PORTLAND PRESS

Research Article

Human microRNA similarity in breast cancer

Ying Jing and 🆔 **Donghai Li**

State Key Laboratory of Pharmaceutical Biotechnology, Jiangsu Engineering Research Center for MicroRNA Biology and Biotechnology, Nanjing Advanced Institute for Life Sciences (NAILS), School of Life Sciences, Nanjing University, Jiangsu 210023, P.R. China

Correspondence: Donghai Li (donghaili@nju.edu.cn)

OPEN ACCESS

MicroRNAs (miRNAs) play important roles in a variety of human diseases, including breast cancer. A number of miRNAs are up- and down-regulated in breast cancer. However, little is known about miRNA similarity and similarity network in breast cancer. Here, a collection of 272 breast cancer-associated miRNA precursors (pre-miRNAs) were utilized to calculate similarities of sequences, target genes, pathways and functions and construct a combined similarity network. Well-characterized miRNAs and their similarity network were highlighted. Interestingly, miRNA sequence-dependent similarity networks were not identified in spite of sequence–target gene association. Similarity networks with minimum and maximum number of miRNAs originate from pathway and mature sequence, respectively. The breast cancer-associated miRNAs were divided into seven functional classes (classes I–VII) followed by disease enrichment analysis and novel miRNA-based disease similarities were found. The finding would provide insight into miRNA similarity, similarity network and disease heterogeneity in breast cancer.

Introduction

MicroRNAs (miRNAs) are endogenous non-coding RNAs with $\sim$22 nucleotides in length and inhibit target genes at the post-transcriptional level. A total of 1917 human miRNA precursors (pre-miRNAs) and 2654 mature miRNAs are available in miRBase sequence database [1]. Many human miRNAs have been recognized as gene regulators in physiological and pathophysiological conditions.

To rapidly identify and characterize novel miRNA function, network analysis is increasingly emerging. Several publications have studied miRNA functional similarity, in which Gene Ontology (GO)-based methods (iRGOFS, MFS_GO), consensus-based network inference method (miRsig), miRNA target site accessibility and topology of target gene network-based method, miRNA-disease association-based method (MISIM), graph theoretic property-based method (miRFunSim), protein–protein interaction (PPI) network-based method in combination with GO and graph theoretic properties (PPImiRFS) have been developed [2–8]. MiRNAs have a pleiotropic effect and synergistically inhibit their target genes. MiRNA synergistic networks have been constructed based on target genes, GO, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, PPI network and transcriptional regulation [9–12]. These methods and networks could shed light on function and synergism of miRNAs.

Breast cancer is a highly heterogeneous disease and the most common cancer in women worldwide. There are four major breast cancer subtypes: HR$^+$/HER2$^-$ ('Luminal A'), HR$^-$/HER2$^-$ ('Triple Negative'), HR$^+$/HER2$^+$ ('Luminal B') and HR$^-$/HER2$^+$ ('HER2-enriched'). MiRNAs serve as oncogenes or tumor suppressor genes, and expression patterns alter across breast cancer subtypes and stages. One of network studies has constructed miRNA–miRNA functional synergistic network and identified key synergistic miRNAs and potentially stronger synergistic effect of miRNA family in breast cancer [9]. Additionally, functional synergistic miRNA–miRNA networks across diverse cancer types including breast invasive cancer have been reported [11]. Apparently, a comprehensive miRNA similarity network could further improve current knowledge.

MiRNAs function in signaling pathway and diseases through binding to their target genes by sequence complementarity. MiRNA diversity and cancer heterogeneity facilitate complexity of miRNA function and miRNA–disease association. In the present study, a combined similarity network was constructed based on mature sequences, target genes, pathways and functions and enabled a classification of breast cancer-associated miRNAs.

Methods
Breast cancer-associated miRNA data sources
Regulatory miRNAs of interest were extracted from miRCancer and PubMed [13–39]. A total of 264 miRNA entries were extracted from miRCancer database, and 108 miRNAs were obtained from PubMed database.

Construction of miRNA similarity networks
miRPathDB calculated similarity scores of miRNA mature sequence, target genes or pathways [40]. Briefly, Hamming distance and Jaccard coefficient were calculated to define similarity of sequence, predicted target and target pathway, respectively. TargetScan, miRTarBase and MiRanda predicted target genes of miRNAs [41–43].

