ADAMTS5 modulates breast cancer development as a diagnostic biomarker and potential tumor suppressor, regulating by BAIAP2-AS1, VTI1B, CRNDE, and hsa-miR-135b-3p: integrated systems biology and experimental approach

Najmeh Tavousi
Isfahan University of Medical Sciences
Qazal Taqizadeh
Zist Fanavari Novin Biotechnology Institute, Isfahan, Iran
Elnaz Nasiriyan
Zist Fanavari Novin Biotechnology Institute, Isfahan, Iran
Parastoo Tabaeian
Zist Fanavari Novin Biotechnology Institute, Isfahan, Iran
Mohammad Rezaei
University of Isfahan
Mansoureh Azadeh (mazadeh@phd.iaurasht.ac.ir)
Islamic Azad University

Article

Keywords: Breast cancer, Microarray, CRNDE, ADAMTS5, Real-time PCR, Bioinformatics

Posted Date: July 27th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1861409/v1

Abstract

Background: A disintegrin and metalloproteinases (ADAMs) are a novel gene family of proteins that share the metalloproteinase domain with matrix metalloproteinases and have sequence similarities to the reprolysin family of snake venomases (MMPs). They are divided into two classes based on their structural similarities: membrane-anchored ADAM and ADAM with thrombospondin motifs (ADAMTS). These molecules have a role in a number of biological processes, including membrane protein shedding, proteolysis, cell adhesion, cell fusion, and cell migration.

Method: In this study, the RNA interaction analysis (miRWalk, lncRRisearch, ENCORI, and lncBase3) and protein interaction analysis (STRING) was performed. Survival, expression, and correlation analyses was performed by ENCORI and GEPIA2. Expression analysis also performed by microarray analysis. GSEA analysis was performed to find top up and down-regulated genes and relevant signaling pathway. Gene ontology analysis was performed by g: Profiler and clusterprofiler. Pathway enrichment was performed by GSEA analysis.

Results

ADAMTS5 has a significant low-expression, based on ENCORI (Fold Change: 0.12, FDR < 0.0001), microarray analysis (logFC: -3.791881, adj. P. Val < 0.0001), and qRT-PCR experiment (logFC: -1.975, FC: 0.025, p. value < 0.0001, patients under 40). ADAMTS5 is a potential diagnostic biomarker of BC, based on ROC curve (AUC: 0.75, *p*. value: 0.0007). mRNA VTI1B, lncRNAs CRNDE and BAIAP2-AS1, and hsa-miR-135a-3b have interaction with ADAMTS5. VTI1B has a significant low-expression (FC: 0.65, FDR: 0.0001). BAIAP2-AS1 is significantly up-regulated (FC: 1.37, FDR: 0.0001). CRNDE is significantly up-regulated (FC: 1.30, FDR: 0.0001). BAIAP2 (log-Rank p: 0.0017, Hazard Ratio: 0.59) and CRNDE (log-Rank p: 0.34, Hazard Ratio: 1.17) are the two potential BC survival predictor.

Conclusion

In this study, we performed a differential expression analysis on the expression level of ADAMTS5, VTI1B, and BAIAP2-AS1 in the BC patients for the first time. ADAMTS5 and its correlated oncogene and tumor suppressor RNAs and proteins are the hub therapeutic target of BC. DEGs of the GSE42568 dataset are involve in PPAR and cell cycle checkpoint signaling pathways and during the regulation of mentioned pathway, regulate the BC development. Targeting the expression and interaction of mentioned RNAs can be helpful for the diagnosis, prognosis, and treatment of BC.

Background

The majority of cancer research has focused on determining how tumor cells differ from normal cells in gene expression over the last 50 years (1). Dysregulation of gene expression level could provide some crucial human abnormalities and diseases, including hepatitis (2), Alzheimer (3), diabetes (4), MS (5), and the various human cancer types, including live cancer (6), colorectal cancer (7), retinoblastoma (8), lung

cancer (9), head and neck cancer (10), and breast cancer (11–14). An in-depth look at gene expression patterns in a variety of pathogenic illnesses, such as breast cancer, can help speed up treatment and predict disease emergence.

Breast cancer is the most common cancer among women, accounting for 1.8 million new cases in 2013, or 12 percent of all cancers. Despite the fact that the incidence of breast cancer is lower among Iranian women, epidemiological studies show that the number of newly diagnosed breast cancer patients has recently increased (15). BC is the most common female malignancy and the leading cause of cancer mortality in women, according to Cancer Statistics 2018, with over 2.1 million females diagnosed each year and over 62,000 deaths. More than 60% of BC deaths occur in developing countries (16). BC also is one of the main reasons of cancer-related death in woman worldwide (17–19). The development of mutations in the genome that result in the gain of function or activation of oncogenes like Human Epidermal Growth Factor Receptor 2 (HER2), mTOR, PI3K, EpCAM, PTBP1, HER4, and WBP2 or the loss of function or deactivation of tumor suppressor genes like p53, BRCA1/2, and p21 is what causes cancer (20). Despite the fact that early detection and recent advances in anti-cancer therapy have improved critical outcomes for BC patients, the disease's recurrence rate remains high (21–23). All in all, measuring the expression of genes linked to breast cancer, identifying diagnostic and prognostic biomarkers, and understanding gene expression patterns in various clinical and pathological situations linked to breast cancer could provide valuable information about the disease and aid in its prevention.

Several techniques for monitoring gene expression have been developed by scientists. Real-time PCR and microarray are two powerful technologies for assessing gene expression. Gene expression profiling and genome-wide gene expression analysis using DNA microarray could provide information on the level of expression and relative expression of genes and RNAs in different groups, such as "tumor" and "control," or "treated" and “untreated” (24).

Among various genes involved in BC, A Disintegrin-like And Metalloproteinase with ThromboSpondin (*ADAMTS5)* was selected as the potential biomarker gene. *ADAMTS5* is a secreted metalloproteinase with numerous proteoglycan substrates that belongs to the A Disintegrin-like And Metalloproteinase with ThromboSpondin Motifs (ADAMTS) family. However, it is widely known for its involvement in cartilage deterioration and arthritis, the prominent role of *ADAMTS5* in cancer (25). Also, although transforming growth factor-β (TGF-β) and proinflammatory cytokines have been shown to promote ADAMTS4 expression in chondrocytes, nothing is known regarding the control of *ADAMTS5* expression (26).

Previous studies revealed that *ADAMTS5* has some potentially essential roles in different cancer types. For example, an investigation in 2019 revealed that *ADAMTS5* could act as a tumor suppressor and have an essential rule in migration, invasion, and angiogenesis in gastric cancer (27). Also, previous studies showed that the expression level of this gene reduced in human Breast cancer samples of Indian women, compared to control (28). In return, high expression of *ADAMTS5* was reported in several investigations about hepatocellular carcinoma (29,30), and colorectal cancer (31). Therefore, investigations about the

role of these genes in different populations could provide important information about the molecular causes of related diseases, including cancer.

LncRNAs are long RNA molecules found in the nucleus or cytoplasm and have a length of more than 200 nucleotides. Only a handful can code for a small number of polypeptides, whereas the others do not. LncRNAs play various biological processes during the pre-and post-transcriptional stages, including tumor invasion, metastasis, and apoptosis (32–34). One of the critical objectives of this study is to investigate the expression pattern of this gene in breast cancer patients. Also, the relationship between the expression of this gene and different clinical and pathological conditions is another vital issue in this study. Another goal of this study is to investigate the expression of lncRNA expression associated with this gene. The expression pattern of the relevant lncRNA in Breast cancer patients in Isfahan and the study of the biomarker potential are also among the research items.

Methods

3.1. Microarray analysis

The datasets related to breast cancer underwent microarray analysis. In order to identify the differentially expressed genes (DEGs) in the breast cancer microarray datasets, GSE42568 (35) was examined. In this collection, 104 BC samples and 17 control samples were examined. This dataset is provided by GPL570 ([HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array). The affy (36) package was used to normalize the raw data after being sent to the R Studio environment (*Read.Affy* command in R) from the GEO online database (https://www.ncbi.nlm.nih.gov/geo/). The limma (37) package carried out the statistical analysis of the microarray dataset. The Bioconductor web database was used to get the Affy and limma packages. (https://www.bioconductor.org/). The significance level for the analysis of microarray data was set at 0.01 (adjusted p value). The ggplot2 and pheatmap packages, which can be acquired from CRAN, were used to construct the microarray data analysis visualizations (https://cran.r-project.org). The expression of 54675 mRNA and lncRNA transcripts was examined in this microarray research. The difference in the expression level of all RNAs was computed following normalizing (quantiles normalization method), logarithmic scaling, and deletion of the transcripts with no expression in the dataset. The logFC − 3 was chosen as the threshold of low expression, and the RNAs with logFC > 3 were chosen as the up-regulated RNAs.

3.2. RNA interaction, expression, survival, and gene ontology analyses

The mRNA – RNA interaction analysis was performed by ENCORI (https://starbase.sysu.edu.cn/) (https://doi.org/10.1093/nar/gkt1248) (38). Expression, correlation and survival analyses was performed by GEPIA2 (39) (http://gepia2.cancer-pku.cn/) and ENCORI online software. RNA interactions were visualized by Cytoscape (3.8.2) software (40). Protein-protein interaction analysis was performed by the STRING online database (41). Gene ontology analysis was performed by gprofiler2 (42) (from CRAN) and

clusterprofiler (43) packages (from Bioconductor). microRNA (miRNA) – mRNA interaction analysis was performed by miRWalk (http://mirwalk.umm.uni-heidelberg.de/) (44–46). miRNA-lncRNA interaction analysis was performed by lncBase 3 (https://diana.e-ce.uth.gr/lncbasev3) (47).

3.3. Gene set enrichment analysis (GSEA)

In the GSE42568 dataset, the samples were divided into the control and tumor samples. To investigate biological function, gene set enrichment analysis (GSEA) (https://software.broadinstitute.org/gsea/) was carried out (48,49). Significant terms were determined to have $P < 0.05$.

3.4. Clinical characteristics of tissue samples

All procedures for the research in this study that included human samples were authorized by the Al-Zahra Hospital Ethics Committee, Isfahan University of Medical Science, and all patients signed written consent forms. In a case-control study, breast cancer and surrounding normal breast tissue samples from 50 breast cancer patients were examined. None of the patients had ever had chemotherapy or radiation treatment. Before being immersed in RNA Later solution (Invitrogen, USA) and immediately frozen in liquid nitrogen for pathologist review, tissue samples were washed in distilled water. Table 1 lists the clinicopathological features of patients with gastric and breast cancer.

Table 1
The clinicopathological characteristics of BC patients.

Characteristic	Status	Number of patients
Stage	I	8 (16%)
	II	15 (30%)
	III	12 (24%)
	IV	9 (18%)
	Unknown	6 (12%)
Age	< 45	20 (40%)
	> 45	24 (48%)
	Unknown	6 (12%)
Lymph node metastasis	Yes	17 (34%)
	No	26 (52%)
	Unknown	7 (14%)
Tumor size (TS)	< 2 cm	15 (30%)
	5 > TS > 2	5 (10%)
	> 5 cm	29 (58%)
	Unknown	1 (2%)
Menopausal status	Yes	12 (24%)
	No	15 (30%)
	Unknown	23 (46%)
ER receptor	Positive	14 (28%)
	Negative	13 (26%)
	Unknown	23 (46%)
PR receptor	Positive	19 (38%)
	Negative	21 (42%)
	Unknown	10 (20%)
HER2/neu receptor	Positive	12 (24%)
	Negative	21 (42%)
	Unknown	17 (34%)

3.5. RNA extraction, cDNA synthesis, and qRT-PCR analyzers

For RNA extraction of both tissues, tumorous and normal, TRIZOL was used (Invitrogen, Carlsbad, CA, USA). After the RNA extraction stages, the synthesis of cDNA (TaKaRa, Tokyo, Japan) was carried out, using the TaKaRa cDNA synthesis kit described by the producer protocol. To perform Real-Time PCR Sybergreen, (Amplicon Company, and Denmark) was used, and RT- qPCR carried out using MIC Real-Time PCR device (Australia). To perform PCR reactions following conducting were set as Initial denaturation, 95º C for 15 minutes, the secondary 95 ºC for 15 seconds, a 60 ºC for 20 seconds, and a 72 ºC for 20 seconds, the number of cycles was 40. The primers were synthesized by TAQ Copenhagen Company (Denmark), and the sequences are shown in Table 1. The associated expression was normalized with the amount of GAPDH as an internal control. Primer sequence is showing in Table 2.

Table 2
Primer sequences of ADAMTS5 and GAPDH as the housekeeping gene.

Gene	Forward/Reverse	Primer sequence
ADAMTS5	F	5'TGCACTTCAGCCACCATCAC3'
ADAMTS5	R	5'CGTACCACAGCACACCACAG3'
GAPDH	F	5'ACAGGGTGGTGGACCTCAT3'
GAPDH	R	5'AGGGGTCTACATGGCAACTG3'

3.6. Statistical analysis

Real-time PCR data and associated graphs were statistically analyzed using the GraphPad Prism program (version 8). The − ΔΔCT approach was used to compare the expression levels between the tumor and control samples in the qRT-PCR results. To determine if the data were normal, the Shapiro-Wilk test was run on the expression data. The expression levels in tumor and control samples were compared using a paired t-test on the − ΔΔCT data. R Studio was used to do DEG analysis on the microarray data (4.1.2). The GraphPad prism performed the recipient operating characteristic (ROC) analysis for the real-time PCR datasets based on sensitivity and specificity. The significance level for this investigation was set at a p-value of less than 0.05. In the ROC analysis, AUC values between 0.7 and 0.8 are considered fair, 0.8 and 0.9 are considered good (signifying a good biomarker), while 0.9 and 1 showed an exceptional biomarker.

Results

4.1. Quality control of microarray samples

The GSE42568 microarray dataset underwent microarray analysis. To assess the quality of the samples and the microarray data, principal component and correlation analyses were performed (Fig. 1, 2). The

unprocessed microarray expression data is shown in Fig. 1A. The normalized expression data was displayed in Fig. 1B following quantile normalization. The microarray control and tumor samples are shown in Figure A's x axis, and the expression of all 54675 represented genes is shown in Figure A's y axis. The samples that are noisy in the aforementioned boxplots are those whose medians deviate from those of the other samples. According to the expression boxplots in Fig. 1, there were no examples of low quality in this dataset. Based on the first and second main components of the control and tumor data, a PCA analysis was carried out to identify the low-quality samples in Fig. 2. Also, there was no bad-quality samples based on the PCA analysis.

4.2. ADAMTS5 has a significant low-expression in the microarray samples, RNA-seq (ENCORI and GEPIA2), and qRT-PCR experiment

Microarray data analysis of GSE42568 revealed that ADAMTS5 has a significant down-regulation in the BC samples, compared to control (logFC: -3.791881, adj. P. Val < 0.0001, Fig. 3). In the microarray analysis of GSE42568, 393 genes are the down-regulated genes and 142 genes are the up-regulated genes. The heatmap of top 50 DEGs are provided in the Fig. 4. Top 50 up and down regulated genes of microarray analysis are provided in the Tables 2 & 3. Based on the differential expression analysis of 1104 tumor and 113 normal RNAseq samples of BC, ADAMTS5 has a significant low-expression in the BC samples (Fold Change: 0.12, FDR < 0.0001, Fig. 5A). Also, based on the GEPIA2 data analysis, ADAMTS5 has a significant low-expression in the BC samples (Fig. 5B). Real-time PCR experiment was performed on *ADAMTS5* to find the changes in the expression level of this gene in the breast cancer samples. Data analysis revealed that *ADAMTS5* (logFC: -1.975, FC: 0.025, p. value < 0.0001) has a significant low-expression in the human breast cancer samples (Fig. 6).

