



This is a repository copy of *Bioinformatics analysis of the potentially functional circRNA-miRNA-mRNA network in breast cancer*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/211659/>

Version: Published Version

---

**Article:**

Erdogan, C. [orcid.org/0000-0001-5495-7754](https://orcid.org/0000-0001-5495-7754), Suer, I. [orcid.org/0000-0003-1954-4190](https://orcid.org/0000-0003-1954-4190), Kaya, M. et al. (3 more authors) (2024) Bioinformatics analysis of the potentially functional circRNA-miRNA-mRNA network in breast cancer. PLOS ONE, 19 (4). e0301995. ISSN 1932-6203

<https://doi.org/10.1371/journal.pone.0301995>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

## RESEARCH ARTICLE

# Bioinformatics analysis of the potentially functional circRNA-miRNA-mRNA network in breast cancer

Cihat Erdogan<sup>1</sup>, Ilknur Suer<sup>2,3</sup>, Murat Kaya<sup>3</sup>, Sukru Ozturk<sup>3</sup>, Nizamettin Aydin<sup>4</sup>, Zeyneb Kurt<sup>5\*</sup>

**1** Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana, United States of America, **2** Department of Medical Genetics, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey, **3** Department of Internal Medicine, Division of Medical Genetics, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey, **4** Department of Computer Engineering, Faculty of Computer and Informatics, Istanbul Technical University, Istanbul, Turkey, **5** Information School, The University of Sheffield, Sheffield, United Kingdom

☞ These authors contributed equally to this work.

\* [z.kurt@sheffield.ac.uk](mailto:z.kurt@sheffield.ac.uk)



## OPEN ACCESS

**Citation:** Erdogan C, Suer I, Kaya M, Ozturk S, Aydin N, Kurt Z (2024) Bioinformatics analysis of the potentially functional circRNA-miRNA-mRNA network in breast cancer. PLoS ONE 19(4): e0301995. <https://doi.org/10.1371/journal.pone.0301995>

**Editor:** Jianhong Zhou, PLOS: Public Library of Science, UNITED STATES

**Received:** February 6, 2023

**Accepted:** March 25, 2024

**Published:** April 18, 2024

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0301995>

**Copyright:** © 2024 Erdogan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All circRNA expression profiles are available from the Gene Expression Omnibus (GEO) database (accession numbers GSE101124 and GSE182471). All miRNA

## Abstract

Breast cancer (BC) is the most common cancer among women with high morbidity and mortality. Therefore, new research is still needed for biomarker detection. GSE101124 and GSE182471 datasets were obtained from the Gene Expression Omnibus (GEO) database to evaluate differentially expressed circular RNAs (circRNAs). The Cancer Genome Atlas (TCGA) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) databases were used to identify the significantly dysregulated microRNAs (miRNAs) and genes considering the Prediction Analysis of Microarray classification (PAM50). The circRNA-miRNA-mRNA relationship was investigated using the Cancer-Specific CircRNA, miRDB, miRTarBase, and miRWalk databases. The circRNA-miRNA-mRNA regulatory network was annotated using Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database. The protein-protein interaction network was constructed by the STRING database and visualized by the Cytoscape tool. Then, raw miRNA data and genes were filtered using some selection criteria according to a specific expression level in PAM50 subgroups. A bottleneck method was utilized to obtain highly interacted hub genes using cytoHubba Cytoscape plugin. The Disease-Free Survival and Overall Survival analysis were performed for these hub genes, which are detected within the miRNA and circRNA axis in our study. We identified three circRNAs, three miRNAs, and eighteen candidate target genes that may play an important role in BC. In addition, it has been determined that these molecules can be useful in the classification of BC, especially in determining the basal-like breast cancer (BLBC) subtype. We conclude that hsa\_circ\_0000515/miR-486-5p/SDC1 axis may be an important biomarker candidate in distinguishing patients in the BLBC subgroup of BC.

and mRNA expression data as well as meta-data of the samples are available on the TCGA database (the breast cancer cohort on TCGA).

**Funding:** The author(s) received no specific funding for this work.

**Competing interests:** The authors have declared that no competing interests exist.

## 1 Introduction

Breast cancer (BC) is a heterogeneous type of malignancy that occurs as a result of distinct molecular alterations in breast tissue [1]. Circular RNAs (circRNAs) are evolutionarily conserved and stable RNA regulators that can behave as microRNA (miRNA) sponges, regulate alternative splicing mechanisms, and take an active role in the expression of the gene in which they are encoded [2]. circRNAs have been shown to play crucial roles in the cell, and in recent years this RNA class has been one of the most important research focuses, particularly in the field of cancer [3]. It has been noted that biological processes in living cells would better be modeled by networks since molecular phenotypes do not operate in isolation, instead their interactions collectively carry out these processes [4]. Hence, a network representation can provide a better understanding of the biological and molecular processes beyond analyzing a single molecule or gene, for instance identified from differentially expressed gene analyses. It is expected that identifying circRNA-miRNA-mRNA connections will be essential in explaining the molecular processes of numerous illnesses, detecting biomarkers for early diagnosis, and expanding therapy choices. A single miRNA has the capacity to target hundreds of genes, while a single circRNA can serve as a sponge for dozens of miRNAs. Using bioinformatics data to simplify the circRNA-miRNA-mRNA interactions, which are comprised of such complicated processes, can shed light on in vitro and in vivo investigations. For example, in the bioinformatics study of Liu et al [5], it was emphasized that *hsa\_circRNA\_0003638* may play a role in the pathogenesis of atrial fibrillation by targeting the *CXCR4* gene via *hsa-miR-1207-3p*. Similarly, Hu et al [6] suggested that the interaction of *hsa\_circ\_0009581/hsa-miR-150-5p*, and *hsa\_circ\_0001947/hsa-miR-454-3p* may play a role in the AML cancer process. The precise biological classification of the BC subtype is critical for predicting the disease's progression. Clinical management of BC is dependent on criteria such as tumor size, age, Estrogen (ER) and Progesterone (PR) expression, and the presence or absence of amplification and concurrent enhanced Human epidermal growth factor receptor 2 (HER2) expression. However, these indicators are currently insufficient for accurately categorizing individuals into sections with a high or low risk of relapse, as well as identifying subgroups resistant to therapy [7]. Technological breakthroughs in recent decades have enabled molecular classification based on distinct global gene expression. mRNA expression patterns assessed using microarrays revealed that BC had distinct intrinsic fingerprints that may be utilized to classify tumors into intrinsic molecular subgroups [8–11]. Despite considerable advances in this field, there is still a need for novel markers to refine categorization, particularly for some subtypes [7]. Studies on circRNA, which is a relatively new field of research area, and its relationship with BC subtypes are still quite insufficient. The identification of new genes with variable expression across different types of BC, as well as the detection of miRNAs and circRNAs associated with these genes, may be critical for cancer categorization and potential treatment. Therefore, in our study we demonstrated the circRNA-miRNA-mRNA regulatory connections in BC subtypes using various databases (S1 Table). The circRNAs were detected using the GSE101124 [12] and GSE182471 [13] datasets. miRNAs with significantly altered expression in Prediction Analysis of Microarray (PAM50) subtypes were identified using The Cancer Genome Atlas (TCGA) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) datasets. The TCGA dataset was also used to identify genes with dramatically changed expressions in BC. The circRNA-miRNA-mRNA relationship was investigated for each PAM50 subtype using the Cancer-Specific CircRNA (CSCD) [14], *miRDB* [15], *miRWalk* [16], and *miRTarBase* [17] databases. Previously, overall BC-associated circRNA-miRNA-mRNA interactions have been found out [18] but this was not investigated for different PAM50 subtypes individually. Also, solely a single dataset for each molecular data type has been used previously.

We identified those interactions for each PAM50 subtype using multiple datasets per data type. Although the PAM50 subtyping was available for the miRNA and gene expression datasets, this information does not exist for the circRNA datasets. We have initiated our examination from the upper stream of the candidate circRNA-miRNA-mRNA axes and the PAM50 subtype-associated shortlisted molecules were mapped to the downstream until the subtype information was not available any further. We have proceeded our further investigation and in-silico confirmation for the differentially regulated miRNAs and genes that are shared across all PAM50 subtypes with a further emphasize on implications of our findings on the basal-like breast cancer (BLBC) molecular subtype since it is the most aggressive one among all with a higher recurrence rate and poorer outcome for 5-year survival [19].

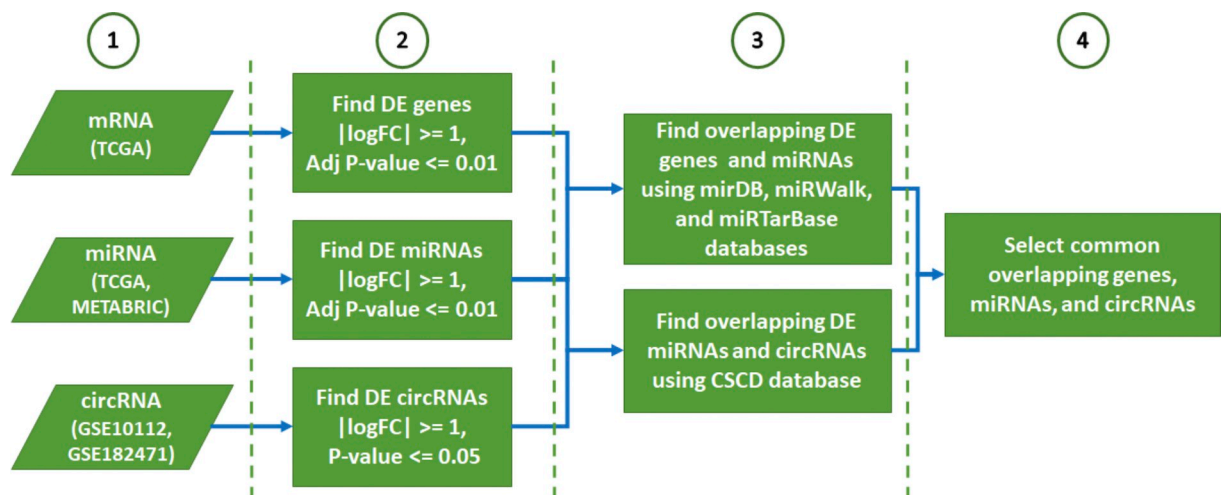
## 2 Materials and methods

### 2.1 Differentially expression analysis of BC datasets

The circRNA, miRNA and mRNA expressions were analyzed using various databases and datasets. The block diagram of our pipeline is illustrated in Fig 1.

**2.1.1 circRNA expression.** The circRNA expression profiles were gathered from the Gene Expression Omnibus (GEO) database with an access code of GSE101124 [12] (the dataset includes four BC cell samples, four triple negative breast cancer (TNBC) and four Luminal-A (LumA) molecular subtype tissue samples and three non-tumorous mammary gland tissue samples) and GSE182471 [13] (the dataset includes five BC samples and five non-tumor samples). Four cancer cell lines in the GSE101124 dataset have been removed and the eight tumorous and three non-tumorous tissue samples have been kept. Molecular subtypes of the samples in the dataset GSE182471 were not annotated. Then, the differentially expressed circRNAs (DECs) were identified by using the *limma* R package (v.3.46.0) with a p-value less than 0.05 and an absolute log<sub>2</sub>-transformed FC (fold change) value of  $\geq 1$ .

**2.1.2 miRNA and mRNA expression.** The miRNA and mRNA expression data as well as meta-data of the samples were downloaded from the TCGA database [20]. The TCGA dataset contains alterations in miRNA and mRNA expressions from 901 BC samples (162 basal-like, 73 HER2-positive, 455 Luminal A, 178 Luminal B, 33 normal-like) and 112 control samples.



**Fig 1. The steps of the BC PAM50 subtype analysis.** The miRNA-mRNA interactions were estimated with *mirDB* (v6), *miRTarBase* (Release 8.0), and *miRWalk* (v3). DE: Differentially expressed, *CSCD* (v2.0): The Cancer-specific circRNAs database, BC: Breast cancer, FC: Fold change, log: logarithm base 2.

