Original Article
Methylation and serum response factor mediated in the regulation of gene ARRDC3 in breast cancer

Sheyu Lin[1*], Ge Zhang[1,2*], Yongfang Zhao[1,3], Danfang Shi[1,2], Qianqian Ye[1,5], Yao Li[1,6], Shuaiyao Wang[1,4]

[1]School of Life Sciences, Nantong University, Nantong 226019, China; [2]Institutes of Basic Medical Science, [3]Institutes of Brain Sciences, [4]School of Life Sciences, Fudan University, Shanghai 200032, China; [5]School of Basic Medical Science, Nanjing Medical University, Nanjing 211166, China; [6]School of Basic Medical Science, Tianjin Medical University, Tianjin 30020, China. *Equal contributors.

Abstract: Breast cancer poses a serious threat to women's life and health and many factors contribute to breast cancer including gene mutation and epigenetics. Gene ARRDC3 was usually repressed in breast cancer and methylation in promoter was reported to be involved in gene ARRDC3 expression regulation. To this end, the methylation status for gene ARRDC3 promoter was assayed by the Massarray quantitative method. The results indicated that different methylation level CpG sites including CpG_6, CpG_13.14, CpG_17.18, and CpG_25 existed between the tumor tissue and the adjacent normal tissue. In order to further verify whether methylation participated in gene ARRDC3 expression, three cell lines were treated with methylation inhibitor Aza-2'-deoxycytidine including A-375, HepG2, and MDA-MB-231. The results revealed that methylation inhibition observably increased ARRDC3 mRNA expression. Then we confirmed the effective length of promoter through the fluorescence report assay used for further analysis. The results showed that the 1746 bp length promoter produced the maximum fluorescence signal. To obtain the direct evidence that methylation in gene ARRDC3 promoter mediated in ARRDC3 expression regulation, the promoter plasmid was methylated by M.SssI enzyme and subjected to the fluorescence report assay. The results showed that methylation in the promoter markedly suppressed relative luciferase activity. In addition, the ecRNA was also analyzed for the methylation regulation and results illustrated that the ecRNA did not regulate ARRDC3 promoter methylation. However, several methylation CpG sites were found to be around CpG_25 site such as TGCATGG, TTGCAA, TTCGTA, and ATAGTT. These sites provide a good clue for further research in methylation for gene ARRDC3 expression regulation. Furthermore, the possible transcription factors involved in the ARRDC3 regulation were investigated by western blot, luciferase activity analysis and ChiP assay. These results documented that gene ARRDC3 expression was improved by SRF and that the methylation affected the interaction between the promoter and SRF. Lastly, the inhibition role of gene ARRDC3 on breast cancer was probed in vivo and in vitro and our results demonstrated that ARRDC3 could inhibit breast cancer growth through the STAT3 signal pathway. In summary, Gene ARRDC3 was inhibited by promoter methylation and was promoted by transcription factor SRF by binding the promoter region and the inhibition on breast cancer growth was exerted by ARRDC3 through STAT3 signal pathway.

Keywords: Breast cancer, ARRDC3, methylation, SRF, STAT3

Introduction

Breast cancer is the most common cause of cancer mortality in female individuals. Its heterogeneity poses immense challenges in deciphering therapeutic strategies. Especially, epigenetic abnormalities play a significant role in the initiation, progression, and metastasis of the disease [1]. Although significant treatment advances, breast cancer remains the second leading cause of cancer-related death in wo-

men [2]. DNA methylation is a key epigenetic change involving the addition of a methyl group to cytosine nucleotides, and this modification is used by living systems to control genes and their genetic programs [3]. The cancer landscape is generally characterized by focal hypermethylation in CpG-rich regions known as CpG islands (CGIs). CGI hypermethylation at promoters represses transcription of genes acting as tumor suppressors, a well-known mechanism operating in cancer. A large fraction of DNA

methylation is also observed in gene body CGIs, with an apparent intriguing positive correlation between methylation and gene expression [4].

ARRDC3, one member of the α-arrestin family, is linked to obesity in men and was recently identified as a regulator of body mass, adiposity, and energy expenditure in mice [5]. Usually, it preferentially lost in a subset of breast cancers [6]. Meanwhile, it was reported to act as a tumor suppressor by facilitating YAP1 degradation in renal cell carcinoma initiation, progression, or metastasis [7]. Furthermore, ARRDC3 could prevent EGF-driven endocytic recycling of ITGβ4 by inducing NEDD4-dependent ubiquitination of ITGβ4 and targeting endosomal ITGβ4 into lysosomes [8]. Additionally, Wang's study indicated that promoter hypermethylation was involved in the inactivation of ARRDC3 in invasive ductal breast carcinoma [9]. And a study said that ecRNA (Extra coding RNA), a newly discovered RNA, had a longer extension at the 3' and 5' end and could regulate the methylation state of the promoter region of the coding gene [10].

Serum response factor (SRF) is a transcription factor and belongs to the MADS box family. Its role was correlated with various cellular processes such as cell proliferation, differentiation, and apoptosis [11]. Generally, it is widely expressed and mediates serum- and growth factor-induced activation of the immediate early genes by acting at serum response elements [12]. Besides, SRF functions in orchestrating disparate programs of gene expression linked to muscle differentiation and cellular growth [13]. SRF depletion could affect the expansion of the high and low differentiation grade HCC cells [14]. Moreover, studies showed that the upregulation of SRF expression in Gastric carcinoma promoted its invasion and metastasis [15].

STAT3 is a key transcription factor that mediates the expressions of a variety of genes in response to cell stimuli and plays a key role in cell survival. It also initiates inflammation, programmed death-ligand 1 (PD-L1) upregulation and cell transformation in numerous malignant cells, and underlays the bad prognosis of the diseases [16]. The JAK/STAT (Janus kinase/signal transducer and activator of transcription) signaling pathways are major pathways for cytokine signaling and regulate important cellular events, such as hematopoiesis and immune development. Aberrant STAT3 signaling is frequently linked to cancer cell proliferation, survival, metastasis, tumor immunosuppression, and angiogenesis. JAK2 is the major kinase of STAT3 and has been reported to be constitutively active in many cancers and other proliferative diseases [17].

In our study, we analyzed the promoter status of gene ARRDC3 by MASSARRAY. Then through luciferase activity experiments, we confirmed that methylation in the promoter was involved in gene ARRDC3 regulation. Next, we confirmed that the transcription factor SRF could promote gene ARRDC3 expression by binding its promoter which was disturbed by methylation. Lastly, through cell behaviors analysis and in vivo nude mice experiment we found that gene ARRDC3 could inhibit breast cancer growth through the STAT3 signal pathway.

Materials and methods

Samples and cell lines

Informed, written consent regarding the use of the tissue samples was obtained from each patient prior to the study. Seventeen breast cancer samples (tumor) and adjacent matched normal tissues (non-tumor) from breast cancer patients were collected at Henan Cancer Hospital (Zhengzhou, China). Ethical approval was obtained from the research ethics committee of the hospital. The cell lines used in this study were obtained from the Shanghai cell bank of the Chinese Academy of Sciences (Shanghai, China) (A-375, MDA-MB-231, MDA-MB-468, Hep G2, and MCF-7).

Reagents and materials

The rabbit polyclonal antibodies from Cell Signaling Technology (Massachusetts, USA) include ARRDC3, Akt, PI3K, MEK1/2, ERK1/2, JAK2, and STAT3. The rabbit polyclonal antibody for GAPDH was purchased from the BOSTER Biological Technology Co.ltd (Wuhan, China). Dulbecco's modified Eagle medium: DMEM/F-12 was supplied by Corning (Grand Island, NY, USA). Fetal bovine serum (FBS) was from Zhejiang Tianhang Biotechnology Co.ltD. Penicillin/streptomycin and trypsin solution were from Thermo Fisher Scientific (Shanghai, China). The decitabine was got from sigma Inc. (Shanghai, China). The kits for cell viability, cell apoptosis, and ChiP assay were all from Be-

Table 1. Primers used for gene ARRDC3 MassArray quantitative methylation analysis

Primers	Sequences (5'-3')	Length (bp)
Meth4s	aggaagagagTGTTTTGAATTTTAGTGTGATGTGA	473
Meth4a	cagtaatacgactcactatagggagagaaggctAAAAACAACCAAATTATTTCCAAAA	
Meth6s	aggaagagagGAGAAAGAAGTTGAGGAGATAGTAAAGTT	386
Meth6a	cagtaatacgactcactatagggagagaaggctCAAAACCAAATTAAAATAAACTACAAC	
Meth8s	aggaagagagGTTGTAGTTTATTTTAATTTGGTTTTGTA	352
Meth8a	cagtaatacgactcactatagggagagaaggctCAACTTTATAAAAAACAACCCCAATAC	
Meth12s	aggaagagagATTGGGGTTGTTTTTTATAAAGTTG	365
Meth12a	cagtaatacgactcactatagggagagaaggctCCCCACTAAAATACACAAAAACATT	
Meth14s	aggaagagagAATGTTTTTGTGTATTTTAGTGGGG	366
Meth14a	cagtaatacgactcactatagggagagaaggctAAAAAATCCAAACCTTCCAACTTTA	

yotime biotechnology. Other chemicals were of the highest grade of purity commercially available.

DNA/RNA extraction

Genomic DNA was isolated using the Genomic DNA Mini Preparation Kit with Spin Column (Beytime biotechnology, Shanghai, China). RNA was gathered by using Trizol reagent (Invitrogen, Shanghai, China). The concentration and quality of the isolated DNA and RNA were measured with the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Montchanin, DE, USA).

Bisulfite conversion and MassArray quantitative methylation analysis

The bisulfite conversion of genomic DNA was done with the EZ DNA Bisulfite Treatment Kit (ZYMO Research, CA, USA). The methylation level was measured following the MassCLEAVE training protocol under the MassArray Compact System (Sequenom, San Diego, CA, USA). The target region in the promoter was from -1223 bp to 721 bp (GRCh37/hg19, UCSC genes) and the primer pairs are indicated in **Table 1**. The resultant methylation calls were analyzed on Epi-Typer software (Sequenom, CA, USA) to generate quantitative CpG methylation results.