Table 3
The top 50 up-regulated genes in the microarray analysis

ID	Gene.symbol	logFC	AveExpr	adj.P.Val
1552797_s_at	PROM2	4.034416	9.615846	7.90E-16
1553986_at	RASEF	3.68242	7.24066	1.84E-12
1559949_at		3.60705	7.174521	1.34E-13
200606_at	DSP	4.970517	9.80584	3.63E-25
201131_s_at	CDH1	4.93688	10.26391	2.23E-18
201286_at	SDC1	4.627368	7.160284	1.91E-20
201287_s_at	SDC1	4.258994	7.332663	1.71E-19
201291_s_at	TOP2A	4.437646	7.620142	1.69E-15
201596_x_at	KRT18	4.963591	10.30379	2.22E-18
201650_at	KRT19	5.519178	11.01424	3.29E-15
201689_s_at	TPD52	3.620107	8.270104	1.80E-14
201690_s_at	TPD52	3.735863	10.34887	3.03E-19
201839_s_at	EPCAM	5.175473	11.24738	3.84E-21
201890_at	RRM2	4.33196	8.367756	1.32E-16
202286_s_at	TACSTD2	4.570748	10.90014	4.21E-14
202376_at	SERPINA3	3.278976	9.958011	1.00E-05
202421_at	IGSF3	3.289024	8.027429	3.92E-14
202454_s_at	ERBB3	3.646062	7.027292	6.72E-15
202489_s_at	FXYD3	3.164143	9.64655	2.78E-11
202503_s_at	KIAA0101	3.737831	8.097359	1.22E-15
202525_at	PRSS8	3.790543	6.957124	2.15E-19
202575_at	CRABP2	3.024231	9.224404	1.21E-08
202597_at	IRF6	3.284609	7.090864	2.09E-15
202870_s_at	CDC20	3.684426	6.747565	5.88E-15
203108_at	GPRC5A	3.516755	7.459268	1.62E-08
203228_at	PAFAH1B3	3.685702	8.123168	3.49E-21
203358_s_at	EZH2	3.657506	6.690811	7.14E-19

ID	Gene.symbol	logFC	AveExpr	adj.P.Val
203453_at	SCNN1A	4.104894	7.50119	2.70E-10
203878_s_at	MMP11	3.30896	6.469012	1.17E-10
203917_at	CXADR	3.171424	8.239799	1.04E-09
203954_x_at	CLDN3	3.161605	8.294682	5.04E-11
203961_at	NEBL	3.081555	8.001078	3.89E-10
204320_at	COL11A1	4.872168	7.58127	5.74E-12
204351_at	S100P	4.267184	7.495964	3.31E-08
204475_at	MMP1	3.120998	5.850599	0.000102
204653_at	TFAP2A	4.636436	9.505486	9.41E-19
204667_at	FOXA1	3.270144	5.845329	1.05E-08
204798_at	MYB	3.623567	6.904626	1.18E-09
205713_s_at	COMP	3.390541	6.369407	1.27E-09
205941_s_at	COL10A1	3.519242	6.483829	2.36E-11
205980_s_at	ARHGAP8	3.732744	6.911901	8.21E-18
206102_at	GINS1	3.046119	6.568088	1.29E-14
207847_s_at	MUC1	3.864021	7.741723	1.86E-09
208079_s_at	AURKA	3.53248	7.384416	3.06E-15
208451_s_at	C4A	4.328142	8.979695	6.85E-10
208650_s_at	CD24	4.458536	10.17113	1.23E-12
208651_x_at	CD24	3.549324	8.195981	1.46E-12
209008_x_at	KRT8	3.909263	10.17412	1.07E-18
209016_s_at	KRT7	3.688189	8.09024	2.57E-07
209173_at	AGR2	4.299528	9.036295	1.60E-06

Table 4
The top 50 down-regulated genes in the microarray analysis

ID	Gene.symbol	logFC	AveExpr	adj.P.Val
1552509_a_at	CD300LG	-4.420	3.46875	1.71E-23
1552519_at	ACVR1C	-5.658	3.422548	2.09E-33
1552616_a_at	ACACB	-3.193	2.87015	1.09E-27
1553102_a_at	CCDC69	-3.522	3.438745	1.24E-21
1553243_at	ITIH5	-3.793	3.481887	1.39E-31
1553583_a_at	THRSP	-3.170	4.653558	7.83E-26
1554044_a_at	MRAP	-4.797	3.288932	5.04E-29
1554062_at	XG	-3.148	3.904648	3.41E-18
1554485_s_at	TMEM37	-4.517	4.193475	6.63E-22
1555854_at	LOC101930400	-4.399	3.181976	5.57E-26
1556427_s_at	LRRN4CL	-4.052	3.852743	3.69E-21
1557832_at	LOC101926960	-4.836	3.240912	1.16E-26
1558420_at	C14orf180	-4.314	3.175471	2.08E-31
1558421_a_at	C14orf180	-5.140	3.488894	7.23E-28
1559942_at	MDFIC	-3.222	3.397004	1.56E-23
1563182_at	ACVR1C	-4.711	4.476274	1.80E-32
200831_s_at	SCD	-3.019	8.493693	4.89E-14
201012_at	ANXA1	-3.178	9.089452	2.93E-10
201348_at	GPX3	-4.835	7.049173	2.48E-13
201425_at	ALDH2	-3.067	9.132746	9.36E-13
201525_at	APOD	-3.064	7.927405	0.000278
201539_s_at	FHL1	-5.453	3.388847	2.66E-31
201540_at	FHL1	-4.577	9.634601	1.09E-14
201843_s_at	EFEMP1	-3.079	6.173931	2.38E-07
201963_at	ACSL1	-3.496	8.548934	4.06E-15
202022_at	ALDOC	-3.787	4.732487	5.44E-17
202345_s_at	FABP5	-3.024	8.453878	3.23E-11

ID	Gene.symbol	logFC	AveExpr	adj.P.Val
202350_s_at	MATN2	-3.146	6.349562	2.17E-07
202768_at	FOSB	-3.778	6.027379	1.94E-09
202884_s_at	PPP2R1B	-3.587	3.272612	2.28E-30
202886_s_at	PPP2R1B	-3.353	4.103628	1.18E-23
202920_at	ANK2	-3.323	3.877897	1.49E-22
202973_x_at	FAM13A	-3.065	8.152986	7.12E-14
202976_s_at	RHOBTB3	-3.034	6.010552	5.40E-09
202990_at	PYGL	-3.294	7.326218	4.51E-15
203065_s_at	CAV1	-4.433	6.667073	6.43E-17
203296_s_at	ATP1A2	-4.441	3.766834	1.69E-21
203305_at	F13A1	-3.349	5.124114	1.56E-14
203323_at	CAV2	-4.491	6.094718	1.85E-15
203324_s_at	CAV2	-4.081	7.228227	8.82E-15
203337_x_at	ITGB1BP1	-4.099	5.107344	5.94E-21
203400_s_at	TF	-4.658	4.571469	7.03E-14
203414_at	MMD	-4.089	7.958242	1.51E-24
203434_s_at	MME	-4.392	4.151296	3.69E-24
203435_s_at	MME	-4.210	4.986593	2.37E-28
203548_s_at	LPL	-6.402	6.112699	4.72E-17
203549_s_at	LPL	-5.905	7.263051	4.20E-15
203571_s_at	ADIRF	-3.565	7.894999	5.18E-11
203680_at	PRKAR2B	-3.808	7.651282	2.36E-13
203851_at	IGFBP6	-3.632	5.482259	2.84E-18

4.3. ADAMTS5 is a potential prognostic and diagnostic biomarker of BC

To find the distinguishing pattern of RNAs expression level in tumor and control samples and find the prognostic and diagnostic biomarker property of these RNAs, ROC and clinicopathological analyses were

performed on the RNA expression data. According to ROC analysis, *ADAMTS5* (AUC: 0.75, *p*. value: 0.0007) could be the potential significant diagnostic biomarker of Breast cancer (Fig. 7).

Clinicopathological analysis was performed on the RNA expression data in different Stage, Age, Lymph node metastasis, Menopausal, Tumor size, ER receptor, PR receptor, and HER2 receptor Statuses. According to these analyses, *ADAMTS5* expression level had a significant low-expression in the patients younger than 45 years (Fig. 8). So *ADAMTS5* could be a potential diagnostic and prognostic biomarker for Isfahan Breast cancer samples.

4.4. Protein and RNA interaction analysis and relevant expression and co-expression analysis

Based on STRING online database, ADAMTS5 has interaction with Thrombospondin 1 (THBS1) and Thrombospondin 2 (THBS2), Fibronectin 1 (FN1), CD36, Aggrecan core protein (ACAN), ADAM Metallopeptidase with Thrombospondin Type 1 Motif 2–5 (ADAMTS2-5), Spondin 1 (SPON1), Beta 3-Glucosyltransferase (B3GLCT), Protein O-Fucosyltransferase 2 (POFUT2), ADAMTS Like 1 (ADAMTSL1), Syndecan 4 (SDC4), and Matrix Metallopeptidase 3 and 13 (MMP3 and MMP13, Fig. 9).

mRNA-mRNA interaction analysis (ENCORI) revealed that ADAMTS5 has a significant RNA interaction with VTI1B (Free Energy: -39.3, Align Score: 25.5). miRNA-mRNA interaction analysis demonstrated the miRNA interaction of ADAMTS5 (Score: 1, region: 3'UTR). Based on the energy of binding, ADAMTS5 is a significant target of hsa-miR-135b-3p. lncBase miRNA-lncRNA revealed that hsa-miR-135-3p affects the expression level of OIP5-AS1, CRNDE, SNHG14, PRKAG2-AS1, POC1B-AS1, and PKMP5. Furthermore, lncRRisearch interaction analysis revealed that ADAMTS5 has RNA interaction with lncRNA CRNDE and lncRNA BAIAP2-AS1 (Fig. 10).

4.5. Expression analysis of VTI2B, CRNDE, and lncRNA BAIAP2-AS1

Relative expression analysis (GEPIA2 and ENCORI) revealed that VTI1B has a significant low-expression in the BC samples (FC: 0.65, FDR < 0.0001, Fig. 11A). BAIAP2-AS1 has a significant up-regulation in the BC samples (FC: 1.37, FDR < 0.0001, Fig. 11B). CRNDE has a significant up-regulation in the BC samples (FC: 1.30, FDR < 0.0001, Fig. 11C).

4.6. BAIAP2-AS1 and lncRNA CRNDE have significant negative co-expression with ADAMTS5

Pearson correlation analysis of RNAseq data was performed by ENCORI and GEPIA2 online databases. Based on mentioned analyses, ADAMTS5 has a significant negative correlation with BAIAP2-AS1 (r: -0.202, *p*-value < 0.0001, Fig. 12B) and lncRNA CRNDE (r: -0.233, *p*-value < 0.0001, Fig. 12C). However, based on ENCORI, VTI1B has a not-significant positive correlation with ADAMTS5 (r: 0.029, *p*-value: 0.34), and based on GEPIA2 correlation analysis VTI1B has a not-significant negative correlation with ADAMTS5 (*p*-value: 0.97, R: -0.0013, Fig. 12A).

4.7. Survival analysis of ADAMTS5, VTI1B, BAIAP2-AS1, and CRNDE in the breast cancer patients with high and low expression levels

ENCORI Survival analysis revealed that higher expression level of BAIAP2-AS1 has a significant correlation with higher survival rate of the BC patients (log-Rank p: 0.0017, Hazard Ratio: 0.59, Fig. 13B). The expression level of lncRNA CRNDE have a not-significant negative correlation with the survival rate of the BC patients (log-Rank p: 0.34, Hazard Ratio: 1.17, Fig. 13D). There was not a significant correlation between expression level of ADAMTS5 and VTI1B with the survival rate of the patients (Fig. 13). GEPIA2 survival analysis revealed that the expression level of BAIAP2-AS1 have a not-significant positive correlation with the survival rate of BC patients (Fig. 14)

4.8. GSEA analysis

GSEA analysis revealed that the up-regulated genes on GSE42568 have a significant role in the cell cycle checkpoints signaling pathway (Fig. 15). Also, the own-regulated genes in this dataset of BC samples have a significant role in the PPAR signaling pathway (Fig. 16).

4.9. Gene ontology analysis

Based on gene ontology analysis, the up-regulated genes of GSE42568 (Table 3) are involved in the anatomical structure development, tissue development, cell differentiation, and cell motility biological processes. These high expressed RNAs (protein coding or non-coding) of BC samples are located in the extracellular space and organelles (Fig. 17). Based on GO analysis, the down-regulated genes of GSE42568 (involves ADAMTS5, Table 4) are involve in the following biological processes: response to stimulus, response to chemical, lipid metabolic process, small molecule metabolic process, and response to peptide. Mentioned genes are regulating signaling receptor binding, integrin binding, and oxidoreductase activity, in the extracellular region and cell periphery (Fig. 18, 19).

Discussion

At the first, we perform an integrated high-throughput bioinformatics analysis to demonstrate significantly biomarkers involved in breast cancer. We carried out a microarray analysis to find the differentially expressed genes that could affect breast cancer development. After finding the DEGs, we choose *ADAMTS5* for further analyses and investigations. Also, we performed following bioinformatics analyses to find a systematic view of the regulatory mechanism of up and down regulated genes in this dataset on the BC development, specially ADANTS5: mRNA – mRNA interaction, mRNA – lncRNA interaction, mRNA – microRNA interaction, miRNA – lncRNA interaction, GSEA, and GO. Based on mentioned analyses, Also, we find that the VTI1B mRNA have a significant RNA interaction with the ADAMTS5, as a down-regulated tumor suppressor gene in the BC. Also, we find that lncRNA BAIAP2-AS1 has a significant interaction with ADAMTS5, as an up-regulated tumor suppressor lncRNA. ADAMTS5 also is a potential target of hsa-miR-181a-5p, as a competitive endogenous RNA (ceRNA) of lncRNA

CRNDE. We find a multi-component ceRNA network that regulated by has-miR-135b-3p. From this network, we selected the *CRNDE* - hsa-miR-135b-3p – *ADAMTS5* axis and predicted that lncRNA *CRNDE* and *ADAMTS5* could indirectly affect expression of each other by changes in expression level and binding affinity of hsa-miR-135b-3p in breast cancer patients.

Based on this hypothesis, we evaluated the expression level of *ADAMTS5* in human BC samples in order to gain some understanding of the function of this network and the changes in gene expression in this axis. Our real-time PCR experiment revealed that the expression level of *ADAMTS5* is decreasing in breast cancer patients. Also, these two RNA could be the diagnostic and prognostic biomarkers, based on ROC and clinicopathological analyses. Also, we find that the lower expression level of lncRNA BAIAP2-AS1 has a significant correlation with the higher death rate of the patients.

Generally, the possible role of Has-miR-135b-3p in some different human diseases have been investigated. Based on previous studies, this micro-RNA could have some effect on Breast cancer (50), Epithelial Ovarian cancer (51), and Adenocarcinoma (52). However, the main characteristics of this miRNA in different cancer types – including Breast cancer - is still unclear. Based on miRWalk and lncBase analysis, we demonstrated that this miRNA could regulate the *ADAMTS5* and lncRAN *CRNDE*. *CRNDE* is the gene symbol for colorectal neoplasia differentially expressed (non-protein-coding), a long non-coding RNA (lncRNA) gene that expresses numerous splice variants and has a tissue-specific expression pattern. According to previous studies, this lncRNA could have some essential role in cancer, neurobiology, and development (53). This non-coding RNA is involved in enteral metabolism by regulating different genes, responding to insulin/IGF signaling. *CRNDE* nuclear transcripts are specifically downregulated by insulin, IGF1, and IGF2, by both the PI3K/Akt/mTOR and Raf/MAPK signaling pathways (54). This lncRNA also could promote glioma cell growth and invasion through mTOR signaling pathway (55). Interestingly, the effect of *CRNDE* on the growth of myeloma cell could perform by suppressing miR-451 (56). About the functions of this lncRNA in several disease, various results have been reported. *CRNDE* by suppressing miR-348 could promote hepatic carcinoma cell migration, invasion, and proliferation (57). Suppressing miR-145 by this lncRNA could promote proliferation of Gastric cancer cells (58). This RNA by suppressing PUMA expression, enhance the cervical cancer progression (59). Specially in breast cancer, previous studies have been investigated about some different networks. For example, previous studies revealed that *CRNDE* by inhibiting miR-136 in BC cell lines could activate Wnt/β-catenin signaling and overexpression of this RNA in Breast cancer is reported (60). Our study discussed about a novel and different ceRNA network and specially, a single *CRNDE* – hsa-miR-135b-3p – *ADAMTS5* axis in a multi-component ceRNA network. Based on our investigation, both *ADAMTS5* and *CRNDE* could have the tumor suppressor effect on the human breast cancer.

