Identifying of a Novel miR-98-5p/IGF1 Axis Contributes the Pathogenesis of Breast Cancer Using Comprehensive Bioinformatic Analyses Methods and Experiments Validation

Dapeng Sun

The First Affiliated Hospital of Jinzhou Medical University

Xigang Luo

The Thirs Affiliated Hospital of Jinzhou Medical University

Lingling Ma

The First Affiliated Hospital of Jinzhou Medical University

Yi Wang

The Third Affiliated Hospital of Jinzhou Medical University

Fengxiang Zhang (✉ fengxiang_zhang@yeah.net)

The First Affiliated Hospital of Jinzhou Medical University <https://orcid.org/0000-0002-7984-2950>

Research article

Keywords: Breast cancer, miRNA, mRNA, has-miR-98-5p, IGF1

Posted Date: July 27th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-46919/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Breast cancer (BC) is a huge threat for the health of women worldwide. Although the numerous microRNAs (miRNA) have been identified to be aberrantly expressed in BC, the construction of a comprehensive miRNA-messenger RNA (mRNA) network is still needed. This study was aimed to identify BC-associated miRNAs through analyzing microarray datasets obtained from GEO database and to construct a miRNA-mRNA network for BC.

Methods: Limma package was used to identify differentially expressed miRNAs (DEMs) in microarray datasets. Genes targeted by DEMs were analyzed with miRTarBase. Gene Ontology and pathway enrichment analysis for the predicted target genes were performed at DAVID. Correlation of DEMs and target genes was analyzed at ENCORI. Based on these results, a miRNA-mRNA regulatory network was constructed.

Results: A total of 17 overlapping DEMs were identified at these two microarray datasets. Expression of DEMs in BC were further validated by ENCORI. By utilizing miRTarBase, a total of 167 target genes for DEMs were obtained. 10 hub genes (AKT1, MYC, VEGFA, CCND1, PTEN, IL6, CASP3, KRAS, IGF1, ESR1) were identified after network analysis at STRING and CytoScape. Through analyzing the effects of hub genes on overall survival of BC patients and correlation of DEMs and hub genes, we found hsa-miR-98-5p/IGF1 axis may play a crucial role in BC progression. The connections of hsa-miR-98-5p and IGF1 were further validated by luciferase activity reporter assay and functional assays.

Conclusion: In this work, a miRNA-mRNA network related to BC progression was built, and identified one important miRNA-mRNA axis in BC.

Background

The estimated newly diagnosed cases for Breast cancer (BC) are 2,088,849 in 2018 at the worldwide range [1]. Hence, BC represents a huge health threat for women worldwide [2]. Cancer is characterized as genome mutation, abnormal cellular growth status, failed to immune response, and so on [3]. As a result, it was believed that identifying molecules related to these processes is crucial to understand the mechanisms behind BC progression for the identification of novel prognosis or treatment biomarkers.

The recent development in sequencing technology and bioinformatic analyses measures has identified numerous aberrantly expressed protein coding genes and non-coding genes related to cancer metastasis, drug resistance, initiation and development [4–6].

As reported, there are only about 2% protein-coding RNAs in genome according to the human genome project, while most of them belong to non-coding RNAs (ncRNAs) [7, 8]. ncRNAs including pseudogene, circle RNA (circRNA), long non-coding RNA (lncRNA), microRNA (miRNA) were used to be regarded as “junk genes” but their roles in cancers have been gradually recognized in recent years [9]. miRNAs are a class of ncRNAs with the length of approximately 18–25 nucleotides that can directly regulate target

gene expression mainly via 3'-untranslated region (3'-UTR) binding [10]. miRNAs were found closely associated with BC cancer cell growth, apoptosis, metastasis, and drug sensitive [11]. Yao et al. identified 218 differentially expressed miRNAs (DEMs) and 2222 differentially expressed messenger RNAs (DEGs) correlated with prognosis of BC patients using the data obtained from The Cancer Genome Atlas (TCGA) database, and based on that a miRNA-mRNA regulatory network was constructed [12]. In addition, He et al. reported 11 highly expressed glycolytic proteins, who have a closely association with poor prognosis of BC [13]. By furthering analyses, they identified miR-140-5p was downregulated in BC and targeted GLUT1 expression to regulate BC glycolytic and proliferation [13].

