of pan-cancer. This study revealed linc00152 as a potential therapy target and may aid in the comprehensive management of pan-cancer.

Acknowledgements

This work was supported by funding from the Heilong Jiang province Health and Family Planning Commission Foundation Grant (Grant Number 2016-087), National Natural Science Foundation of China (Grant Number 81602323), China Postdoctoral Science Foundation Grant (Grant Number 2016M600262), Heilong Jiang postdoctoral Foundation Grant (Grant Number LBH-Z16163)、Wu Lien-teh Science Foundation of Harbin Medical University (Grant Number WLD-QN1706), Young Elite Training Foundation Grant of Harbin Medical University Cancer Hospital (Grant Number JY2016-02).

Disclosure Statement

The authors have nothing to disclose.

References

1 DeSantis CE, Fedewa SA, Goding Sauer A, Kramer JL, Smith RA, Jemal A: Breast cancer statistics, 2015: Convergence of incidence rates between black and white women. CA Cancer J Clin 2016;66:31-42.
2 Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F: Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359-386.
3 Mercer TR, Dinger ME, Mattick JS: Long non-coding RNAs: insights into functions. Nat Rev Genet 2009;10:155-159.
4 Chen J, Lin C, Yong W, Ye Y, Huang Z: Calycosin and genistein induce apoptosis by inactivation of HOTAIR/p-Akt signaling pathway in human breast cancer MCF-7 cells. Cell Physiol Biochem 2015;35:722-728.
5 Li C, Chen J, Zhang K, Feng B, Wang R, Chen L: Progress and Prospects of Long Noncoding RNAs (lncRNAs) in Hepatocellular Carcinoma. Cell Physiol Biochem 2015;36:423-434.
6 Liu H, Song G, Zhou L, Hu X, Liu M, Nie J, Lu S, Wu X, Cao Y, Tao L, Chen L, Qian L: Compared analysis of LncRNA expression profiling in pdk1 gene knockout mice at two time points. Cell Physiol Biochem 2013;32:1497-1508.
7 Liu Y, Zhou D, Li G, Ming X, Tu Y, Tian J, Lu H, Yu B: Long non coding RNA-UCA1 contributes to cardiomyocyte apoptosis by suppression of p27 expression. Cell Physiol Biochem 2015;35:1986-1998.
8 Luo G, Wang M, Wu X, Tao D, Xiao X, Wang L, Min F, Zeng F, Jiang G: Long Non-Coding RNA MEG3 Inhibits Cell Proliferation and Induces Apoptosis in Prostate Cancer. Cell Physiol Biochem 2015;37:2209-2220.
9 Xu J, Zhang R, Zhao J: The Novel Long Noncoding RNA TUSC7 Inhibits Proliferation by Sponging MiR-211 in Colorectal Cancer. Cell Physiol Biochem 2017;41:635-644.
10 Huang JL, Zheng L, Hu YW, Wang Q: Characteristics of long non-coding RNA and its relation to hepatocellular carcinoma. Carcinogenesis 2014;35:507-514.
11 Huang X, Zhi X, Gao Y, Ta N, Jiang H, Zheng J: LncRNAs in pancreatic cancer. Oncotarget 2016;7:57379-57390.
12 Li Q, Jia H, Li H, Dong C, Wang Y, Zou Z: LncRNA and mRNA expression profiles of glioblastoma multiforme (GBM) reveal the potential roles of lncRNAs in GBM pathogenesis. Tumour Biol 2016;37:14537-14552.
13 Yu X, Li Z: Long non-coding RNA growth arrest-specific transcript 5 in tumor biology. Oncol Lett 2015;10:1953-1958.
14 Zhang K, Li Q, Kang X, Wang Y, Wang S: Identification and functional characterization of lncRNAs acting as ceRNA involved in the malignant progression of glioblastoma multiforme. Oncol Rep 2016;36:2911-2925.
15 Xu CZ, Jiang C, Wu Q, Liu L, Yan X, Shi R: A Feed-Forward Regulatory Loop between HuR and the Long Noncoding RNA HOTAIR Promotes Head and Neck Squamous Cell Carcinoma Progression and Metastasis. Cell Physiol Biochem 2016;40:1039-1051.
16 Tusher VG, Tibshirani R, Chu G: Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98:5116-5121.
17 Ning LF, Yu YQ, GuoJi ET, Kou CG, Wu YH, Shi JP, Ai LZ, Yu Q: Meta-analysis of differentially expressed genes in autism based on gene expression data. Genet Mol Res 2015;14:2146-2155.
18 Eisen MB, Spellman PT, Brown PO, Botstein D: Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863-14868.
19 Huang da W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44-57.
20 Zhang B, Horvath S: A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol 2005;4:Article17.
21 Yue B, Cai D, Liu C, Fang C, Yan D: Linc00152 Functions as a Competing Endogenous RNA to Confer Oxaliplatin Resistance and Holds Prognostic Values in Colon Cancer. Mol Ther 2016;24:2064-2077.
22 Ji J, Tang J, Deng L, Xie Y, Jiang R, Li G, Sun B: LINC00152 promotes proliferation in hepatocellular carcinoma by targeting EpCAM via the mTOR signaling pathway. Oncotarget 2015;6:42813-42824.