Based on genecards.com, VTI1B is a protein-coding gene that codes for vesicle transport through interaction with T-SNAREs 1B. Hermansky-Pudlak syndrome and familial hemophagocytic lymphohistiocytosis 4 are two conditions linked to VTI1B. Response to increased platelet cytosolic Ca2+, and wtCFTR as well as late endosome and lysosome, are some of its linked processes (norm and CF). In this study, we performed the DE analysis of the VTI1B in the BC samples for the first time. There was no

previous research about the correlation of VTI1B's expression level with the BC development. About lncRNA BAIAP2-AS1, previous studies revealed that BAIAP2-AS1 activated by E2F1 and HepG2 and PLC5 cell growth and metastasis were suppressed by BAIAP2-AS1 silencing. In general, BAIAP2-AS1 altered the miR-361-3p/SOX4 axis to encourage the growth of HCC (61). Xianli Gong et al. in a same gene expression profiling demonstrated that BAIAP2-AS1 is a differentially expressed RNA in the HBV-related HCC (62). Furthermore, Xiaogang Mao et al.in 2018 revealed that BAIAP2-AS1 is a potential cervical squamous cell carcinoma biomarker and can be considered as a novel predictor of the patients' survival (63). However, there was no previous study about the effect of changes in the expression level of BAIAP2-AS1 in the BC. In this study, we find that BAIAP2-AS1 has a significant high-expression in the BC patients. Also, based on our bioinformatics investigation, there was a significant positive correlation between the expression of BAIAP2-AS1 and the survival rate of the patients. So, this lncRNA could be a potential predictor of the BC survival. This lncRNA has a significant negative co-expression with ADAMTS5. So, this lncRNA could have a significant suppressor effect on the ADAMTS5's expression level.

To finding more accurate and valid results, it is suggested that the same experiments be perform on the samples from different populations and examine the possibility of being biomarker in the different disease. This study demonstrates a possible ceRNA network that could have effect on Breast cancer. Additionally, as we select and analyzed the has-miR-135b-3p completely bioinformatically, we suggest that the expression level of this micro-RNA be analyzed in next investigations and the biomarker capability of this miRNA be evaluated. High or low expression of this miRNA could have some essential information about the oncogenic or tumor suppressor effect of this miRNA on Breast cancer, especially in Isfahan. Further studies could make these results clearer and have more accurate information about this network and the correlation of it in different cancer types, including BC. Furthermore, we highly recommended that the expression level of BAIAP2-AS1, VTI1B, and lncRNA CRNDE be evaluated by qRT-PCR experiment on the human clinical tumor samples, or in the different BC cell lines. Also, we suggest that the RNA interaction of BAIAP2-AS1 with ADAMTS5 be evaluated by luciferase assay experiments for validation of this study.

Conclusion

In this study, we demonstrated that ADAMTS5 can be regulated by lncRNAs CRNDE, BAIAP2-AS1, mRNA VTI1B. Dysregulation of mentioned RNAs can lead ADAMTS5 expression level to the abnormality (in this study, low-expression). In this study, we demonstrated that low-expression of ADAMTS5 is a potential cause of the malignancy. There was no previous study about the difference in the expression of ADAMTS5 in BC patients. Furthermore, for the first time, we demonstrated that VTI1B and lncRNA CRNDE are the two potential oncogene of BC (up-regulation). Also, we find that lncRNA BAIAP2-AS1 is a novel tumor suppressor RNA in the BC samples. lncRNA BAIAP2-AS1 and lncRNA CRNDE can predict the BC survival rate. Also, ADAMTS5 have a significant negative correlation with lncRNA BAIAP2-AS1 and lncRNA CRNDE.

Declarations

Ethics approval and consent to participate: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of the Ethics Committee of Isfahan University of Medical Sciences (approval number: 3838988).

Consent for publication: Informed consent was obtained from all individual participants included in the study.

Availability of data and materials: The datasets generated or analyzed during the current study are available in the GEO repository, GSE42568.

Competing interests: The authors declare that they have no competing interests.

Funding: Not applicable.

Authors' contributions: NT, QT, EN, and PT performed the experiments, data analysis, and visualization. MA is the corresponding authors and have revised the manuscript and validated the experiment. MR have revised the manuscript, validated the results and the experiment and designed the project as the supervisor of this research.

Acknowledgment: Not applicable.

References

1. Zhang L, Zhou W, Velculescu VE, Kern SE, Hruban RH, Hamilton SR, et al. Gene expression profiles in normal and cancer cells. Science [Internet]. 1997 May 23 [cited 2022 Jun 29];276(5316):1268–72. Available from: https://pubmed.ncbi.nlm.nih.gov/9157888/
2. Asselah T, Bièche I, Sabbagh A, Bedossa P, Moreau R, Valla D, et al. Gene expression and hepatitis C virus infection. Gut [Internet]. 2009 Jun [cited 2022 Jun 29];58(6):846. Available from: /pmc/articles/PMC2673514/
3. Theuns J, Van Broeckhoven C. Transcriptional regulation of Alzheimer's disease genes: implications for susceptibility. Hum Mol Genet [Internet]. 2000 [cited 2022 Jun 29];9(16):2383–94. Available from: https://pubmed.ncbi.nlm.nih.gov/11005793/
4. Das UN, Rao AA. Gene expression profile in obesity and type 2 diabetes mellitus. Lipids Health Dis [Internet]. 2007 [cited 2021 May 13];6:35. Available from: /pmc/articles/PMC2242786/
5. Tajouri L, Fernandez F, Griffiths L. Gene Expression Studies in Multiple Sclerosis. Curr Genomics [Internet]. 2007 May 31 [cited 2021 May 13];8(3):181–9. Available from: /pmc/articles/PMC2435352/
6. Chen X, Cheung ST, So S, Fan ST, Barry C, Higgins J, et al. Gene Expression Patterns in Human Liver Cancers. Mol Biol Cell [Internet]. 2002 Jun [cited 2021 May 13];13(6):1929–39. Available from: http://genome-www.stanford.edu/hcc/

7. Kapatai G, Brundler MA, Jenkinson H, Kearns P, Parulekar M, Peet AC, et al. Gene expression profiling identifies different sub-types of retinoblastoma. Br J Cancer [Internet]. 2013 Jul 23 [cited 2021 May 13];109(2):512–25. Available from: /pmc/articles/PMC3721394/
8. Kheirelseid EAH, Miller N, Chang KH, Nugent M, Kerin MJ. Clinical applications of gene expression in colorectal cancer. J Gastrointest Oncol [Internet]. 2013 [cited 2021 May 13];4(2):144–57. Available from: /pmc/articles/PMC3635192/
9. Petty RD, Nicolson MC, Kerr KM, Collie-Duguid E, Murray GI. Gene expression profiling in non-small cell lung cancer: From molecular mechanisms to clinical application [Internet]. Vol. 10, Clinical Cancer Research. Clin Cancer Res; 2004 [cited 2021 May 13]. p. 3237–48. Available from: https://pubmed.ncbi.nlm.nih.gov/15161676/
10. Nagai MA. Genetic alterations in head and neck squamous cell carcinomas. Brazilian J Med Biol Res [Internet]. 1999 [cited 2021 May 13];32(7):897–904. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-879X1999000700015&lng=en&nrm=iso&tlng=en
11. Bao T, Davidson NE. Gene Expression Profiling of Breast Cancer [Internet]. Vol. 42, Advances in Surgery. NIH Public Access; 2008 [cited 2021 May 13]. p. 249–60. Available from: /pmc/articles/PMC2775529/
12. Arpino G, Generali D, Sapino A, Lucia DM, Frassoldati A, de Laurentis M, et al. Gene expression profiling in breast cancer: A clinical perspective. Vol. 22, Breast. Churchill Livingstone; 2013. p. 109–20.
13. Guler EN. Gene Expression Profiling in Breast Cancer and Its Effect on Therapy Selection in Early-Stage Breast Cancer. Eur J Breast Heal [Internet]. 2017 Sep 27 [cited 2021 May 13];13(4):168–74. Available from: /pmc/articles/PMC5648272/
14. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: Classification, prognostication, and prediction [Internet]. Vol. 378, The Lancet. Elsevier B.V.; 2011 [cited 2021 May 14]. p. 1812–23. Available from: http://www.thelancet.com/article/S0140673611615390/fulltext
15. Assad Samani L, Javadirad SM, Parsafar S, Tabatabaeian H, Ghaedi K, Azadeh M. TP53 rs1625895 is Related to Breast Cancer Incidence and Early Death in Iranian Population. Indian J Clin Biochem [Internet]. 2019 Oct 1 [cited 2021 May 23];34(4):485–9. Available from: https://link.springer.com/article/10.1007/s12291-018-0774-6
16. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin [Internet]. 2018 Nov [cited 2021 May 14];68(6):394–424. Available from: https://pubmed.ncbi.nlm.nih.gov/30207593/
17. Rouigari M, Dehbashi M, Tabatabaeian H, Ghaedi K, Mohammadynejad P, Azadeh M. Evaluation of the Expression Level and Hormone Receptor Association of miR-126 in Breast Cancer. Indian J Clin Biochem [Internet]. 2019 Oct 1 [cited 2022 Jun 29];34(4):451. Available from: /pmc/articles/PMC6801275/

18. Azadeh M, Salehzadeh A, Ghaedi K, Talesh Sasani S. NEAT1 can be a diagnostic biomarker in the breast cancer and gastric cancer patients by targeting XIST, hsa-miR-612, and MTRNR2L8: integrated RNA targetome interaction and experimental expression analysis. Genes Environ 2022 441 [Internet]. 2022 May 17 [cited 2022 Jun 29];44(1):1–18. Available from: https://genesenvironment.biomedcentral.com/articles/10.1186/s41021-022-00244-3
19. Azadeh M, Salehzadeh A, Ghaedi K, Sasani ST. Decreased expression level of long non-coding RNA CCAT1, was observed in breast cancer tissue of an Isfahanian population (Iran). Gene Reports. 2021 Jun 1;23:101154.
20. Fattahi Dolatabadi N, Dehghani A, Shahand E, Yazdanshenas M, Tabatabaeian H, Zamani A, et al. The interaction between MALAT1 target, miR-143-3p, and RALGAPA2 is affected by functional SNP rs3827693 in breast cancer. Hum Cell 2020 334 [Internet]. 2020 Sep 3 [cited 2022 Jun 29];33(4):1229–39. Available from: https://link.springer.com/article/10.1007/s13577-020-00422-x
21. Kim YA, Cho H, Lee N, Jung SY, Sim SH, Park IH, et al. Doxorubicin-induced heart failure in cancer patients: A cohort study based on the Korean National Health Insurance Database. Cancer Med [Internet]. 2018 Dec 1 [cited 2021 May 14];7(12):6084–92. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/cam4.1886
22. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2020. CA Cancer J Clin [Internet]. 2020 Jan 1 [cited 2021 May 14];70(1):7–30. Available from: https://acsjournals.onlinelibrary.wiley.com/doi/full/10.3322/caac.21590
23. Li N, Deng Y, Zhou L, Tian T, Yang S, Wu Y, et al. Global burden of breast cancer and attributable risk factors in 195 countries and territories, from 1990 to 2017: Results from the Global Burden of Disease Study 2017. J Hematol Oncol [Internet]. 2019 Dec 21 [cited 2021 May 14];12(1):1–12. Available from: https://doi.org/10.1186/s13045-019-0828-0
24. Dufva M. Introduction to microarray technology. [Internet]. Vol. 529, Methods in molecular biology (Clifton, N.J.). Methods Mol Biol; 2009 [cited 2021 May 14]. p. 1–22. Available from: https://pubmed.ncbi.nlm.nih.gov/19381982/
25. Kumar S, Sharghi-Namini S, Rao N, Ge R. ADAMTS5 functions as an anti-angiogenic and anti-tumorigenic protein independent of its proteoglycanase activity. Am J Pathol. 2012 Sep 1;181(3):1056–68.
26. Kobayashi H, Hirata M, Saito T, Itoh S, Chung U Il, Kawaguchi H. Transcriptional induction of ADAMTS5 protein by nuclear factor-κB (NF-κB) family member RelA/p65 in chondrocytes during osteoarthritis development. J Biol Chem. 2013 Oct 4;288(40):28620–9.
27. Huang J, Sun Y, Chen H, Liao Y, Li S, Chen C, et al. ADAMTS5 acts as a tumor suppressor by inhibiting migration, invasion and angiogenesis in human gastric cancer. Gastric Cancer [Internet]. 2019 Mar 15 [cited 2021 May 27];22(2):287–301. Available from: https://pubmed.ncbi.nlm.nih.gov/30105548/
28. Malvia S, Bagadi SAR, Pradhan D, Chintamani C, Bhatnagar A, Arora D, et al. Study of Gene Expression Profiles of Breast Cancers in Indian Women. Sci Rep [Internet]. 2019 Dec 1 [cited 2021

May 27];9(1):1–15. Available from: https://doi.org/10.1038/s41598-019-46261-1

29. Zhu Z, Xu J, Wu X, Lin S, Li L, Ye W, et al. In Silico Identification of Contradictory Role of ADAMTS5 in Hepatocellular Carcinoma. Technol Cancer Res Treat [Internet]. 2021 Jan 1 [cited 2021 May 27];20. Available from: https://pubmed.ncbi.nlm.nih.gov/33522433/
30. Li C, Xiong Y, Yang X, Wang L, Zhang S, Dai N, et al. Lost expression of ADAMTS5 protein associates with progression and poor prognosis of hepatocellular carcinoma. Drug Des Devel Ther [Internet]. 2015 Mar 24 [cited 2021 May 27];9:1773–83. Available from: https://pubmed.ncbi.nlm.nih.gov/25848214/
31. Haraguchi N, Ohara N, Koseki J, Takahashi H, Nishimura J, Hata T, et al. High expression of ADAMTS5 is a potent marker for lymphatic invasion and lymph node metastasis in colorectal cancer. Mol Clin Oncol [Internet]. 2017 Jan [cited 2021 May 27];6(1):130–4. Available from: https://pubmed.ncbi.nlm.nih.gov/28123746/
32. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: Insights into functions [Internet]. Vol. 10, Nature Reviews Genetics. Nat Rev Genet; 2009 [cited 2021 May 14]. p. 155–9. Available from: https://pubmed.ncbi.nlm.nih.gov/19188922/
33. Fang Y, Fullwood MJ. Roles, Functions, and Mechanisms of Long Non-coding RNAs in Cancer [Internet]. Vol. 14, Genomics, Proteomics and Bioinformatics. Beijing Genomics Institute; 2016 [cited 2021 May 14]. p. 42–54. Available from: https://pubmed.ncbi.nlm.nih.gov/26883671/
34. Bin X, Hongjian Y, Xiping Z, Bo C, Shifeng Y, Binbin T. Research progresses in roles of LncRNA and its relationships with breast cancer [Internet]. Vol. 18, Cancer Cell International. BioMed Central Ltd.; 2018 [cited 2021 May 14]. Available from: https://pubmed.ncbi.nlm.nih.gov/30459529/
35. Clarke C, Madden SF, Doolan P, Aherne ST, Joyce H, O'Driscoll L, et al. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. Carcinogenesis [Internet]. 2013 Oct [cited 2022 Jun 29];34(10):2300–8. Available from: https://pubmed.ncbi.nlm.nih.gov/23740839/
36. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics [Internet]. 2004 Feb 12 [cited 2021 Dec 17];20(3):307–15. Available from: https://academic.oup.com/bioinformatics/article/20/3/307/185980
37. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res [Internet]. 2015 Apr 20 [cited 2022 Jun 29];43(7):e47–e47. Available from: https://academic.oup.com/nar/article/43/7/e47/2414268
38. Li JH, Liu S, Zhou H, Qu LH, Yang JH. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res [Internet]. 2014 Jan 1 [cited 2022 Jun 29];42(D1):D92–7. Available from: https://academic.oup.com/nar/article/42/D1/D92/1063720
39. Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. Nucleic Acids Res [Internet]. 2019 Jul 1 [cited 2022 Jun