In light of this, we aimed to build a miRNA-mRNA regulatory network in BC based on the datasets obtained from Gene Expression Omnibus (GEO). By analyzing the targets for miRNAs, we constructed miRNA-mRNA network related to the progression of BC. Finally, we explored the correlation of miRNA and mRNA in BC to provide more evidence to strength the findings in this work.

Methods

Microarray data analysis

Two microarray datasets were downloaded from GEO (<https://www.ncbi.nlm.nih.gov/gds>), a database contains gene expression data. Dataset GSE83270 contains gene profile data from 6 health control tissues and 6 BC patients' tissues was analyzed at 7th generation of miRCURYTM LNA Array (v.18.0, Exiqon) and submitted by Ya-Wen Wang from Shandong University. Dataset GSE68085 contains gene expression data from 104 BC patients' tissues and 11 healthy normal tissues was analyzed at GPL10999 platform and submitted by Sambasivarao Damaraju from University of Alberta.

Identification of DEMs

Limma package in R software was employed to identify DEMs between cancer tissues and normal tissues at both datasets. Cut-off values were set as $|\log_2 \text{fold change (FC)}| > 2$ and $P < 0.01$. Venn diagram was utilized to identify the overlapping DEMs in these two datasets.

Validation of DEMs expression at ENCORI

The expression of these overlapping DEMs in BC patients was validated using ENCORI, a public accessed platform for studying RNA interactions [14].

miRNA targets prediction

The mRNA targets for DEMs were analyzed at miRTarBase, a database contains experimental validated targets [15]. Only the targets validated by strong methods including luciferase activity reporter assay, western blot, or quantitative real-time PCR (RT-qPCR) were selected for followingly analyses. Based on these results, a miRNA-mRNA regulatory network related to the development of BC was established and visualized at Cytoscape V_3.6.0 software [16].

Protein-protein interaction (PPI) network

PPI network for the above predicted miRNA targets was analyzed at Search Tool for the Retrieval of Interacting Genes (<https://string-db.org>) and visualized at Cytoscape V_3.6.0 software. Hub genes in this network were analyzed at CytoHubba with the cut-off value of degree above 10.

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis

Database for annotation, visualization, and integrated discovery (<https://david.ncifcrf.gov/>) was employed for GO and KEGG analyses of the predicted miRNA targets [17]. P value less than 0.05 was believed as significant enrichment.

Validation of hub gene expression at ENCORI, GEPIA, and UALCAN

The expression level of these 10 identified hub genes in BC was analyzed at both ENCORI, GEPIA, and UALCAN [18,19]. The genes that revealed to be abnormally expressed in these three databases were selected for following analyses.

Survival analysis using Kaplan-Meier Plotter

The effects of hub genes on the overall survival of BC patients were analyzed at Kaplan-Meier Plotter [20]. The median value was used as cut-off value to classify patients into high or low expression groups.

Analysis of miRNA and mRNA correlation at ENCORI

Subsequently, hub genes after expression and prognostic value validation were selected to analyze its connection of miRNA at ENCORI. The pairs that inversely correlated were selected for following analyses.

Cell line

Human BC cell lines (MCF7 and MDA-MB-415) and mammary epithelial cell (MCF10A) were incubated in DMEM containing 10 % FBS (Invitrogen, Thermo Fisher Scientific, Inc., Waltham, MA, USA). Cells cultivation atmosphere was maintained at 37 °C containing 5 % of CO₂.

RNA extraction and RT-qPCR analysis

Total RNA was extracted from cultured cells with Trizol reagent obtained from Invitrogen (Thermo Fisher Scientific, Inc.). Complementary DNA (cDNA) was synthesized from RNA using miScript Reverse Transcription Kit (Takara, Dalian, Liaoning, China). RT-qPCR was performed at ABI 7500 system (Applied Biosystems, Foster City, CA, USA) using SYBR Green Mix (Takara) to measure the expression level of miR-98-5p or IGF1. The primers used were as follows: hsa-miR-98-5p: forward 5'-TGAGGTAGTAGTTTGTGCTGTT-3', reverse 5'-GCGAGCACAGAATTAATACGAC-3'; U6 snRNA: forward 5'-CTCGCTTCGGCAGCACA-3', reverse 5'-AACGCTTCACGAATTTGCGT-3'; IGF1: forward 5'-

CACATCATGTCTGCTTTCACACC-3', reverse 5'-GGAAGCAACACTCATCCACAATG-3'; GAPDH: forward 5'-CTGGGCTACACTGAGCACC-3', reverse 5'-AAGTGGTCGTTGAGGGCAATG-3'. $2^{-\Delta\Delta C_t}$ method was utilized to analyze relative expression levels of hsa-miR-98-5p or IGF1 using U6 snRNA or GAPDH as internal control, respectively.