MISIM web server calculated miRNA functional similarity based on positive and negative miRNA–disease associations [44]. 'ALL vs ALL Similarity' option calculated paired functional similarity values for miRNAs. For all similarity analyses, threshold is 0.5.

Similarity networks was visualized, merged and analyzed by Cytoscape [45]. Venn diagram was done by FunRich [46].

Disease enrichment analysis
miEAA 2.0 performed enrichment analysis (analysis method: over-representation analysis, adjusted P-value <0.05) [47]. Represented were top five diseases plus breast cancer according to adjusted P-values.

Results
Collection of breast cancer-associated miRNAs
A total of 372 miRNA entries were evidenced by experimental data related to breast cancer. PubMed and miRCancer were the key miRNA data sources. Regardless of miRNA expression, the extraction had only one criterion – miRNA of interest regulated breast cancer, as demonstrated by the experiments. Notably, miRNA expression was considered for subsequent functional similarity analysis. Some reviews have summarized breast cancer-associated miRNAs which have been covered by the present miRNA list. Fifty predicted breast neoplasm-associated miRNA were found using heterogeneous information network with GraRep embedding model, and half of them almost overlapped with the present miRNA list [48]. Together, this suggested that breast cancer-associated miRNAs were successfully collected.

To calculate similarity score and construct similarity network, specially functional similarity network, the miRNA list was further screened and narrowed down to 272 pre-miRNAs. Well-characterized cancer-associated miRNAs include miR-17-92 cluster (oncogene/tumor suppressor gene), miR-21 (oncogene), miR-221/222 (oncogene), let-7 (tumor suppressor gene), miR-15/16 (tumor suppressor gene), miR-200 (tumor suppressor gene) and miR-34 (tumor suppressor gene). To simplify and optimize similarity networks, these miRNA-associated networks were emphasized and analyzed [49].

Sequence, target gene, pathway and function similarities and similarity networks
Mature sequence of miRNA is of importance to regulate its targets through sequence complementarity of target mRNA 3′-UTR. Therefore, mature sequence was used to calculate sequence similarity score based on Hamming distance method. As shown in Figure 1, sequence similarity network is the largest one comprising four subnetworks and 102 miRNAs.

Target genes were predicted by three target prediction servers and utilized to calculate similarity scores of target genes and target pathways. Similarity networks of target genes and pathways are composed of seven subnetworks (27 miRNAs) and three subnetworks (8 miRNAs), respectively. The smallest similarity network is pathway-associated.