29];47(W1):W556–60. Available from: https://pubmed.ncbi.nlm.nih.gov/31114875/

40. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res [Internet]. 2003 Nov [cited 2022 Jun 29];13(11):2498. Available from: /pmc/articles/PMC403769/
41. Jensen LJ, Kuhn M, Stark M, Chaffron S, Creevey C, Muller J, et al. STRING 8–a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res [Internet]. 2009 [cited 2022 Jun 29];37(Database issue). Available from: https://pubmed.ncbi.nlm.nih.gov/18940858/
42. Peterson H, Kolberg L, Raudvere U, Kuzmin I, Vilo J. gprofiler2 – an R package for gene list functional enrichment analysis and namespace conversion toolset g: Profiler. F1000Research [Internet]. 2020 [cited 2022 Jun 29];9. Available from: https://pubmed.ncbi.nlm.nih.gov/33564394/
43. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innov [Internet]. 2021 Aug 28 [cited 2022 Jun 29];2(3):100141. Available from: http://www.cell.com/article/S2666675821000667/fulltext
44. H D, C S, P P, N G. miRWalk–database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform [Internet]. 2011 Oct [cited 2021 Oct 31];44(5):839–47. Available from: https://pubmed.ncbi.nlm.nih.gov/21605702/
45. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: An online resource for prediction of microRNA binding sites. PLoS One [Internet]. 2018 Oct 1 [cited 2022 Jan 14];13(10):e0206239. Available from: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0206239
46. H D, N G. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods [Internet]. 2015 Jul 30 [cited 2021 Jul 29];12(8):697. Available from: https://pubmed.ncbi.nlm.nih.gov/26226356/
47. Karagkouni D, Paraskevopoulou MD, Tastsoglou S, Skoufos G, Karavangeli A, Pierros V, et al. DIANA-LncBase v3: indexing experimentally supported miRNA targets on non-coding transcripts. Nucleic Acids Res [Internet]. 2020 Jan 8 [cited 2021 Aug 4];48(D1):D101–10. Available from: https://academic.oup.com/nar/article/48/D1/D101/5626521
48. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. From the Cover: Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A [Internet]. 2005 Oct 10 [cited 2022 Jun 29];102(43):15545. Available from: /pmc/articles/PMC1239896/
49. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, et al. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat Genet 2003 343 [Internet]. 2003 Jun 15 [cited 2022 Jun 29];34(3):267–73. Available from: https://www.nature.com/articles/ng1180
50. Lü L, Sun J, Shi P, Kong W, Xu K, He B, et al. Identification of circular RNAs as a promising new class of diagnostic biomarkers for human breast cancer. Oncotarget [Internet]. 2017 [cited 2021 May 27];8(27):44096–107. Available from: https://pubmed.ncbi.nlm.nih.gov/28484086/

51. Wang L, Zhu MJ, Ren AM, Wu HF, Han WM, Tan RY, et al. A ten-microRNA signature identified from a genome-wide microRNA expression profiling in human epithelial ovarian cancer. PLoS One [Internet]. 2014 May 9 [cited 2021 May 27];9(5). Available from: https://pubmed.ncbi.nlm.nih.gov/24816756/
52. Hu Z, Wang X, Yang Y, Zhao Y, Shen Z, Huang Y. MicroRNA expression profiling of lung adenocarcinoma in Xuanwei, China: A preliminary study. Med (United States) [Internet]. 2019 May 1 [cited 2021 May 27];98(21). Available from: https://pubmed.ncbi.nlm.nih.gov/31124951/
53. Ellis BC, Molloy PL, Graham LD. CRNDE: A long non-coding RNA involved in CanceR Neurobiology, and DEvelopment [Internet]. Vol. 3, Frontiers in Genetics. Front Genet; 2012 [cited 2021 May 27]. p. 1–15. Available from: https://pubmed.ncbi.nlm.nih.gov/23226159/
54. Ellis BC, Graham LD, Molloy PL. CRNDE, a long non-coding RNA responsive to insulin/IGF signaling, regulates genes involved in central metabolism. Biochim Biophys Acta - Mol Cell Res. 2014 Feb 1;1843(2):372–86.
55. Wang Y, Wang Y, Li J, Zhang Y, Yin H, Han B. CRNDE, a long-noncoding RNA, promotes glioma cell growth and invasion through mTOR signaling. Cancer Lett. 2015 Oct 28;367(2):122–8.
56. Meng Y Bin, He X, Huang YF, Wu QN, Zhou YC, Hao DJ. Long noncoding RNA CRNDE promotes multiple myeloma cell growth by suppressing miR-451. Oncol Res [Internet]. 2017 [cited 2021 May 27];25(7):1207–14. Available from: https://pubmed.ncbi.nlm.nih.gov/28276319/
57. Ji D, Jiang C, Zhang L, Liang N, Jiang T, Yang B, et al. LncRNA CRNDE promotes hepatocellular carcinoma cell proliferation, invasion, and migration through regulating miR-203/ BCAT1 axis. J Cell Physiol. 2019 May 1;234(5):6548–60.
58. Hu CE, Du PZ, Zhang HD, Huang GJ. Long noncoding RNA CRNDE promotes proliferation of gastric cancer cells by targeting miR-145. Cell Physiol Biochem [Internet]. 2017 Jun 1 [cited 2021 May 27];42(1):13–21. Available from: https://pubmed.ncbi.nlm.nih.gov/28490034/
59. Zhang JJ, Fan LP. Long non-coding RNA CRNDE enhances cervical cancer progression by suppressing PUMA expression. Biomed Pharmacother. 2019 Sep 1;117:108726.
60. Huan J, Xing L, Lin Q, Xui H, Qin X. Long noncoding RNA CRNDE activates Wnt/β-catenin signaling pathway through acting as a molecular sponge of microRNA-136 in human breast cancer - PubMed [Internet]. American journal of translational research. 2017 [cited 2021 May 27]. Available from: https://pubmed.ncbi.nlm.nih.gov/28469804/
61. Yang Y, Ge H, Li DQ, Xu AX. E2F1-Induced lncRNA BAIAP2-AS1 Overexpression Contributes to the Malignant Progression of Hepatocellular Carcinoma via miR-361-3p/SOX4 Axis. Dis Markers [Internet]. 2021 [cited 2022 Jun 29];2021. Available from: https://pubmed.ncbi.nlm.nih.gov/34616498/
62. Gong X, Wei W, Chen L, Xia Z, Yu C. Comprehensive analysis of long non-coding RNA expression profiles in hepatitis B virus-related hepatocellular carcinoma. Oncotarget [Internet]. 2016 Jun 7 [cited 2022 Jun 29];7(27):42422–30. Available from: https://pubmed.ncbi.nlm.nih.gov/27285756/
63. Mao X, Qin X, Li L, Zhou J, Zhou M, Li X, et al. A 15-long non-coding RNA signature to improve prognosis prediction of cervical squamous cell carcinoma. Gynecol Oncol [Internet]. 2018 Apr 1 [cited

2022 Jun 29];149(1):181–7. Available from: http://www.gynecologiconcology-online.net/article/S0090825817315949/fulltext

Figures

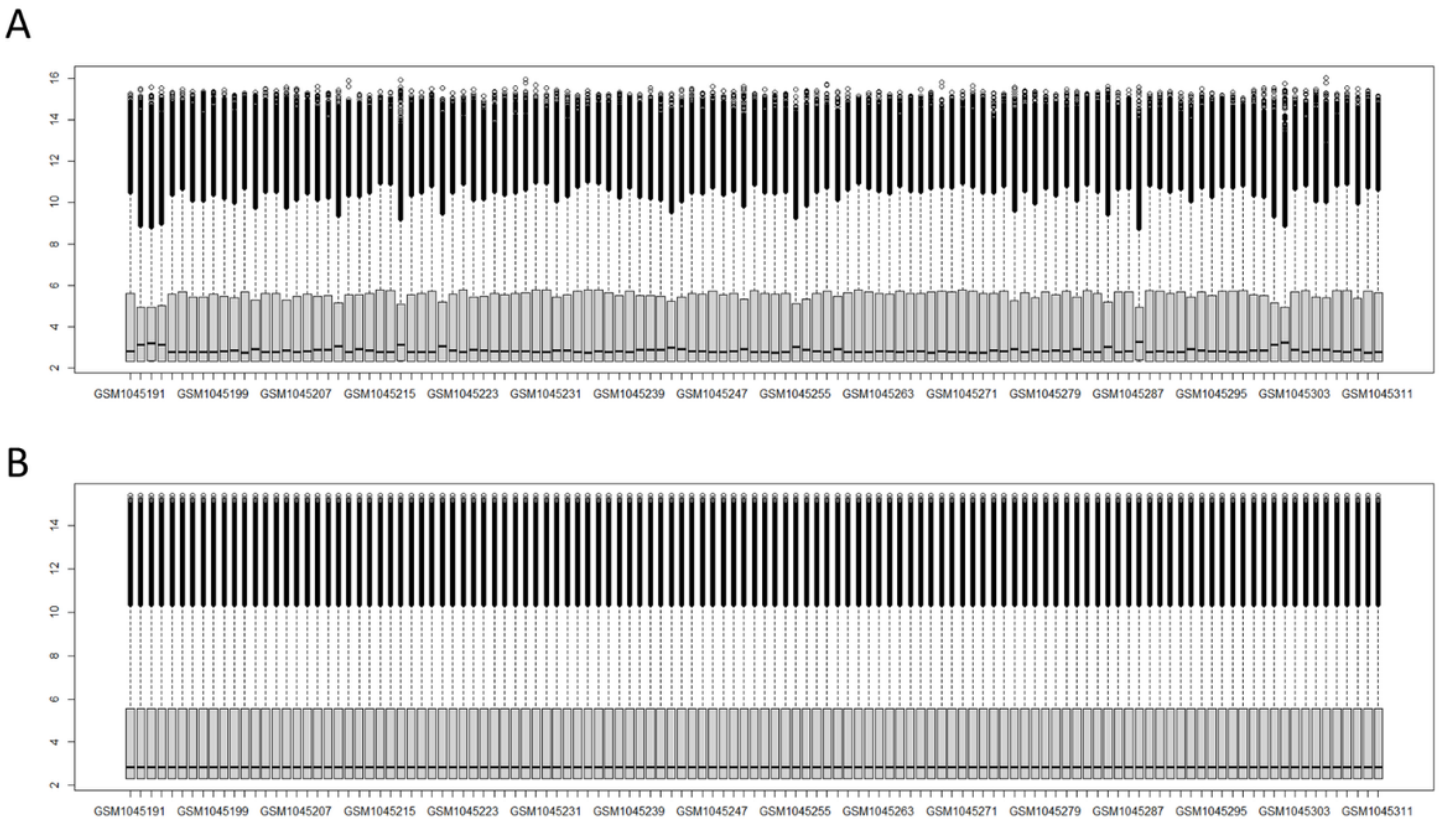


Figure 1

Boxplot of the raw and normalized expression level of 54675 genes in the microarray samples. A) Boxplot of raw microarray samples. B) Boxplot of the normalized expression data of microarray samples.

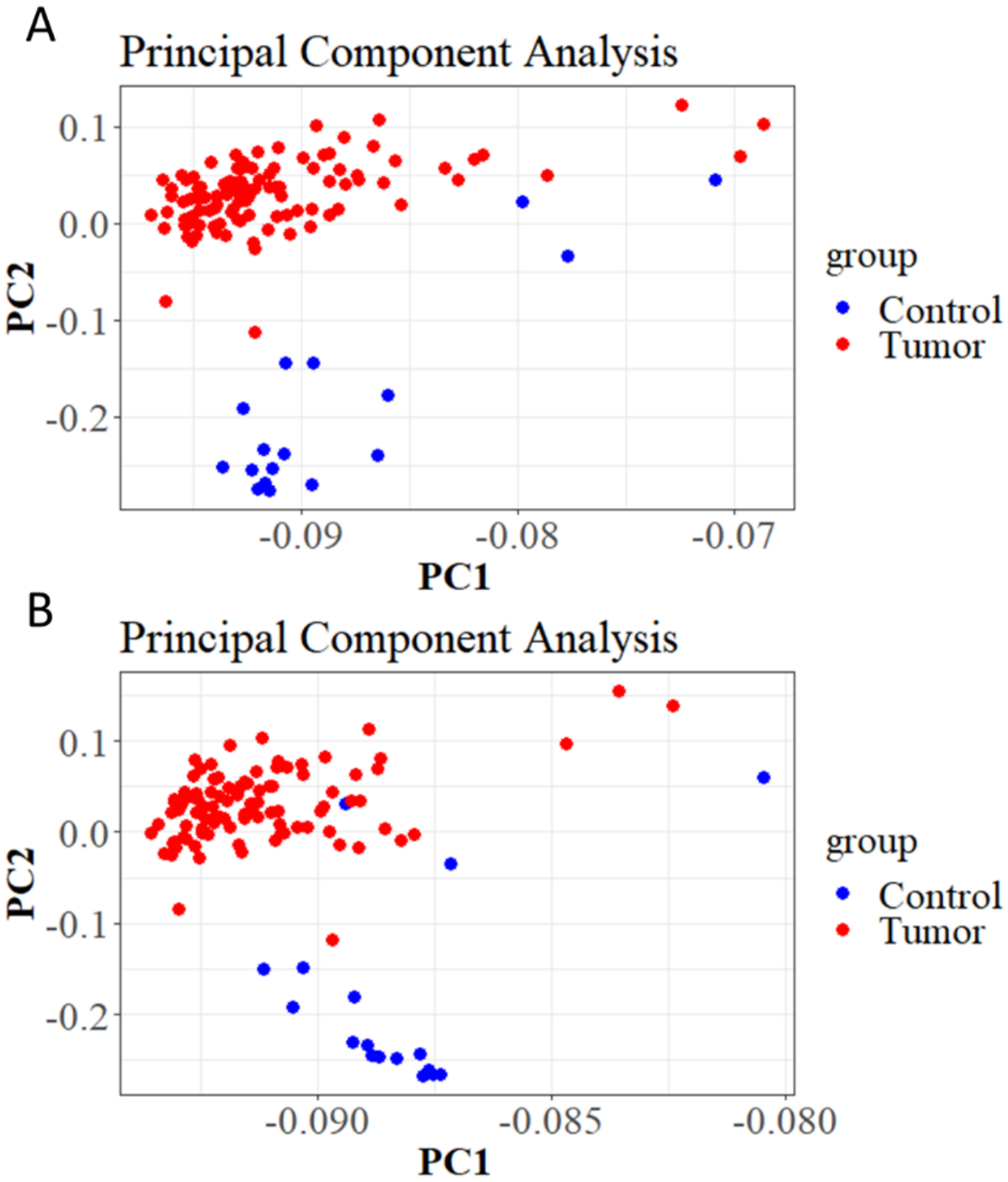


Figure 2

Principal component analysis (PCA) of microarray samples. A) PCA analysis based on raw microarray expression data. B) PCA analysis of normalized expression data.