<https://doi.org/10.1371/journal.pone.0301995.g001>

Also, the METABRIC (University of Cambridge) dataset [21] was obtained from the European Genome-Phenome Archive (EGA-S00000000122) to validate our findings. The METABRIC dataset includes 1,301 BC samples (198 basal-like, 161 HER2-positive, 461 Luminal A, 370 Luminal B, 99 normal-like, and 12 unknown tumors) and 116 control samples of the miRNA expression analyses. We used the *DESeq2* R package (v.1.28.1) to determine differentially expressed mRNAs (DEGs) and miRNAs (DEMs) with a set of criteria, False Discovery Rate (FDR) < 0.01 and an absolute log<sub>2</sub>FC value of  $\geq 1$  (for both mRNAs and miRNAs), for each group of the PAM50 classification across the TCGA samples in comparison to the healthy tissue samples. Since there are no control samples in the mRNA expression dataset from the METABRIC database, we conducted the DEG analysis only on the TCGA.

Since the raw data was not provided, but only the normalized data were available in the METABRIC database, the DEMs were determined with the same criteria given above by using the *limma* R package. Overlapping mRNA and miRNAs from differential expression analyses were curated for further analysis. All FDR values were obtained by the Benjamini-Hochberg method.

## 2.2 Predicting the associated biological features

In order to predict circRNAs and miRNAs interactions, the most significantly altered 13 DEGs (both down- and up-regulated) were chosen via the *CSCD* v2.0 database. On the other hand, the *miRDB*, *miRWalk*, and *miRTarBase* databases were used to find the interactions between the DEGs and DEMs across all of the PAM50 classes. Hence, the knowledge base-driven target mRNAs of the up-regulated miRNAs in all PAM50 classes were searched among the down-regulated mRNAs, whereas the down-regulated miRNAs' targets were searched among the up-regulated mRNAs. Similarly, knowledge base-driven target miRNAs of the up-regulated circRNAs were searched among the down-regulated miRNAs, whereas the targets of the down-regulated circRNAs were searched among the up-regulated miRNAs. After that, the interactions of circRNAs, miRNAs and mRNAs, which were found to have the most significant expression change, were investigated from the literature.

## 2.3 Selection criteria for filtered candidate miRNAs and mRNAs

The selected miRNAs and mRNAs should have a strong association with both BC and other cancers in the literature (keywords such as "gene name, miRNA name, cancer, breast cancer, breast" were searched in the PubMed database),

A distinct altered expression level of these mRNAs should be detected across all of the PAM50 subgroups from normal-like to the BLBC with an increasing trend in the form of a pan flute. Genes that do not meet this criterion will be eliminated and genes with the top significant log fold change (log<sub>2</sub>FC) values will be kept,

The selected genes should be associated with poor OS and DFS in BC.

## 2.4 Survival analysis

The survival data of the TCGA dataset was obtained from the Pan-Cancer Clinical publication [22]. Survival curves were obtained according to the Kaplan-Meier method (*surv\_plot* function) from the *finalfit* (v1.0.3) R package [23], and differences between survival distributions were assessed by log-rank test. The patients were divided into two risk groups as high and low according to their normalized median expression values. The normalized expression values were obtained using *voom* function from *limma* (v3.46.0) R package [24]. For analysis of relationships between the selected gene and BC, univariate models were fitted using cox proportional hazard regression (*coxph* function) from the *survival* R package [25]. Furthermore, we



used the GSE25066 dataset from GEO to validate our survival analysis findings [26]. The GSE25066 dataset contains 508 BC (189 BLBC, 37 HER2-positive, 160 LumA, 78 LumB, 44 normal-like) mRNA samples.

## 2.5 The relation of circRNA, miRNA and mRNA

The circRNA-miRNA-mRNA regulatory network was established using the *Cytoscape* tool (v3.9.0) [27] based on the interactions between circRNA, miRNA, and mRNA obtained from *CSCD*, *mirDB*, *miRTarBase*, and *miRWalk* databases. After finding the DECs shared between the two circRNA datasets and DEMs shared between METABRIC and TCGA, we searched for the presence of the circRNA-miRNA pairs from the *CSCD* v2.0 database. We also checked the direction of the regulation of the DE miRNAs and DE circRNAs, since they are expected to be inversely related. This election process has shortlisted the number of DECs and their matching DEMs as described in the Results. Similarly, the presence of the matching miRNA-mRNA pairs was searched from the *mirDB*, *miRTarBase*, and *miRWalk* databases and this process has led to narrowing down the DEG list.

## 2.6 Correlation analysis between the selected miRNA and mRNAs

Spearman correlation was used to measure the correlation between the selected miRNA and mRNA expressions in the TCGA dataset. The *corrplot* function from the *corrplot* (v 0.92) [28] R package was used to visualize the correlation heatmap.

## 2.7 Analysis of the protein-protein interaction (PPI) network

The PPI network was created by using the *STRING 2021* [29] database with a minimum required interaction score of  $> 0.4$  and visualized by the *Cytoscape* tool (v3.9.0).

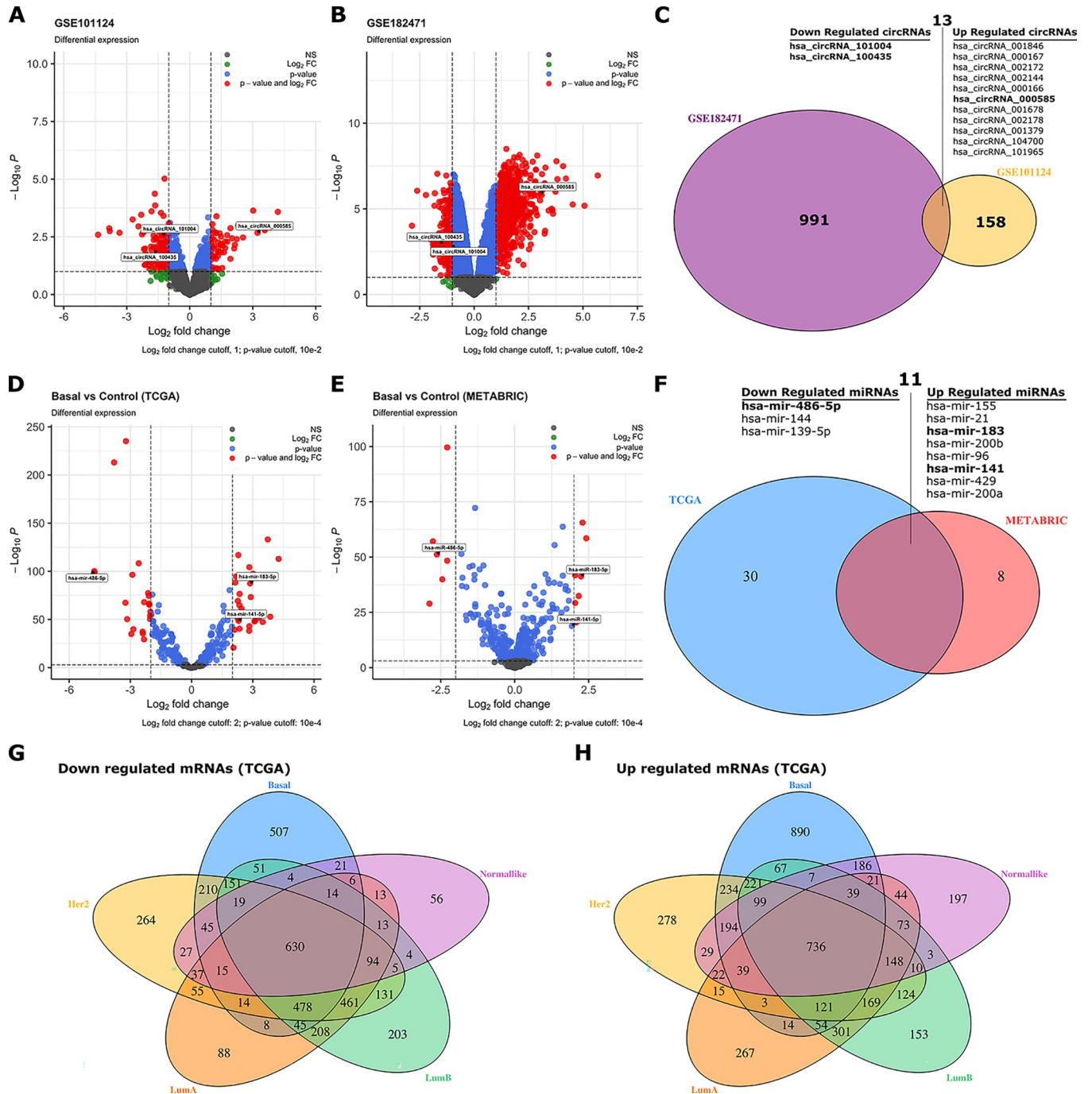
## 2.8 Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses

The gene set enrichment analyses were obtained by using *Enrichr* [30] web tool with the criterion of FDR value lower than 0.05 according to *GO* annotation and *KEGG* Pathway. *Enrichr* is a gene list enrichment analysis tool that is frequently used in the literature and allows querying on hundreds of gene sets such as *KEGG*, *GO*, *Reactome* [31], and *DisGeNet* [32]. The p-value, provided by *Enrichr*, as a result of the enrichment analysis is determined by Fisher's exact test (hypergeometric test), which is a binomial proportionality test that assumes the binomial distribution and independence for the probability of any gene set. Also, the FDR value, provided by *Enrichr*, is calculated using the Benjamini-Hochberg method to adjust the multiple hypothesis testing.

# 3 Results and discussion

## 3.1 Determination of DECs, DEMs and DEGs

**3.1.1 DECs.** We observed that 149 circRNAs (94 of them were up-regulated and 55 were down-regulated) in GSE101124 (Fig 2A) and 993 circRNA (665 of them were up-regulated and 328 were down-regulated) in GSE182471 (Fig 2B) in BC tumor samples were differentially expressed when compared to the control samples. Furthermore, we obtained 13 down- and up-regulated circRNAs in total that are shared between the GSE101124 and GSE182471 datasets. The overlapped 11 up- and two down-regulated circRNAs are listed in Fig 2C. The expression of the overlapped up- and down-regulated circRNAs in each dataset is demonstrated in S1 Fig of S1 File.



**Fig 2. The volcano plot for DECs in BC based on the two microarray datasets from GEO and intersected up- and down regulated circRNAs.** The volcano plot for DEMs in BC based on the two microarray datasets from TCGA and EGA, and intersected up- and down regulated miRNAs. (A): GSE101124, (B): GSE182471, (C): The intersected up- and down-regulated circRNAs between the GSE101124 and the GSE182471 datasets. (D): TCGA, (E): METABRIC, (F): The intersected up- and down-regulated miRNAs from shared miRNAs in the TCGA and the METABRIC datasets, The intersection of the down-regulated mRNAs (G) and the up-regulated mRNAs (H) between PAM50 subtypes, DECs: differentially expressed circRNAs, BC: Breast cancer, DEMs: Differentially expressed miRNAs, EGA: European Genome-phenome Archive, BC: Breast cancer, hsa: Homo-sapiens.

<https://doi.org/10.1371/journal.pone.0301995.g002>

**3.1.2 DEMs.** Regarding the miRNAs, in the TCGA dataset 133 miRNAs (66 of them were up-regulated and 67 were down-regulated) in BLBC samples (Fig 2D, shown as an example subtype), 114 miRNAs (65 up- and 49 down-regulated) in HER2-positive samples, 105 miRNAs (49 up- and 56 down-regulated) in LumA group, 133 miRNAs (69 up- and 64 down-regulated) in LumB group, and 78 miRNAs (41 up- and 37 down-regulated) in normal-like tumor group were differentially expressed, when compared to the control samples.