Functional similarity network was constructed based on miRNA–disease associations. Forty-seven miRNAs formed five subnetworks by similarity score. Besides sequence similarity, these miRNAs mainly showed certain functional similarity.
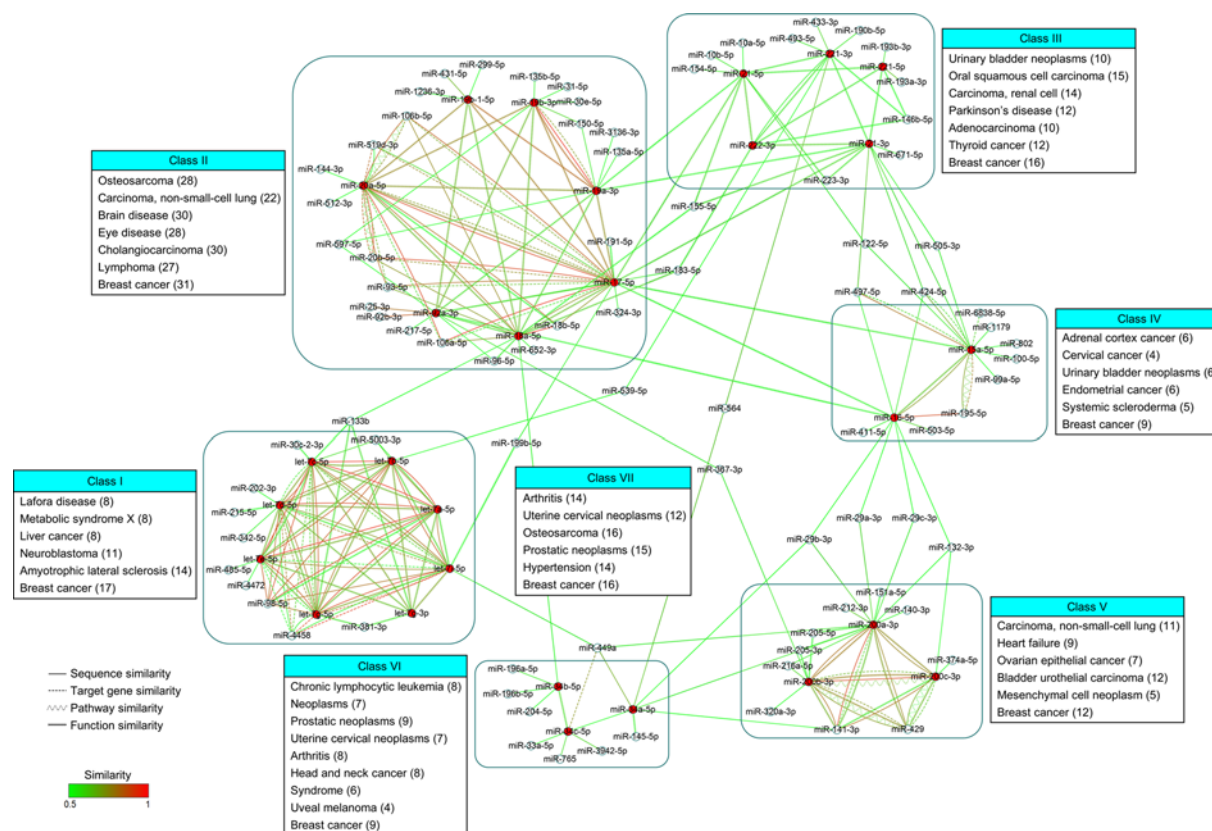
Figure 1. Similarity network, classification and enrichment analsysis of breast cancer-associated miRNAs
Emphasized cancer-associated miRNAs are displayed in red. Values in parentheses represent numbers of observed miRNAs in relevant diseases.

Although well-characterized miRNAs were emphasized and their similarity networks were extracted, only five shared miRNAs (miR-15a-5p, miR-200b-3p, miR-200c-3p, miR-20a-5p, miR-429) were found in similarity network of sequences, target genes, pathways and functions (Supplementary Figure S1). Otherwise, sequence, target gene and pathway similarity networks shared three miRNAs; sequence, target gene and function similarity networks consisted of 11 shared miRNAs. Of these shared miRNAs, five miRNAs (miR-20b-5p, miR-106b-5p, miR-141-3p, miR-195-5p, miR-429) did not belong to above-mentioned emphasized miRNAs but contributed to cancer-associated KEGG pathways, including pathways in cancer, miRNA in cancer and cell cycle (Supplementary Table S1, miRPathDB, $P < 0.05$).

Similarity network-based classification and disease enrichment analysis

To clearly annotate breast cancer-associated miRNAs, the combined similarity network was artificially classified into seven functional classes (classes I–VII) based on edge density and similarity (Figure 1). The classes contain the following emphasized miRNAs: let-7a-5p, let-7b-5p, let-7c-5p, let-7d-5p, let-7e-5p, let-7g-5p, let-7g-3p, let-7i-5p (class I), miR-17-5p, miR-18a-5p, miR-19a-3p, miR-19b-3p, miR-19b-1-5p, miR-20a-5p, miR-92a-3p (class II), miR-21-5p, miR-21-3p, miR-221-5p, miR-221-3p, miR-222-3p (class III), miR-15a-5p, miR-16-5p (class IV), miR-200a-3p, miR-200b-3p, miR-200c-3p (class V) and miR-34a-5p, miR-34b-5p, miR-34c-5p (class VI). Exceptionally, miRNAs in class VII are devoid of emphasized miRNAs and become connectors between other classes. Key genes in significant pathways of breast cancer subtypes and stages were used to inversely predict their regulatory miRNAs [50]. Surprisingly, partial miRNA connectors in class VII were related to breast cancer subtypes and stages (data not shown).

