CircRNA expression profiles of breast cancer and construction of a circRNA-miRNA-mRNA network

Liping Xu1,4, Mengmeng Lyu2,4, Sujin Yang3, Jian Zhang3 & Dandan Yu1,*

Circular RNAs (circRNA), a group of endogenous small noncoding RNAs, were identified in the 1990s and initially thought to be useless products of RNA splicing. With the development of high-throughput sequencing, an increasing number of functional circRNAs have been discovered. CircRNAs are characterized by closed loops formed by covalent bonds and are stable enough to resist degradation by RNA exonucleases. Current knowledge shows that circRNAs function by three mechanisms: acting as competing endogenous RNAs (ceRNA), regulating gene transcription in the nucleus, and binding and coding proteins. CircRNAs are attracting growing attention because of their multiple functions in various cancers. For example, circIRAK3 regulates malignant progression of breast cancer through sequestering miR-3607, and circ-FBXW7 represses glioma tumorigenesis by encoding novel proteins.

Currently, only a few studies have reported the functions of circRNAs in breast cancer. Herein, our team acquired the circRNA expression profile in breast cancer tissues with LNM and in matched paracancer tissues using high-throughput RNA sequencing. We selected 136 significantly differentially expressed circRNAs in breast cancer and paracancer tissues, which may be involved in the invasion and metastasis of breast cancer. We also collected the miRNA and mRNA expression profiles in breast cancer and paracancer tissues from the Gene

1Department of Radiation Oncology, The First Affiliated Hospital of Nanjing Medical University, Guangzhou Road 300, Nanjing 210029, People’s Republic of China. 2Department of Gynecologic Oncology, Jiangsu Cancer Hospital and Jiangsu Institute of Cancer Research and The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, People’s Republic of China. 3Department of General Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing 210029, People’s Republic of China. 4These authors contributed equally: Liping Xu and Mengmeng Lyu. *email: njmuyudandan@163.com

Breast cancer is the most commonly occurring cancer and the dominant induction of cancer-associated death in females globally. There were about 2.3 million new cases and 685,000 deaths of breast cancer worldwide in 2020. Breast cancer can be treated by various approaches, including surgery, radiotherapy, and chemotherapy. Molecular biomarkers of breast cancer include ER, PR, Her-2, EFGR and so on. Therapies targeting hormone and her-2 receptors become standard breast cancer treatments. These effective treatment advances in recent years have brought remarkable survival gains for breast cancer patients. Nevertheless, the survival rate of distant-stage breast cancer is much lower than that of regional-stage one. Distant metastasis and lymph node metastasis (LNM) will worsen the prognosis and survival of breast cancer patients. The progression of breast cancer metastasis is complicated and involves various molecules and pathways. Therefore, searching for new biomarkers of breast cancer metastasis may provide new treatments of metastatic breast cancer.

Breast cancer is the most commonly occurring cancer and the dominant induction of cancer-associated death in females globally. There were about 2.3 million new cases and 685,000 deaths of breast cancer worldwide in 2020. Breast cancer can be treated by various approaches, including surgery, radiotherapy, and chemotherapy. Molecular biomarkers of breast cancer include ER, PR, Her-2, EFGR and so on. Therapies targeting hormone and her-2 receptors become standard breast cancer treatments. These effective treatment advances in recent years have brought remarkable survival gains for breast cancer patients. Nevertheless, the survival rate of distant-stage breast cancer is much lower than that of regional-stage one. Distant metastasis and lymph node metastasis (LNM) will worsen the prognosis and survival of breast cancer patients. The progression of breast cancer metastasis is complicated and involves various molecules and pathways. Therefore, searching for new biomarkers of breast cancer metastasis may provide new treatments of metastatic breast cancer.