Furthermore, in the METABRIC dataset 69 miRNAs (34 of them were up-regulated and 35 were down-regulated) in BLBC samples (Fig 2E, shown as an example subtype), 66 miRNAs (31 up- and 35 down-regulated) in HER2-positive samples, 51 miRNAs (28 up- and 23 down-regulated) in LumA group, 70 miRNAs (31 up- and 39 down-regulated) in LumB group, and 31 miRNAs (17 up- and 14 down-regulated) in normal-like tumor group were differentially expressed, when compared to the control samples.

Finally, we obtained the overlapped DEMs between the TCGA and METABRIC datasets and that are shared miRNAs across all PAM50 subtypes. The determined miRNAs (three down- and eight up-regulated miRNAs) are given in Fig 2F.

**3.1.3 DEGs.** We observed that 5,143 genes (2,925 of them were up-regulated and 2,218 were down-regulated) in BLBC samples, 5,078 genes (2,442 up- and 2,636 down-regulated) in HER2-positive samples, 4,245 genes (2,066 up- and 2,179 down-regulated) in LumA group, 4,836 genes (2,325 up- and 2,511 down-regulated) in LumB group, and 2,850 genes (1,847 up- and 1,003 down-regulated) in normal-like tumor group were differentially expressed, when compared to the control samples.

Among the down-regulated genes (Fig 2G), 630 were shared across all five PAM50 classes, whereas 736 up-regulated genes (Fig 2H) were shared across all five PAM50 classes in the TCGA dataset.

## 3.2 Determining the relationship between the detected DECs, DEMs and DEGs

Knowledge-driven investigation between the shared mRNA and miRNAs, across all five PAM50 classes, revealed that 188 up-regulated genes are associated with the 3 down-regulated miRNAs; whereas 317 down-regulated genes are found to be associated with the 8 up-regulated miRNAs, based on the *miRDB*, *miRWalk*, and *miRTarBase* databases.

## 3.3 Identification of the circRNA-miRNA interactions

The overlapped DECs were selected for further analysis. To indicate whether the 13 circRNAs (described in Section 3.1.1) play a significant role in BC, we gathered their potential target miRNAs from the CSCD v2.0 online databases. In the GSE101124 dataset, compared to the normal group, *hsa\_circRNA\_100435* and *hsa\_circRNA\_101004* were down-regulated in the TNBC group, while *hsa\_circRNA\_000585* was up-regulated ( $p < 0.05$ ;  $\log_{2}FC > 1.5$ ). In total, three circRNA-miRNA interactions including three circRNAs (*hsa\_circRNA\_000585*, *hsa\_circRNA\_101004*, and *hsa\_circRNA\_100435*) and three miRNAs (*miR-486-5p*, *miR-141-5p*, and *miR-183-5p*) were identified in the database. *MIENTURNET* [33] was used to investigate the signaling pathways (*KEGG*, *Reactome*, *WikiPathways*, and *Disease Ontology*) in which the three miRNAs may be involved according to *miRTarBase* database. As shown in S2 Fig of S1 File, all three miRNAs were associated with some cancer-related pathways.

The basic features of the three circRNAs are displayed in Table 1. The main structural models of the three circRNAs are given in S3 Fig of S1 File. The unpaired two-samples Wilcoxon test results according to tumor and control samples of the selected three DECs are given in violin plots in Fig 3 for each circRNA dataset separately.



Table 1. The basic features of the three DECs.

circRNA	Alias	circRNA type	Position (HG19)	Position (HG38)	Strand	Regulation	Gene symbol
<i>hsa_circRNA_100435</i>	<i>hsa_circ_0016201</i>	exonic	chr1:205156545 205156934	chr1:205187417 205187806	-	Down	DSTYK
<i>hsa_circRNA_101004</i>	<i>hsa_circ_0000375</i>	exonic	chr12:6657590 6657991	chr12:6548424 6548825	-	Down	IFFO1
<i>hsa_circRNA_000585</i>	<i>hsa_circ_0000515</i>	sense overlapping	chr14:20811305 20811534	chr14:20343146 20343375	-	Up	RPPH1

The unpaired two-samples t-test results according to basal and control samples of the selected three (*miR-141-5p*, *miR-183-5p*, and *miR-486-5p*) DEMs are given as violin plot in Fig 4 for each miRNA dataset separately.

<https://doi.org/10.1371/journal.pone.0301995.t001>

### 3.4 Identification of circRNA-miRNA-mRNA association

We investigated the miRNA and mRNA associations by using the *miRDB* (v6), *miRTarBase* (r8.0), *miRWalk* (v3) databases for the intersected miRNAs and shared DEGs with an absolute log<sub>2</sub>FC values greater and equal than 1. Then, we combined the circRNA-miRNA interactions and miRNA-mRNA interactions to identify the circRNA-miRNA-mRNA associations. Finally, we established a circRNA-miRNA-mRNA network, which ensures a preliminary insight into the links between the three circRNAs (*hsa\_circRNA\_000585*, *hsa\_circRNA\_101004*, and *hsa\_circRNA\_100435*), the three miRNAs (*miR-486-5p*, *miR-141-5p*, and *miR-183-5p*) and the 339 mRNAs. The constructed network can be seen in Fig 5.

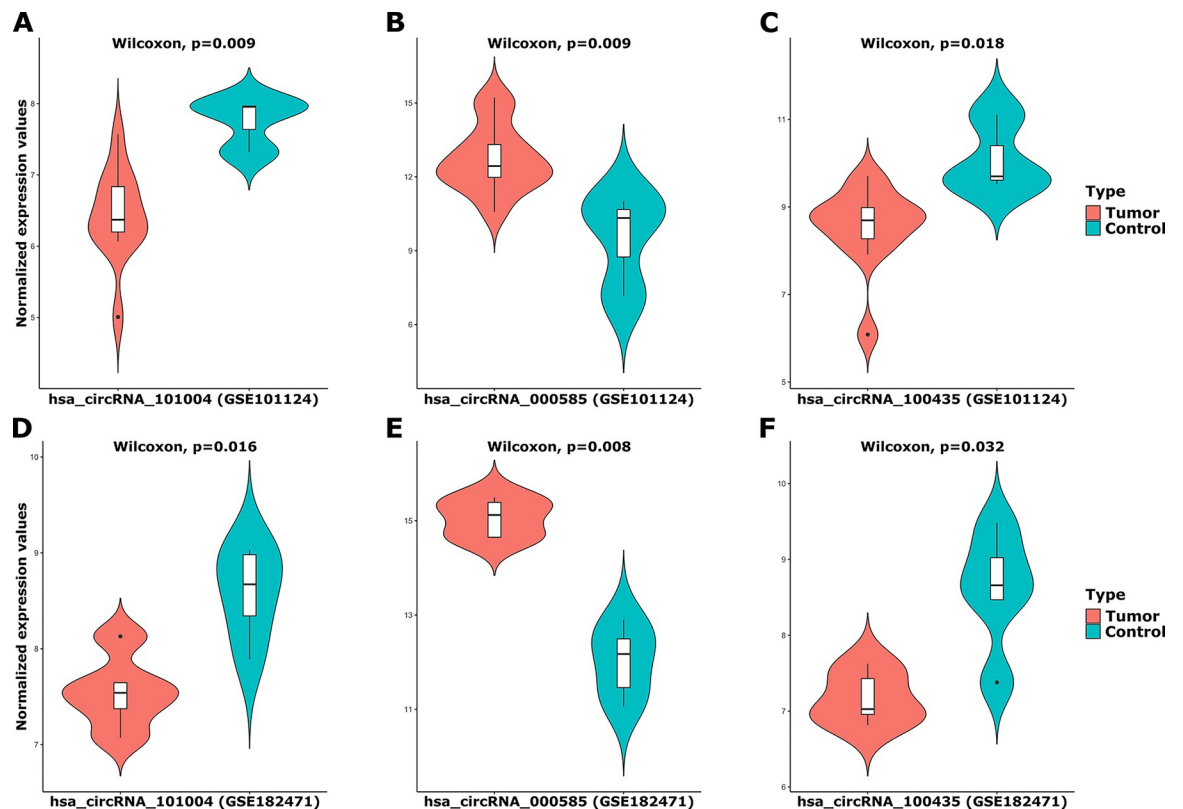
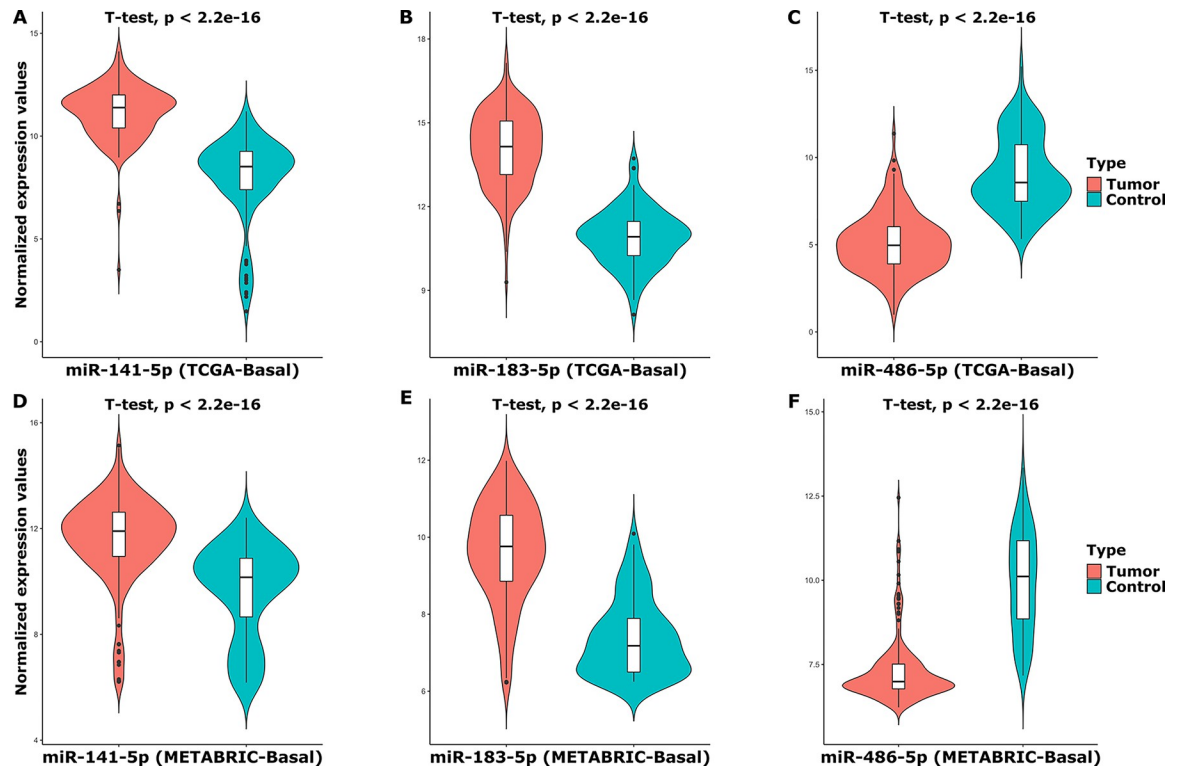


Fig 3. The combined violin and box plots for the normalized expression values of *hsa\_circRNA\_101004*, *hsa\_circRNA\_000585*, *hsa\_circRNA\_100435* in GSE101124 and GSE182471 datasets by the unpaired two-samples Wilcoxon test according to tumor and control samples. (A): *hsa\_circRNA\_101004* in GSE101124, (B): *hsa\_circRNA\_000585* in GSE101124, (C): *hsa\_circRNA\_100435* in GSE101124, (D): *hsa\_circRNA\_101004* in GSE182471, (E): *hsa\_circRNA\_000585* in GSE182471, (F): *hsa\_circRNA\_100435* in GSE182471.