Disease enrichment analysis of miRNAs in classes I–VII showed that most miRNAs in all classes were enriched in breast cancer, and other diseases in classes I–VII were mainly implicated in cancers. Top 1 non-cancer diseases (Lafora disease, arthritis) were found in classes I and VII. Disease–symptoms network analysis has demonstrated

that there is expected similarity between Lafora disease/arthritis and cancer based on symptom similarity and protein interaction [51]. Lafora disease is a neurodegenerative disease. Neuronatin aggregation within cortical neurons is a hallmark of Lafora disease. MiR-708, a member of 272 pre-miRNAs, has been reported to target neuronatin and regulate metastatic breast cancer [52]. Arthritis is a chronic condition involved in joint pain or joint disease. A potential arthritis–breast cancer association is that arthritis drugs could prevent breast cancer bone metastasis [53].

Discussion

MiRNA similarity network is useful for functional annotation of novel miRNA, miRNA–disease association and classification in a unique disease. To better recognize similarity network in single cancer, a combined similarity network was constructed based on mature sequences, target genes, pathways and functions of breast cancer-associated miRNAs. Classification and enrichment analysis of these miRNAs were further performed.

Target prediction algorithms of miRNAs primarily pay attention to miRNA mature sequences, especially seed sequences. These miRNA of interest mostly showed the similarity of mature sequences establishing a maximum similarity network. However, the sequence similarity did not lead to identical similarity of target genes, pathways and functions. An possible explanation for this is that different similarity algorithm may result in inconsistent similarity scores. For functional analysis of miRNAs, miRNA–disease association is successfully used to calculate functional similarity of miRNAs. To date, several publications have proposed miRNA–disease association methods, including combinatorial prioritization-based method, network similarity integration method, multiple features-based method, network-based inference method and its combinational method with disease phenotypic similarity-based inference [54–56]. In non-model species without disease database, miRNA similarity method has been also developed [8]. To minimize the biases and represent a high-quality similarity network, similarity comparisons of mature sequences, target genes, pathways and functions of breast cancer-associated miRNAs were performed and applied for constructing a combined similarity network. Combination of four independent subnetworks provided an overview of miRNA similarity network and highlighted the relationships between individual similarity networks.

Network analysis is able to investigate molecular and biomedical networks, including PPI network, function linkage network, transcriptional regulatory network, metabolic network, phenotype network and drug–target network. In 2015, a knowledge-based computational framework 'network fingerprint' was introduced to analyze biomedical networks [57]. Disease network analysis identified disease–disease and disease–signaling pathway associations. As published previously, a combinational non-miRNA method of semantic and gene function association (SemFunSim) has been put forward and compared with other multiple methods [58]. A transcriptomics and epigenomics data driven approach has identified disease similarity at the molecular level [59]. A biological network (disease–miRNA, miRNA–gene, disease–gene and protein–protein interactome)-based method (mpDisNet) has been established to infer disease–disease relationship [60]. As shown in Figure 1, enrichment analyses of classes I–VII miRNAs pointed out possible miRNA-based disease similarity. Previous studies on disease similarity are partially consistent with the present miRNA-based disease similarity such as breast cancer-associated similarities of adenocarcinoma, adrenal cortex neoplasms, amyotrophic lateral sclerosis, arthritis, brain disease, carcinoma (non-small-cell lung), carcinoma (renal cell), cholangiocarcinoma, endometrial neoplasms, head and neck neoplasms, heart failure, hypertension, leukemia (lymphocytic), liver neoplasms, lymphoma, melanoma, metabolic syndrome X, neoplasms, neuroblastoma, osteosarcoma, ovarian neoplasmas, Parkinson's disease, prostatic neoplasms, systemic scleroderma, thyroid neoplasms, urinary bladder neoplasms, uterine cervical neoplasms and various syndromes [51,58,60]. Although eye disease was found in class II, similarity between breast cancer and eye disease had been not definitely demonstrated. Twenty-eight breast cancer-associated miRNAs in class II contributed to eye disease enrichment and were also found in ncRNAomics-eye disease database Nc2Eye [61]. Most miRNAs in class II have a role in age-related macular degeneration. Another supporting result was from mpDisNet analysis, which showed breast cancer-associated similarities of cataract, diabetic retinopathy, glaucoma, Graves' disease and retinoblastoma. Other cancers are class III-associated oral squamous cell carcinoma, class IV-associated cervical cancer and class V-associated bladder urothelial carcinoma, mesenchymal cell neoplasm and ovarian epithelial cancer. Otherwise, classes I–VII are not independent and linked by similarity of sequence and function derived from miRNA–disease association. Functional similarity is a unique linker between classes II and III, and classes II and VI, suggesting that miRNAs in classes II–VI have disease-based functional similarity and may co-regulate diseases.