Promoter composition analysis

The transcription factor binding sites in the SASH1 promoter were predicted by using the TFSEARCH software (http://mbs.cbrc.jp/research/db/TFSEARCH.html) and WWW promoter scan software (http://www-bimas.cit.nih.gov/molbio/proscan/index.html) and confirmed by the high quality transcription factor binding profile database.

Real-time PCR

The cDNA was reversely transcribed by PrimerScriptRT Reagent kit (Takara Bio Inc., Otsu, Japan). The ARRDC3 gene was co-amplified with a fragment of the glyceraldehyde 3-phosphate dehydrogenase (GAPDH) gene, which served as an internal standard. Quantitative PCR was conducted by amplifying 1.0 µL cDNA with the SYBR Premix Ex Taq kit (Takara Bio Inc.) on the ABI 7500 Fast Real-Time PCR system (Life Technologies, Carlsbad, CA, USA). The cycling conditions of forty cycles of PCR were 95°C/5 sec, 55°C/30 sec and 72°C/30 sec. The amount of specific mRNA was quantified by determining the point at which the fluorescence accumulation entered the exponential phase (Ct), and the Ct ratio of the target gene to GAPDH was calculated for each sample. Each sample was run in four repeats and all the PCR data were analyzed with the ABI 7500 system software version 2.0.4.

Cell culture

The cell lines were all grown in DMEM/F12 basal medium containing 10% fetal bovine serum, 100 units/ml penicillin and 100 units/ml streptomycin. The flasks were incubated in a humidified incubator under the condition of 37°C with 5% CO_2. The plates were observed under a phase contrast microscope for cell adherence and morphology. In the methylation analysis experiment, the concentration of decitabine was at 50 µM.

Vector construction

The gene ARRDC3 full length and ecRNA fragments (1100 bp and 400 bp) for gene ARRDC3 were synthesized and inserted into pCDH plasmid. The synthesized interference fragments

were annealed to double strands and inserted into pGreen vector. Besides, three promoters for gene ARRDC3 (-1377-369, -882-369 and -559-369) were amplified and cloned into the pGL3-Basic plasmid. All the inserts were verified by DNA sequencing analysis.

Bisulfite sequencing

The extraction and bisulfite conversion for genomic DNA were as above. Bisulfite-treated DNA was amplified using two pairs of specific primers covering the region from -186 to 348 in gene ARRDC3 promoter. The sequences for the two pairs of primers were as follows: 5'-ATAATGAAGAGAAGT TAGGTTTAGT-3' (forward) and 5'-ACTTTATAAAAAACAACCCCAATAC-3' (reverse); 5'-GGGGTTGTTTT TTATAAAGTTG-3' (forward) and 5'-CCCACTAAAATACACAAAAACA-3' (reverse). The PCR products were cloned into T/A vector and twenty monoclonal vectors were selected for sequencing in each group.

Western blot

The whole protein was obtained from lysed cells using RIPA buffer with PMSF (1 mM). The extracted proteins were quantified by the BCA assay (Haimen, China) and analyzed by Western blot. The protein was separated by SDS polyacrylamide gel electrophoresis and transferred to a nitrocellulose membrane (GE Healthcare Life Sciences, China), which was then blocked overnight with TBST containing 5% skimmed milk powder. The membranes were rinsed with TPBS and treated with polyclonal primary antibody and with the HRP-conjugated secondary antibody. Lastly, protein bands were visualized by a chemiluminescent detection system (ECL, BOSTER, China). Densitometry was performed with Image Lab™ software (Bio-Rad Laboratories, CA, USA), and the band signal was normalized to the level of GAPDH protein.

Cell viability

The cell viability was measured by MTT assay. The cells were seeded into 96-well plates at the density of 3×10^5/mL in three repeats with the volume 100 μL/well and allowed to adhere overnight. Forty-eight hours later, 100 μL MTT (0.5 mg/mL) was added and the cells were incubated for another 4 h. Then, the medium was removed and 100 μL dimethylsulfoxide was added and incubated for another 2 h. After the formazan crystal dissolved, OD values were measured at 570 nm wavelength with En-Spire-2300 multimode plate reader (Perkin Elmer, Norwalk, USA). The experiment was repeated three times.

Cell scratch

The breast cell lines were seeded on 60-mm culture dishes and incubated at 37°C with 5 % CO_2. After 24 h, the attached cells were scratched three times in parallel with a 100 μl pipette tip. The incubation continued for another 24 h and the images for wound healing were captured with a 10 × objective lens using a Nikon camera (Nikon Corporation, Tokyo, Japan).

Fluorescent system assay

The fluorescent system assay was conducted as follows: first, the Cells (4×10^4 cells/well in 24-well plate) were transiently transfected with 0.1 μg of pRL-Tk and the target luciferase constructs. 48 h after transfection, the cells were harvested with lysis buffer and the luciferase activities of the cell extracts were measured by the dual luciferase assay system (Promega, Beijing, China). The luciferase activity was normalized for transfection efficiency with Renilla luciferase activity.

Chromatin immunoprecipitation assay

Chromatin immunoprecipitation assay was performed using the ChiP Assay Kit from beyotime Biotechnology (Shanghai, China) following the manufacturer's instructions. Briefly, the cells were seeded at 3×10^5 cells per 100-mm dish 48 h prior to treatment. Then the cells were fixed with formaldehyde (1% final concentration, Sigma-Aldrich) and the Glycine solution was added. Next, the cells were rinsed with cold PBS and collected in a cell lysis buffer. The resulting chromatin was sheared on ice for 15 cycles of 30 s ON/40 s OFF with the ultrasonic homogenizer (Scientz, Ningbo, China) to yield an average length between 200 and 1000 bp. The sheared cross-linked chromatin was diluted in the dilution buffer and incubated with the immunoprecipitating antibody (SRF) and Protein A/G beads overnight at 4°C with rotation. The protein/DNA complex was washed with wash buffer and reverse cross-linked with 5 M NaCl by incubating at 65°C for 4 h. Then the DNA was purified by the DNA purification kit and used for quantitative Realtime PCR. The

primer sequences were SRF-ChIP-F: 5'-TTTTA-GCCGCCCGTCTTTG-3' and SRF-ChIP-R: 5'-TCA-CCTTTCCCAGC ACCAT-3'. The qPCR data were calculated as fold enrichment. The irrelevant rabbit antibody serum was used as the control.

The establishment for ARRDC3 or ecRNA stable cell line

Stable cell line expressing ARRDC3 protein or ecRNA or corresponding interference plasmids were created as the following steps: firstly, the gene full sequence or interference sequence were synthesized and inserted into corresponding plasmids pCDH or pGreen; secondly, the lentivirus was produced by transiently transfecting the 293T cells with the target vector and corresponding helper plasmids; lastly, the clones were introduced into cell line MDA-MB-231 via lentiviral transfection. Cells were expanded under GFP selection to be stored at liquid nitrogen.

Animal model

Four-week female BALB/c nude mice were purchased from Shanghai SLRC laboratory animal Co., Ltd and maintained in the animal facilities. Five female BALB/c nude mice per group were subcutaneously injected at the armpit and hind leg area with MDA-MB-231 cells stably transfected with pCDH-ARRDC3, pCDH or pGreen-interference fragments respectively (1×10^8 cells in 100 μl of basic medium). Then the mice were observed every week and after 8 weeks, sacrificed to recover the tumors. The wet weight and the volume of each tumor were determined.

Statistics analysis

The methylation rates in two independent sample groups were compared using the Mann-Whitney test. One-way ANOVA or student's t-test was used to analyze the corresponding data. All *p*-values were two-sided and P<0.05 was considered statistically significant. All statistical analysis was carried out with SPSS version 15.0 (SPSS Inc, Chicago, IL, USA). The cases' hierarchical cluster analysis clustered the 32 CpG sites in the ARRDC3 promoter based on Euclidean distances and the average linkage clustering algorithm. This clustering was implemented using Cluster 3.0 and viewed on Java Treeview.

Results

The methylation in the promoter and the expression of gene ARRDC3 in breast cancer

According to MassArray quantitative methylation analysis in 17 breast cancer samples, the mean methylation level of each CpG site was used to be compared between nontumor and tumor tissues (**Figure 1A**, Supplementary Figure 1). With the statistics analysis by t-test, significant differences (P<0.05) were revealed at the following CpG sites: CpG_6, CpG_13.14, CpG_17.18, CpG_25. Among them, at CpG_6 and CpG_25 CpG sites, the mean methylation level in tumor tissue was higher than that in the nontumor tissue, but the other two CpG sites have the opposite trend.

Additionally, the promoter region from CpG_6 to CpG_25 was analyzed with WWW promoter scan software and TFSEARCH software for searching the transcription factors binding sites. Combining with the different CpG sites, CpG_6, CpG_13.14, CpG_17.18, CpG_25 and further determining by the high quality transcription factor binding profile database, five sites were selected as the potential transcription factor binding sites which were possibly involved in the gene ARRDC3 regulation as follows sp1, C-myb, GRalpha, GRbeta and SRF (**Figure 1B**, only shown SRF).

Then, the general methylation feature across the whole promoter region was analyzed by unsupervised clustering and the results showed that, based on the ratios of methylation level (tumor/nontumor), the methylation levels were up-regulated mainly centralized around CpG_25 site (from CpG_22 to CpG_29, represented in red in **Figure 1C**) compared with nontumor tissues.