CircRNAs are a group of endogenous small noncoding RNAs that are involved in multiple diseases including cancers. At present, the functions of circRNAs in breast cancer need to be further explored. In this study, 3 pairs of breast cancer and paracancer tissues with axillary lymph node metastasis were collected for circRNA high-throughput sequencing. We have identified 17,966 distinct circRNA candidates. Significant differential expressions were found in 136 circRNAs in breast cancer tissues relative to the matched paracancer tissues. We also identified differentially expressed 156 miRNAs and 1105 mRNAs in breast cancer tissues and normal breast tissues from public databases. Then we constructed a regulatory ceRNA network. 12 mRNAs were associated with prognosis of breast cancer. We also constructed a circRNAs-mediated subnetwork which might be related to prognosis of breast cancer. This article provides a better understanding of circRNAs-mediated ceRNA regulatory network by which circRNAs compete for endogenous RNAs in breast cancer.
Identification of differentially expressed circRNAs, miRNAs and mRNAs. We used paired t-test and Limma package of R/Bioconductor to identify differentially expressed circRNAs, miRNAs and mRNAs between breast cancer tissues and normal breast tissues. Differentially expressed circRNAs were filtered by absolute value of log (fold change (FC)) > 2 and adj. p value < 0.05. Differentially expressed miRNAs were filtered by absolute value of log (fold change (FC)) > 1 and adj. p value < 0.05. Differentially expressed mRNAs were filtered by absolute value of log (fold change (FC)) > 1 and adj. p value < 0.05. Heatmap package of R was used to conduct heatmaps of differentially expressed circRNAs, miRNAs and mRNAs.
Construction of the circRNA-miRNA-mRNA network. We got information about gene locations, source genes and lengths of circRNAs from circBase website (http://www.circbase.org/). We predicted target miRNAs which may bind to differentially expressed circRNAs using the cancer-specific circRNA database (CSCD) (http://gb.whu.edu.cn/cscd/). We further intersected these target miRNAs with differentially expressed miRNAs in GSE143564. Then we used miRDB and TargetScan databases to predict target mRNAs of the miRNAs intersection. We intersected these target mRNAs with differentially expressed mRNAs in GSE50428. We constructed the circRNA-miRNA-mRNA network according to circRNA-miRNA and miRNA-mRNA correspondence. The expression of miRNAs in the network was opposite to that of circRNAs and mRNAs. We used Cytoscape 3.8.0 to visualize the ceRNA network.

GO and KEGG pathway enrichment analysis. The potential functions of the differentially expressed circRNAs were further studied by exploring the target mRNAs in the ceRNA network. GO and KEGG pathway enrichment of target mRNAs were analyzed and visualized by using the org.hs.eg.db, clusterProfiler, enrichplot and ggplot2 packages of R/Bioconductor. GO is mainly used to analyze mRNA functions enriched in cell component (CC), molecular function (MF) and biological process (BP). KEGG is used to analyze mRNA functions involved in multiple pathways.

Survival analysis. To analyze the relationship between mRNAs in the ceRNA network and overall survival of breast cancer patients, we downloaded mRNA expression data and prognostic information from TCGA database. We preformed the overall survival analysis using survival and survminer packages of R/Bioconductor. The mRNAs which were significantly associated with survival time (p < 0.5) were selected to construct a subnetwork associated with prognosis.

Ethics approval and consent to participate. Our study was approved by the First Affiliated Hospital of Nanjing Medical University Ethical Committee (2021-SR-408). All included patients gave their written informed consent. All methods in our study were carried out in accordance with Declaration of Helsinki. All experimental protocols were approved by the First Affiliated Hospital of Nanjing Medical University Ethical Committee.

Results

CircRNA expression profiles of breast cancer and matched paracancer tissues. We acquired rRNA and linear-depleted RNA sequencing data of 3 pairs of breast cancer and paracancer tissues. The sequencing process and construction of ceRNA network are shown in Fig. 1. We identified 17,966 circRNA candidates with at least 2 unique backspliced reads (Fig. 2A). RPM distribution of all samples is displayed in Fig. 2B. The number and distribution of circRNAs on each chromosome are shown in Fig. 2C,D. There is no obvious difference in chromosome distribution between breast cancer tissues and normal tissues. Totally 6591 existing circRNAs in circBase (36.69%) and 11,375 novel circRNAs (63.31%) were identified. The circRNAs are mainly 100 to 1000 nucleotides long (Fig. 2E). About 10,813 circRNAs are annot_exons (60.2%) and other types are evenly distrib-
CircRNA expression data and differentially expressed circRNAs. A total of 12,676 circRNAs were identified in the first pair of samples (11,329 upregulated and 1347 downregulated). 3660 circRNAs were identified in the second pair of samples (3340 upregulated and 320 downregulated). 8690 circRNAs were identified in the third pair of samples (7545 upregulated and 1145 downregulated). The intersections of upregulated and downregulated circRNAs in 3 pairs of samples are shown in Venn diagram (Fig. 3A,B). To acquire potential functional circRNAs associated with breast cancer metastasis, we selected 136 circRNAs with significant differential expression using paired-samples t-test with adjusted $P$ value < 0.05 and fold change > 2. These 136 differentially expressed circRNAs are all upregulated in breast cancer tissues and presented in Fig. 3C and Table S1.