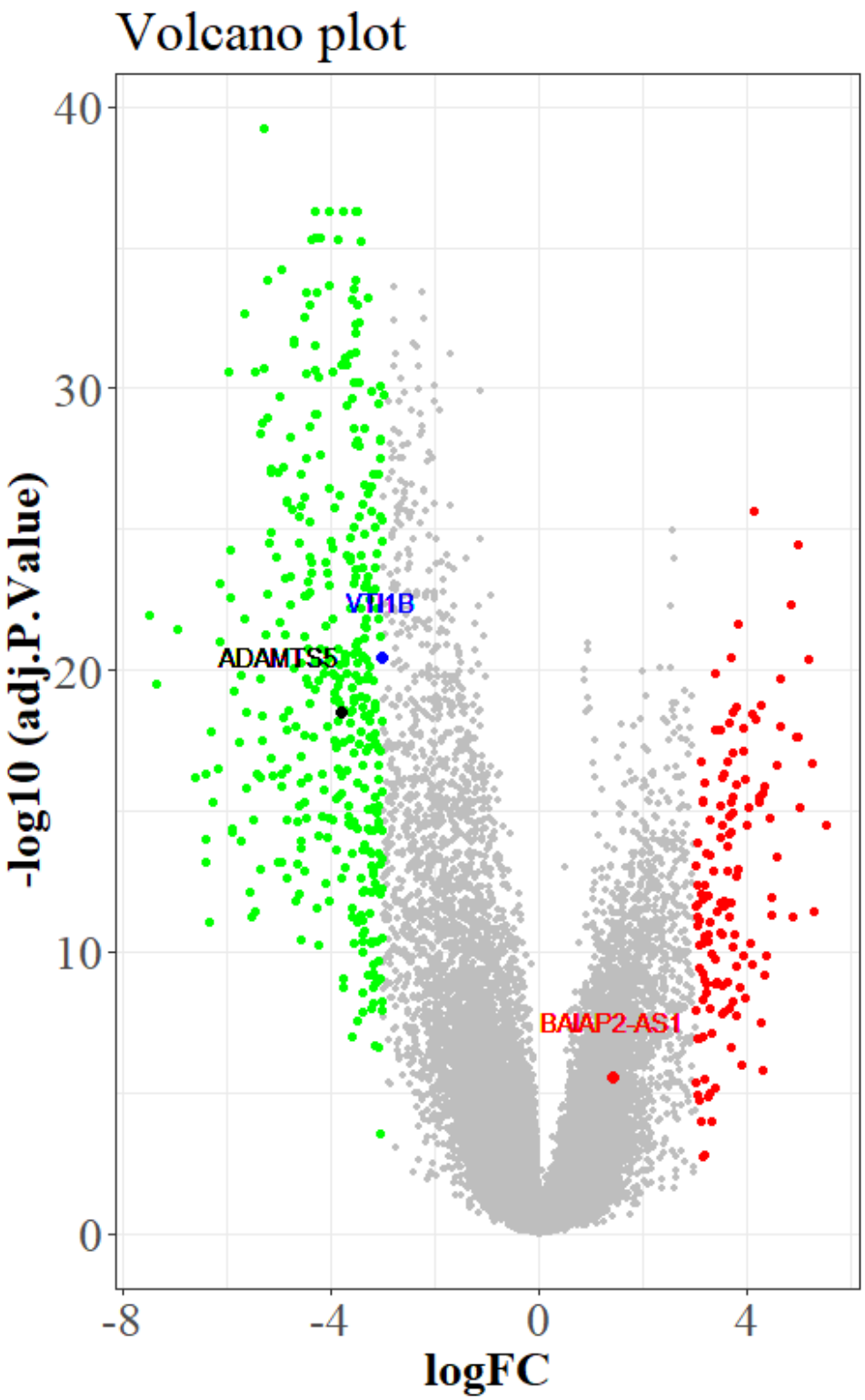


Figure 3

Volcano plot showing the differentially expressed genes in the microarray BC samples. Red color indicates the significant up-regulated genes and the green color indicates the significant low-expressed genes in the GSE42568. ADAMTS5, VTI1B, and BAIAP2-AS1 are showing in the plot by black, blue, and red color.

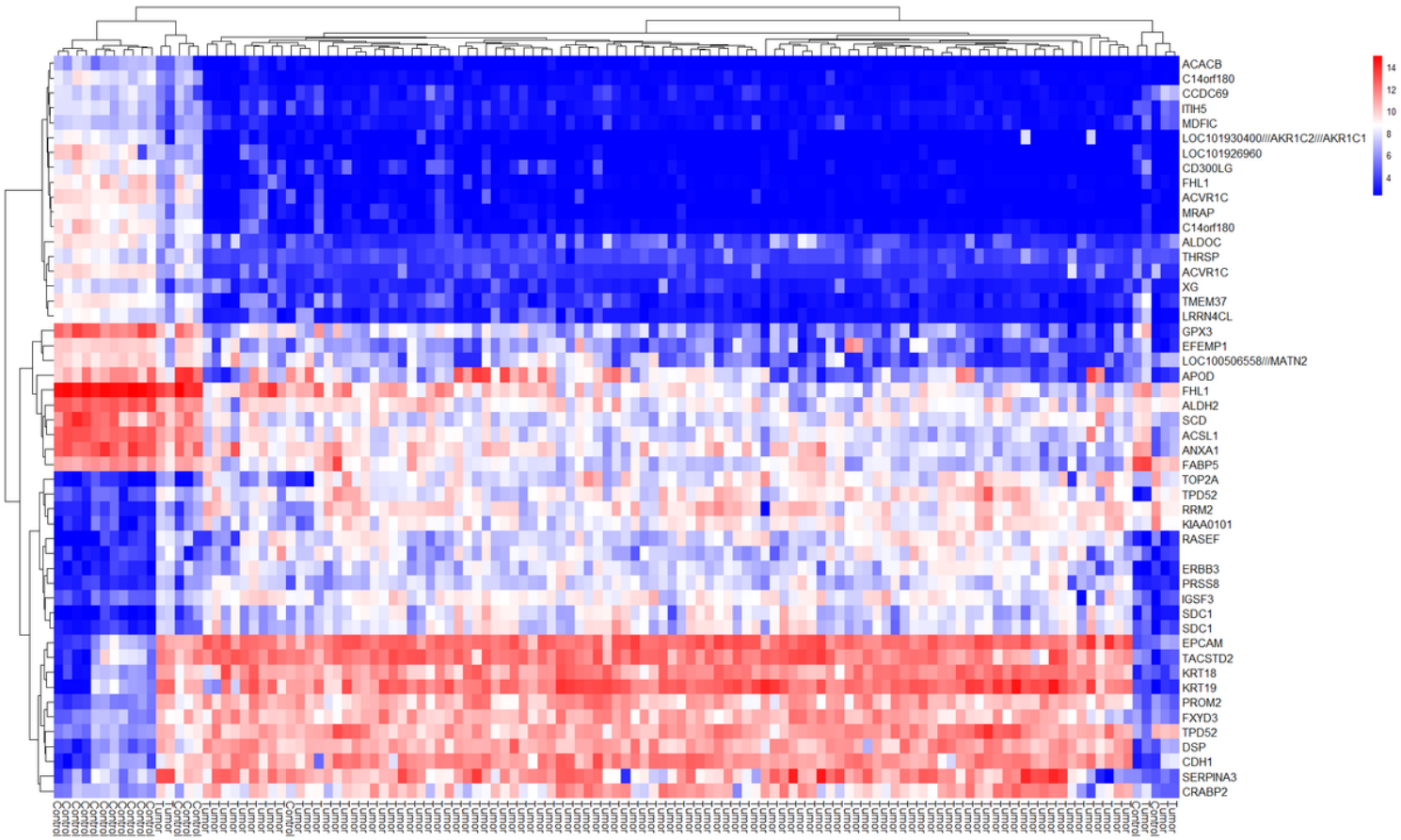


Figure 4

The heatmap of top 50 differentially expressed genes in the GSE42568 microarray dataset.

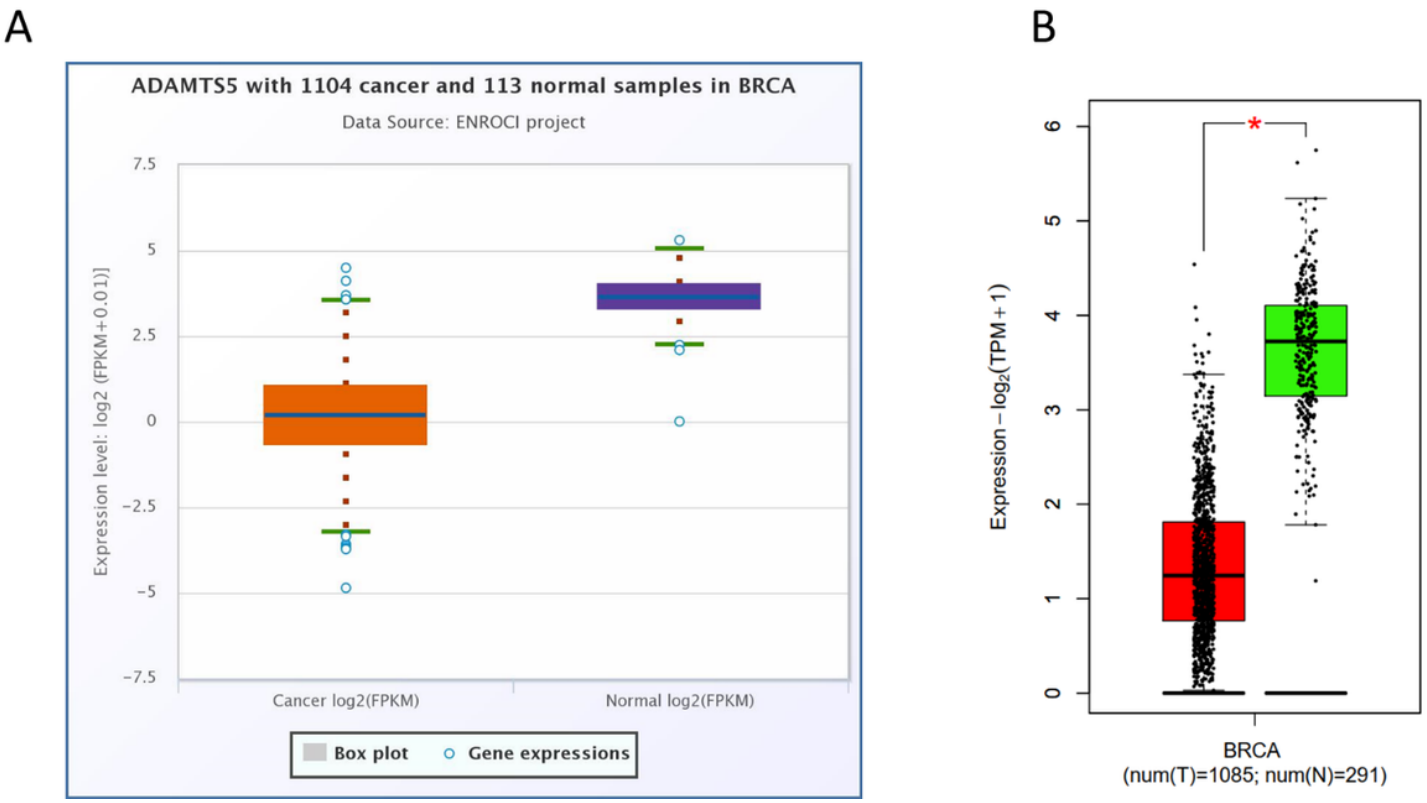


Figure 5

ENCORI and GEPIA2 RNA-seq data analysis revealed that ADAMTS5 have a significant down-regulation in the breast cancer samples. A) The expression analysis of ENCORI online software. B) The differential expression analysis of ADAMTS5 in the BC samples, performed by GEPIA2.

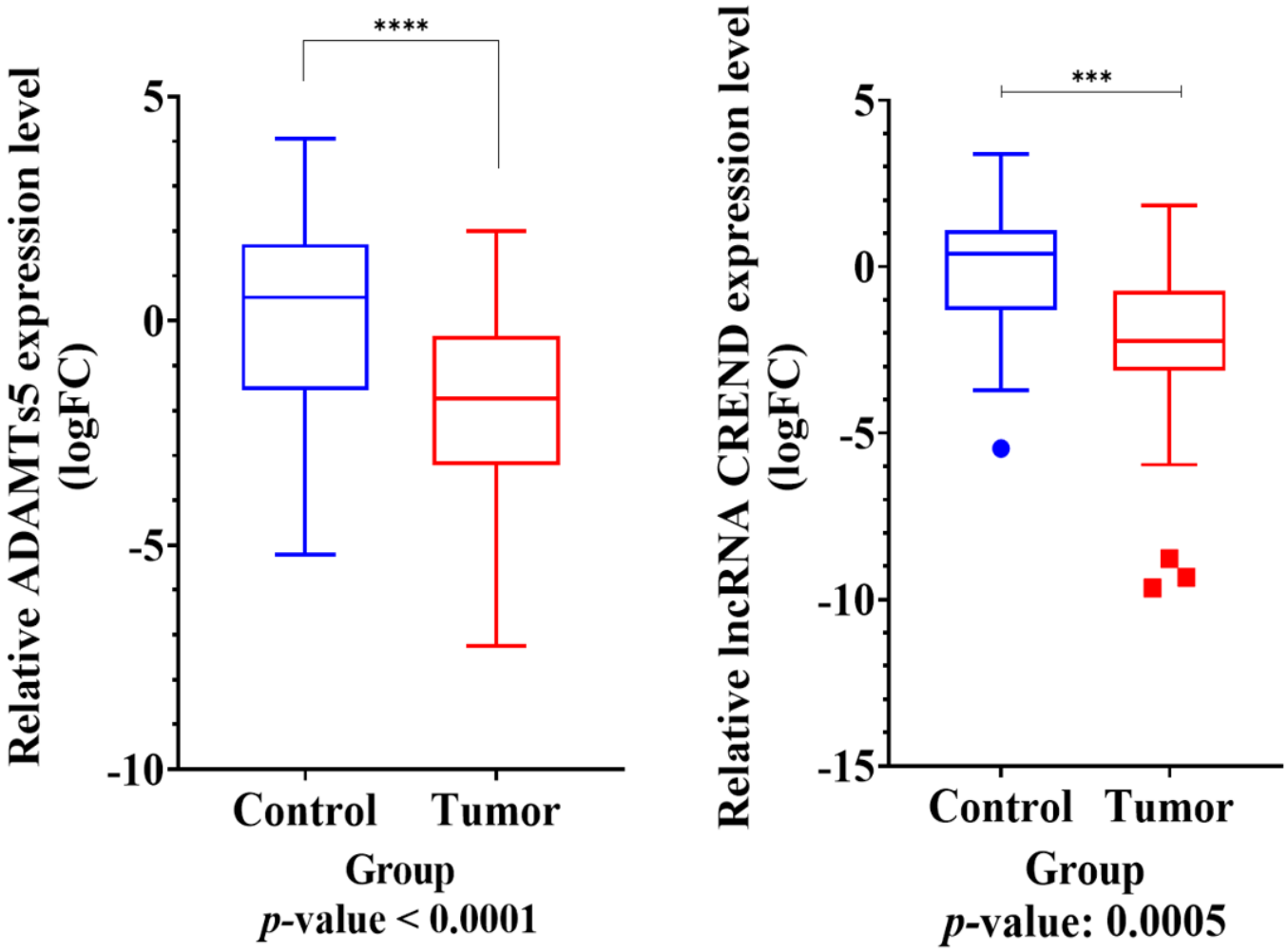


Figure 6

The boxplot of the real-time PCR data of ADAMTS5 expression level, based on the logFC data. ADAMTS5 has a significant low-expression in the human clinical breast cancer samples.