Furthermore, the gene ARRDC3 expression was also measured by Realtime-PCR and the results indicated that the mRNA level in the tumor group was significantly lower than that in nontumor group (**Figure 1D**). In addition, the ARRDC3 expression was analyzed by the online tool UALCAN (http://ualcan.path.uab.edu/index.html) and the same results were produced in the TAGC database (Supplementary Figure 1). Consequently, together with the methylation results, we speculated that the methylation around the TSS sites perhaps participated in the gene ARRDC3 expression regulation.
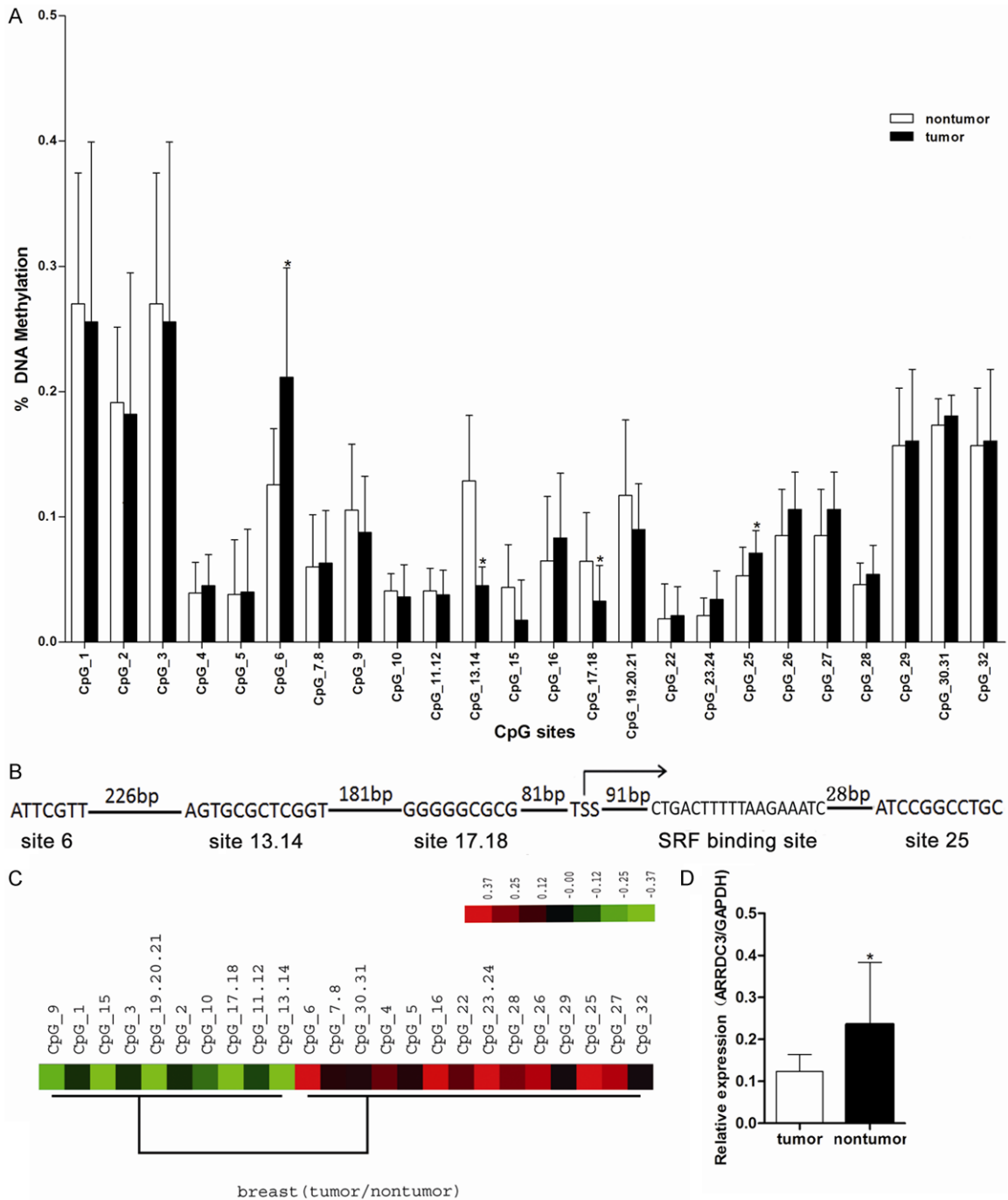
Figure 1. The methylation status of gene ARRDC3 promoter and its expression in breast cancer tissue. A. Comparison of mean methylation for each CpG site between nontumor and tumor. The X-axis represents 32 informative CpG sites within 5 MassArray amplicons for ARRDC3 promoter; the Y-axis shows the average methylation value of each CpG site (or clusters of CpG sites). Error bars = SD. Significant differences were indicated as *(P<0.05) by the t-test. The CpG_6, CpG_13.14, CpG_17.18, CpG_25 emerged with a significant difference between the notumor group and the tumor group. B. 5' end of ARRDC3 gene, indicating the position of significant difference CpG sites, transcriptional start site from the UC Santa Cruz Genome Browser shown with bent right arrows and transcription factor SRF binding site. C. The hierarchical cluster analysis of methylation patterns of 32 CpG sites measured on 17 samples. The methylation level (subtracting the general mean value) of each CpG site within each sample is presented in the plot with color ranging from green (indicating low methylation) to red (indicating high methylation). The high methylation region was found to be mainly centralized around the CpG_25 site (from CpG_22 to CpG_29). D. Relative mRNA expression level of ARRDC3 in tumor or non-tumor tissue detected by real-time quantitative RT-PCR (*P<0.05). The tumor tissue showed the lower ARRDC3 level related to that in the normal tissue. The statistics were performed with the t-test.

The relationship between the methylation and gene ARRDC3 expression

To verify the relationship between ARRDC3 promoter methylation and its expression, mRNA expression levels were measured before and after treatment with 5-AzadC in A375, HepG2 and MDA-MB-231 cell lines. The results indicated that the gene ARRDC3 expression was increased in all three cell lines after treatment with 5-Aza-dC, and that especially, in cell line MDA-MB-231, the improvement is about 4 folds (**Figure 2A**). These data implied that methylation was closely correlated with the inhibition of gene ARRDC3 expression.

In order to confirm the effective length of the ARRDC3 promoter involved in the gene ARRDC3 expression regulation, we constructed three different lengths of promoter reporter vectors that is, 1746 bp, 1251 bp, and 928 bp. Then the luciferase report assays were performed with the three reporter vectors and the results indicated that 1746 bp promoter possessed the best initiating transcription function (**Figure 2B**).

However, whether the methylation participated in the gene ARRDC3 expression regulation still needed to be answered. Then the 1746 bp promoter was used as the reporter system in the following experiments based on the above results. The 1746 bp plasmid was methylated by M.SssI enzyme and transfected into cell line MDA-MB-231 and the luciferase report assay was conducted. The results revealed that the methylation treatment significantly reduced the fluorescence ratio of firefly to Renilla. The data showed that methylation in the promoter mediated in gene ARRDC3 expression regulation (**Figure 2C**).

Additionally, we had known that ecRNA regulated the promoter methylation by blocking the interaction between DNMT1 and gene promoter region. Hence, we constructed the ecRNA overexpression and interference vectors such as ecRNA1100, ecRNA400 and ecRNAi. By transfection with the corresponding lentivirus, we get the corresponding stable cell lines. And the ecRNA expression was verified by realtime-PCR. The results indicated that the ecRNA level in the overexpression group was higher than that in the other two groups and no difference existed between the interference group and the control group (**Figure 2D**).

In order to ascertain whether the ecRNA influence the promoter methylation level of gene ARRDC3, the genome from the four stable cell lines were extracted and were bisulfite converted. Afterward, the BSP assay was executed. The BSP results indicated no significant difference on the methylation level at different CpG sites. But we found many methylation CpG sites and methylation non-CpG sites which were displayed at the bottom of **Figure 2**.

The methylation in the promoter and its effect on the binding for transcription factor SRF

In order to confirm which transcription factor was involved in the gene ARRDC3 expression regulation, we transiently transfected the cell line MDA-MB-231 with the five different transcription factors overexpressed plasmids including C-myb, Sp1, GRalpha, GRbeta, and SRF. Then the expression of gene ARRDC3 expression was measured by western blot. The results indicated that transcription factor SRF could effectively promote gene ARRDC3 expression compared to other transcription factor transfection groups (**Figure 3A**).

However, whether the transcription factor SRF could directly bind the promoter of gene ARRDC3 still needed to be verified. So the experiments of the transcription factors SRF plasmids and the promoter plasmids co-transfection into the MDA-MB-231 cells were performed. Then the fluorescence report assay was conducted and the results illustrated that transcription factor SRF could significantly increase the fluorescence ratio of firefly to Renilla. These data suggested that the transcription factor SRF could directly bind the promoter of gene ARRDC3 (**Figure 3B**).

Whether methylation was involved in the gene ARRDC3 promoter binding by transcription factor SRF needed to be confirmed, so the fluorescence report analysis was performed after cell line MDA-MB-231 co-transfected with SRF plasmid and ARRDC3 promoter plasmid methylated or not. The results indicated that the methylation could reduce the fluorescence ratio of firefly to Renilla. These data demonstrated that methylation could affect the gene ARRDC3 expression regulation by blocking the binding between the transcription factor SRF and the promoter (**Figure 3C**).