Luciferase activity reporter assay

The 3'-UTR contains the binding site for hsa-miR-98-5p in IGF1 was amplified from genome and cloned into pGL3 vector (Promega, Madison, WI, USA) and named as wt-IGF1. The mutant IGF1 luciferase vector (mt-IGF1) was constructed from wt-IGF1 using site-direct mutagenesis kit (Takara). Cells were co-transfected with wt-IGF1 or mt-IGF1 and hsa-miR-98-5p mimic or negative control miRNA (miR-NC) using Lipofectamine 2000 (Invitrogen). The hsa-miR-98-5p mimic or miR-NC were purchased from RiboBio (Guangzhou, China). After co-transfection for 48 h, relative luciferase activity was measured using dual-luciferase activity reporter system (Promega).

Cell transfection

pcDNA3.1 contains the coding sequence of IGF1 (pIGF1) was designed by GenScript (Nanjing, Jiangsu, China). For miRNAs or siRNAs transfection, Lipofectamine 2000 obtained from Invitrogen was used according to the manufacturer's protocol.

Cell counting kit-8 assay

Cell Counting Kit-8 (CCK-8) bought from Beyotime (Haimen, Jiangsu, China) was used to measure cell proliferation rate. Briefly, 2,000 cells were cultured into 96-well plates and incubated for indicated time. Then, CCK-8 reagent was added to each well and the detected the optical density value at 450 nm after further incubation for 2 h.

Wound-healing assay

5×10^5 cells were plated in 6-well plates. After wound generation at cell surface, cells were washed with PBS and incubated in serum-free medium. After 48 h, cell images were captured to measure the migration rate.

Transwell invasion assay

Cell invasion ability was detected with Matrigel coated transwell chamber (8- μ m pore size). Cells in serum-free medium were plated to upper chamber, while the serum contained medium was filled to lower chamber. Invaded cells were stained by crystal violet and counted under microscope.

Tumor formation in BALB/c nude mice

BALB/c athymic nude mice (female, 4-6 weeks old and 16-20 g) purchased from Beijing Vital River Laboratory Animal Technology Co., Ltd. were used for animal experiments. MCF7 cells stable expressed

has-miR-98-5p inhibitor and NC-inhibitor suspended in PBS were inoculated subcutaneously into nude mice (6 per group). Tumor length and width was measured every 5 days. The mice were euthanized via overdose of pentobarbital, and the tumors were weighed. Tumor size was calculated using the formula: $(L \times W^2)/2$ (L represents Length; W represents Width). Animal experiments were approved by animal care committee of The First Affiliated Hospital of Jinzhou Medical University.

Statistical analysis

Data obtained in this work were analyzed using GraphPad software (San Diego, CA, USA) and presented as mean $\pm$ SD. Differences in groups were analyzed with Student's t-test (two groups) or ANOVA and post-hoc test (for three groups). P value less than 0.05 was considered as statistically significant.

Results

Screening and analyzing DEMs between BC tumor tissues and normal tissues

Through limma package analyses, we identified 46 DEMs in GSE83270 dataset. In the meantime, we found 297 DEMs in GST68085 dataset. Using Venn diagram, we identified 17 overlapping DEMs in both datasets (Fig. 1a). The detailed expression information of these 17 DEMs in these two datasets were summarized in Table 1. Moreover, the expression of these 17 DEMs in BC patients was further validated at ENCORI and we showed 13 of them were abnormally expressed (Fig. 1b). Moreover, 11 of these 13 miRNAs fit the expression trend as indicated in two GEO datasets (Fig. 1b).