23 Marques AC, Hughes J, Graham B, Kowalczyk MS, Higgs DR, Ponting CP: Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. Genome Biol 2013;14:R131.
24 Chapuy B, McKeown MR, Lin CY, Monti S, Roemer MG, Qi J, Rahl PB, Sun HH, Yeda KT, Doench JG, Reichert E, Kung AL, Rodig SJ, Young RA, Shipp MA, Bradner JE: Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. Cancer Cell 2013;24:777-790.
25 Wei Y, Zhang S, Shang S, Zhang B, Li S, Wang X, Wang F, Su J, Wu Q, Liu H, Zhang Y: SEA: a super-enhancer archive. Nucleic Acids Res 2016;44:D172-179.
26 Chen H, Peng H, Liu W, Sun Y, Su N, Tang W, Zhang X, Wang J, Cui L, Hu P, Liu S: Silencing of plasminogen activator inhibitor-1 suppresses colorectal cancer progression and liver metastasis. Surgery 2015;158:1704-1713.
27 Lara PC, Lloret M, Valenciano A, Clavo B, Pinar B, Rey A, Henriquez-Hernandez LA: Plasminogen activator inhibitor-1 (PAI-1) expression in relation to hypoxia and oncoproteins in clinical cervical tumors. Strahlenther Onkol 2012;188:1139-1145.
28 Fortenberry YM, Brandal SM, Carpentier G, Hemani M, Pathak AP: Intracellular Expression of PAI-1 Specific Aptamers Alters Breast Cancer Cell Migration, Invasion and Angiogenesis. PLoS One 2016;11:e0164288.
29 Mashiko S, Kitatani K, Toyoshima M, Ichimura A, Dan T, Usui T, Ishibashi M, Shigeta S, Nagase S, Miyata T, Yaegashi N: Inhibition of plasminogen activator inhibitor-1 is a potential therapeutic strategy in ovarian cancer. Cancer Biol Ther 2015;16:253-260.
30 Vyas AR, Singh SV: Functional relevance of D,L-sulforaphane-mediated induction of vimentin and plasminogen activator inhibitor-1 in human prostate cancer cells. Eur J Nutr 2014;53:843-852.
31 Fu Y, Chen J, Pang B, Li C, Zhao J, Shen K: EZH2-induced H3K27me3 is associated with epigenetic repression of the ARHI tumor-suppressor gene in ovarian cancer. Cell Biochem Biophys 2015;71:105-112.
32 Hibino S, Saito Y, Muramatsu T, Otani A, Kasai Y, Kimura M, Saito H: Inhibitors of enhancer of zeste homolog 2 (EZH2) activate tumor-suppressor microRNAs in human cancer cells. Oncogenesis 2014;3:e104.
33 Kunderfranco P, Mello-Grand M, Cangemi R, Pellini S, Mensah A, Albertini V, Malek A, Chiorino G, Catapano CV, Carbone GM: ETS transcription factors control transcription of EZH2 and epigenetic silencing of the tumor suppressor gene Nkx3.1 in prostate cancer. PLoS One 2010;5:e10547.
34 Ler LD, Ghosh S, Chai X, Thike AA, Heng HL, Siew EY, Dey S, Koh LK, Lim JQ, Lim WK, Myint SS, Loh JL, Ong P, Sam XX, Huang D, Lim T, Tan PH, Nagarajan S, Cheng CW, Ho H, Ng LG, Yuen J, Lin PH, Chuang CK, Chang YH, Weng WH, Rozen SG, Tan P, Creasy CL, Pang ST, McCabe MT, Poon SL, Teh BT: Loss of tumor suppressor KDM6A amplifies PRC2-regulated transcriptional repression in bladder cancer and can be targeted through inhibition of EZH2. Sci Transl Med 2017;9(378).
35 Ren G, Baritaki S, Marathe H, Feng J, Park S, Beach S, Bazeley PS, Beshir AB, Fenteany G, Mehra R, Daignault S, Al-Mulla F, Keller E, Bonavida B, de la Serna I, Yeung KC: Polycomb protein EZH2 regulates tumor invasion via the transcriptional repression of the metastasis suppressor RKIP in breast and prostate cancer. Cancer Res 2012;72:3091-3104.
36 Cogill SB, Wang L: Co-expression Network Analysis of Human lncRNAs and Cancer Genes. Cancer Inform 2014;13:49-59.
37 Zhou J, Zhi X, Wang L, Wang W, Li Z, Tang J, Wang J, Zhang Q, Xu Z: Linc00152 promotes proliferation in gastric cancer through the EGFR-dependent pathway. J Exp Clin Cancer Res 2015;34:135.
38 Cai Q, Wang ZQ, Wang SH, Li C, Zhu ZG, Quan ZW, Zhang WJ: Upregulation of long non-coding RNA LINC00152 by SP1 contributes to gallbladder cancer cell growth and tumor metastasis via PI3K/AKT pathway. Am J Transl Res 2016;8:4068-4081.