<https://doi.org/10.1371/journal.pone.0301995.g003>



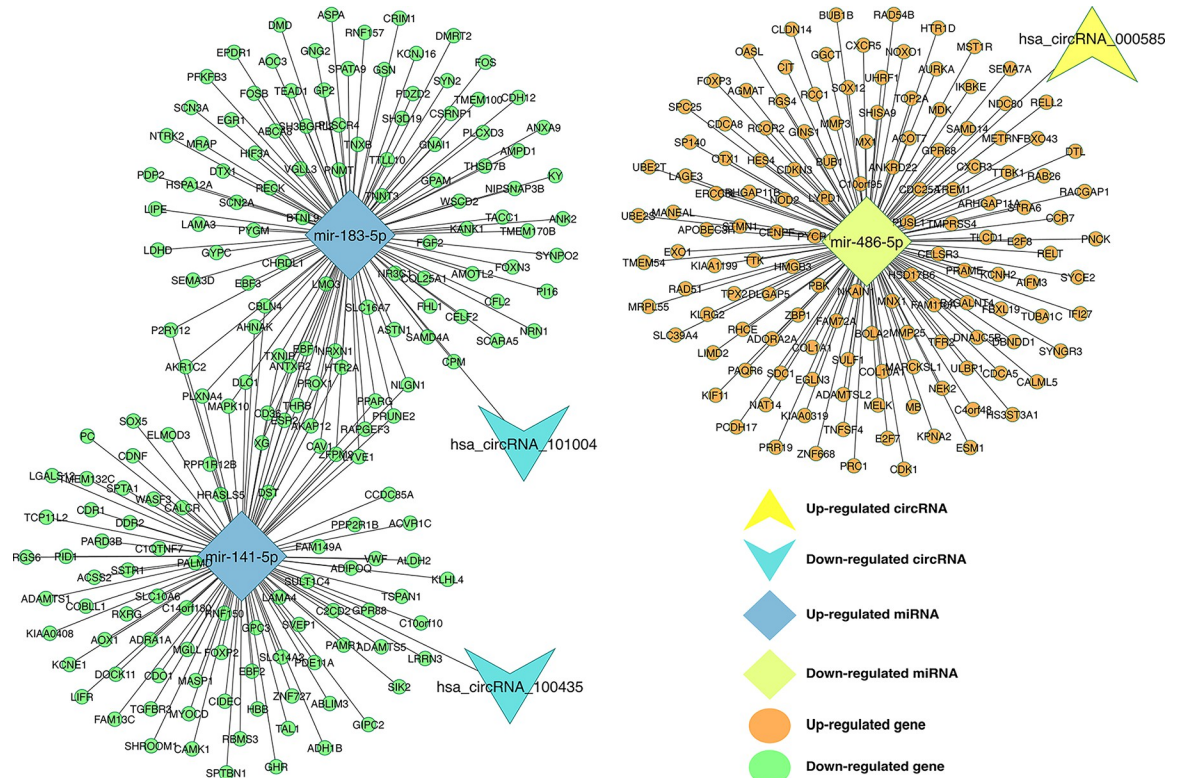
**Fig 4.** The combined violin and box plots for the normalized expression values of *miR-141-5p*, *miR-183-5p*, and *miR-486-5p* in TCGA and METABRIC datasets by the unpaired two-samples t-test according to basal and control samples. (A): *miR-141-5p* in TCGA, (B): *miR-183-5p* in TCGA, (C): *miR-486-5p* in TCGA, (D): *miR-141-5p* in METABRIC, (E): *miR-183-5p* in METABRIC, (F): *miR-486-5p* in METABRIC.

<https://doi.org/10.1371/journal.pone.0301995.g004>

Among the DE 339 mRNAs, we selected the top 10 genes for each one of the three selected miRNAs individually according to the log<sub>2</sub>FC values and that are the target of the shortlisted three miRNAs. After that, we only kept the mRNAs which demonstrate an increasing log<sub>2</sub>FC trend in the form of a pan flute from normal-like to the BLBC (Fig 6). Hence, according to the criteria we defined (criterion I and II), a distinct altered expression level of 18 genes (*SDC1*, *PRAME*, *MELK*, *NEK2*, *EXO1*, *TPX2*, *BUB1*, *DLGAP5*, *CIDEA*, *ADH1B*, *TMEM132C*, *ACVR1C*, *LIPE*, *ABCA8*, *BTNL9*, *TNXB*, *GPAM*, and *AOC3*) were detected in PAM50 sub-groups from the normal-like group to the BLBC group (expression levels shown in S4 Fig of S1 File). We demonstrated the Log<sub>2</sub>FC values of these 18 DEGs in S2 Table and in the form of barcharts in Fig 6. This splits the genes into three groups according to their target miRNAs. BLBC has the greatest Log<sub>2</sub>FC values for 13 out of 18 DEGs (almost all target genes of *miR-486-5p* and *miR-183-5p*) and the bar charts represents almost like a pan flute form, with the Log<sub>2</sub>FC values represent a decreasing trend from BLBC towards HER2, LumB, LumA, and Normal-like subtype. We also showed the correlation patterns between shortlisted three DEMs and their target shortlisted 18 DEGs (Fig 7 represents correlation patterns (A) across all samples and (B) across solely BLBC-subtype samples).

### 3.5 Survival analysis

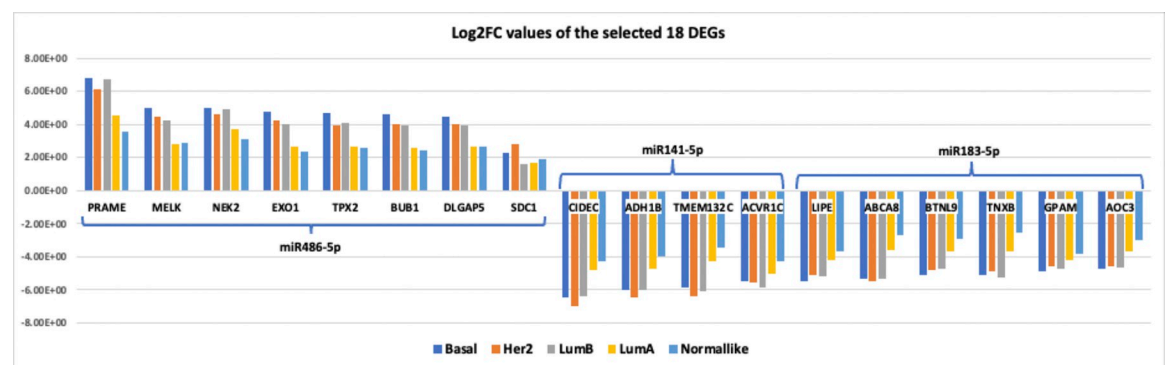
We performed survival analysis on the shortlisted 18 genes (S5 and S6 Figs in S1 File for overall and disease-free survival, respectively, in the TCGA dataset). We observed that only four genes (*SDC1*, *DLGAP5*, *PRAME*, and *EXO1*) demonstrated a significant survival outcome in both



**Fig 5. circRNA-miRNA-mRNA regulatory network.** The network consisting of three circRNAs (*hsa\_circRNA\_000585*, *hsa\_circRNA\_101004*, and *hsa\_circRNA\_100435*), three miRNAs (*miR-486-5p*, *miR-141-5p*, and *miR-183-5p*) and 339 genes was generated by Cytoscape 3.9.0.

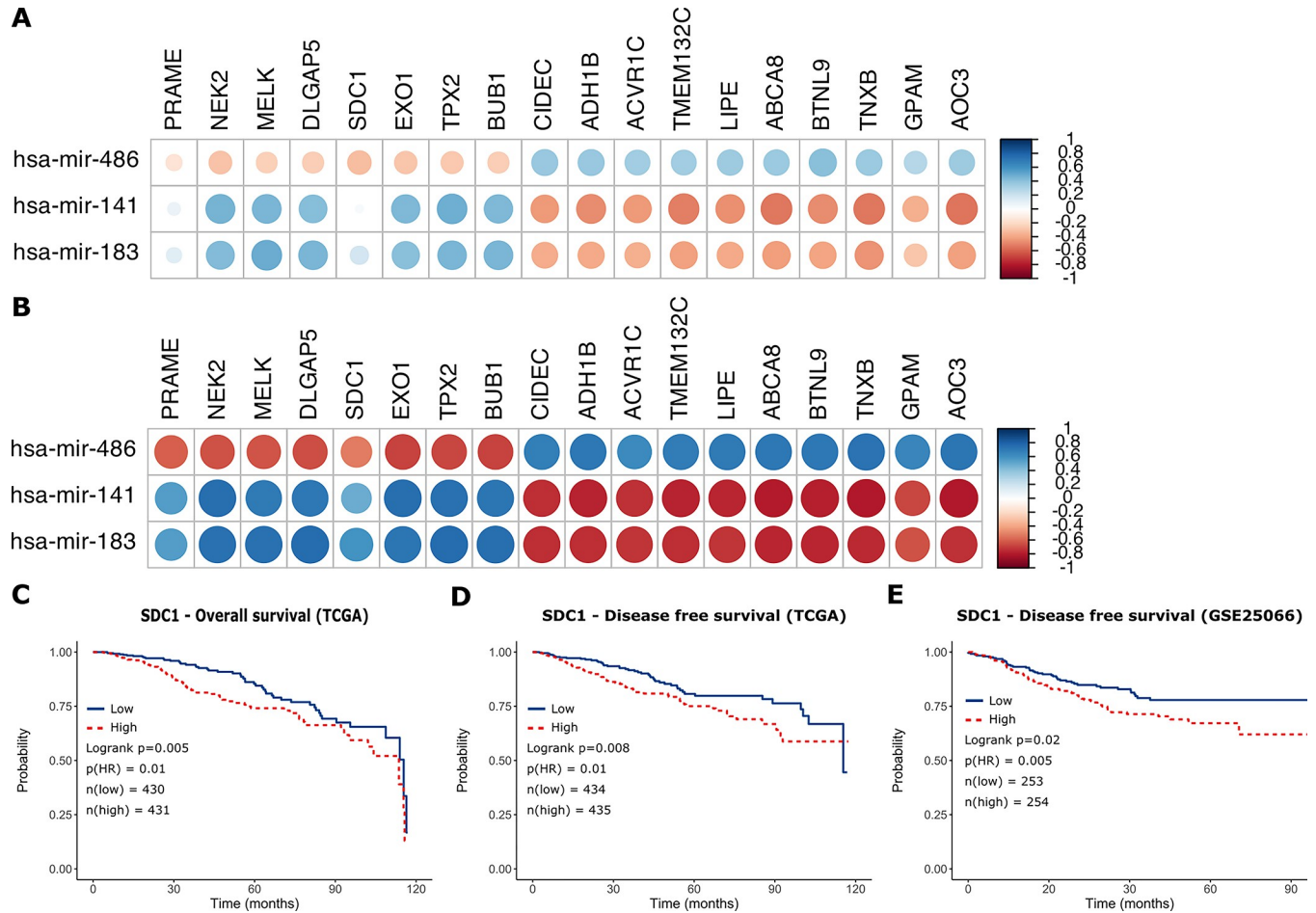
<https://doi.org/10.1371/journal.pone.0301995.g005>

TCGA and GSE25066 datasets. The increased expression of the *SDC1* gene had a poor disease-free survival (DFS) and overall survival (OS) and over-expressions of *DLGAP5*, *PRAME*, and *EXO1* genes had a poor DFS (Fig 7C-7E and S7A-S7C Fig in S1 File) in the TCGA dataset. We also found that highly expressed *SDC1* had a poor DFS, and over-expressions of *DLGAP5*, *PRAME*, and *EXO1* genes had a poor DFS (S7D-S7F Fig in S1 File) in the GSE25066 [26] dataset. The expression distributions of the *SDC1*, *PRAME*, *EXO1*, and *DLGAP5* genes are shown in S8 Fig of S1 File.