Table 1
Expression of 17 overlapping DEMs in GSE83270 and GSE68085

Name	GSE83270		GSE68085	
	log2FC	P-value	log2FC	P-value
hsa-miR-323b-3p	2.128601	0.0007409	2.8354	3.82E-05
hsa-miR-598-3p	2.114422	0.0010003	2.2093	0.0083164
hsa-miR-192-5p	2.241504	0.0010774	2.1266	0.00021483
hsa-miR-374b-5p	2.217938	0.0021335	3.8403	6.41E-09
hsa-let-7f-2-3p	2.183516	0.0027173	3.4297	8.66E-06
hsa-miR-362-3p	2.368714	0.0028746	2.4015	0.00057224
hsa-miR-20a-3p	2.466883	0.0029047	3.6082	0.0078811
hsa-miR-26b-5p	2.98713	0.0037988	2.9919	5.45E-08
hsa-miR-96-5p	2.671409	0.0045554	2.7382	7.01E-07
hsa-miR-374a-3p	2.925296	0.0057706	2.7373	1.56E-07
hsa-miR-30e-3p	2.170472	0.0058891	2.8748	0.0082688
hsa-miR-98-5p	2.732916	0.0065607	4.5715	4.51E-09
hsa-let-7f-5p	2.777093	0.0070485	2.0709	3.73E-10
hsa-miR-454-3p	3.348565	0.0073194	2.1667	7.55E-08
hsa-miR-195-5p	2.536921	0.0081288	2.7727	0.0012129
hsa-miR-126-3p	2.819163	0.0085849	2.3322	0.00019853
hsa-miR-424-5p	2.294141	0.0087538	2.2342	1.20E-09
DEM: differentially expressed miRNAs; FC: fold change.				

Targets prediction for DEMs

Subsequently, we searched the direct targets of these miRNAs in miRTarBase. The detailed targets information for miRNAs were summarized in Table 2. A total of 167 targets for these 10 miRNAs were then identified because targets for hsa-let-7f-2-3p were not available. After that, a miRNA-mRNA regulatory network was constructed (Fig. 2).

Table 2
The targets for DEMs as predicted by miRTarBase

miRNA	Target genes
hsa-miR-192-5p	ALCAM, BCL2, CDC7, CUL5, DLG5, DTL, ERCC3, HOXA10, HRH1, LMNB2, MAD2L1, MCM10, MIS12, KIF20B, PIM1, PRPF38A, RACGAP1, SEPT10, SMARCB1, TRAPPC2B, WNK1, ACVR2B, RB1, ERCC4, ATP1B1, NOB1, XIAP, ITGB1, ITGAV, DHFR, DICER1, CAV1, SCN5A, BMI1, NID1, ESR1, AKT1, ITGB3, CD47, H3F3A, SERPINE1
hsa-miR-374b-5p	RECK, VEGFA, WNT16, AKT1
hsa-miR-26b-5p	PTGS2, EPHA2, CDK6, CCNE1, ABCA1, ARL4C, GATA4, CHORDC1, PLOD2, NR2C2, TAB1, EZH2, USP9X, KPNA2, RB1, NAMPT, IGF1R, PTEN, COX2, MIEN1, ULK2, TRAF5, SMAD1, HAS2, COL1A2, CTGF, TLR4, IGF1, ST8SIA4, PDE4A, FH, HGF, JAG1, LARP1
hsa-miR-96-5p	FOXO1, MITF, ADCY6, HTR1B, CDKN1A, PRMT5, FOXO3, KRAS, REV1, RAD51, SCARB1, ALK, GSK3B, ZEB1, SNAI2, SLC1A1, SLC6A6, IARS, RECK, TRIB3, SYCP1, PTPN9, MTSS1, TP53INP1, TIMP1, BDNF, GDNF
hsa-miR-374a-3p	AXIN2
hsa-miR-30e-3p	NFKBIA
hsa-miR-98-5p	E2F2, HMGA2, MYC, CASP3, TUSC2, SOCS4, CCL5, EZH2, IL13, CCNJ, IL6, PGRMC1, HK2, NRAS, ITGB3, CYP19A1, CCND2, E2F5, EDN1, CBL, SNX6, PBX3, IGF1, CHUK, PAK1, CCL2, SALL4, CTNND1, SLC31A1, BCL2
hsa-miR-454-3p	SMAD4, CXCL12
hsa-miR-362-3p	E2F1, USF2, PTPN1, CD82, BCAR1
hsa-miR-424-5p	CCNF, CDC14A, CDC25A, CHEK1, KIF23, ATF6, WEE1, ANLN, PLAG1, CCNE1, CCND3, CDK6, CCND1, MAP2K1, FGFR1, NFIA, MYB, SIAH1, HIF1A, CUL2, SPI1, FASN, SOCS6, CDX2, FGF2, PTCH1, SMAD7, TGFB3, SOCS2, KDM5B, SMAD3
DEMs: differentially expressed miRNAs.	