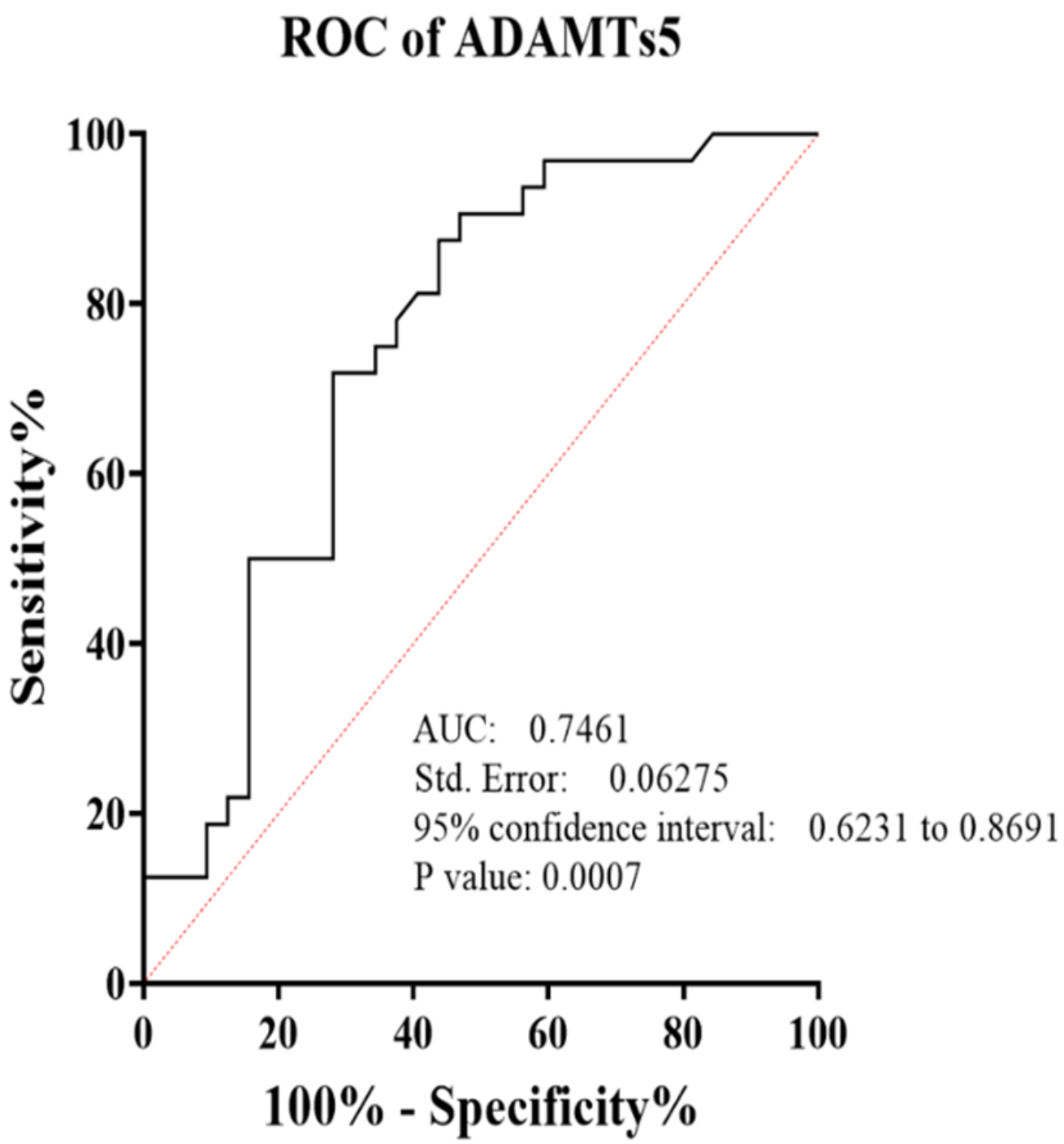


Figure 7

ROC analysis based on the expression data of ADAMTS5 revealed that ADAMTS5 is a potential diagnostic biomarker of breast cancer samples.

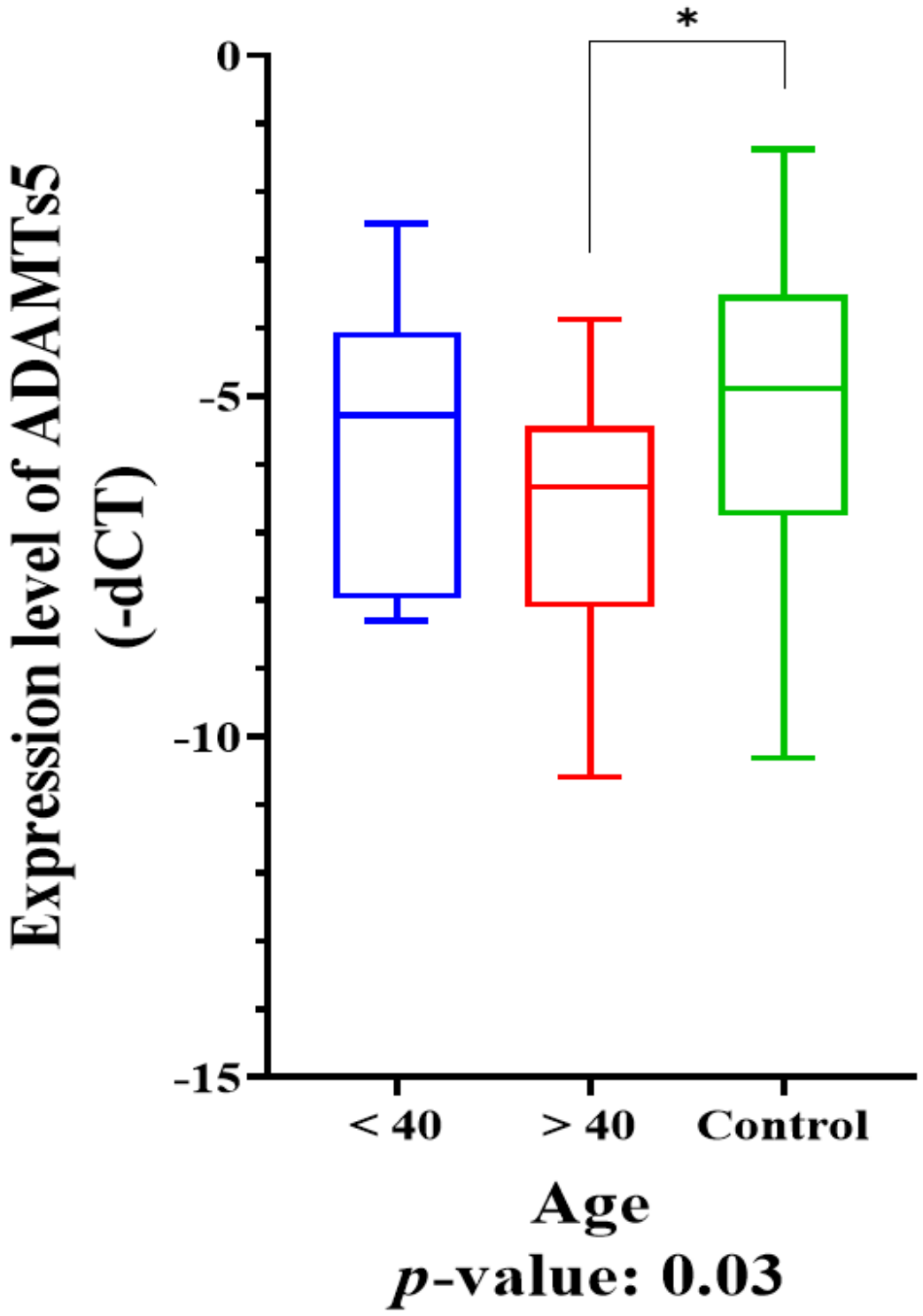


Figure 8

The clinicopathological analysis of ADAMTS5 revealed that ADAMTS5 have a significant down-regulation in the patients older than 40.

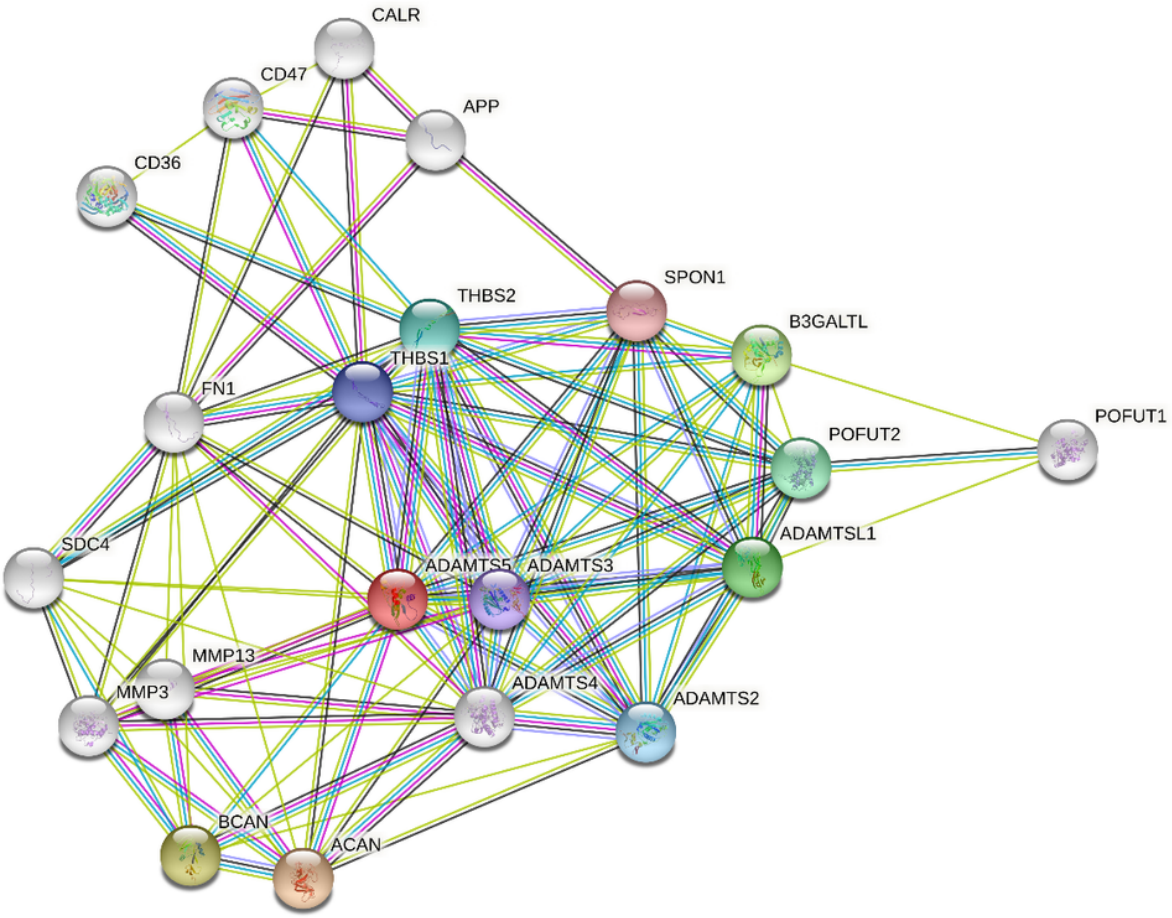


Figure 9

Protein-protein interaction of ADAMTS5, generated by STRING online database.

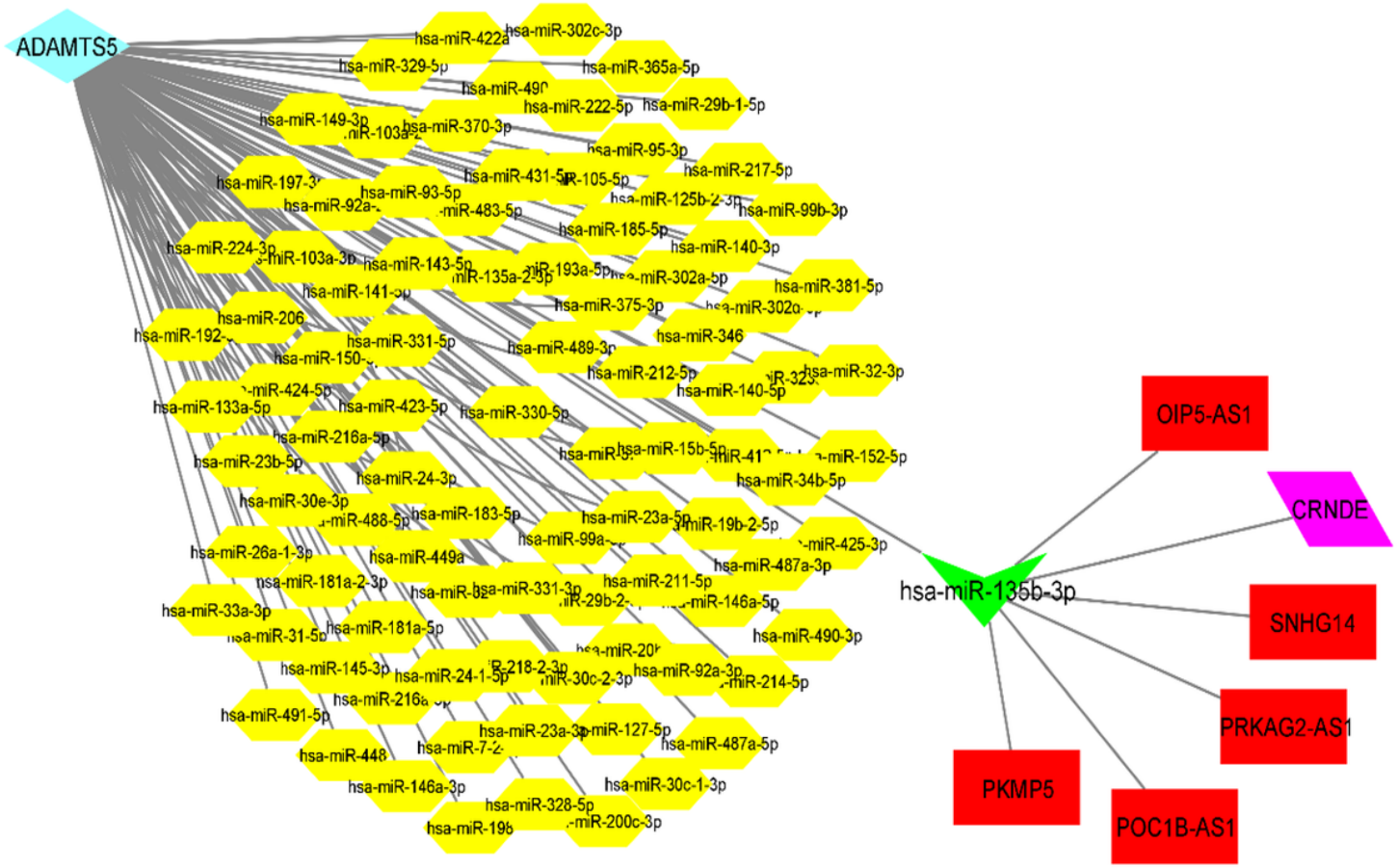


Figure 10

The RNA interaction network of ADAMTS5 gene revealed the hub miRNA and lncRNAs correlated to the ADAMTS5 (hsa-miR-135b-3p, lncRNA CRNDE), directly or indirectly.