Significance of miRNAs as diagnostic, prognostic biomarkers and therapeutic targets to clinic has been evaluated in basic research or clinical trial [14]. Partial miRNAs in each miRNA class are potential diagnostic, prognostic biomarkers. MiRNA signatures are essential for stage and subtype identifications in breast cancer [62–65]. MiRNAs in classes II–V and classes I–IV, VII contributed to stage and subtype of breast cancer, respectively. In contrast with

predicted result, few experimental data have shown that miRNA connectors in class VII significantly refer to stage and subtype in breast cancer. Oncogenes and tumor suppressor genes could be therapeutic targets. These potentially therapeutic targets are distributed across all classes. Oncology network analysis would help to understand molecular mechanism underlying tumorigenesis and cancer progression. Different from similarity network analysis, a study focused on the change in miRNA–target regulation across subtypes in non-small cell lung cancer to construct a miRNA dysregulational synergistic network, discover miRNA dysregulatory modules and select potential biomarker [66]. For a combined similarity network, mapping of miRNAs with novel physiological and pathophysiolgical roles could infer molecular mechanism underlying breast cancer and potential application for biomarker and therapeutic target based on well-characterized breast cancer-associated miRNAs. Apparently, functional complementarity or redundancy of miRNAs in classes I–VII would influence treatment of breast cancer. In-depth experiments are also needed to support the similarity network and elucidate molecular mechanisms of breast cancer.

Brown adipose tissue (BAT) with a hallmark of mitochondrial uncoupling protein 1 acts as a key therapeutical target to mitigate obesity via energy expenditure. It is suggested that white adipose tissue (WAT) could enhance thermogenesis by the process of 'browning'. MiRNAs also modulate brite/brown adipogenesis [67], intriguingly, in which regulatory miRNAs are almost encompassed in breast cancer-associated miRNA list. It suggests that adipose tissue regulation and breast cancer have specific similarity at regulatory miRNA level. Also, obesity is one of risk factors for breast cancer and menopausal status, and breast cancer is a factor associated with higher incidence of BAT activation.

In summary, a combined similarity network was constructed to elucidate similarity of sequences, target genes, pathways and functions of breast cancer-associated miRNAs. The similarity network analysis may promote the understanding of sequence, target gene, pathway, function and disease associations of miRNAs.

Data Availability

The detailed methods and datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing Interests

The authors declare that there are no competing interests associated with the manuscript.

Funding

CRediT Author Contribution

Ying Jing: Investigation. **Donghai Li:** Conceptualization, Methodology, Investigation, Writing–review & editing, Supervision, Project administration, Funding acquisition.

Abbreviations

BAT, brown adipose tissue; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; miRNA, microRNA; PPI, protein–protein interaction; pre-miRNA, microRNA precursor; WAT, white adipose tissue.

References

1 Kozomara, A., Birgaoanu, M. and Griffiths-Jones, S. (2019) miRBase: from microRNA sequences to function. *Nucleic Acids Res.* **47**, D155–D162, https://doi.org/10.1093/nar/gky1141

2 Yang, Y., Fu, X., Qu, W., Xiao, Y. and Shen, H.B. (2018) MiRGOFS: a GO-based functional similarity measurement for miRNAs, with applications to the prediction of miRNA subcellular localization and miRNA-disease association. *Bioinformatics* **34**, 3547–3556, https://doi.org/10.1093/bioinformatics/bty343

3 Nalluri, J.J., Barh, D., Azevedo, V. and Ghosh, P. (2017) miRsig: a consensus-based network inference methodology to identify pan-cancer miRNA-miRNA interaction signatures. *Sci. Rep.* **7**, 39684, https://doi.org/10.1038/srep39684

4 Wang, D., Wang, J., Lu, M., Song, F. and Cui, Q. (2010) Inferring the human microRNA functional similarity and functional network based on microRNA-associated diseases. *Bioinformatics* **26**, 1644–1650, https://doi.org/10.1093/bioinformatics/btq241