Differentially expressed miRNAs and mRNAs. We identified 156 miRNAs and 1105 mRNAs with differentially expression from miRNA and mRNA datasets (GSE143564 and GSE50428). The top 50 upregulated and downregulated miRNAs and mRNAs are shown in Fig. 4A,B. We predicted that 925 miRNAs might bind to the top 20 differentially expressed circRNAs from CSCD. Then we got 43 intersecting miRNAs of differentially expressed miRNAs and predicted miRNAs (Fig. 4C). We obtained 9448 potential target mRNAs of 43 intersecting miRNAs by miRDB and TargetScan databases. Then we got 453 intersecting mRNAs of differentially expressed mRNAs and target mRNAs (Fig. 4D).

Construction of a circRNA-miRNA-mRNA network. We constructed the network using 20 circRNAs, 43miRNAs and 453 mRNAs and leave out the RNAs which can't form a complete circRNA-miRNA-mRNA axis.

Figure 2. Statistics of identified circRNAs. (A) Number and backspliced reads of the circRNAs in 3 paired breast cancer tissues and matched paracancer tissues. (B) Horizontal axis: expression levels of circRNAs computed using RPM. Density is the number of genes on the horizontal axis divided by total number of expressed genes. (C,D) Number and distribution of circRNAs on each chromosome (count: number of circRNAs distributed on each chromosome). (E) Length distribution of circRNAs. (F) Type distribution of circRNAs. Compositions: Annot_exon: completely exons; Exon_intron: exons and intron; Intronic: intrones; One_exon: one exon. And locations: Antisense: on antisense chains of genes; Intergenic: on intergenic region.
Finally, there are 13 circRNAs, 14 miRNAs and 72 mRNAs in this network (Fig. 5A) and the expression of these RNAs were shown if Fig. 5B–D.

**GO and KEGG analysis.** The potential functions of the 13 selected circRNAs were further studied by exploring the 72 mRNAs through GO annotation and KEGG pathway enrichment analyses. The top 30 enriched GO terms of each class are illustrated in Fig. 6A–C. The mRNAs were partly enriched in 'non-canonical Wnt signaling pathway', 'transforming growth factor beta-activated receptor activity' and 'protein kinase regulator activity'. The KEGG pathway analysis is illustrated in Fig. 6D. The mRNAs were partly enriched in 'Chemical carcinogenesis—reactive oxygen species', 'Oxidative phosphorylation' and 'Cellular senescence'. It has been reported that oxidative phosphorylation is involved in drug resistance and metastasis of breast cancer. Alteration of the cellular senescence pathway through multiple molecular mechanisms may play an important role in the progression of breast cancer.

**Survival analysis and construction of survival-related subnetwork.** We downloaded mRNA expression data of 1053 breast cancer samples and prognostic information from TCGA database. We conducted survival analysis on the 72 mRNAs in the ceRNA network. There were 12 mRNAs significantly associated with prognosis in breast cancer patients (partly displayed in Fig. 7A). Prognostic ceRNA subnetwork was constructed with these survival-associated mRNAs (Fig. 7B).

**Discussion**

Breast cancer patients with distant metastasis and LNM are faced with worse survival prognosis. Therefore, our study focuses on metastatic breast cancer and explores the detailed mechanisms. Misregulation of transcriptome contributes to multiple diseases, including cancers. Despite many studies on linear transcriptomes, accumulating evidence indicates circRNAs participate in cancer progression. Many circRNAs have been demonstrated by the advancement of high-throughput RNA sequencing to be dysregulated in cancers, such as gastric cancer, bladder cancer and hepatocellular cancer.

Herein, we acquired the circRNA expressions of breast cancer tissues with LNM and of matched paracancer tissues using RNA sequencing. We identified 17,966 circRNA candidates, including 6591 existing circRNAs and 11,375 novel circRNAs. These novel circRNAs are of particular interest for further study. The total expression of circRNAs in breast cancer tissues was upregulated. Reportedly, however, total expression of circRNAs was downregulated in glioblastoma compared with normal brain tissues. Most circRNAs are entirely composed of exons. Then we acquired 136 significantly differential expressed circRNAs using paired-samples t-test. We note...
Figure 4. (A) Heatmap analysis of top 50 upregulated and downregulated miRNAs from miRNA dataset (GSE143564). (B) Heatmap analysis of top 50 upregulated and downregulated mRNAs from mRNA dataset (GSE50428). (C) Venn diagram showed intersecting miRNAs of differentially expressed miRNAs and predicted miRNAs. (D) Venn diagram showed intersecting mRNAs of differentially expressed mRNAs and target mRNAs.
Figure 5. (A) CeRNA network of circRNA-miRNA-mRNA interactions in breast cancer. Red indicates circRNAs, blue indicates miRNAs, and yellow indicates mRNAs. (B) Expression of the 13 circRNAs in 3 pairs of breast cancer and paracancer tissues. (C) Expression of the 14 miRNAs in GSE143564 dataset containing 3 breast cancer tissues and 3 adjacent tissues. (D) Expression of the 72 mRNAs in GSE50428 dataset containing 5 normal breast tissues and 26 breast cancer tissues.
that the significantly upregulated circRNAs are much more than the downregulated ones in breast cancer tissues compared with normal breast tissues in this study.