In order to further confirm the interaction between the transcription factor SRF and the
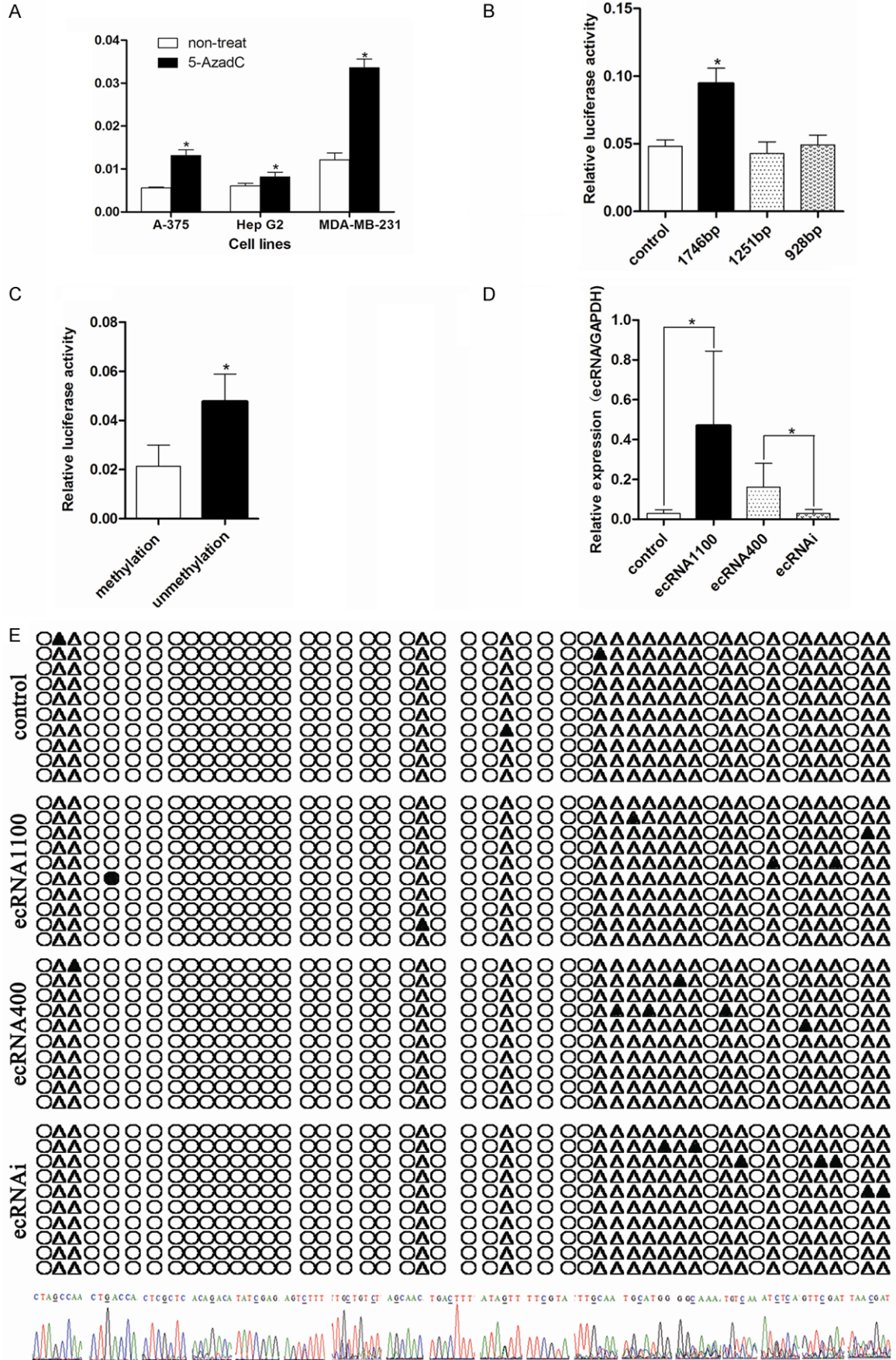
Figure 2. Promoter methylation of gene ARRDC3 was involved in its expression regulation. A. Quantitative analysis of ARRDC3 mRNA in three cell lines before and after treatment with 5-AzadC, the results were expressed as the ratio of copies of the target gene relative to GAPDH. Error bars = SD. The statistics analysis was performed with a t-test and significant differences are indicated by * (P<0.05). ARRDC3 could be inhibited significantly by the reagent 5'-Aza-dC. B. The fluorescence report assay in cell line MDA-MB-231 transfected with the three different length pGL3-ARRDC3 promoter plasmids and pGL3 as control. After statistics analysis with one way ANOVA method, we found the longest promoter plasmid had the maximum transcription activity. C. The fluorescence report assay in cell line MDA-MB-231 transfected with the 1746 bp length pGL3-ARRDC3 promoter plasmid which was methylated or not. The experiment indicated that methylation mediated in the gene ARRDC3 expression. D. The ARRDC3 ecRNA expression level in different stable cell lines detected by real-time quantitative RT-PCR. The statistics analysis was performed with one way ANOVA method and the results demonstrated that these stable cell lines were successfully established. E. Bisulfite sequencing results for different stable cell line (from top to bottom was: control group, ecRNA 1100 group, ecRNA 400 group, and ecRNAi group). Filled circles: methylated CpG sites; open circles: unmethylated CpG sites; filled triangles: methylated CpG sites; open triangles: unmethylated non CpG sites. And the bottom of graph E was the specific sequence representing the methylated sites including the CpG sites or non-CpG sites. We found that ecRNA was not correlated with the methylation of ARRDC3 promoter region.
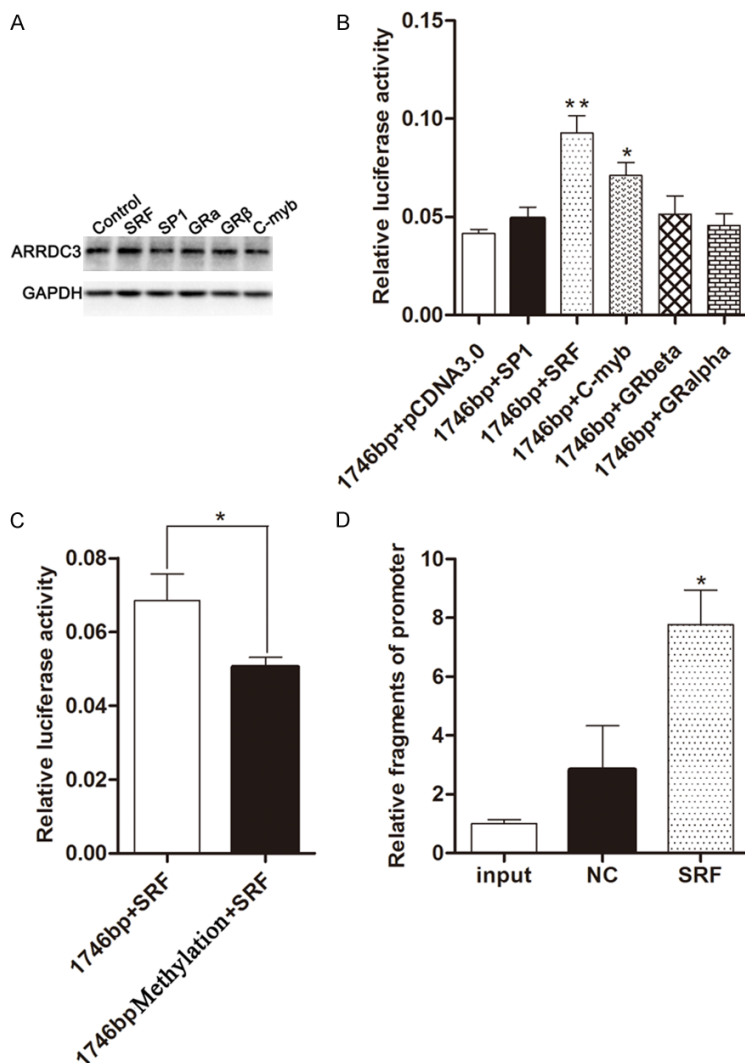

SP1, GRalpha, GRbeta, and C-myb. After statistics analysis performed with one way ANOVA method, we found that SRF and C-myb could interact with the ARRDC3 promoter. C. The fluorescence report assay in cell line MDA-MB-231 co-transfected transcription factor SRF plasmid and 1746 bp length pGL3-ARRDC3 promoter plasmid methylated or not respectively. The experiment revealed that methylation of the promoter region could influence the interaction between the transcription SRF and ARRDC3 promoter region. D. The affinity of SRF to the promoter of ARRDC3 was determined by ChIP assay and qPCR, and qPCR quantification was performed by specific ChIP primers. Quantification of the qPCR data in the immunoprecipitates is shown as fold enrichment relative to input treatment. The results are expressed as the mean ± SEM of three independent experiments. The statistics analysis was performed with one way ANOVA method. Transcription factor SRF was further proved to be involved in the ARRDC3 promoter binding.

Figure 3. Transcription factor SRF could promote gene ARRDC3 expression by binding its promoter. A. Western blot showed ARRDC3 expression in cell line MDA-MB-231 transfected with different transcription factors such as SRF, SP1, GRalpha, GRbeta, and C-myb. And SRF was found to be involved in the increase of ARRDC3 expression. B. The fluorescence report assay in cell line MDA-MB-231 co-transfected with 1746 bp length pGL3-ARRDC3 promoter plasmid and different transcription factor plasmids such as SRF,

promoter of gene ARRDC3 in vivo, the ChiP and Realtime-PCR experiments were performed and the results indicated that the SRF group recruited the highest amount of promoter fragments among the three groups. The data illustrated that transcription factor SRF could improve the gene AR-RDC3 expression by binding the gene ARRDC3 promoter (**Figure 3D**).
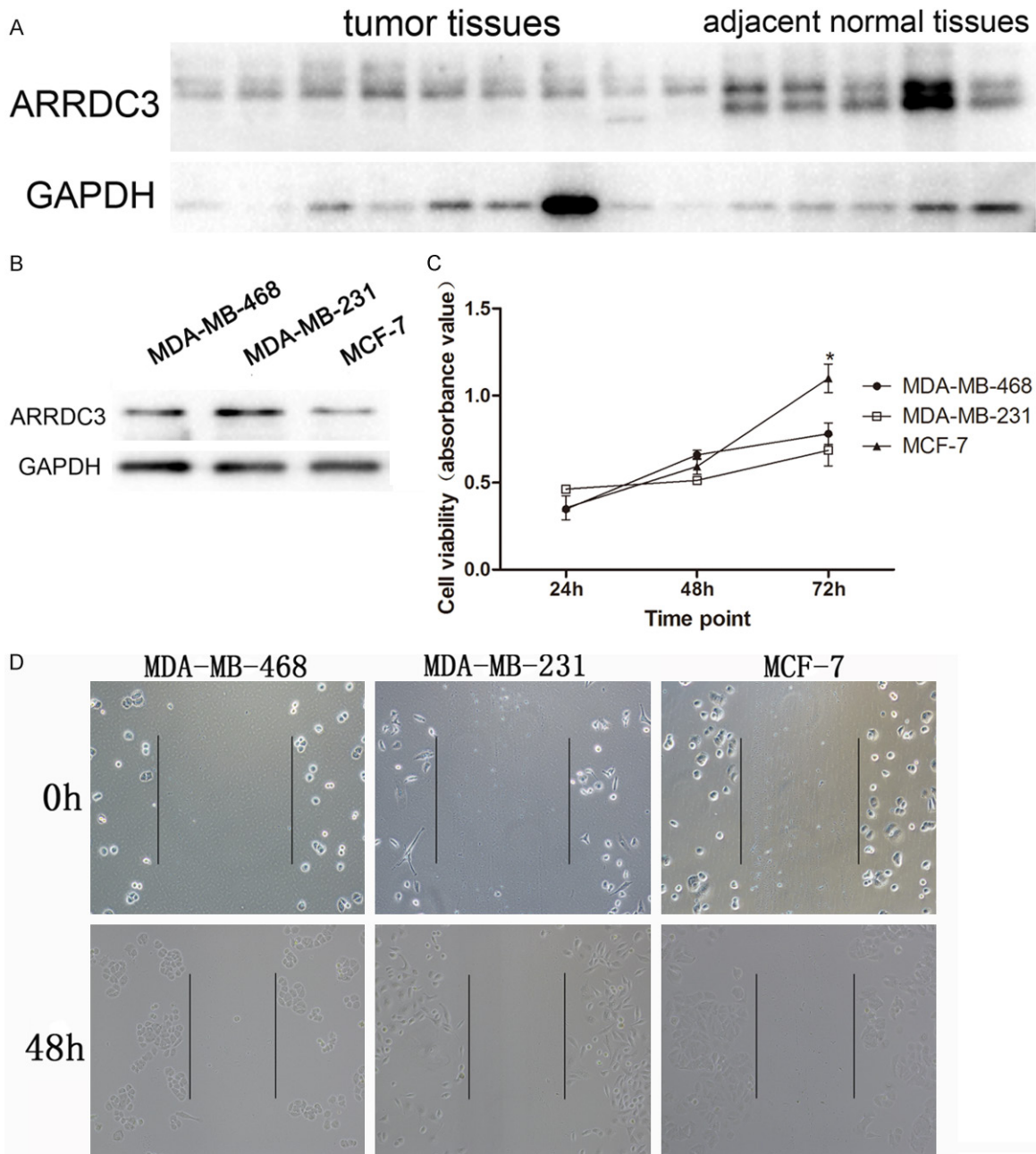
Figure 4. The correlation of expression of ARRDC3 in different cell lines with their behaviors. A. The protein level of ARRDC3 in tumor tissue and adjacent normal tissue assayed by western blot. After statistics analysis by t-test, we found that tumor tissue presented a lower protein level of ARRDC3 compared with that in normal tissue. B. ARRDC3 protein was measured in three different cell lines by western blot. C. The cell viability was assayed in different cell lines at different time points such as 24 h, 48 h, and 72 h. After statistics analysis by one way ANOVA method, ARRDC3 was found to be involved in cell proliferation. D. The cell scratch experiment was performed in different cell lines at 24 h and 48 h time point. No obvious difference was found among the three groups.