5 Sun, J., Zhou, M., Yang, H., Deng, J., Wang, L. and Wang, Q. (2013) Inferring potential microRNA-microRNA associations based on targeting propensity and connectivity in the context of protein interaction network. *PLoS ONE* **8**, e69719, https://doi.org/10.1371/journal.pone.0069719

6 Meng, J., Liu, D. and Luan, Y. (2015) Inferring plant microRNA functional similarity using a weighted protein-protein interaction network. *BMC Bioinformatics* **16**, 361, https://doi.org/10.1186/s12859-015-0789-4

7   Luo, J., Dai, D., Cao, B. and Yin, Y. (2016) Inferring human miRNA functional similarity based on gene ontology annotations. *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pp. 1407–1413, IEEE, Changsha, China

8   Xu, Y., Guo, M., Liu, X., Wang, C. and Liu, Y. (2014) Inferring the soybean (*Glycine max*) microRNA functional network based on target gene network. *Bioinformatics* **30**, 94–103, https://doi.org/10.1093/bioinformatics/btt605

9   Geng, X., Kong, X., Chen, Q., Su, S. and Hu, Y. (2017) Analysis of miRNA functional synergistic network in breast cancer. *ICBBS '17: 6th International Conference on Bioinformatics and Biomedical Science*, pp. 22–29, Association for Computing Machinery, New York, NY, United States

10 Xu, J., Li, C.X., Li, Y.S., Lv, J.Y., Ma, Y., Shao, T.T. et al. (2011) MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. *Nucleic Acids Res.* **39**, 825–836, https://doi.org/10.1093/nar/gkq832

11 Meng, X., Wang, J., Yuan, C., Li, X., Zhou, Y., Hofestadt, R. et al. (2015) CancerNet: a database for decoding multilevel molecular interactions across diverse cancer types. *Oncogenesis* **4**, e177, https://doi.org/10.1038/oncsis.2015.40

12 Zhang, J., Duy Le, T., Liu, L., He, J. and Li, J. (2016) Identifying miRNA synergistic regulatory networks in heterogeneous human data via network motifs. *Mol. Biosyst.* **12**, 454–463, https://doi.org/10.1039/C5MB00562K

13 Xie, B., Ding, Q., Han, H. and Wu, D. (2013) miRCancer: a microRNA-cancer association database constructed by text mining on literature. *Bioinformatics* **29**, 638–644, https://doi.org/10.1093/bioinformatics/btt014

14 Abolghasemi, M., Tehrani, S.S., Yousefi, T., Karimian, A., Mahmoodpoor, A., Ghamari, A. et al. (2020) MicroRNAs in breast cancer: Roles, functions, and mechanism of actions. *J. Cell. Physiol.* **235**, 5008–5029, https://doi.org/10.1002/jcp.29396

15 Cava, C., Novello, C., Martelli, C., Lodico, A., Ottobrini, L., Piccotti, F. et al. (2020) Theranostic application of miR-429 in HER2+ breast cancer. *Theranostics* **10**, 50–61, https://doi.org/10.7150/thno.36274

16 Purohit, P.K., Edwards, R., Tokatlidis, K. and Saini, N. (2019) MiR-195 regulates mitochondrial function by targeting mitofusin-2 in breast cancer cells. *RNA Biol.* **16**, 918–929, https://doi.org/10.1080/15476286.2019.1600999

17 Xu, F., Li, H. and Hu, C. (2019) MiR-202 inhibits cell proliferation, invasion, and migration in breast cancer by targeting ROCK1 gene. *J. Cell. Biochem.* **120**, 16008–16018, https://doi.org/10.1002/jcb.28879

18 Qiu, C., Huang, F., Zhang, Q., Chen, W. and Zhang, H. (2019) miR-205-3p promotes proliferation and reduces apoptosis of breast cancer MCF-7 cells and is associated with poor prognosis of breast cancer patients. *J. Clin. Lab. Anal.* **33**, e22966, https://doi.org/10.1002/jcla.22966

19 Cui, Y., Wang, J., Liu, S., Qu, D., Jin, H., Zhu, L. et al. (2019) miR-216a promotes breast cancer cell apoptosis by targeting PKCα. *Fundam. Clin. Pharmacol.* **33**, 397–404, https://doi.org/10.1111/fcp.12481