To explore the role of circRNAs in breast cancer, we obtained the miRNA and mRNA expression profiles from GEO and identified the miRNAs and mRNAs with differential expressions. Then we constructed a circRNA-mediated ceRNA network. We analyzed the mRNAs in the network by performing GO and KEGG pathway enrichment analyses.

**Figure 6.** GO and KEGG enrichment analyses of the 72 mRNAs in the ceRNA network: (A) biological process, (B) cellular component and (C) molecular function. (D) Enrichment analysis of KEGG pathways.

**Figure 7.** TCGA database was used to analyze survival prognosis on the 72 mRNAs in the ceRNA network. (A) The PAAM1, TGFBR1, TP53 and VBP1 gene high expression groups in breast cancer samples had worse prognosis (P < 0.05). (B) Construction of prognostic ceRNA subnetwork with these survival-associated mRNAs.
enrichment tests. Results showed the mRNAs were involved in some cancer-related GO terms, such as 'non-canonical Wnt signaling pathway' and 'transforming growth factor beta-activated receptor activity'. Some evidence shows that cancer cells acquire metastatic and migratory capacity through hijacking the non-canonical Wnt signaling pathway. For example, melanoma and gastric cancer cells acquired increased migration and metastasis while Wnt5a was overexpressed.\textsuperscript{27,28} Non-canonical Wnt/ Ca\textsuperscript{2+} signaling pathway also plays a role in cisplatin resistance of ovarian cancer.\textsuperscript{29} These results demonstrate that circRNAs in this network might play a critical role in progression of breast cancer.

To further explore the role of circRNAs in breast cancer prognosis, we carried out survival analysis on the 72 mRNAs in the ceRNA network based on TCGA database. We found that 12 mRNAs were associated with prognosis of breast cancer. Then we constructed a survival-associated ceRNA subnetwork. There were 9 circRNAs, 9miRNAs and 12 mRNAs in the network.

**Conclusions**

We showed circRNA expression profile of breast cancer tissues and identified 136 circRNAs differentially expressed between breast cancer tissue and paracancer tissues. We also identified 156 miRNAs and 1105 mRNAs with differentially expression from public database. Then we constructed a circRNAs-miRNA-mRNA network. 12 mRNAs were associated with prognosis of breast cancer based on TCGA database. We also constructed a circRNAs-mediated subnetwork which might be related to prognosis of breast cancer. The results in this study provide a feasible molecular mechanism of metastasis of breast cancer. Further studies revealing the functions of these circRNAs, which are involved in breast cancer invasion and metastasis, will provide a new perspective for treatment of metastatic breast cancer.

**Data availability**

The circRNA sequences data reported in this study was archived in the Sequence Read Archive (SRA) with the accession number PRJNA822662. Data could be available at: https://www.ncbi.nlm.nih.gov/bioproject/PRJNA822662.

Received: 24 July 2022; Accepted: 5 October 2022
Published online: 22 October 2022