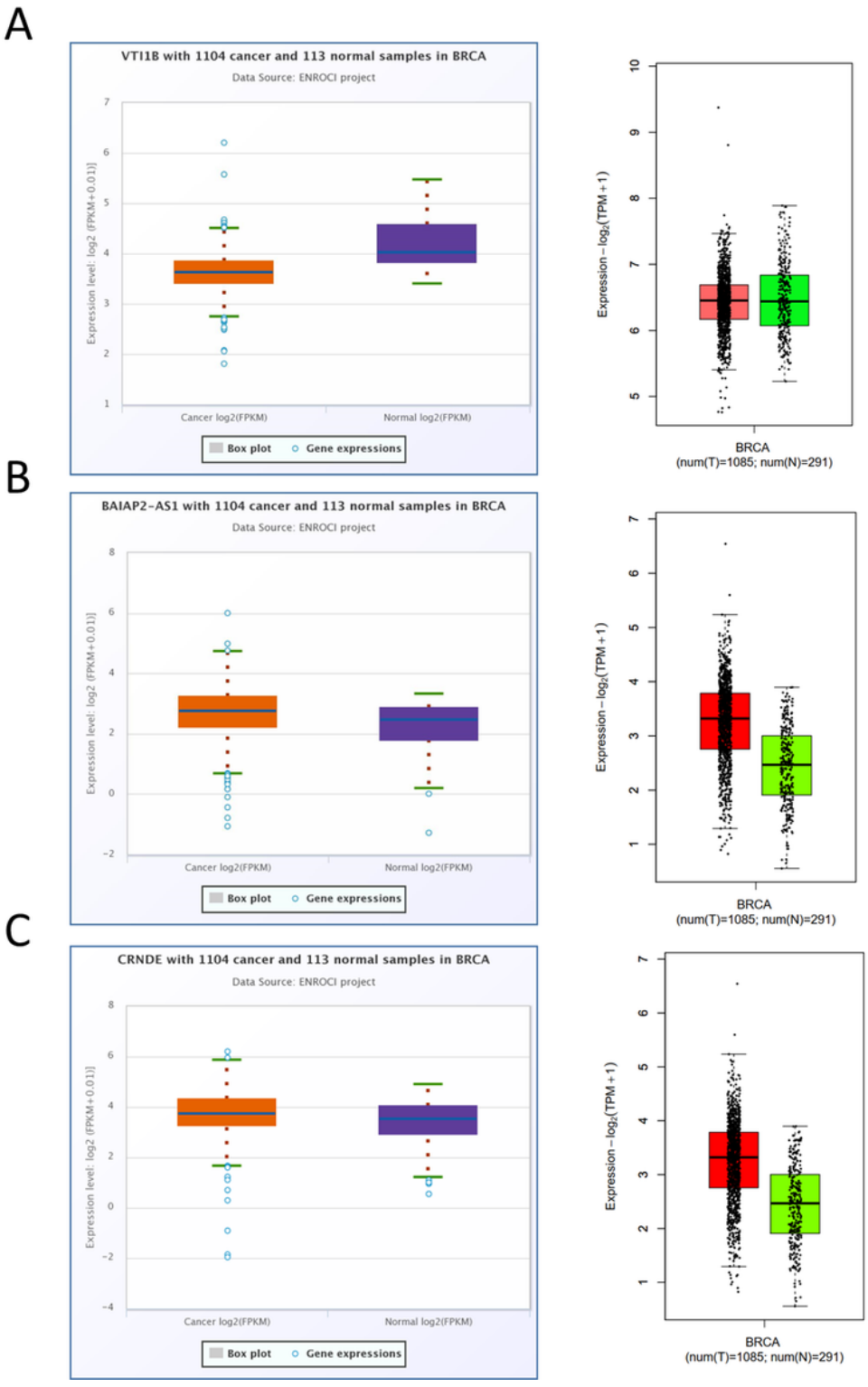


Figure 11

Relative expression analysis of VTI1B, BAIAP2-AS1, and lncRNA CRNDE in the RNAseq breast cancer samples. A) Relative expression analysis of VTI1B revealed that VTI1B has a significant down-regulation in the BC samples. B) Relative expression analysis of B lncRNAAIAP2-AS1 has a significant high expression in the breast cancer samples, compared to control. C) Relative expression analysis of lncRNA CRNDE has a significant up-regulation in the BC samples, compared to control.

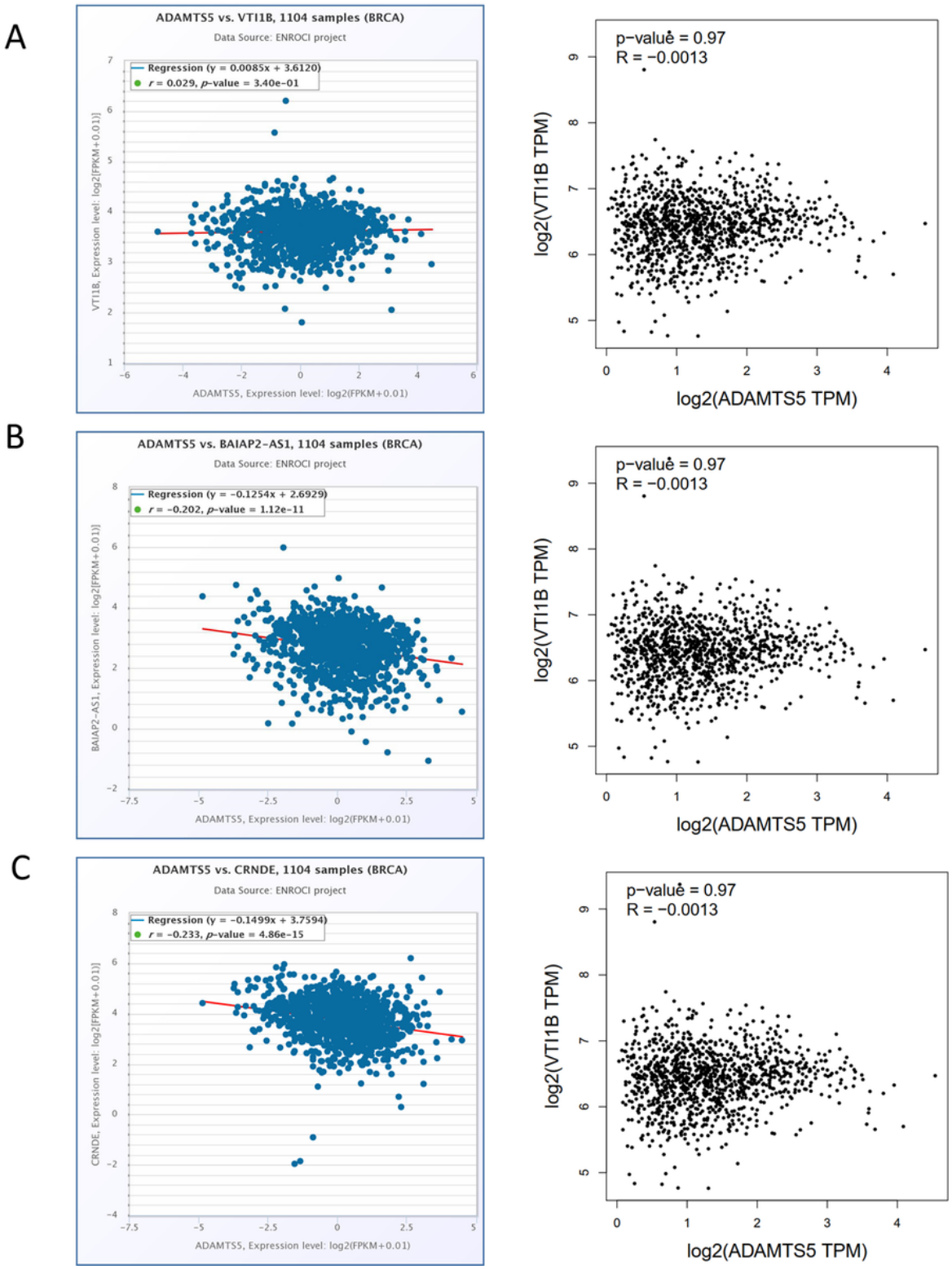


Figure 12

Pearson correlation analysis of the VTI1B, BAIAP2-AS1, and lncRNA CRNDE with ADAMTS5's expression level in the RNAseq samples of breast cancer patients. The right column is the plots of GEPIA2 online software, and the left column is the plots of ENCORI online software. A) ADAMTS5 have no significant

co-expression with the VTI1B. B) lncRNA BAIAP2-AS1 has a significant co-expression with ADAMTS5. C) lncRNA CRNDE has a significant correlation in the expression level with ADAMTS5.

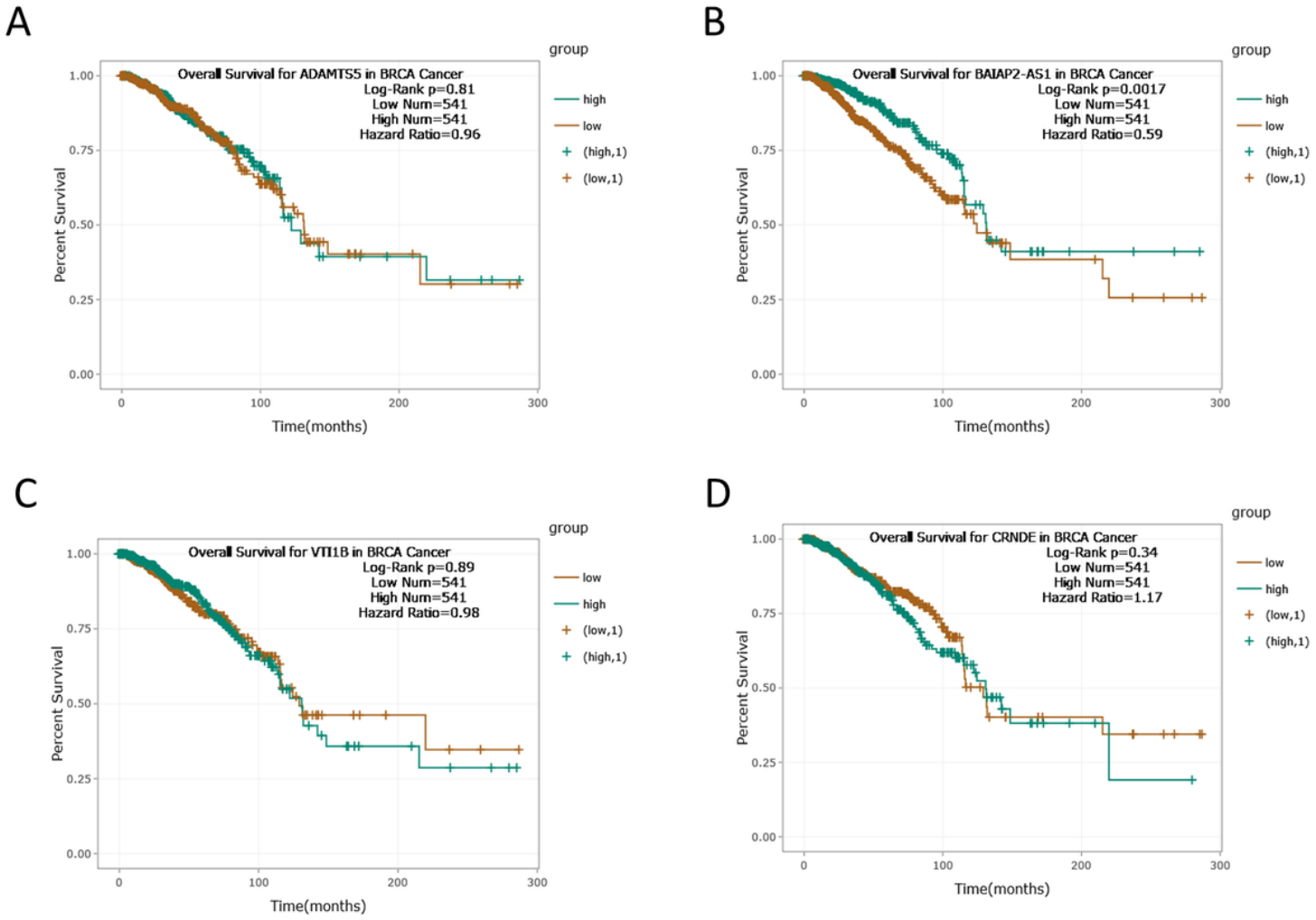


Figure 13

Survival analysis of ENCORI online software revealed the negative and positive correlation of BAIAP2-AS1 and lncRNA CRNDE with the survival rate of BC patients.

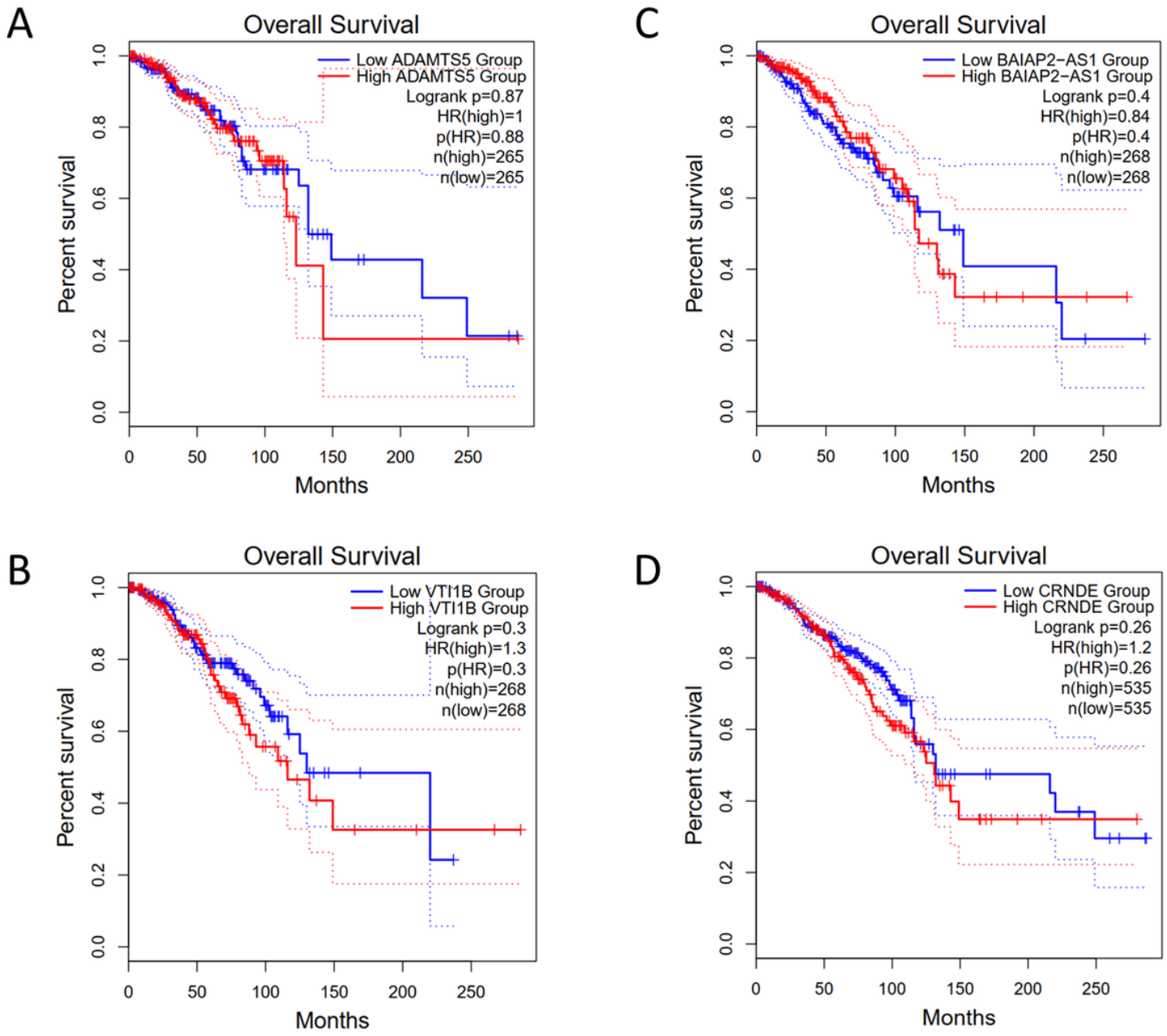


Figure 14

Survival analysis of GEPIA online software.