20 Lindholm, E.M., Leivonen, S.K., Undlien, E., Nebdal, D., Git, A., Caldas, C. et al. (2019) miR-342-5p as a potential regulator of HER2 breast cancer cell growth. *MicroRNA* **8**, 155–165, https://doi.org/10.2174/2211536608666181206124922

21 Rodriguez-Barrueco, R., Nekritz, E.A., Bertucci, F., Yu, J., Sanchez-Garcia, F., Zeleke, T.Z. et al. (2017) miR-424(322)/503 is a breast cancer tumor suppressor whose loss promotes resistance to chemotherapy. *Genes Dev.* **31**, 553–566, https://doi.org/10.1101/gad.292318.116

22 Wang, L. and Wang, J. (2012) MicroRNA-mediated breast cancer metastasis: from primary site to distant organs. *Oncogene* **31**, 2499–2511, https://doi.org/10.1038/onc.2011.444

23 Li, L., Xiao, B., Tong, H., Xie, F., Zhang, Z. and Xiao, G.G. (2012) Regulation of breast cancer tumorigenesis and metastasis by miRNAs. *Expert Rev. Proteomics* **9**, 615–625, https://doi.org/10.1586/epr.12.64

24 Corcoran, C., Friel, A.M., Duffy, M.J., Crown, J. and O'Driscoll, L. (2011) Intracellular and extracellular microRNAs in breast cancer. *Clin. Chem.* **57**, 18–32, https://doi.org/10.1373/clinchem.2010.150730

25 Greene, S.B., Herschkowitz, J.I. and Rosen, J.M. (2010) Small players with big roles: microRNAs as targets to inhibit breast cancer progression. *Curr. Drug Targets* **11**, 1059–1073, https://doi.org/10.2174/138945010792006762

26 Wu, J., Miao, J., Ding, Y., Zhang, Y., Huang, X., Zhou, X. et al. (2019) MiR-4458 inhibits breast cancer cell growth, migration, and invasiveness by targeting CPSF4. *Biochem. Cell Biol.* **97**, 722–730, https://doi.org/10.1139/bcb-2019-0008

27 Liang, T.C., Fu, W.G. and Zhong, Y.S. (2019) MicroRNA-1236-3p inhibits proliferation and invasion of breast cancer cells by targeting ZEB1. *Eur. Rev. Med. Pharmacol. Sci.* **23**, 9988–9995

28 Zhang, N., Zeng, X., Sun, C., Guo, H., Wang, T., Wei, L. et al. (2019) LncRNA LINC00963 promotes tumorigenesis and radioresistance in breast cancer by sponging miR-324-3p and inducing ACK1 expression. *Mol. Ther. Nucleic Acids* **18**, 871–881, https://doi.org/10.1016/j.omtn.2019.09.033

29 Zhang, L., Liu, Y., Song, F., Zheng, H., Hu, L., Lu, H. et al. (2011) Functional SNP in the microRNA-367 binding site in the 3'UTR of the calcium channel ryanodine receptor gene 3 (RYR3) affects breast cancer risk and calcification. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 13653–13658, https://doi.org/10.1073/pnas.1103360108

30 Zhao, J., Zou, H., Han, C., Ma, J., Zhao, J. and Tang, J. (2018) Circlular RNA BARD1 (Hsa_circ_0001098) overexpression in breast cancer cells with TCDD treatment could promote cell apoptosis via miR-3942/BARD1 axis. *Cell Cycle* **17**, 2731–2744, https://doi.org/10.1080/15384101.2018.1556058

31 Zhao, L., Feng, X., Song, X., Zhou, H., Zhao, Y., Cheng, L. et al. (2016) miR-493-5p attenuates the invasiveness and tumorigenicity in human breast cancer by targeting FUT4. *Oncol. Rep.* **36**, 1007–1015, https://doi.org/10.3892/or.2016.4882

32 Zhao, P., Guan, H., Dai, Z., Ma, Y., Zhao, Y. and Liu, D. (2019) Long noncoding RNA DLX6-AS1 promotes breast cancer progression via miR-505-3p/RUNX2 axis. *Eur. J. Pharmacol.* **865**, 172778, https://doi.org/10.1016/j.ejphar.2019.172778