**References**

1. Sung, H., Ferlay, J. & Siegel, R. L. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**(3), 209–249 (2021).
2. Berry, D. A. et al. Effect of screening and adjuvant therapy on mortality from breast cancer. *N. Engl. J. Med.* **353**(17), 1784–1792 (2005).
3. Munoz, D. et al. Effects of screening and systemic adjuvant therapy on ER-specific US breast cancer mortality. *J. Natl. Cancer Inst.* **101**(11), dju289. https://doi.org/10.1093/jnci/dju289 (2014).
4. De Santis, C. E. et al. Breast cancer statistics, 2017, racial disparity in mortality by state. *CA Cancer J. Clin.* **67**(6), 439–448 (2017).
5. Qu, S. et al. Circular RNA: A new star of noncoding RNAs. *Cancer Lett.* **365**(2), 141–148 (2015).
6. Barrett, S. P. & Salzman, J. Circular RNAs: Analysis, expression and potential functions. *Development* **143**(11), 1838–1847 (2016).
7. Li, X., Yang, L. & Chen, L. L. The biogenesis, functions, and challenges of circular RNAs. *Mol. Cell* **71**(3), 428–442 (2018).
8. He, R. et al. circGIRK1 and GIRK1 act as ceRNAs in triple negative breast cancer by regulating miR-34a. *J. Exp. Clin. Cancer Res.* **36**(1), 145 (2017).
9. Wu, J. et al. CircIRAK3 sponges miR-3607 to facilitate breast cancer metastasis. *Cancer Lett.* **430**, 179–192 (2018).
10. Zhou, J. et al. Downregulation of hasa_circ_0011946 suppresses the migration and invasion of the breast cancer cell line MCF-7 by targeting RFX3. *Cancer Manag. Res.* **10**, 535–544 (2018).
11. Liu, Y. et al. Circular RNA-MTO1 suppresses breast cancer cell viability and reverses monastrol resistance through regulating the TRA4/Eg5 axis. *Int. J. Oncol.* **53**(4), 1752–1762 (2018).
12. Yang, Y. et al. Novel role of FBXW7 circular RNA in repressing glioma tumorigenesis. *J. Natl. Cancer Inst.* **110**(3), 304–315 (2018).
13. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**(4), 357–359 (2012).
14. Kim, D. et al. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**(4), R36 (2013).
15. Memczak, S. et al. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**(7441), 333–338 (2013).
16. Evans, K. W. et al. Oxidative phosphorylation is a metabolic vulnerability in chemotherapy-resistant triple-negative breast cancer. *Cancer Res.* **81**(21), 5572–5581 (2021).
17. Ramchandani, D. et al. Copper depletion modulates mitochondrial oxidative phosphorylation to impair triple negative breast cancer metastasis. *Nat. Commun.* **12**(1), 7311 (2021).
18. Yang, D. et al. Wogonin induces cellular senescence in breast cancer via suppressing TXNRD2 expression. *Arch. Toxicol.* **94**(10), 3433–3447 (2020).
19. Ossey-Alaoui, K. et al. The Kindlin2-p53-SerpinB2 signaling axis is required for cellular senescence in breast cancer. *Cell Death Dis.* **10**(8), 539 (2019).
20. Kanekri, M. et al. KEGG: Integrating viruses and cellular organisms. *Nucleic Acids Res.* **49**(D1), D545–D551 (2021).
21. Kanekri, M. Toward understanding the origin and evolution of cellular organisms. *Protein Sci.* **28**(11), 1947–1951 (2019).
22. Kanekri, M. & Goto, S. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**(1), 27–30 (2000).
23. Lee, T. I. & Young, R. A. Transcriptional regulation and its misregulation in disease. *Cell* **152**(6), 1237–1251 (2013).
24. Xie, F. et al. Circular RNA BCRC-3 suppresses bladder cancer proliferation through miR-182-5p/p27 axis. *Mol. Cancer* **17**(1), 144 (2018).
25. Meng, J. et al. Twist1 regulates vimentin through Cul2 circular RNA to promote EMT in hepatocellular carcinoma. *Cancer Res.* **78**(15), 4150–4162 (2018).
26. Liu, H. et al. Circular RNA YAP1 inhibits the proliferation and invasion of gastric cancer cells by regulating the miR-367-5p/p27 (Kip1) axis. *Mol. Cancer* **17**(1), 151 (2018).
27. Kurayoshi, M. et al. Expression of Wnt-5a is correlated with aggressiveness of gastric cancer by stimulating cell migration and invasion. *Cancer Res.* **66**(21), 10439–10448 (2006).
28. Weeraratna, A. T. et al. Wnt5a signaling directly affects cell motility and invasion of metastatic melanoma. *Cancer Cell* **1**(3), 279–288 (2002).
29. Huang, L. et al. Role of Wnt/β-catenin, Wnt/c-Jun N-terminal kinase and Wnt/Ca(2+) pathways in cisplatin-induced chemoresistance in ovarian cancer. *Exp. Ther. Med.* **12**(6), 3851–3858 (2016).

Acknowledgements
We thank ShanLiang Zhong and Sujin Yang for useful discussions and help in revision of the present paper.

Author contributions
L.X., M.L. and D.Y. wrote the main manuscript text. J.Z. was responsible for samples collection. S.Y. prepared figures and tables. All coauthors have reviewed and approved of the manuscript prior to submission.

Funding
This work was supported by National Natural Science Foundation of China (grant number 81702613), the National Key Research and Development Program of China (No. 2016YFC0905900), National Natural Science Foundation of China (grant number 81702895) and Basic Research Program of Jiangsu Province (Natural Science Foundation, No. BK20191078).

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-21877-y.

Correspondence and requests for materials should be addressed to D.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2022