**Fig 6. Log2FC values of the selected 18 DEGs.**

<https://doi.org/10.1371/journal.pone.0301995.g006>



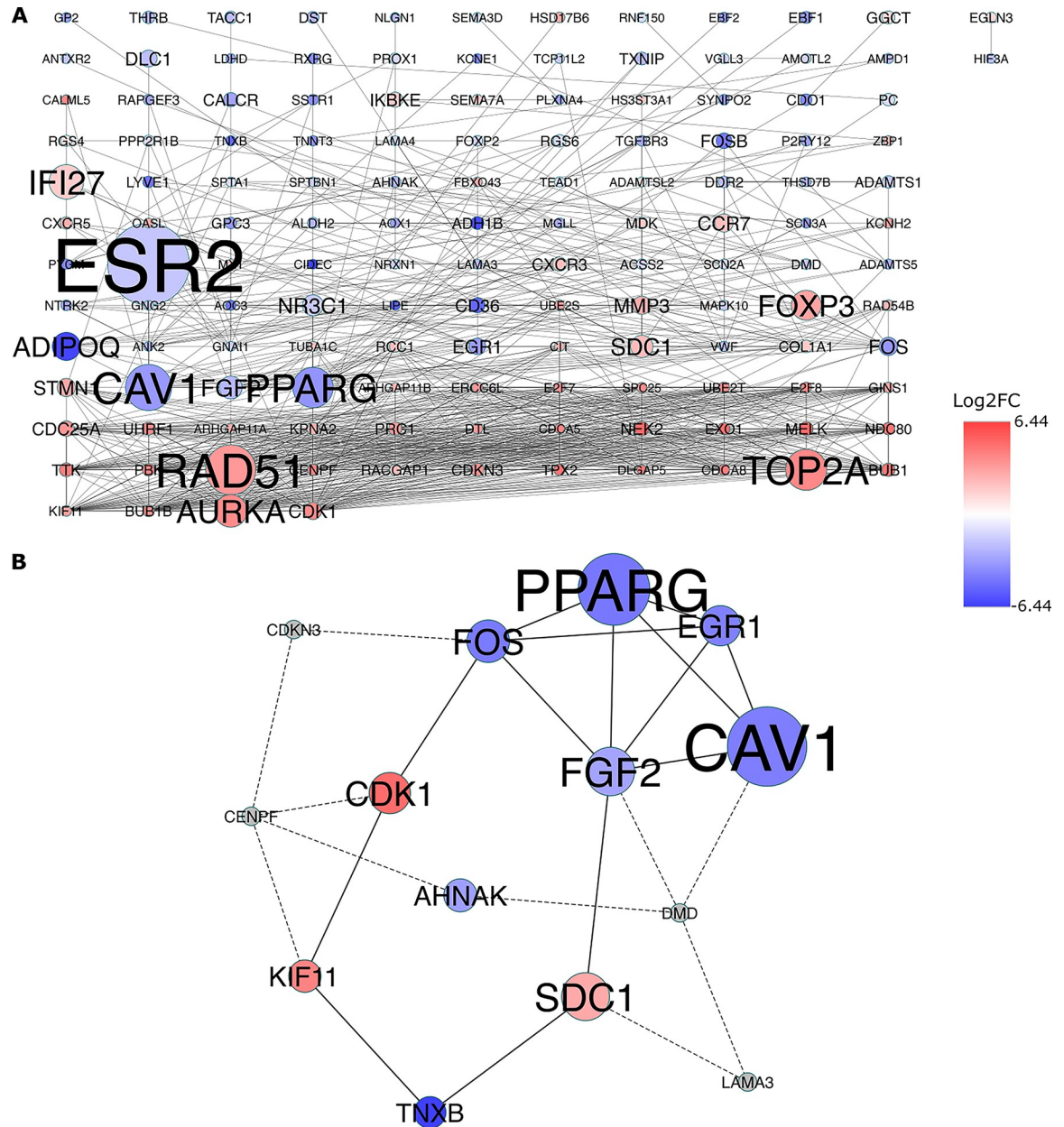
**Fig 7.** The correlation heatmap of the selected mRNA and miRNAs (A):All PAM50 groups, (B): Only BLBC subgroup, Survival analysis of *SDC1*. (C): Overall survival of *SDC1* in TCGA dataset, (D): Disease-free survival of *SDC1* in TCGA dataset, (E): Disease-free survival of *SDC1* in GSE25066 dataset.

<https://doi.org/10.1371/journal.pone.0301995.g007>

### 3.6 Identification of the hub genes with bottleneck algorithm from the PPI network

Using the genes in Fig 5, after removing the isolated nodes, we constructed a PPI network, which consists of 138 nodes and 780 edges, to observe the interactions among the 339 mRNAs (Fig 8A). Considering the importance of the hub gene in a network, we utilized the bottleneck algorithm to screen hub-genes from the PPI network. The subnetwork with 14 nodes (10 hub genes and 4 extended genes) and 24 (14 between hub genes and 10 between extended genes) edges was determined (Fig 8B), which uncover the crucial roles of the ten genes (*AHNAK*, *CAV1*, *CDK1*, *EGR1*, *FGF2*, *FOS*, *KIF11*, *PPARG*, *SDC1*, and *TNXB*) in BC. A circRNA-miRNA-hub gene network was then built to describe the links among the DECs, DEMs and hub genes (Fig 9). Thirteen circRNA-miRNA-mRNA regulatory modules, including *hsa\_circRNA\_100435/ miR-141-5p/ AHNAK* regulatory axis, *hsa\_circRNA\_100435/ miR-141-5p/ PPARG* regulatory axis, *hsa\_circRNA\_100435/ miR-141-5p/ CAV1* regulatory axis, *hsa\_circRNA\_101004/ miR-183-5p/ AHNAK* regulatory axis, *hsa\_circRNA\_101004/ miR-183-5p/ PPARG* regulatory axis, *hsa\_circRNA\_101004/ miR-183-5p/ CAV1* regulatory axis, *hsa\_circRNA\_101004/ miR-183-5p/ FGF2* regulatory axis, *hsa\_circRNA\_101004/ miR-183-5p/*



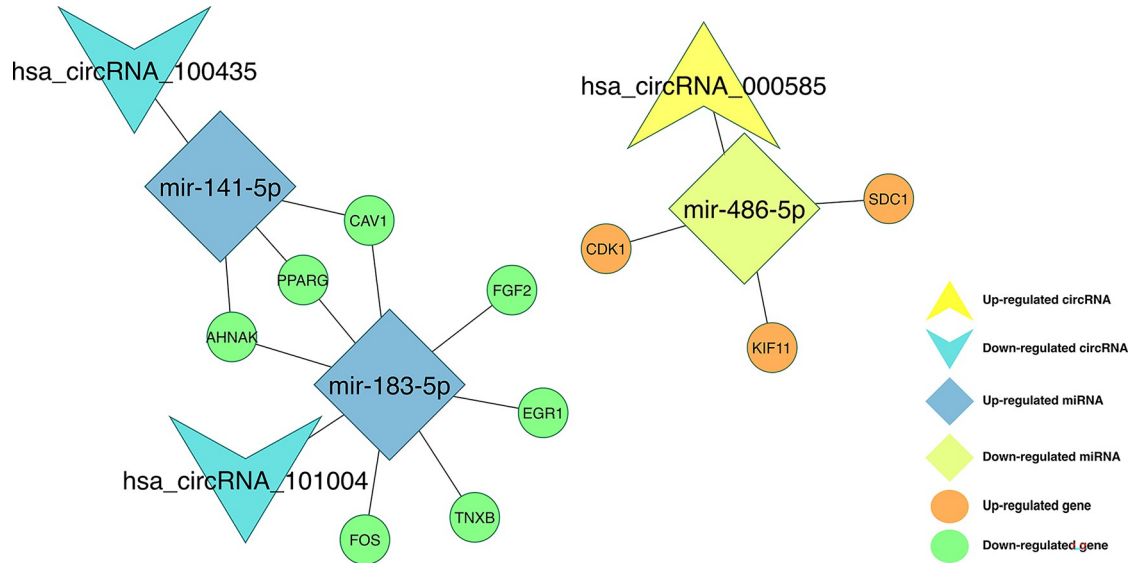


**Fig 8. Identification of hub genes from the PPI network by bottleneck algorithm using the cytoHubba Cytoscape plugin.** The node color changes gradually from blue to red in ascending order according to the log2 (fold change) of genes. (A): A PPI network of the 339 target genes playing crucial roles in BC. This network consists of 138 nodes and 780 edges. The node size changes gradually from small to large in ascending order according to the number of the PMIDs from DisGeNET per gene. (B): A PPI network consist of the ten hub genes (colored blue and red) and 4 extended genes (colored gray) extracted from a. This network consists of 14 (10 hub genes and 4 extended genes) nodes and 24 (14 between hub genes and 10 between extended genes) edges. PPI protein–protein interaction, BC: Breast Cancer.

<https://doi.org/10.1371/journal.pone.0301995.g008>

*EGR1* regulatory axis, *hsa\_circRNA\_101004/ miR-183-5p/ TNXB* regulatory axis, *hsa\_circRNA\_101004/ miR-183-5p/ FOS* regulatory axis, *hsa\_circRNA\_000585/ miR-486-5p/ CDK1* regulatory axis, *hsa\_circRNA\_000585/ miR-486-5p/ KIF11* regulatory axis, and *hsa\_circRNA\_000585/ miR-486-5p/ SDC1* regulatory axis, were found from the network.



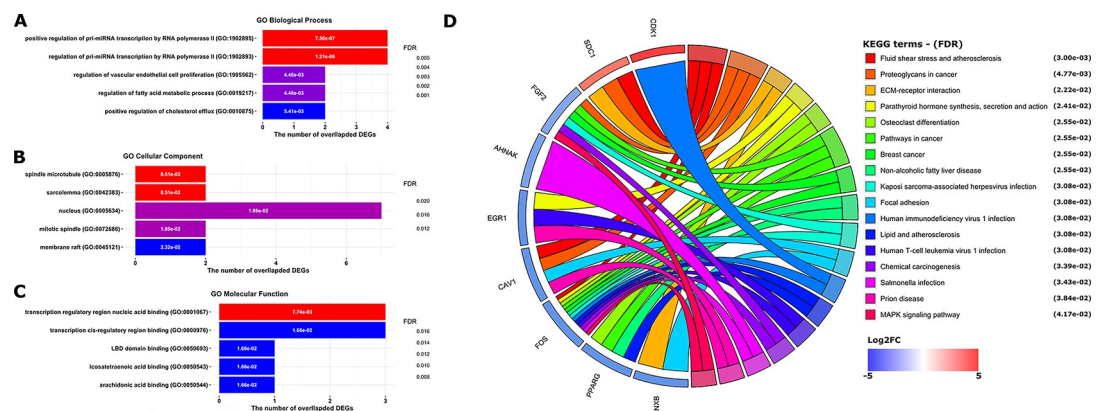


**Fig 9. CircRNA-miRNA-hubgene network.** The network consisting of three circRNAs (*hsa\_circRNA\_000585*, *hsa\_circRNA\_101004*, and *hsa\_circRNA\_100435*), three miRNAs (*miR-486-5p*, *miR-141-5p*, and *miR-183-5p*) and 10 hub genes (*AHNAK*, *CAV1*, *CDK1*, *EGR1*, *FGF2*, *FOS*, *KIF11*, *PPARG*, *SDC1*, and *TNXB*) was generated by Cytoscape 3.9.0.