GO and KEGG analyses

Next, GO and KEGG analyses were performed on these 167 mRNAs to understand the biological functions of these genes. GO analysis revealed that these genes were mostly located at nucleoplasm and cytoplasm (Fig. 3a). Molecular functions (MF) enrichment showed that these genes were associated with

protein binding, protein kinase binding, and protein kinase activity (Fig. 3b). In the meantime, the enriched biological processes (BP) were those associated with regulation of cellular processes (Fig. 3c). KEGG analysis showed these genes were most enriched in pathways related to cancer including Pathways in cancer, Prostate cancer, Melanoma, Proteoglycans in cancer, Cell cycle, Small cell lung cancer, PI3K-Akt signaling pathway, Hepatitis B, HTLV-I infection, and Chronic myeloid leukemia (Fig. 3d).

PPI network construction and hub gene identification

To investigate whether these genes were essential for the development of BC, PPI network consists 177 nodes and 176 edges was built (Fig. 4a). Top 10 hub genes according to degree method were selected using CytoHubba and the genes were: AKT1, MYC, VEGFA, CCND1, PTEN, IL6, CASP3, KRAS, IGF1, and ESR1 (Fig. 4b). Network analysis showed the PPI network fits the characters of small-world network (Fig. 5).

Prognostic value of hub genes

In order to explore the significance of these hub genes, we analyzed their effects on overall survival of BC patients. We found MYC, VEGFA, CCND1, IL6, CASP3, KRAS, IGF1, and ESR1 were indicators for the poor overall survival of BC patients (Fig. 6).

Hub gene validation

Then, the expression level of hub genes in BC tissues was validated using ENCORI, GEPIA, and UALCAN. The genes that validated to be aberrantly expressed in all these three databases were selected for analyses. It was found MYC, CCND1, IL6, and IGF1 were abnormally expressed in BC tissues compared with normal tissues (Fig. 7). Moreover, miRNAs that can target these four genes were summarized in Table 3.

Table 3	
miRNAs that can target MYC, CCND1, IL6, and IGF1 as predicted by miRTarBase	
Gene name	miRNA
MYC	hsa-miR-98-5p
CCND1	hsa-miR-424-5p
IL6	hsa-miR-98-5p
IGF1	hsa-miR-26b-5p, hsa-miR-98-5p
miRNA: microRNA.	

Correlation of miRNAs and hub genes in BC

The connection of miRNA and hub genes in BC was then analyzed at ENCORI. As presented in Fig. 8, we found hsa-miR-98-5p and IGF1 was negatively correlated in BC tissues ($P = 1.58 \times 10^{-9}$).

Validation of hsa-miR-98-5p and IGF1 axis in BC

At length, we explored the expression level of hsa-miR-98-5p and IGF1 in BC cells. As presented in Fig. 9a, we showed hsa-miR-98-5p expression level was significantly increased in BC cells compared with the normal cell line. In the meantime, we found IGF1 expression level was significantly decreased in BC cells in comparison with normal cell line (Fig. 9b). The binding region between hsa-miR-98-5p and IGF1 was presented in Fig. 9c. Dual-luciferase activity reporter assay revealed that hsa-miR-98-5p mimic transfection decreased luciferase activity in cells harbor wt-IGF1 (Fig. 9d). These results validated IGF1 was a direct target for hsa-miR-98-5p.

hsa-miR-98-5p regulates the BC cell proliferation, migration, and invasion via targeting IGF1

Then, functional assays were conducted to evaluate the biological roles of hsa-miR-98-5p and IGF1 in regulating BC cell behaviors. We found IGF1 expression was decreased by hsa-miR-98-5p mimic and increased by pIGF1 (Fig. 10a). Functional assays showed cell proliferation, migration, and invasion abilities were stimulated by hsa-miR-98-5p mimic and inhibited by pIGF1 (Fig. 10b-10d). Moreover, rescue experiments indicated that pIGF1 transfection partially abolished the effects of hsa-miR-98-5p on BC cell behaviors (Fig. 10b-10d).

hsa-miR-98-5p regulates tumor growth in vivo

To validate the roles of hsa-miR-98-5p *in vivo*, we established nude mouse xenograft model. As expected, both the tumor size and weight were decreased with hsa-miR-98-5p inhibitor relative to those treated with NC-inhibitor (Fig. 11a and 11b).