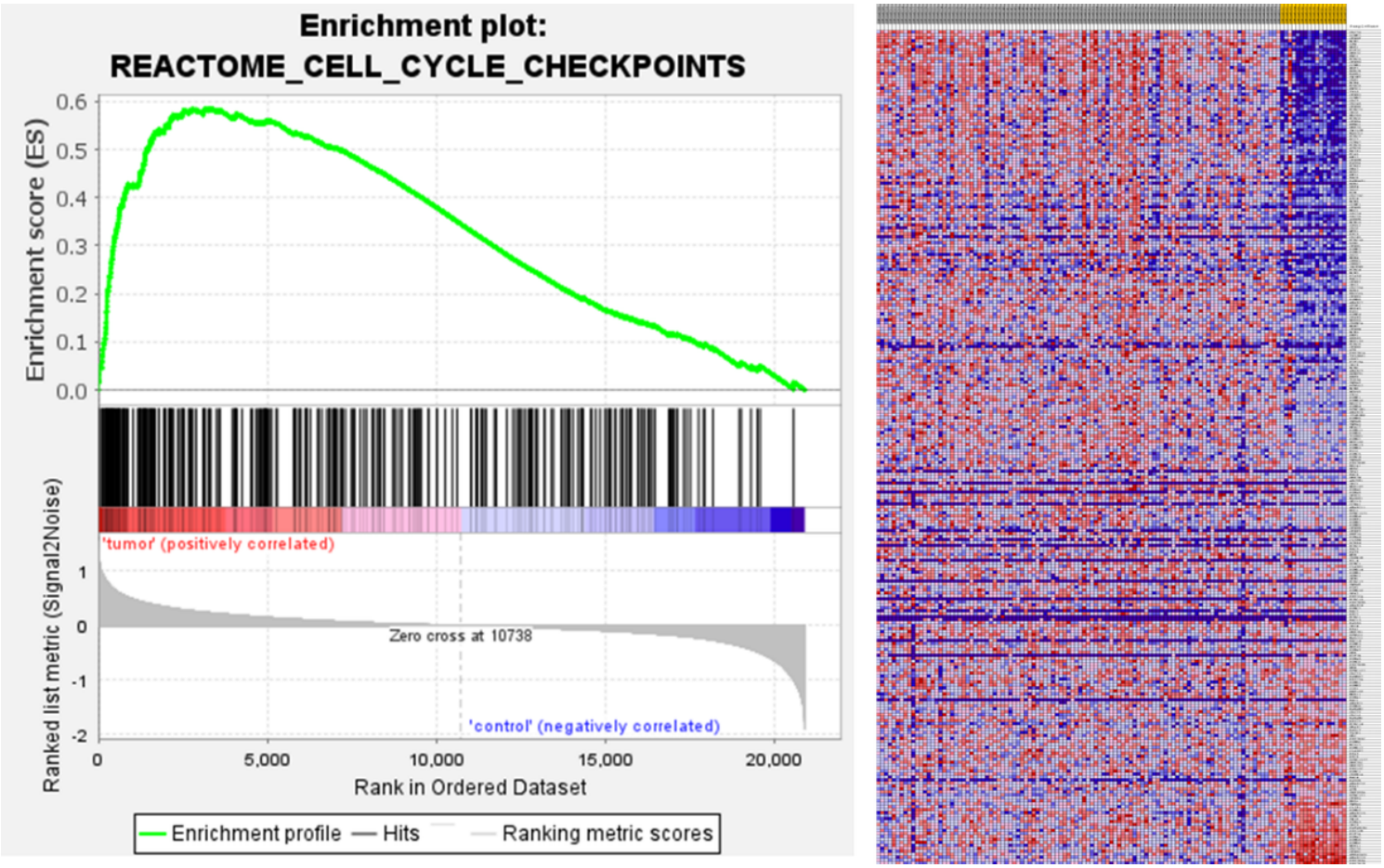


Figure 15

The GSEA data analysis revealed that the up-regulated genes of GSE42568 have a significant role in the cell cycle checkpoints signaling pathway.

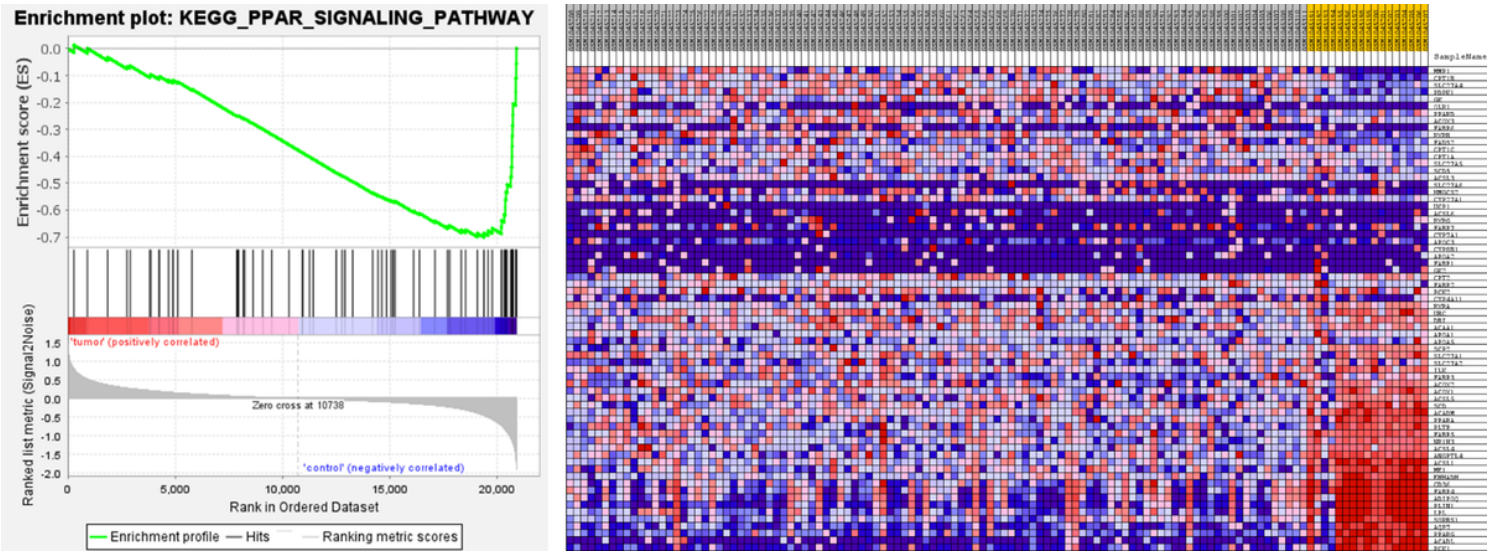
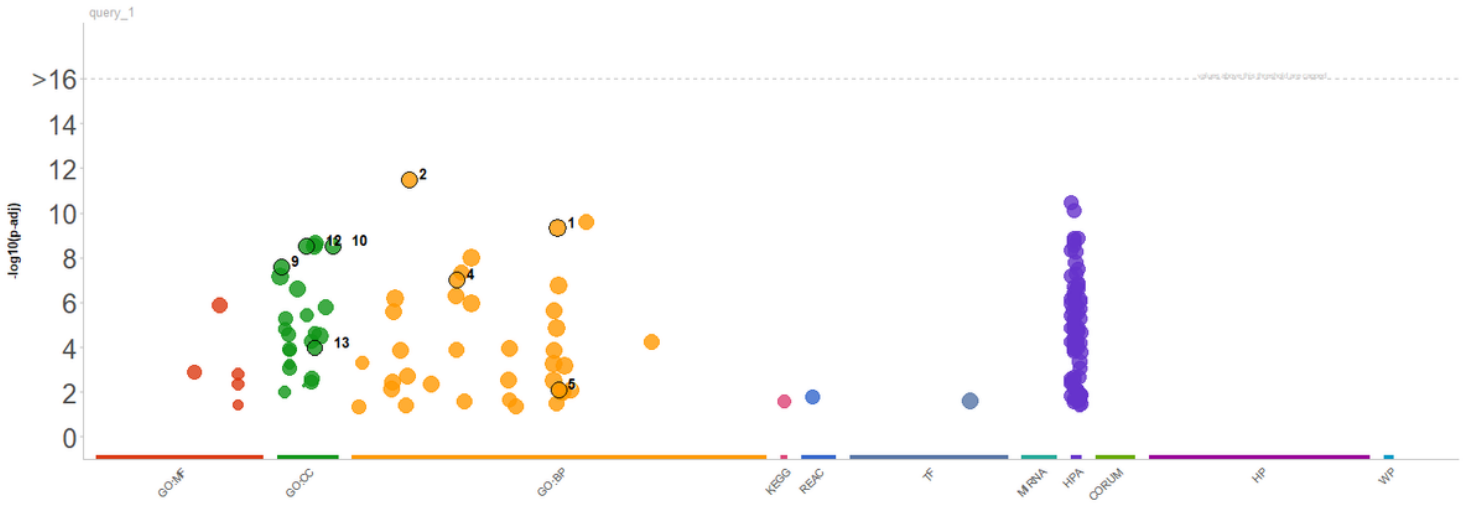


Figure 16

Based on the GSEA data analysis on the GSE42568 dataset, the down-regulated genes in the mentioned BC dataset have a significant role in the PPAR signaling pathway.

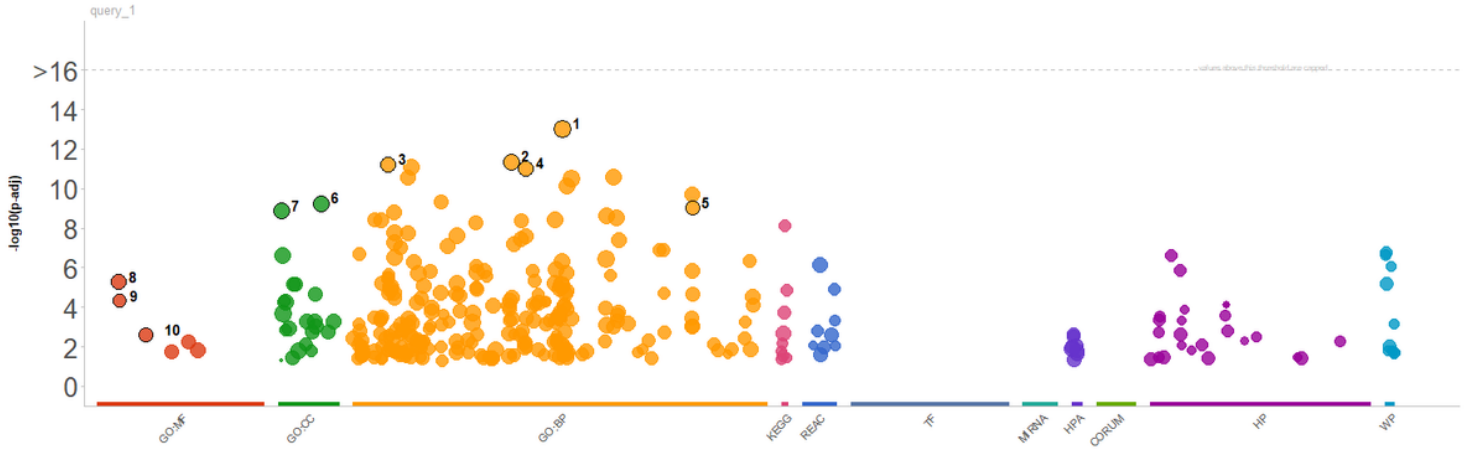


id	source	term_id	term_name	term_size	p_value
1	GO:BP	GO:0048856	anatomical structure development	5718	4.7e-10
2	GO:BP	GO:0009888	tissue development	1912	3.3e-12
4	GO:BP	GO:0030154	cell differentiation	4128	9.3e-08
5	GO:BP	GO:0048870	cell motility	1645	7.7e-03
9	GO:CC	GO:0005615	extracellular space	3353	2.6e-08
10	GO:CC	GO:1903561	extracellular vesicle	2132	2.9e-09
12	GO:CC	GO:0043230	extracellular organelle	2133	2.9e-09
13	GO:CC	GO:0070161	anchoring junction	839	1.1e-04

g:Profiler (biit.cs.ut.ee/gprofiler)

Figure 17

Gene ontology of the up-regulated genes in the GSE42568 BC dataset.



id	source	term_id	term_name	term_size	p_value
1	GO:BP	GO:0050896	response to stimulus	8815	9.6e-14
2	GO:BP	GO:0042221	response to chemical	4316	4.4e-12
3	GO:BP	GO:0006629	lipid metabolic process	1345	6.1e-12
4	GO:BP	GO:0044281	small molecule metabolic process	1779	1.0e-11
5	GO:BP	GO:1901652	response to peptide	478	9.8e-10
6	GO:CC	GO:0071944	cell periphery	6181	5.8e-10
7	GO:CC	GO:0005576	extracellular region	4303	1.4e-09
8	GO:MF	GO:0005102	signaling receptor binding	1546	5.5e-06
9	GO:MF	GO:0005178	integrin binding	147	5.1e-05
10	GO:MF	GO:0016491	oxidoreductase activity	756	2.5e-03

g:Profiler (biit.cs.ut.ee/gprofiler)

Figure 18

Gene ontology of the low expressed genes in the GSE42568 microarray dataset.

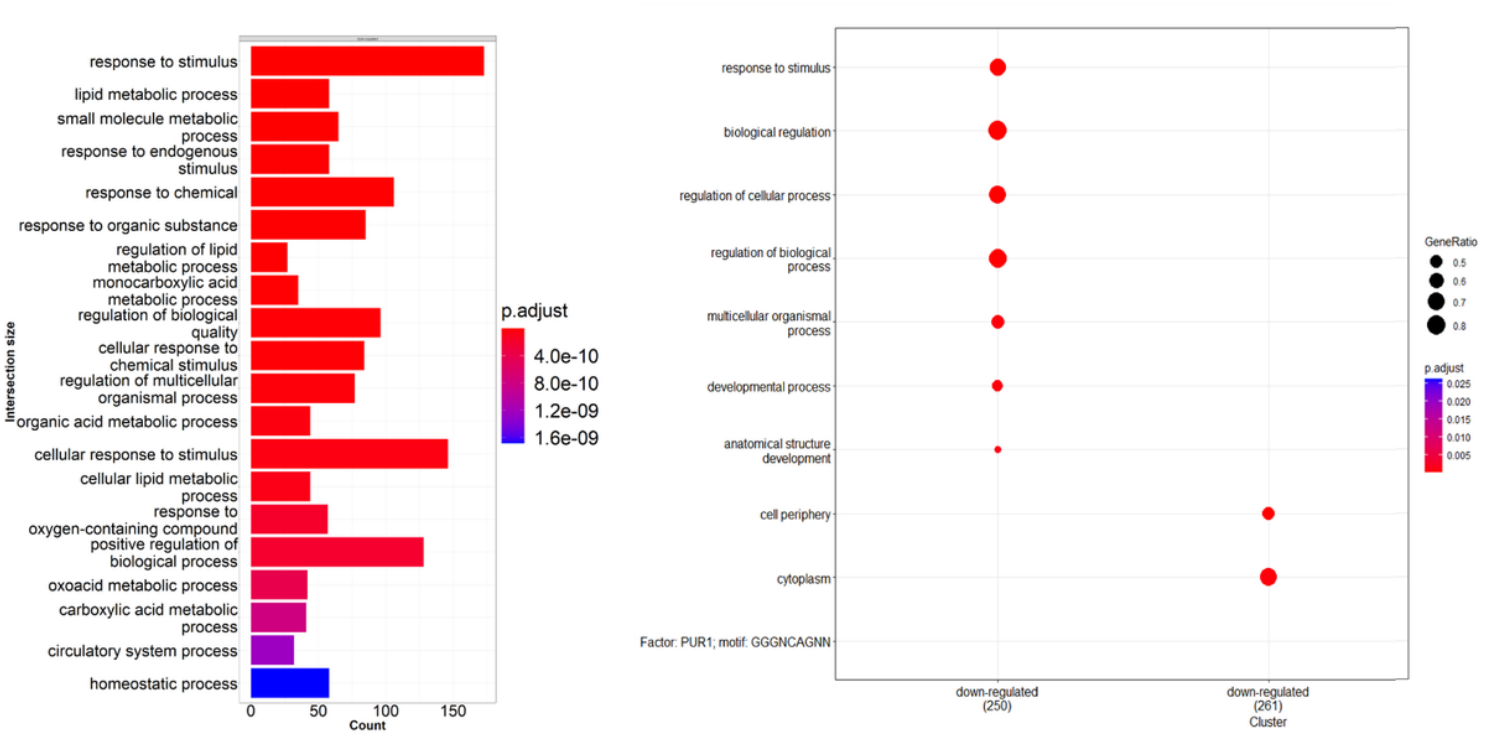


Figure 19

Gene ontology bar plot and dot plot of the down-regulated genes in the GSE42568 microarray dataset. Red color indicates more significant gene ontology terms. Size of the dots in the right chart indicates the gene ratio (number of genes involve in the relevant process or component).