The association between the gene ARRDC3 and cell line behaviors

Although the ARRDC3 expression in protein was reported to be reduced in breast cancer, we verified its expression again in breast cancer tissue and adjacent normal tissue. Our results indicated that the protein level of gene ARRDC3 in the tumor tissue was significantly lower than that in the adjacent normal tissue (**Figure 4A**).

However, could different breast cancer cell lines display a different expression of gene

ARRDC3? To this end, three breast cancer cell lines including MDA-MB-468, MDA-MB-231, and MCF-7 were tested. The results demonstrated that the highest expression of gene ARRDC3 emerged in cell line MDA-MB-231 and the lowest expression emerged in cell line MCF-7 (**Figure 4B**).

According to the above results of different expression status in three cell lines, we intend to explore the cell proliferation and migration abilities by using the three cell line models. The proliferation assay showed that the cell line MCF-7 had the highest cell viability at 72 h time point. And no difference was found between cell line MDA-MB-468 and MDA-MB-231. In addition, from time point 24 h to 72 h, the proliferation curve of MCF-7 possessed the biggest slope among the three curves (**Figure 4C**). Furthermore, the migration was analyzed by cell scratch assay and the results manifested no significant difference in migration among them (**Figure 4D**).

The role of gene ARRDC3 in breast cancer growth and its mechanism

To further investigate the role of gene ARRDC3 in breast neoplasm growth, the stable cell lines were established including ARRDC3 overexpression, ARRDC3 interference and control groups infected with the overexpression lentivirus, the interference lentivirus and pCDH blank lentivirus. Then the three stable cell lines were affirmed by western blot. The results proved that the three stable cell lines were produced successfully (**Figure 5A**).

Because the above proliferation assay of different cell lines have suggested that gene ARRDC3 was possibly involved in the breast neoplasm growth, so the proliferation assay was performed in the three stable cell lines. The results revealed that the interference cell line presented the highest proliferation ability among them and that the overexpression cell line had the lowest proliferation ability (**Figure 5B**).

Besides, the apoptosis experiment for the three stable cell lines was also carried out. The results documented that the inhibition of gene ARRDC3 expression significantly reduced the apoptosis rate and that overexpression of gene ARRDC3 obviously increased the apoptosis rate among the three stable cell lines (**Figure 5C**).

In order to explore the mechanism behind the inhibition of the growth by gene ARRDC3, several signal pathways were probed in the three stable cell lines including Akt-PI3K, MEK1/2-ERK1/2 and JAK2-STAT3. The results showed that in the gene Akt, PI3K, MEK1/2, ERK1/2 and JAK2, they all displayed no difference among the three groups. But in gene STAT3, the significant difference was found among them (**Figure 5D**). In addition, these signal pathways were also analyzed in the TCGA and CPTAC databases with the online tool UALCAN (Supplementary Figures 2, 3, 4). The results indicated that except genes Akt and STAT3, other genes possessed the same variation trend in mRNA and protein level. Based on the results, we concluded that the three signal pathways were all involved in breast cancer occurrence and development. Furthermore, that the ARRDC3 and STAT3 had the same variation trend in mRNA level indirectly implied that ARRDC3 was possibly involved in the STAT3 expression regulation.

However, whether gene ARRDC3 could exert its effect on the cell proliferation in vivo needed to be solved, the nude mouse model was established by injecting corresponding stable cell line respectively. The results indicated that the tumor size in the overexpression group was significantly smaller than that in the control group or the interference group (P<0.05) (**Figure 5E**). In addition, the average tumor weight in the overexpression group was also significantly lower than that in the other two groups and the interference group displayed the inverse effect (**Figure 5F**). The volume also had the same trend as the weight (**Figure 5G**). These findings further demonstrated that ARRDC3 was involved in the inhibition of proliferation for breast cancer.

Discussion

DNA methylation is a risk factor for breast cancer. Several genome-wide studies of DNA methylation have found evidence that increased methylation levels within functional promoters have been associated with an increased risk of breast cancer. The peripheral blood DNA methylation at the BRCA1 promoter was found to be associated with an estimated 3.5-fold (95% CI, 1.4-10.5) increased risk of breast cancer diagnosed before the age of 40 years [18]. And gene ARRDC3 has been found to be decreased
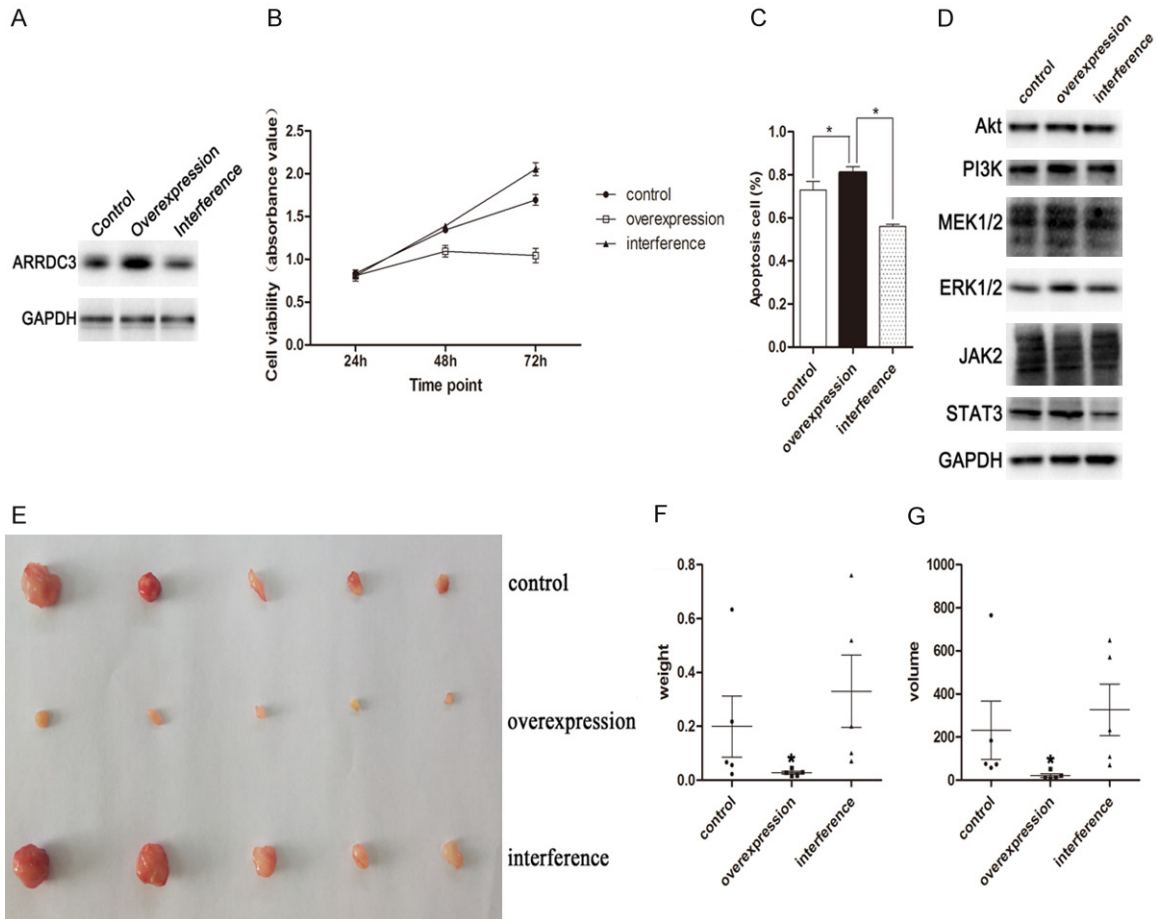
Figure 5. STAT3 pathway was involved in the inhibition of proliferation by gene ARRDC3. A. The protein levels of ARRDC3 in three stable cell lines were measured by western blot. The plot indicated that the three stable cell lines were established successfully. B. The cell viability was assayed in different cell lines at different time points such as 24 h, 48 h, and 72 h. The results analyzed by one way ANOVA method indicated that ARRDC3 possessed a negative effect on cell proliferation. C. Apoptosis analysis in the three stable cell lines represented by the percent of apoptosis cells. Gene ARRDC3 was found to participate in the cell apoptosis after one way ANOVA method analysis. D. The protein level of Akt, PI3K, MEK1/2, ERK1/2, JAK2, and STAT3 were measured by western blot in the three stable cell lines including ARRDC3 overexpression, interference or control group. STAT3 was found to be regulated by gene ARRDC3. E. The tumor tissues from nude mice indicated the inhibition effect on the proliferation by gene ARRDC3. F. The column graph for tumor weight from nude mice. The overexpression of ARRDC3 could reduce tumor tissue weight. G. The column graph for tumor size from nude mice. The statistics analysis was performed with one way ANOVA method and the overexpression of ARRDC3 was found to reduce tumor tissue volume.

in breast cancer and methylation was possibly involved in the inhibition of gene ARRDC3 [6, 9]. Hence, the methylation status in the promoter region of gene ARRDC3 was analyzed in our study (from -1223 bp to 721 bp, GRCh37/hg19, UCSC genes) by MassArray Compact System and our results indicated that totally 32 CpG sites were measured across the promoter region and 6 CpG sites emerged with significant differences between the tumor and nontumor group. Among them, only CpG 6 and CpG 25 sites had a higher methylation level in the tumor group than that in the nontumor group. Then the clustering was performed and the

results revealed that many CpG sites around CpG 25 site possessed higher mean methylation level trends in the tumor group than that in the nontumor group. In addition, the expression of gene ARRDC3 was also investigated and results demonstrated that the expression was inhibited in the mRNA level in the breast cancer tissues relative to that in the adjacent normal tissues. All these data suggested that methylation may play a role in gene ARRDC3 expression regulation.