<https://doi.org/10.1371/journal.pone.0301995.g009>

### 3.7 GO annotation and KEGG pathway analyses of the ten hub genes

GO analysis was performed to demonstrate the functional annotations of the ten hub genes. The top five highly enriched GO terms related to biological process (BP), cellular component (CC) and molecular function (MF) are shown in Fig 10A. The most enriched GO terms in BP was “positive regulation of pri-miRNA transcription by RNA polymerase II (GO:1902895)” (FDR = 7.56E-07), that in CC was “sarcolemma (GO:0042383)” (FDR = 8.51E-03), and that in MF was “transcription regulatory region nucleic acid binding (GO:0001067)” (FDR = 7.74E-03). KEGG pathway analysis was carried out to determine the signaling cascade in which ten genes are involved. With an FDR < 0.05, 17 significantly enriched pathways were determined



**Fig 10. Top five Gene Ontology (GO) enrichment annotations of the ten hub genes:** (A): biological process, (B): cellular component, (C): molecular function. (D)The significantly enriched Kyoto Encyclopedia of hub-genes and genomes (KEGG) pathways with a FDR < 0.05. The results of the GO and the KEGG analyses were obtained from the ‘Enrichr’ web tool (<https://maayanlab.cloud/Enrichr/>) and visualized by R package ‘ggplot2’. Cohort plot shows that the ten hub genes are correlated via ribbons with their assigned KEGG terms. FDR: False discovery rate, is calculated using the Benjamini-Hochberg method to adjust the multiple hypothesis testing.

<https://doi.org/10.1371/journal.pone.0301995.g010>

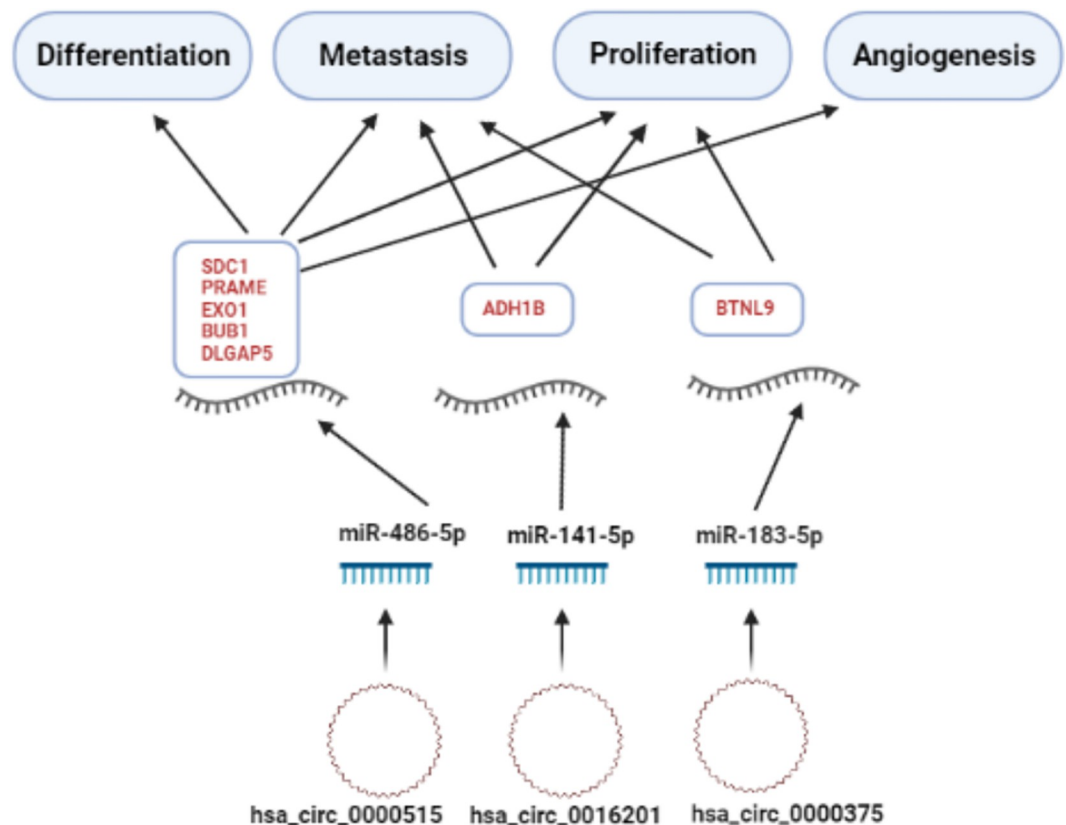
(Fig 10B). Among the 17 pathways, “Proteoglycans in cancer pathway” and “Breast cancer pathway” are linked with the BC progression [34, 35]. In addition, some of the pathways such as “Pathways in cancer”, “Chemical carcinogenesis”, and “Non-alcoholic fatty liver disease” were also tumor-related pathways.

### 3.8 Summary of the filtered circRNA/miRNA/gene axis

The DE circRNAs with Log2FC value  $\geq 1$  were intersected in both datasets, GSE101124 and GSE182471, and three circRNA remained (*hsa\_circ\_0000515*, *hsa\_circ\_0016201*, and *hsa\_circ\_0000375*) according to their existence in the CSCD database. Similarly, overlapping DE miRNAs with Log2FC value  $\geq 2$  in both TCGA and METABRIC, were kept. Among them, three miRNAs (*miR-486-5p*, *miR-141-5p*, and *miR-183-5p*) were presented in the CSCD database as the targets of the three shortlisted circRNAs.

Then, the 18 DE genes with a Log2FC value of  $\geq 2$  which were reported to be strongly associated with BC, and could be targeted by the shortlisted miRNAs, were determined. Those genes were found to be differentially expressed for all PAM50 groups but most significantly in the BLBC subtype (as shown in Fig 6 the logFC values are highest for the BLBC group). Additionally, as shown in Fig 7, correlations between the mRNAs and the target miRNAs are stronger in the BLBC-subtype samples (Fig 7B) compared to the correlation patterns obtained by using all tumor samples together (Fig 7A).

The possible circRNA-miRNA-mRNA interaction, which was detected to play a role in the cellular processes of BC, is shown in Fig 11.



**Fig 11.** The summary of the possible role of circRNA/miRNA/gene axis in BC pathogenesis according to our study.

<https://doi.org/10.1371/journal.pone.0301995.g011>

### 3.9 Discussion

BC is the most frequently identified tumor among women around the world and more than 90% of BC deaths are related to metastasis. Existing treatment approaches for metastatic BC have been inadequate, compounded by a lack of early prognosis/ predictive criteria for estimating which body parts are most susceptible to metastasis. Although there are many new developments in the fields such as chemotherapy, endocrine treatment and targeted therapy for BC in recent years, this cancer type is still the most common cancer in women with high morbidity and mortality [36]. Subtypes in BC are heterogeneous and treatment practices are determined according to these subtypes. From better to the worst, the aggressiveness of the BC subgroups are generally in the following order: Normal breast-like, LumA, LumB, HER2-positive, BLBC [37, 38]. It is clearly known that the OS of cases with HER2-positive and BLBC groups are worst in PAM50 subtypes [39]. The cells are "basal-like," which implies they match the basal cells which line the breast ducts. It is strongly associated with TNBC appearance described by the deficiency of expression of ER, PR, and HER2-positive. BLBC, which is more associated with distant metastasis, has an extremely poor prognosis compared to other intrinsic BC groups, and the success in its treatment is currently limited [40, 41]. This knowledge has substantially advanced our understanding of BC's heterogeneity and the several biological processes that the disease employs. In 2009, Parker et al. defined a minimum gene set, PAM50, for categorizing these intrinsic subgroups [42]. Because the biology of all these intrinsic subgroups indicates changes in incidence, responsiveness to therapy and survival, unique genes for each subtype may be evaluated as markers to direct potential treatments. In this respect, it is crucial to elucidate novel circRNA-miRNA-mRNA relationships in the determination of these subgroups [43–45]. In our study, the expression states of miRNAs and genes in datasets were classified according to the five molecular subtypes classification.

circRNAs, which are a new class of endogenous evolutionarily conserved RNAs, have a stable structure and they are stated to serve as vital regulators in the various cellular activities. According to studies conducted so far, it has been understood that the main reason for circRNAs to act as critical regulators in cells is their relationship with target miRNAs [46]. In recent years, it has been suggested that miRNAs act as a bridge in the realization of the role of circRNAs in the regularization of cellular events [47]. circRNAs change gene expression by acting as miRNA sponges with their binding sites [3, 48]. As increased expression rates of circRNAs in the cell may contribute to decreased expressions of target miRNAs and the increased expressions of target genes. circRNA-miRNA-mRNA interactions, which are the focus of this work, are very new to the scientific world but experiments have shown that these relationships could be beneficial for the detection of novel biomarkers for cancer [49, 50]. The studies on circRNAs about the determination of subtypes of BC are limited. The study by Nair et al. in 2016 is one of the first studies showing circRNAs may be useful in identifying subtypes of BC [51]. In the study of Darbeheshti et al. in 40 TNBC, 20 Lum A, 18 Lum B and 17 HER2-positive tumor samples, it was determined that *hsa\_circ\_0044234* has a distinct molecular signature as a potential *GATA3* regulator in TNBC [52]. In another study, *circ-PGAP3* was shown to increase TNBC proliferation and invasion via *miR-330-3p/Myc* axis [53]. Sheng et al have also found out overall BC-associated circRNA-miRNA-mRNA interactions but they have not investigated those interactions for PAM50 [18] subtypes which is different from our study design.

As a result of our study, many circRNAs, miRNAs and genes that may be associated with BC have been identified. We found that three circRNAs (*hsa\_circ\_0016201*, *hsa\_circ\_0000375*, and *hsa\_circ\_0000515*), three miRNAs (*mir-183-5p*, *miR-141-5p*, and *miR-486-5p*) and 18 genes (*CIDEA*, *ADH1B*, *TMEM132C*, *ACVR1C*, *LIPE*, *ABCA8*, *BTNL9*, *TNXOCB3*, *GPAM*,

*PRAME*, *MELK*, *NEK2*, *EXO1*, *TPX2*, *BUB1*, *DLGAP5*, and *SDC1*) may be important in BC, especially in a basal-like group, by applying filters as described in the material method section. Analysis of TNBC versus normal tissue samples on the GSE101124 dataset results may indicate that these three circRNAs highlighted in the study may play an important role in BLBC. Moreover, although the expression alteration of genes, targeted by miRNAs that sponged via selected circRNAs, was more prominent in the BLBC and Her2-positive subgroups, it was much more limited in the normal-like and Lum-A groups (an expression alteration was detected as similar to a pan flute, Fig 6).

Expressions of *miR-141-5p* and *mir-183-5p*, which are known to be dysregulated in many cancers including BC (46–50), were found to be significantly increased in our study in all dataset samples from all PAM50 groups. Possible target genes that may contribute to the cancer progression and in which *miR-141-5p* and *mir-183-5p* could alter their expression in BC are shown in S1 Table. According to the criteria we determined, the possible paired targets of *miR-141-5p/ADH1B* and *mir-183-5p/BTLN9*, may be related to the BC process. In the dataset we examined, it was identified that *hsa\_circ\_0016201*, which is among the circRNAs whose expression was significantly decreased, could have a role as a sponge for *miR-141-5p* and *hsa\_circ\_0000375* could be acted as a sponge for *miR-183-5p*. Therefore, we would like to emphasize that the relationship between *hsa\_circ0016201/miR-141-5p/ADH1B* and *hsa\_0000375/miR-183-5p/BTLN9* should be investigated at the cellular functional level. It could be substantial to examine this relationship with conventional molecular genetic techniques in both BC cells and tumor tissue.