Discussion

The 5-year overall survival for BC patients is still far away from satisfactory even though the recently improvements in identifying aberrantly expressed genes in BC [12, 13, 20, 21]. For instance, CYP4Z1 and pseudogene CYP4Z2P was reported to target CDK3 expression to regulate the responses of BC cell to tamoxifen [20]. Besides that, lncRNA SNHG7 expression was reported to be elevated in BC tissues compared with normal tissues [21]. Moreover, the knockdown of SNHG7 was revealed to suppress BC cell

proliferation and invasion via regulating miR-381 [21]. Hence, we believe it is essential to further investigate the mechanisms behind BC development.

Owing to the appearance of new measures to analyze RNA-seq data and microarray-based expression profiling data, we are now more likely to understand the changes in molecular level during cancer progression process. In this study, we downloaded two datasets contains data from a total of 137 of either BC tumor tissues or normal tissues from GEO. We found that 17 overlapping miRNAs were aberrantly expressed in BC tumor tissues compared with the normal tissues. Combined with the targets of these miRNAs, we found 167 mRNAs that were further used for functional enrichment analyses. We showed these mRNAs were mostly enriched in pathways related to cancer progression and able to regulate numerous cellular processes.

Generally, the analysis of interaction connection of genes can be used to identify key genes in that network. Here, using the cut-off criteria value of degree above 10, we identified 10 hub genes in the PPI network. Of these 10 hub genes, 8 of them were revealed could be regarded as predictor for the prognosis of BC patients. By combining the expression results in ENCORI, GEPIA, and UALCAN, we validated 4 genes (MYC, CCND1, IL6, and IGF1) were abnormally expressed in BC. Through analyzing the miRNA and mRNA expression connection in BC patients, we identified a novel hsa-miR-98-5p/IGF1 axis that may play a crucial role in BC. hsa-miR-98-5p has a controversial role in human cancers as it can function as either tumor suppressor or oncogene [22, 23]. For example, Fu and co-workers showed that downregulation of hsa-miR-98-5p could promote the development of pancreatic ductal adenocarcinoma via inversely regulating the expression of MAP4K4, indicating the tumor suppressive role of hsa-miR-98-5p [22]. On the contrary, Wang et al. showed hsa-miR-98-5p was enriched expression in epithelial ovarian cancer, and regulated the sensitive of cancer cells to cisplatin [23].

Subsequently, we validated the hsa-miR-98-5p/IGF1 axis in BC using *in vitro* experiments. Firstly, we showed hsa-miR-98-5p expression was increased, while IGF1 expression was decreased in BC cell lines compared with normal cell line, which is in consistent with the bioinformatic analyses results presented previously. Moreover, the luciferase activity reporter assay was used to validate the direct connection of hsa-miR-98-5p and IGF1. Not surprisingly, we found the introduction of hsa-miR-98-5p mimic could decreased luciferase activity in BC cells harboring wt-IGF1, indicating the direct connection of hsa-miR-98-5p and IGF1. Functional assays showed hsa-miR-98-5p overexpression could increase, while IGF1 overexpression could decrease BC cell proliferation, migration, and invasion. The overexpression of IGF1 could reversed the effects of hsa-miR-98-5p overexpression on BC cell behaviors, indicating IGF1 was a functional target for hsa-miR-98-5p. *In vivo* experiments validated the oncogenic roles of has-miR-98-5p in regulating BC progression.

Conclusion

To summary, our results demonstrated that 13 miRNAs were abnormally expressed in BC tissues compared with normal tissues. Moreover, we identified a novel hsa-miR-98-5p/IGF1 axis that may play

crucial roles in BC development through bioinformatic analyses and experiment validation. Our work presented will help us to understand the molecular mechanisms of BC and aid to develop novel therapeutic target against BC.

Declarations

Acknowledgements

None.