In order to further confirm the methylation in gene ARRDC3 promoter region directly mediat-

ed in its expression, we constructed three different length promoters (-1377-369, -882-369 and -559-369). The fluorescence reporter assays revealed that the longest promoter had the highest activity in initiating the transcription. Then the reporter plasmid was methylated by M.SssI enzyme for further fluorescence reporter assay and the result indicated that the methylation group was decreased in the fluorescence signal. Based on the above data, we concluded that the methylation in gene ARRDC3 promoter region directly participated in its expression regulation.

A study has showed that ecRNA could regulate the methylation in the promoter [10], therefore, the different lengths (1100 bp and 400 bp) ecRNAs were overexpressed in the cell line MDA-MB-231. And we found that ecRNAi stable cell line almost has no effect in the interference expression of ecRNA. We speculated that possibly the cell line MDA-MB-231 had a pretty low amount of ecRNA. Next, the ecRNA effect on the promoter methylation was measured by BSP sequencing from -186 to 348 and no significant difference of methylated CpG sites or non-CpG sites was found among the four groups. The results meant that the ecRNA did not participate in gene ARRDC3 promoter methylation regulation. However, many methylated CpG sits or non-CpG sites were found in promoter region which possibly were involved in the gene ARRDC3 expression, especially, TGCATGG, TTGCAA, TTCGTA and ATAGTT sites located at 17 bp, 6 bp, -17 bp and -22 bp position relative to the CpG_25 site. These sites all around the CpG-25 site means they were more likely to participate in the gene ARRDC3 expression regulation. Of course, more researches needed to be done for verifying their role in expression regulation for gene ARRDC3. In addition, the phenomenon that many CpG sites in this region were not methylated and some non-CpG sites were methylated suggested that the methylated CpG sites were just a small part of the promoter which possibly cooperated with the non-CpG sites together to be involved in the gene ARRDC3 expression regulation.

SRF, a member of the MADS box family of transcription factors, could bind the CArG box or CArG box-like motif, an essential cis-element present in muscle-specific proteins such as SM a-actin, SM22, SM myosin, b-tropomyosin and caldesmon, and stimulates their transcription [19]. And in cancer, a study revealed that SRF driven transcription modulates the expression of genes involved in cytoskeletal regulation, cell adhesion, migration and invasion [20]. The region from CpG-6 site to CpG-25 site in gene ARRDC3 promoter was predicted for the transcription binding sites and five transcription factor sites were found. Next, five transcription factors were overexpressed in the cell line MDA-MB-231, which was used for confirming the effect on the gene ARRDC3 expression. The results showed that the transcription factor SRF could promote ARRDC3 expression. Then we co-transfected five transcription factors and the promoter reporter plasmid respectively (**Figure 3B**) for further verifying the active transcription factor and direct interaction between them. Our results substantiated that SRF was an effective factor in promoting gene ARRDC3 expression by binding its promoter.

However, whether the methylation in the promoter obstructs the SRF interaction with the promoter remained unknown, the SRF plasmid together with the methylated promoter plasmid or non-methylated plasmid were co-transfected into the cell and results indicated that the methylation in the promoter could reduce the transcription activity by disturbing the interaction between them. Lastly, the ChiP experiment was performed for proving the factor SRF interacting with the gene ARRDC3 promoter and results showed that SRF was significantly recruited into the gene ARRDC3 promoter.

In order to further probe the function of gene ARRDC3, the protein level of gene ARRDC3 in the breast cancer tissue was verified and results showed that the ARRDC3 was repressed in the tumor tissue relative to that in the adjacent normal tissue. Then the expression of gene ARRDC3 was measured in three different breast cancer cell lines and the differential expression was found among them. In order to explore the ARRDC3 on the cell behavior of different breast cancer cell lines, the cell viability and cell scratch experiments were performed and no significant difference was found among them. The results suggested that cell behavior was determined by many factors. Thereupon, cell line MDA-MB-231 was selected as a model for gene ARRDC3 overexpression or interference. Next, the cell ability and apoptosis were measured and an obvious difference was found among them (except the cell viability at 48 h time point). These data suggested that gene ARRDC3 acted as a tumor suppressor in breast cancer.

How does gene ARRDC3 execute its tumor suppressor in breast cancer? This made us investigate the mechanism under the gene ARRDC3. We have known that many signal pathways were widely involved in various cell behaviors including cell proliferation, apoptosis, and invasion. For example, Akt/PI3K pathway usually has frequent aberrations and was involved in anoikis resistance in breast cancer [21, 22]. MEK/ERK pathway was the most commonly aberrantly activated pathway in human cancers and often has crosstalk with the Akt/PI3K pathway [23]. In addition, gene expression profiling revealed that JAK/STAT3 pathway was one of the most differentially modulated pathways in basal-like breast cancer cells [24]. Based on these, we measured the expression status of the three signal pathways in gene ARRDC3 overexpression or interference stable cell lines and an obvious difference in gene STAT3 was found among them. Of course, how does gene ARRDC3 play its tumor repressor role in breast cancer still needs more research to be done. Furthermore, the in vivo experiments on the gene ARRDC3 inhibition effect were performed and results indicated that gene ARRDC3 could significantly inhibit cell growth in nude mice.

Acknowledgements

Disclosure of conflict of interest

None.

Address correspondence to: Sheyu Lin, Life Sciences School, Nantong University, 9# Seyuan Road, Nantong, China. E-mail: linsheyu@yeah.net
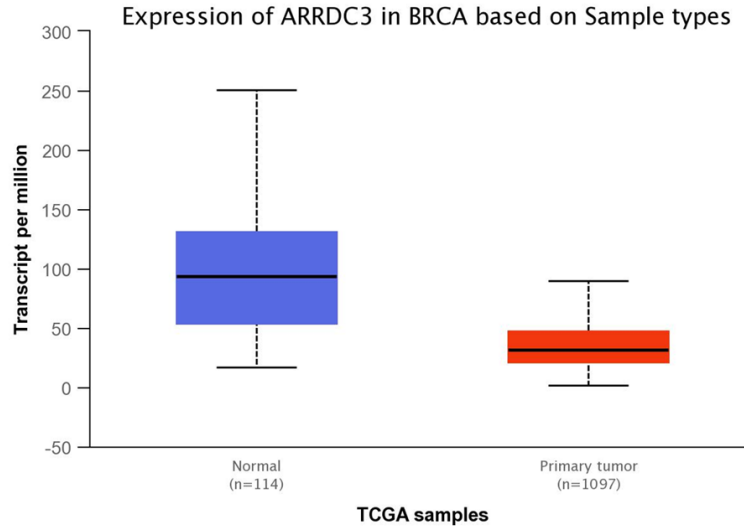
References

[1] Adhikary S, Chakravarti D, Terranova C, Sengupta I, Maitituoheti M, Dasgupta A, Srivastava DK, Ma J, Raman AT, Tarco E, Sahin AA, Bassett R, Yang F, Tapia C, Roy S, Rai K and Das C. Atypical plant homeodomain of UBR7 functions as an H2BK120Ub ligase and breast tumor suppressor. Nat Commun 2019; 10: 1398.

[2] Formisano L, Lu Y, Servetto A, Hanker AB, Jansen VM, Bauer JA, Sudhan DR, Guerrero-Zotano AL, Croessmann S, Guo Y, Ericsson PG, Lee KM, Nixon MJ, Schwarz LJ, Sanders ME, Dugger TC, Cruz MR, Behdad A, Cristofanilli M, Bardia A, O'Shaughnessy J, Nagy RJ, Lanman RB, Solovieff N, He W, Miller M, Su F, Shyr Y, Mayer IA, Balko JM and Arteaga CL. Aberrant FGFR signaling mediates resistance to CDK4/6 inhibitors in ER+ breast cancer. Nat Commun 2019; 10: 1373.

[3] Sina AA, Carrascosa LG, Liang Z, Grewal YS, Wardiana A, Shiddiky MJA, Gardiner RA, Samaratunga H, Gandhi MK, Scott RJ, Korbie D and Trau M. Epigenetically reprogrammed methylation landscape drives the DNA self-assembly and serves as a universal cancer biomarker. Nat Commun 2018; 9: 4915.

[4] Arechederra M, Daian F, Yim A, Bazai SK, Richelme S, Dono R, Saurin AJ, Habermann BH and Maina F. Hypermethylation of gene body CpG islands predicts high dosage of functional oncogenes in liver cancer. Nat Commun 2018; 9: 3164.

[5] Patwari P and Lee RT. An expanded family of arrestins regulate metabolism. Trends Endocrinol Metab 2012; 23: 216-222.

[6] Draheim KM, Chen HB, Tao Q, Moore N, Roche M and Lyle S. ARRDC3 suppresses breast cancer progression by negatively regulating integrin beta4. Oncogene 2010; 29: 5032-5047.