More importantly, the expression of *miR-486-5p*, which is an essential tumor suppressor miRNA in BC and many other cancer types [54–57], was significantly decreased in all PAM50 groups examined in our study. It was determined that *hsa\_circ\_0000515*, one of the circRNAs whose expression was significantly increased in the dataset we detected, could act as a sponge for *miR-486-5p*. In addition, we determined that the increased expression of *SDC1*, *PRAME*, *EXO1*, *BUB1*, and *DLGAP5* genes could be more strongly associated targets of *miR-486-5p* in BC. The overexpressed *SDC1* gene was found to lead a significantly poor OS and DFS and overexpressed *PRAME*, *EXO1*, *BUB1* and *DLGAP5* genes were found to lead a significantly poor DFS in BC (Fig 7C–7E and S8 Fig in S1 File) (criterion III). *miR-486-5p* has been notified as an important tumor suppressor miRNA in various cancers, including BC. It has been reported that *miR-486-5p* which could be found exosomal miRNA in BC inhibits epithelial-mesenchymal transition (EMT) by targeting Dock1 and suppresses cancer cell proliferation by targeting the *PIM-1* oncogene in BC, can be used as a biomarker in the prediction of BC recurrence [56–58]. Valuable studies are showing that *miR-486-5p* may be associated with different circRNAs. The importance of *circNFIB1/miR-486-5p/PIK3R1/VEGF-C* axis in lymphatic system metastasis in pancreatic cancer [59]; *hsa\_circ\_0016788/miR-486-5p/CDK4* pathway in hepatocellular carcinoma tumorigenesis [60]; *circHUWE1/miR-486-5p* in colorectal cancer migration and invasion [61] and *Circ-TCF4.85/miR-486-5p/ABCF2* in hepatocellular carcinoma progression [62] are reported in the literature. However, as far as we know the relationship between *miR-486-5p* and circRNAs has not yet been reported in BC. Syndecan-1 (*SDC1*, *CD138*) is a critical cell surface adhesion molecule required for cell morphology and impact on the natural microenvironment. *SDC1* dysregulation enhances cancer development by increasing cell proliferation, angiogenesis, invasion, and metastasis and is linked to chemo-resistance. *SDC1* expression is also correlated to chemotherapy responses and prognosis in numerous solid and/ or hematological cancers, including BC [63, 64]. It has been suggested that *SDC1* could be a new molecular marker that alters the phenotype of cancer stem cells through the *IL-6/STAT3*, Notch, and EGFR signaling pathways in triple-negative inflammatory BC [65]. Induction of *SDC1* in the lung microenvironment may promote the formation of breast tumor

metastasis [66]. *SDC1* has been found to have a prominent role in the process of BC metastasis to the brain. *SDC1* has been shown to increase BC cell migration across the blood-brain barrier via modulating cytokines, which may alter the blood-brain barrier [67]. *SDC1* overexpression in BC is associated with various miRNAs [68, 69]. However, the relationship between *miR-486-5p/ SDC1* and BC is not yet known. It has been reported that *SDC1* expression can also be indirectly altered by circRNAs as it has been demonstrated that *circCEP128* is associated with bladder cancer progression via the *miR-515-5p/ SDC1* axis [70]. According to our bioinformatics study findings, we recommend further investigation of the *SDC1* gene together with the *hsa\_circ\_0000515/ miR-486-5p* axis when conducting circRNA/miRNA/gene functional research in BC. The *SDC1* hub gene, which is targeted by miR-486-5p, was found to be highly compatible with the selection criteria we applied in our study. We propose that *hsa\_circ\_0000515/ miR-486-5p/ SDC1* axis may be an important biomarker candidate in distinguishing patients especially in the BLBC group, according to the PAM50 classification of BC.

## 4 Conclusion

Finding new biomarkers to clearly classify subtypes of BC could be quite crucial in the battle against cancer. To identify novel biomarkers and new therapeutics, a deeper understanding of the mechanisms underlying BC metastasis is extremely important. According to our study results, we suggest various DE mRNAs, miRNAs and circRNAs that may be important in the onco-transcriptomic cascade for BC. The interrelationships of these molecules can be potential diagnostic biomarkers or therapeutic targets. Therefore, functional experiments such as proliferation, apoptosis, invasion, and metastasis on BC cells should be studied to elucidate these circRNA-miRNA-mRNA relationships in the future.

## Supporting information

**S1 File. Contains all the supporting figures.**

(DOCX)

**S1 Table. Details of databases, tools, and R packages used in this study.**

(XLSX)

**S2 Table. Log-2 transformed fold change (Log2FC) values of the 18 DEGs.**

(XLSX)

## Acknowledgments

Additional information

More detailed information about the manuscript can be found on the GitHub repository.

([https://github.com/cihaterdogan/Breast\\_Cancer\\_RN](https://github.com/cihaterdogan/Breast_Cancer_RN)).

## Author Contributions

**Conceptualization:** Zeyneb Kurt.

**Data curation:** Cihat Erdogan, Ilknur Suer, Murat Kaya.

**Formal analysis:** Cihat Erdogan, Ilknur Suer, Murat Kaya.

**Investigation:** Cihat Erdogan, Ilknur Suer, Murat Kaya.

**Methodology:** Zeyneb Kurt.

**Project administration:** Zeyneb Kurt.



**Supervision:** Sukru Ozturk, Nizamettin Aydin.

**Writing – original draft:** Cihat Erdogan, Ilknur Suer, Murat Kaya.

**Writing – review & editing:** Sukru Ozturk, Nizamettin Aydin, Zeyneb Kurt.

## References

1. Feng Y, Spezia M, Huang S, Yuan C, Zeng Z, Zhang L, et al. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. *Genes & diseases*. 2018; 5(2):77–106. <https://doi.org/10.1016/j.gendis.2018.05.001> PMID: 30258937
2. Wang J, Zhang Q, Zhou S, Xu H, Wang D, Feng J, et al. Circular RNA expression in exosomes derived from breast cancer cells and patients. *Epigenomics*. 2019; 11(4):411–21. <https://doi.org/10.2217/epi-2018-0111> PMID: 30785332
3. Wan B, Liu B, Lv C. Progress of research into circular RNAs in urinary neoplasms. *PeerJ*. 2020; 8: e8666. <https://doi.org/10.7717/peerj.8666> PMID: 32140311
4. Kelly J, Berzuini C, Keavney B, Tomaszewski M, Guo H. A review of causal discovery methods for molecular network analysis. *Molecular Genetics & Genomic Medicine*. 2022; 10(10):e2055. <https://doi.org/10.1002/mgg3.2055> PMID: 36087049
5. Liu X, Zeng Y, Liu Z, Li W, Wang L, Wu M. Bioinformatics analysis of the circRNA-miRNA-mRNA network for atrial fibrillation. *Medicine*. 2022; 101(34). <https://doi.org/10.1097/MD.00000000000030221> PMID: 36042613
6. Hu L, Zheng B, Yang Y, Chen C, Hu M. Construction of circRNA-miRNA-mRNA Network Reveal Functional circRNAs and Key Genes in Acute Myeloid Leukemia. *International Journal of General Medicine*. 2023:1491–504. <https://doi.org/10.2147/IJGM.S402671> PMID: 37123885
7. Søskilde R, Persson H, Ehinger A, Pirona AC, Fernö M, Hegardt C, et al. Refinement of breast cancer molecular classification by miRNA expression profiles. *BMC genomics*. 2019; 20:1–12.
8. Haibe-Kains B, Desmedt C, Piette F, Buyse M, Cardoso F, Van't Veer L, et al. Comparison of prognostic gene expression signatures for breast cancer. *BMC genomics*. 2008; 9:1–9.
9. Bartlett JM, Bayani J, Marshall A, Dunn JA, Campbell A, Cunningham C, et al. Comparing breast cancer multiparameter tests in the OPTIMA prelim trial: no test is more equal than the others. *JNCI: Journal of the National Cancer Institute*. 2016; 108(9):djw050. <https://doi.org/10.1093/jnci/djw050> PMID: 27130929
10. Dowsett M, Sestak I, Lopez-Kowles E, Sidhu K, Dunbier AK, Cowens JW, et al. Comparison of PAM50 risk of recurrence score with onco type DX and IHC4 for predicting risk of distant recurrence after endocrine therapy. *Journal of Clinical Oncology*. 2013; 31(22):2783–90. <https://doi.org/10.1200/JCO.2012.46.1558> PMID: 23816962
11. Van't Veer LJ, Dai H, Van De Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. *nature*. 2002; 415(6871):530–6. <https://doi.org/10.1038/415530a> PMID: 11823860
12. Xu J-Z, Shao C-C, Wang X-J, Zhao X, Chen J-Q, Ouyang Y-X, et al. circTADA2As suppress breast cancer progression and metastasis via targeting miR-203a-3p/SOCS3 axis. *Cell death & disease*. 2019; 10(3):175. <https://doi.org/10.1038/s41419-019-1382-y> PMID: 30787278
13. Huang Y, Qian M, Chu J, Chen L, Jian W, Wang G. Identification of circRNA-miRNA-mRNA network in luminal breast cancers by integrated analysis of microarray datasets. *Frontiers in Molecular Biosciences*. 2023; 10:1162259. <https://doi.org/10.3389/fmolb.2023.1162259> PMID: 37187897
14. Feng J, Chen W, Dong X, Wang J, Mei X, Deng J, et al. CSCD2: an integrated interactional database of cancer-specific circular RNAs. *Nucleic Acids Research*. 2022; 50(D1):D1179–D83. <https://doi.org/10.1093/nar/gkab830> PMID: 34551437
15. Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic acids research*. 2020; 48(D1):D127–D31. <https://doi.org/10.1093/nar/gkz757> PMID: 31504780
16. Sticht C, De La Torre C, Parveen A, Gretz N. miRWalk: an online resource for prediction of microRNA binding sites. *PloS one*. 2018; 13(10):e0206239. <https://doi.org/10.1371/journal.pone.0206239> PMID: 30335862
17. Huang H-Y, Lin Y-C-D, Li J, Huang K-Y, Shrestha S, Hong H-C, et al. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic acids research*. 2020; 48(D1): D148–D54. <https://doi.org/10.1093/nar/gkz896> PMID: 31647101

18. Sheng H, Pan H, Yao M, Xu L, Lu J, Liu B, et al. Integrated analysis of circular RNA-associated ceRNA network reveals potential circRNA biomarkers in human breast cancer. *Computational and Mathematical Methods in Medicine*. 2021; 2021. <https://doi.org/10.1155/2021/1732176> PMID: 34966440
19. Tobin N, Harrell J, Lötvot J, Brage SE, Stolt MF, Carlsson L, et al. Molecular subtype and tumor characteristics of breast cancer metastases as assessed by gene expression significantly influence patient post-relapse survival. *Annals of oncology*. 2015; 26(1):81–8. <https://doi.org/10.1093/annonc/mdl498> PMID: 25361981
20. Tomczak K, Czerwińska P, Wiznerowicz M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary Oncology/Współczesna Onkologia*. 2015; 2015(1):68–77.
21. Dvinge H, Git A, Gräf S, Salmon-Divon M, Curtis C, Sottoriva A, et al. The shaping and functional consequences of the microRNA landscape in breast cancer. *Nature*. 2013; 497(7449):378–82. <https://doi.org/10.1038/nature12108> PMID: 23644459
22. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell*. 2018; 173(2):400–16. e11. <https://doi.org/10.1016/j.cell.2018.02.052> PMID: 29625055
23. Harrison E, Drake T, Ots R. finalfit: Quickly Create Elegant Regression Results Tables and Plots when Modelling. 2023.
24. 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 research*. 2015; 43(7):e47–e. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792
25. Therneau TM. A Package for Survival Analysis in R. 2023.
26. Hatzis C, Pusztai L, Valero V, Booser DJ, Esserman L, Lluch A, et al. A genomic predictor of response and survival following taxane-anthracycline chemotherapy for invasive breast cancer. *Jama*. 2011; 305(18):1873–81. <https://doi.org/10.1001/jama.2011.593> PMID: 21558518
27. 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 research*. 2003; 13(11):2498–504. <https://doi.org/10.1101/gr.1239303> PMID: 14597658
28. Wei T, Simko V. R package 'corrplot': Visualization of a Correlation Matrix. 2021.
29. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic acids research*. 2021; 49(D1):D605–D12. <https://doi.org/10.1093/nar/gkaa1074> PMID: 33237311
30. Xie Z, Bailey A, Kuleshov MV, Clarke DJ, Evangelista JE, Jenkins SL, et al. Gene set knowledge discovery with Enrichr. *Current protocols*. 2021; 1(3):e90. <https://doi.org/10.1002/cpz1.90> PMID: 33780170
31. Fabregat A, Sidiropoulos K, Viteri G, Forner O, Marin-Garcia P, Arnaiz V, et al. Reactome pathway analysis: a high-performance in-memory approach. *BMC bioinformatics*. 2017; 18(1):1–9.
32. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, et al. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic acids research*. 2020; 48(D1):D845–D55. <https://doi.org/10.1093/nar/gkz1021> PMID: 31680165
33. Licursi V, Conte F, Fisco G, Paci P. MIENURNET: an interactive web tool for microRNA-target enrichment and network-based analysis. *BMC bioinformatics*. 2019; 20(1):1–10.
34. Brigham, Hospital Ws, 13 MSCLPPJKR, 25 GdaBCoMCCJDLA, Ilya IFSBRSKRBBBBRETLJTVZWS. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012; 490(7418):61–70. <https://doi.org/10.1038/nature11412> PMID: 23000897
35. Bourguignon LY, Singleton PA, Zhu H, Diedrich F. Hyaluronan-mediated CD44 interaction with Rho-GEF and Rho kinase promotes Grb2-associated binder-1 phosphorylation and phosphatidylinositol 3-kinase signaling leading to cytokine (macrophage-colony stimulating factor) production and breast tumor progression. *Journal of Biological Chemistry*. 2003; 278(32):29420–34. <https://doi.org/10.1074/jbc.M301885200> PMID: 12748184
36. Momenimovahed Z, Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer: Targets and Therapy*. 2019; 15:1–64. <https://doi.org/10.2147/BCTT.S176070> PMID: 31040712
37. Van Nguyen C, Nguyen QT, Vu HTN, Pham KH, Phung HT. Molecular classification predicts survival for breast cancer patients in Vietnam: a single institutional retrospective analysis. *International Journal of Clinical and Experimental Pathology*. 2021; 14(3):322. PMID: 33786149
38. Voduc KD, Cheang MC, Tyldesley S, Gelmon K, Nielsen TO, Kennecke H. Breast cancer subtypes and the risk of local and regional relapse. *Journal of clinical oncology*. 2010; 28(10):1684–91. <https://doi.org/10.1200/JCO.2009.24.9284> PMID: 20194857