Authors' contributions

All authors contributed to data analysis, drafting or revising the article, gave final approval of the version to be published, and agree to be accountable for all aspects of the work.

Funding

This work was supported by Natural Fund of Liaoning Province (No. 20180550601, 20180550461, 2019-ZD-0806, and 2019-ZD-0833).

Availability of data and materials

Data are available upon request.

Ethics approval and consent to participate

Animal experiments were approved by animal care committee of The First Affiliated Hospital of Jinzhou Medical University.

Consent for publication

Informed consent was obtained from all patients.

Competing interests

The authors declare that they have no competing interests.

References

1. 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*. 2018;68(6):394-424.
2. Woolston C. Breast cancer. *Nature*. 2015;527(7578):S101.
3. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell*. 2011;144(5):646-674.

4. Zhao X, Tang DY, Zuo X, Zhang TD, Wang C. Identification of lncRNA-miRNA-mRNA regulatory network associated with epithelial ovarian cancer cisplatin-resistant. *J Cell Physiol.* 2019;234(11):19886-19894.
5. Tang D, Zhao X, Zhang L, Wang Z, Wang C. Identification of hub genes to regulate breast cancer metastasis to brain by bioinformatics analyses. *J Cell Biochem.* 2019;120(6):9522-9531.
6. Ye Y, Li SL, Wang SY. Construction and analysis of mRNA, miRNA, lncRNA, and TF regulatory networks reveal the key genes associated with prostate cancer. *PLoS One.* 2018;13(8):e0198055.
7. Wright JC, Mudge J, Weisser H, Barzine MP, Gonzalez JM, Brazma A, Choudhary JS, Harrow J. Improving GENCODE reference gene annotation using a high-stringency proteogenomics workflow. *Nat Commun.* 2016;7:11778.
8. Ezkurdia I, Juan D, Rodriguez JM, Frankish A, Diekhans M, Harrow J, Vazquez J, Valencia A, Tress ML. Multiple evidence strands suggest that there may be as few as 19,000 human protein-coding genes. *Hum Mol Genet.* 2014;23(22):5866-5878.
9. Diederichs S, Bartsch L, Berkmann JC, Fröse K, Heitmann J, Hoppe C, Iggena D, Jazmati D, Karschnia P, Linsenmeier M, Maulhardt T, Möhrmann L, Morstein J, Paffenholz SV, Röpenack P, Rückert T, Sandig L, Schell M, Steinmann A, Voss G, Wasmuth J, Weinberger ME, Wullenkord R. The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. *EMBO Mol Med.* 2016;8(5):442-457.
10. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell.* 2005;120(1):15–20.
11. O'Day E, Lal A. MicroRNAs and their target gene networks in breast cancer. *Breast Cancer Res.* 2010;12(2):201.
12. Yao Y, Liu R, Gao C, Zhang T, Qi L, Liu G, Zhang W, Wang X, Li J, Li J, Sun C. Identification of prognostic biomarkers for breast cancer based on miRNA and mRNA co-expression network. *J Cell Biochem.* 2019;120(9):15378-15388.
13. He Y, Deng F, Zhao S, Zhong S, Zhao J, Wang D, Chen X, Zhang J, Hou J, Zhang W, Ding L, Tang J, Zhou Z. Analysis of miRNA-mRNA network reveals miR-140-5p as a suppressor of breast cancer glycolysis via targeting GLUT1. *Epigenomics.* 2019;11(9):1021-1036.
14. 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.* 2014;42:D92-D97.
15. Chou CH, Shrestha S, Yang CD, Chang NW, Lin YL, Liao KW, Huang WC, Sun TH, Tu SJ, Lee WH, Chiew MY, Tai CS, Wei TY, Tsai TR, Huang HT, Wang CY, Wu HY, Ho SY, Chen PR, Chuang CH, Hsieh PJ, Wu YS, Chen WL, Li MJ, Wu YC, Huang XY, Ng FL, Buddhakosai W, Huang PC, Lan KC, Huang CY, Weng SL, Cheng YN, Liang C, Hsu WL, Huang HD. miRTarBase update 2018: A resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 2018;46(D1):D296-D302.
16. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: A software environment for integrated models of biomolecular interaction networks.

Genome Res. 2003;13(11):2498-2504.

17. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 2003;4(5):P3.
18. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* 2017;45(W1):W98-W102.
19. Nagy A, Lánckzy A, Menyhárt O, Gyórfy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. *Sci Rep.* 2018;8:9227.
20. Zheng L, Li X, Meng X, Chou J, Hu J, Zhang F, Zhang Z, Xing Y, Liu Y, Xi T. Competing endogenous RNA networks of CYP4Z1 and pseudogene CYP4Z2P confer tamoxifen resistance in breast cancer. *Mol Cell Endocrinol.* 2016;427:133-142.
21. Gao YT, Zhou YC. Long non-coding RNA (lncRNA) small nucleolar RNA host gene 7 (SNHG7) promotes breast cancer progression by sponging miRNA-381. *Eur Rev Med Pharmacol Sci.* 2019;23(15):6588-6595.
22. Fu Y, Liu X, Chen Q, Liu T, Lu C, Yu J, Miao Y, Wei J. Downregulated miR-98-5p promotes PDAC proliferation and metastasis by reversely regulating MAP4K4. *J Exp Clin Cancer Res.* 2018;37(1):130.
23. Wang Y, Bao W, Liu Y, Wang S, Xu S, Li X, Li Y, Wu S. miR-98-5p contributes to cisplatin resistance in epithelial ovarian cancer by suppressing miR-152 biogenesis via targeting Dicer1. *Cell Death Dis.* 2018;9(5):447.

Figures

a

b

Figure 1

Screening abnormally expressed miRNAs in BC tissues compared with normal tissues. a Identification of overlapping miRNAs between dataset GSE68085 and GSE83270. b Validation the expression of overlapping miRNAs at ENCORI database. miRNA: microRNA; BC: breast cancer.

Figure 2

The miRNA-mRNA interaction network in BC. miRNA: microRNA; mRNA: messenger RNA; BC: breast cancer.

Figure 3

Functional enrichment of targets for miRNAs. a Cellular component, b Molecular function, c Biological processes, and d KEGG enrichment results for these miRNA targets. miRNA: microRNA; KEGG: Kyoto Encyclopedia of Genes and Genomes.

a

b

Figure 4

PPI network of the 167 targets for miRNAs. a Using STRING, a PPI network was constructed for these 167 genes. b Top 10 hub genes derived from degree method were selected. PPI: protein-protein interaction; miRNA: microRNA.

Figure 5

PPI network analysis of targets for miRNAs for characteristics of small-world network including distribution of degrees, the proximity to the center, average aggregation coefficient, and distribution of the shortest path. PPI: protein-protein interaction; miRNA: microRNA

Figure 6

Prognostic values of these hub genes on the overall survival of BC patients. BC: breast cancer.

Figure 7

Validation of hub genes expression on BC tissues using the ENCORI, GEPIA, and UALCAN databases.

Figure 8

Connection of miRNA and the predicted targets in BC tissues as analyzed in ENCORI. miRNA: microRNA; BC: breast cancer.

Figure 9

Validation the hsa-miR-98-5p and IGF1 axis in BC cells. a hsa-miR-98-5p expression in BC cells and normal cell line. b IGF1 expression in BC cells and normal cell line. c Binding region between hsa-miR-98-5p and IGF1. d Luciferase activity in BC cells with synthetic miRNAs or luciferase vectors transfection.

hsa-miR-98-5p: microRNA-98-5p; IGF1: insulin like growth factor 1; BC: breast cancer; miR-NC: negative control miRNA; wt: wild-type; mt: mutant.

Figure 10

hsa-miR-98-5p regulates BC cell proliferation, migration, and invasion via targeting IGF1. a hsa-miR-98-5p expression, b Cell proliferation rate, c Cell migration ability, and d Cell invasion ability in BC cells with has-miR-98-5p mimic+pcDNA3.1, has-miR-98-5p mimic+pIGF1, miR-NC+pcDNA3.1, or miR-NC+pIGF1

transfection. hsa-miR-98-5p: microRNA-98-5p; IGF1: insulin like growth factor 1; BC: breast cancer; miR-NC: negative control miRNA.

Figure 11

Inhibition of has-miR-98-5p expression suppresses tumor growth in vivo. a Tumor size, and b Tumor weight in xenograft mouse model. hsa-miR-98-5p: microRNA-98-5p.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [NC3RsARRIVEGuidelinesChecklist2014.docx](#)