[7] Xiao J, Shi Q, Li W, Mu X, Peng J, Li M, Chen M, Huang H, Wang C, Gao K and Fan J. ARRDC1 and ARRDC3 act as tumor suppressors in renal cell carcinoma by facilitating YAP1 degradation. Am J Cancer Res 2018; 8: 132-143.

[8] Soung YH, Ford S, Yan C and Chung J. The role of arrestin domain-containing 3 in regulating endocytic recycling and extracellular vesicle sorting of integrin beta4 in breast cancer. Cancers (Basel) 2018; 10.

[9] Wang D, Yang PN, Chen J, Zhou XY, Liu QJ, Li HJ and Li CL. Promoter hypermethylation may be an important mechanism of the transcriptional inactivation of ARRDC3, GATA5, and ELP3 in invasive ductal breast carcinoma. Mol Cell Biochem 2014; 396: 67-77.

[10] Di Ruscio A, Ebralidze AK, Benoukraf T, Amabile G, Goff LA, Terragni J, Figueroa ME, De Figueiredo Pontes LL, Alberich-Jorda M, Zhang P, Wu M, D'Alo F, Melnick A, Leone G, Ebralidze KK, Pradhan S, Rinn JL and Tenen DG. DNMT1-interacting RNAs block gene-specific DNA methylation. Nature 2013; 503: 371-376.

[11] Kim HJ, Kim KR, Park HS, Jang KY, Chung MJ, Shong M and Moon WS. The expression and role of serum response factor in papillary carcinoma of the thyroid. Int J Oncol 2009; 35: 49-55.

[12] Lin K, Wang D and Sadee W. Serum response factor activation by muscarinic receptors via RhoA. Novel pathway specific to M1 subtype involving calmodulin, calcineurin, and Pyk2. J Biol Chem 2002; 277: 40789-40798.

[13] Miano JM. Serum response factor: toggling between disparate programs of gene expression. J Mol Cell Cardiol 2003; 35: 577-593.

[14] Farra R, Dapas B, Pozzato G, Giansante C, Heidenreich O, Uxa L, Zennaro C, Guarnieri G and Grassi G. Serum response factor depletion affects the proliferation of the hepatocellular carcinoma cells HepG2 and JHH6. Biochimie 2010; 92: 455-463.

[15] Qiao J, Liu Z, Yang C, Gu L and Deng D. SRF promotes gastric cancer metastasis through stromal fibroblasts in an SDF1-CXCR4-dependent manner. Oncotarget 2016; 7: 46088-46099.

[16] Song S, Min H, Niu M, Wang L, Wu Y, Zhang B, Chen X, Liang Q, Wen Y, Wang Y, Yi L, Wang H and Gao Q. S1PR1 predicts patient survival and promotes chemotherapy drug resistance in gastric cancer cells through STAT3 constitutive activation. EBioMedicine 2018; 37: 168-176.

[17] Liu M, Xiao C, Sun M, Tan M, Hu L and Yu Q. Parthenolide inhibits STAT3 signaling by covalently targeting janus kinases. Molecules 2018; 23.

[18] Joo JE, Dowty JG, Milne RL, Wong EM, Dugue PA, English D, Hopper JL, Goldgar DE, Giles GG and Southey MC. Heritable DNA methylation marks associated with susceptibility to breast cancer. Nat Commun 2018; 9: 867.

[19] Yang Y, Beqaj S, Kemp P, Ariel I and Schuger L. Stretch-induced alternative splicing of serum response factor promotes bronchial myogenesis and is defective in lung hypoplasia. J Clin Invest 2000; 106: 1321-1330.

[20] Muller S, Glass M, Singh AK, Haase J, Bley N, Fuchs T, Lederer M, Dahl A, Huang H, Chen J, Posern G and Huttelmaier S. IGF2BP1 promotes SRF-dependent transcription in cancer in a m6A- and miRNA-dependent manner. Nucleic Acids Res 2019; 47: 375-390.

[21] Basho RK, Gilcrease M, Murthy RK, Helgason T, Karp DD, Meric-Bernstam F, Hess KR, Herbrich SM, Valero V, Albarracin C, Litton JK, Chavez-MacGregor M, Ibrahim NK, Murray JL 3rd, Koenig KB, Hong D, Subbiah V, Kurzrock R, Janku F and Moulder SL. Targeting the PI3K/AKT/mTOR pathway for the treatment of mesenchymal triple-negative breast cancer: evidence from a phase 1 trial of mTOR inhibition in combination with liposomal doxorubicin and bevacizumab. JAMA Oncol 2017; 3: 509-515.

[22] Luey BC and May FE. Insulin-like growth factors are essential to prevent anoikis in oestrogen-responsive breast cancer cells: importance of the type I IGF receptor and PI3-kinase/Akt pathway. Mol Cancer 2016; 15: 8.

[23] Hampsch RA, Shee K, Bates D, Lewis LD, Desire L, Leblond B, Demidenko E, Stefan K, Huang YH and Miller TW. Therapeutic sensitivity to Rac GTPase inhibition requires consequential suppression of mTORC1, AKT, and MEK signaling in breast cancer. Oncotarget 2017; 8: 21806-21817.

[24] Chang R, Song L, Xu Y, Wu Y, Dai C, Wang X, Sun X, Hou Y, Li W, Zhan X and Zhan L. Loss of Wwox drives metastasis in triple-negative breast cancer by JAK2/STAT3 axis. Nat Commun 2018; 9: 3486.
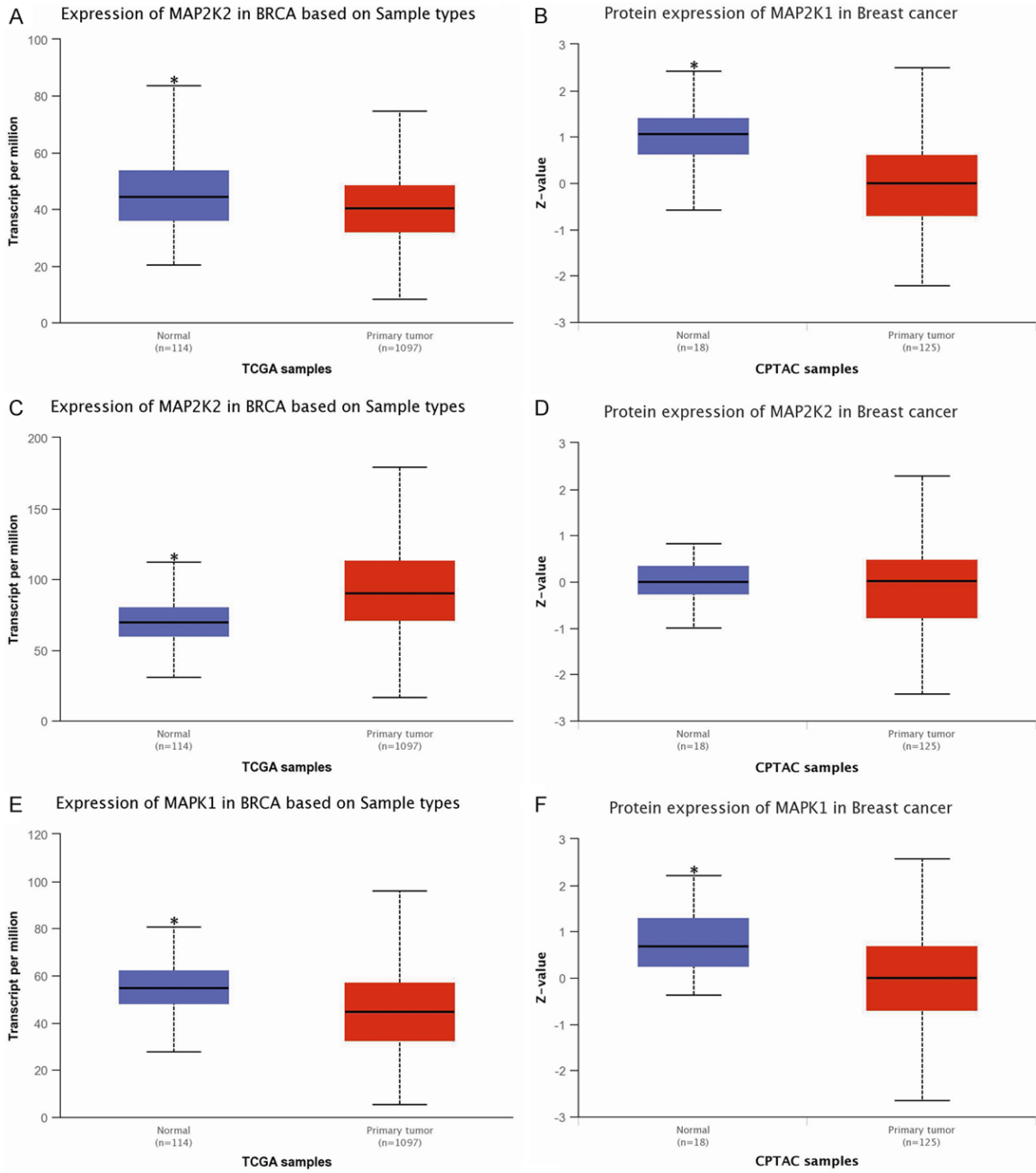
Supplementary Figure 1. Expression of ARRDC3 in breast invasive carcinoma from TCGA database under NIH national cancer institute. The result showed that ARRDC3 was inhibited in primary tumor tissue related to that in normal tissue.
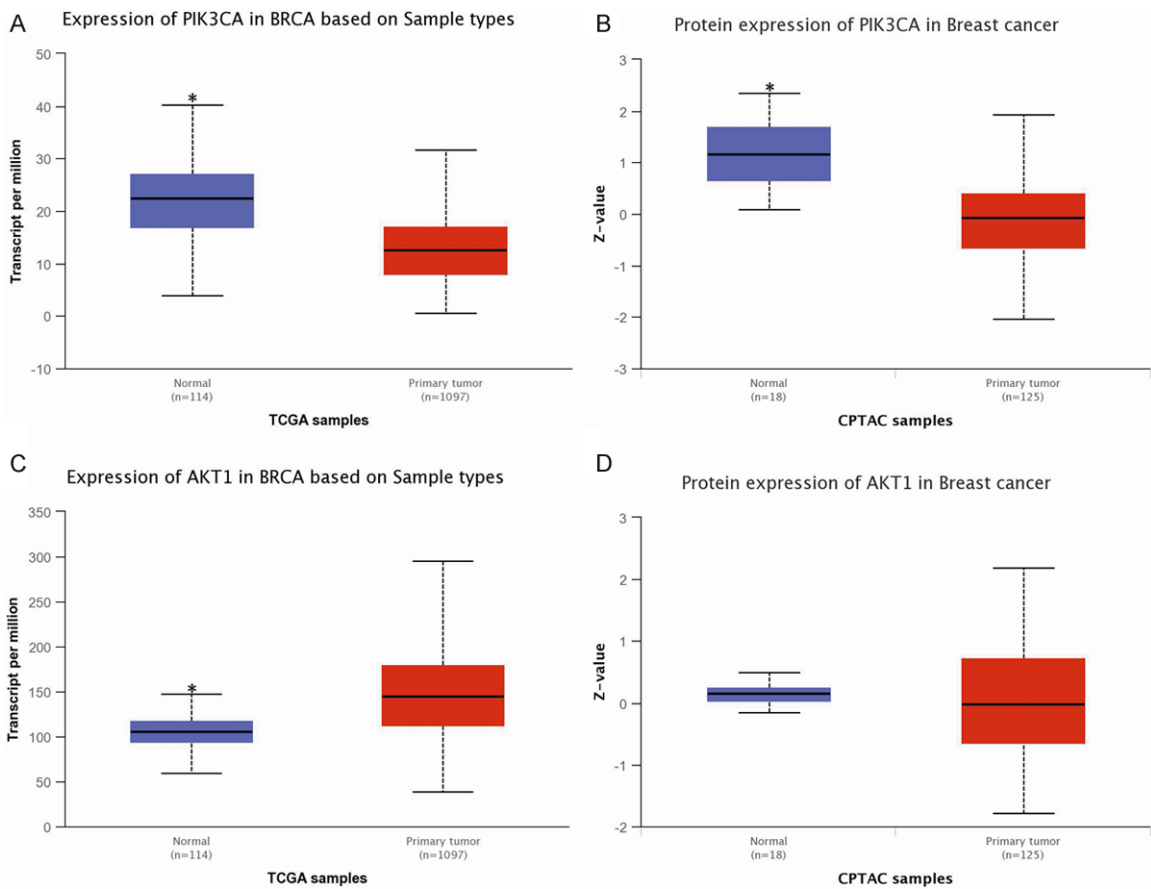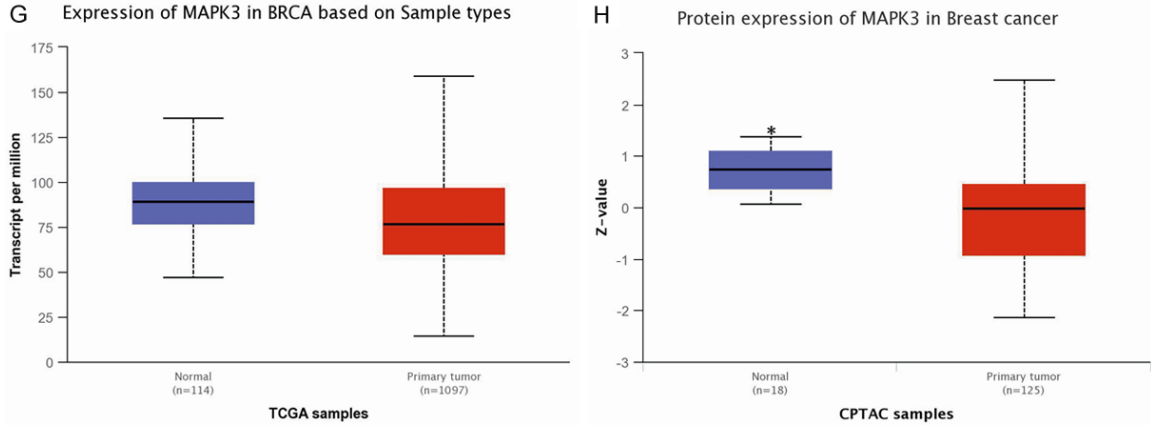