39. Soliman NA, Yussif SM, Shebl AM. Syndecan-1 could be added to hormonal receptors and HER2/neu in routine assessment of invasive breast carcinoma, relation of its expression to prognosis and clinicopathological parameters. *Pathology-Research and Practice*. 2019; 215(5):977–82. <https://doi.org/10.1016/j.prp.2019.02.003> PMID: 30738694
40. Conway ME, McDaniel JM, Graham JM, Guillen KP, Oliver PG, Parker SL, et al. STAT3 and GR cooperate to drive gene expression and growth of basal-like triple-negative breast cancer. *Cancer research*. 2020; 80(20):4355–70. <https://doi.org/10.1158/0008-5472.CAN-20-1379> PMID: 32816914
41. Badowska-Kozakiewicz A, Budzik M. Immunohistochemical characteristics of basal-like breast cancer. *Contemporary Oncology/Współczesna Onkologia*. 2016; 20(6):436–43. <https://doi.org/10.5114/wo.2016.56938> PMID: 28239279
42. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *Journal of clinical oncology*. 2009; 27(8):1160. <https://doi.org/10.1200/JCO.2008.18.1370> PMID: 19204204
43. Kim S, Dougherty ER, Shmulevich I, Hess KR, Hamilton SR, Trent JM, et al. Identification of combination gene sets for glioma classification. *Molecular Cancer Therapeutics*. 2002; 1(13):1229–36. PMID: 12479704
44. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, et al. MicroRNA expression profiles classify human cancers. *nature*. 2005; 435(7043):834–8. <https://doi.org/10.1038/nature03702> PMID: 15944708
45. Yerukala Sathipati S, Ho S-Y. Identifying a miRNA signature for predicting the stage of breast cancer. *Scientific reports*. 2018; 8(1):16138. <https://doi.org/10.1038/s41598-018-34604-3> PMID: 30382159
46. Li B, Xie F, Zheng F-x, Jiang G-s, Zeng F-q, Xiao X-y. Overexpression of CircRNA BCRC4 regulates cell apoptosis and MicroRNA-101/EZH2 signaling in bladder cancer. *Current Medical Science*. 2017; 37:886–90. <https://doi.org/10.1007/s11596-017-1822-9> PMID: 29270748
47. Xu S, Guo J, Zhang W. lncRNA PCAT19 promotes the proliferation of laryngocarcinoma cells via modulation of the miR-182/PDK4 axis. *Journal of cellular biochemistry*. 2019; 120(8):12810–21. <https://doi.org/10.1002/jcb.28552> PMID: 30868666
48. Abdelmohsen K, Panda AC, Munk R, Grammatikakis I, Dudekula DB, De S, et al. Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA biology*. 2017; 14(3):361–9. <https://doi.org/10.1080/15476286.2017.1279788> PMID: 28080204
49. Xu T, Yu Y, Guo S, He L, Mao Z. CircPSMC3 suppresses migration and invasion of non-small cell lung cancer cells via miR-182-5p/NME2 axis. *Medical Science Monitor: International Medical Journal of Experimental and Clinical Research*. 2020; 26:e924134–1. <https://doi.org/10.12659/MSM.924134> PMID: 32386284
50. Sang Y, Chen B, Song X, Li Y, Liang Y, Han D, et al. circRNA\_0025202 regulates tamoxifen sensitivity and tumor progression via regulating the miR-182-5p/FOXO3a axis in breast cancer. *Molecular Therapy*. 2019; 27(9):1638–52. <https://doi.org/10.1016/j.ymthe.2019.05.011> PMID: 31153828
51. Nair AA, Niu N, Tang X, Thompson KJ, Wang L, Kocher J-P, et al. Circular RNAs and their associations with breast cancer subtypes. *Oncotarget*. 2016; 7(49):80967. <https://doi.org/10.18632/oncotarget.13134> PMID: 27829232
52. Darbeheshti F, Zokaei E, Mansoori Y, Emadi Allahyari S, Kamaliyan Z, Kadkhoda S, et al. Circular RNA hsa\_circ\_0044234 as distinct molecular signature of triple negative breast cancer: a potential regulator of GATA3. *Cancer cell international*. 2021; 21(1):1–12.
53. He D, Yang X, Kuang W, Huang G, Liu X, Zhang Y. The novel circular RNA Circ-PGAP3 promotes the proliferation and invasion of triple negative breast cancer by regulating the miR-330-3p/Myc axis. *Oncotargets and therapy*. 2020; 10:149–59. <https://doi.org/10.2147/OTT.S274574> PMID: 33116597
54. Wang W, Liu B, Sun S, Lan L, Chen Y, Han S, et al. Downregulation of miR-486-5p enhances the anti-tumor effect of 5-fluorouracil on pancreatic cancer cells. *Oncotargets and therapy*. 2020; 1649–59. <https://doi.org/10.2147/OTT.S231153> PMID: 32158231
55. Tian F, Wang J, Ouyang T, Lu N, Lu J, Shen Y, et al. MiR-486-5p serves as a good biomarker in non-small cell lung cancer and suppresses cell growth with the involvement of a target PIK3R1. *Frontiers in genetics*. 2019; 10:688. <https://doi.org/10.3389/fgene.2019.00688> PMID: 31402930
56. Li H, Mou Q, Li P, Yang Z, Wang Z, Niu J, et al. MiR-486-5p inhibits IL-22-induced epithelial-mesenchymal transition of breast cancer cell by repressing Dock1. *Journal of Cancer*. 2019; 10(19):4695. <https://doi.org/10.7150/jca.30596> PMID: 31528235
57. Zhang G, Liu Z, Cui G, Wang X, Yang Z. MicroRNA-486-5p targeting PIM-1 suppresses cell proliferation in breast cancer cells. *Tumor Biology*. 2014; 35(11):11137–45. <https://doi.org/10.1007/s13277-014-2412-0> PMID: 25104088

58. Sueta A, Yamamoto Y, Tomiguchi M, Takeshita T, Yamamoto-Ibusuki M, Iwase H. Differential expression of exosomal miRNAs between breast cancer patients with and without recurrence. *Oncotarget*. 2017; 8(41):69934. <https://doi.org/10.18632/oncotarget.19482> PMID: 29050253
59. Kong Y, Li Y, Luo Y, Zhu J, Zheng H, Gao B, et al. circNFIB1 inhibits lymphangiogenesis and lymphatic metastasis via the miR-486-5p/PIK3R1/VEGF-C axis in pancreatic cancer. *Molecular Cancer*. 2020; 19:1–17.
60. Guan Z, Tan J, Gao W, Li X, Yang Y, Li X, et al. Circular RNA hsa\_circ\_0016788 regulates hepatocellular carcinoma tumorigenesis through miR-486/CDK4 pathway. *Journal of Cellular Physiology*. 2019; 234(1):500–8.
61. Chen H-Y, Li X-N, Ye C-X, Chen Z-L, Wang Z-J. Circular RNA circHUWE1 is upregulated and promotes cell proliferation, migration and invasion in colorectal cancer by sponging miR-486. *OncoTargets and therapy*. 2020:423–34. <https://doi.org/10.2147/OTT.S233338> PMID: 32021287
62. Gao J, Dai C, Yu X, Yin XB, Zhou F. Circ-TCF4. 85 silencing inhibits cancer progression through microRNA-486-5p-targeted inhibition of ABCF2 in hepatocellular carcinoma. *Molecular Oncology*. 2020; 14(2):447. <https://doi.org/10.1002/1878-0261.12603> PMID: 31758671
63. Aki MR, Nagpal P, Ayoub NM, Prabhu SA, Gliksman M, Tai B, et al. Molecular and clinical profiles of syndecan-1 in solid and hematological cancer for prognosis and precision medicine. *Oncotarget*. 2015; 6(30):28693. <https://doi.org/10.18632/oncotarget.4981> PMID: 26293675
64. Leivonen M, Lundin J, Nordling S, Von Boguslawski K, Haglund C. Prognostic value of syndecan-1 expression in breast cancer. *Oncology*. 2004; 67(1):11–8. <https://doi.org/10.1159/000080280> PMID: 15459490
65. Ibrahim SA, Gadalla R, El-Ghonaimy EA, Samir O, Mohamed HT, Hassan H, et al. Syndecan-1 is a novel molecular marker for triple negative inflammatory breast cancer and modulates the cancer stem cell phenotype via the IL-6/STAT3, Notch and EGFR signaling pathways. *Molecular cancer*. 2017; 16:1–19.
66. Chute C, Yang X, Meyer K, Yang N, O'Neil K, Kasza I, et al. Syndecan-1 induction in lung microenvironment supports the establishment of breast tumor metastases. *Breast cancer research*. 2018; 20:1–12.
67. Sayyad MR, Puchalapalli M, Vergara NG, Wangenstein SM, Moore M, Mu L, et al. Syndecan-1 facilitates breast cancer metastasis to the brain. *Breast Cancer Research and Treatment*. 2019; 178:35–49. <https://doi.org/10.1007/s10549-019-05347-0> PMID: 31327090
68. Ibrahim SA, Yip GW, Stock C, Pan JW, Neubauer C, Poeter M, et al. Targeting of syndecan-1 by microRNA miR-10b promotes breast cancer cell motility and invasiveness via a Rho-GTPase-and E-cadherin-dependent mechanism. *International journal of cancer*. 2012; 131(6):E884–E96. <https://doi.org/10.1002/ijc.27629> PMID: 22573479
69. Uen Y, Wang J-W, Wang C, Jhang Y, Chung J-Y, Tseng T, et al. Mining of potential microRNAs with clinical correlation-regulation of syndecan-1 expression by miR-122-5p altered mobility of breast cancer cells and possible correlation with liver injury. *Oncotarget*. 2018; 9(46):28165. <https://doi.org/10.18632/oncotarget.25589> PMID: 29963269
70. Cao G, Zhang C, Tian X, Jing G, Zhou X, Yan T. circCEP128 knockdown suppresses bladder cancer progression via regulating microRNA-515-5p/SDC1 axis. *Cancer Management and Research*. 2021:2885–96. <https://doi.org/10.2147/CMAR.S288229> PMID: 33833571