Supplementary Figure 2. The status of the PI3K/Akt signal pathway in breast cancer. A. Expression of PIK3CA in breast invasive carcinoma from the TCGA database. The box plot indicated that PIK3CA mRNA was inhibited in primary tumor tissue than that in normal tissue. B. Protein expression of PIK3CA in breast cancer from the CPTAC database (Clinical Proteomic Tumor Analysis Consortium) under NIH national cancer institute. The plot showed that the protein of PIK3CA was inhibited in the primary tumor tissue than that in normal tissue. C. Expression of AKT1

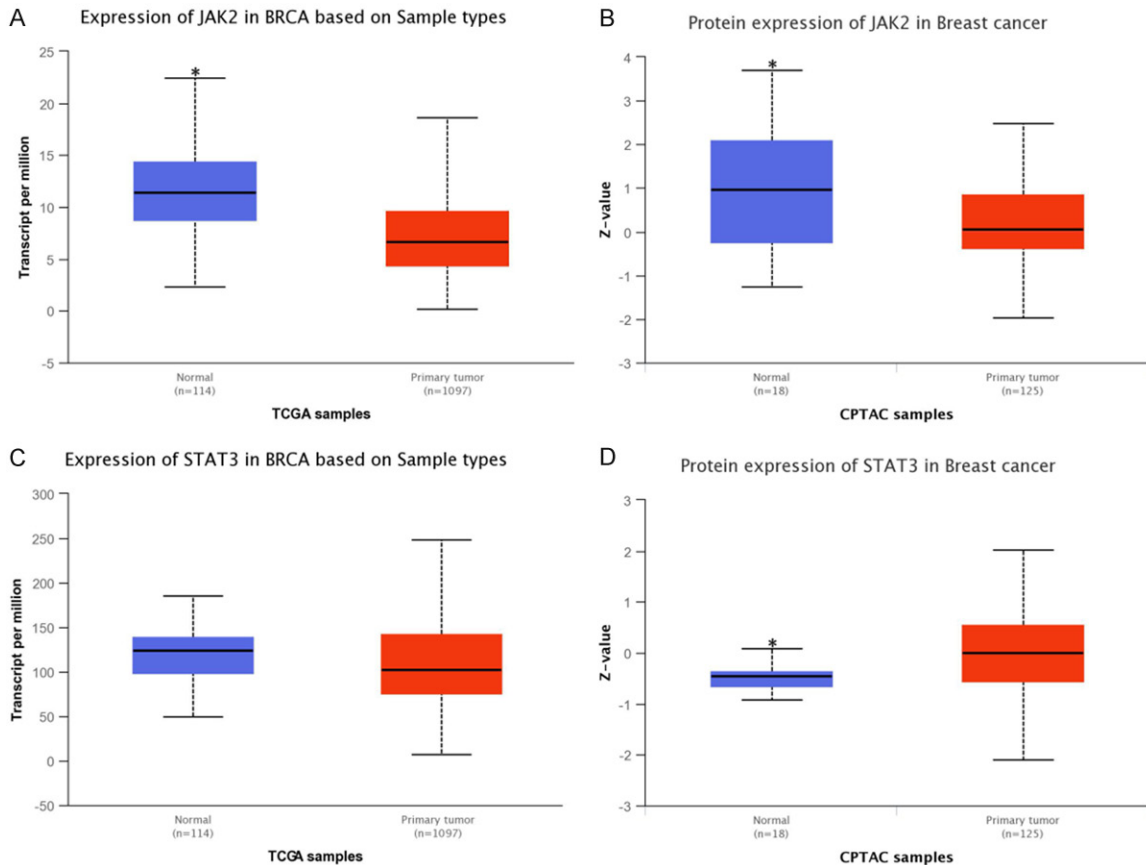# ARRDC3 was regulated by methylation and SRF factor and inhibited breast cancer

in breast invasive carcinoma from the TCGA database. The box plot indicated that AKT1 mRNA was enhanced in primary tumor tissue than that in normal tissue. D. Protein expression of AKT1 in breast cancer from the CPTAC database. No significant difference was found between them.

Supplementary Figure 3. The MEK1/2-ERK1/2 signal pathway was analyzed in breast cancer. A. Expression of MAP2K1 in breast invasive carcinoma from the TCGA database. The result indicated that MAP2K1 mRNA was reduced in primary tumor tissue than that in normal tissue. B. Protein expression of MAP2K1 in breast cancer from the CPTAC database. The plot revealed that the protein of MAP2K1 was decreased in primary tumor tissue than that in normal tissue. C. Expression of MAP2K2 in breast invasive carcinoma from the TCGA database. The box plot demonstrated that MAP2K2 mRNA was increased in primary tumor tissue than that in normal tissue. D. Protein expression of MAP2K2 in breast cancer from the CPTAC database. No significant difference was found between them. E. Expression of MAPK1 in breast invasive carcinoma from the TCGA database. The box plot indicated that MAPK1 mRNA was inhibited in primary tumor tissue than that in normal tissue. F. Protein expression of MAPK1 in breast cancer from the CPTAC database. The plot showed that the protein of MAPK1 was inhibited in the primary tumor tissue than that in normal tissue. G. Expression of MAPK3 in breast invasive carcinoma from the TCGA database. No significant difference was found between them. H. Protein expression of MAPK3 in breast cancer from the CPTAC database. The box plot indicated that MAPK3 was down-regulated in primary tumor tissue than that in normal tissue.

Supplementary Figure 4. The status of the JAK2/STAT3 signal pathway in breast cancer. A. Expression of JAK2 in breast invasive carcinoma from the TCGA database. The box plot indicated that JAK2 mRNA was inhibited in primary tumor tissue than that in normal tissue. B. Protein expression of JAK2 in breast cancer from the CPTAC database. The plot showed that the protein of JAK2 was inhibited in the primary tumor tissue than that in normal tissue. C. Expression of STAT3 in breast invasive carcinoma from the TCGA database. No significant difference was found between them. And the P = 0.105, the result revealed the same variation trend as our results. D. Protein expression of STAT3 in breast cancer from the CPTAC database. The box plot indicated that STAT3 was enhanced in primary tumor tissue than that in normal tissue. The trend was opposite to the above plot. Therefore, based on the two plots, we could speculate that STAT3 was involved in breast cancer development.