33 Duan, W.J., Bi, P.D., Ma, Y., Liu, N.Q. and Zhen, X. (2020) MiR-512-3p regulates malignant tumor behavior and multi-drug resistance in breast cancer cells via targeting Livin. *Neoplasma* **67**, 102–110, https://doi.org/10.4149/neo˙2019˙190106N18

34 Liu, G., Wang, P. and Zhang, H. (2019) MiR-6838-5p suppresses cell metastasis and the EMT process in triple-negative breast cancer by targeting WNT3A to inhibit the Wnt pathway. *J. Gene Med.* **21**, e3129, https://doi.org/10.1002/jgm.3129

35 Jiao, Y., Yuan, C., Wu, H., Li, X. and Yu, J. (2020) Oncogenic microRNA-765 promotes the growth and metastasis of breast carcinoma by directly targeting ING4. *J. Cell. Biochem.* **121**, 3887–3900, https://doi.org/10.1002/jcb.29545

36 Li, M., Pan, M., You, C., Zhao, F., Wu, D., Guo, M. et al. (2020) MiR-7 reduces the BCSC subset by inhibiting XIST to modulate the miR-92b/Slug/ESA axis and inhibit tumor growth. *Breast Cancer Res.* **22**, 26, https://doi.org/10.1186/s13058-020-01264-z

37 Pu, T., Shen, M., Li, S., Yang, L., Gao, H., Xiao, L. et al. (2019) Repression of miR-135b-5p promotes metastasis of early-stage breast cancer by regulating downstream target SDCBP. *Lab. Invest.* **99**, 1296–1308, https://doi.org/10.1038/s41374-019-0258-1

38 Sun, T., Song, Y., Yu, H. and Luo, X. (2019) Identification of lncRNA TRPM2-AS/miR-140-3p/PYCR1 axis's proliferates and anti-apoptotic effect on breast cancer using co-expression network analysis. *Cancer Biol. Ther.* **20**, 760–773, https://doi.org/10.1080/15384047.2018.1564563

39 Zhang, Z.J. and Ma, S.L. (2012) miRNAs in breast cancer tumorigenesis. *Oncol. Rep.* **27**, 903–910, https://doi.org/10.3892/or.2011.1611

40 Kehl, T., Kern, F., Backes, C., Fehlmann, T., Stockel, D., Meese, E. et al. (2020) miRPathDB 2.0: a novel release of the miRNA Pathway Dictionary Database. *Nucleic Acids Res.* **48**, D142–D147, https://doi.org/10.1093/nar/gkz1022

41 Agarwal, V., Bell, G.W., Nam, J.W. and Bartel, D.P. (2015) Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**, e05005, https://doi.org/10.7554/eLife.05005

42 Chou, C.H., Shrestha, S., Yang, C.D., Chang, N.W., Lin, Y.L., Liao, K.W. et al. (2018) miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* **46**, D296–D302, https://doi.org/10.1093/nar/gkx1067

43 Enright, A.J., John, B., Gaul, U., Tuschl, T., Sander, C. and Marks, D.S. (2003) MicroRNA targets in Drosophila. *Genome Biol.* **5**, R1, https://doi.org/10.1186/gb-2003-5-1-r1

44 Li, J., Zhang, S., Wan, Y., Zhao, Y., Shi, J., Zhou, Y. et al. (2019) MISIM v2.0: a web server for inferring microRNA functional similarity based on microRNA-disease associations. *Nucleic Acids Res.* **47**, W536–W541, https://doi.org/10.1093/nar/gkz328

45 Kohl, M., Wiese, S. and Warscheid, B. (2011) Cytoscape: software for visualization and analysis of biological networks. *Methods Mol. Biol.* **696**, 291–303, https://doi.org/10.1007/978-1-60761-987-1˙18

46 Pathan, M., Keerthikumar, S., Ang, C.S., Gangoda, L., Quek, C.Y., Williamson, N.A. et al. (2015) FunRich: an open access standalone functional enrichment and interaction network analysis tool. *Proteomics* **15**, 2597–2601, https://doi.org/10.1002/pmic.201400515

47 Kern, F., Fehlmann, T., Solomon, J., Schwed, L., Grammes, N., Backes, C. et al. (2020) miEAA 2.0: integrating multi-species microRNA enrichment analysis and workflow management systems. *Nucleic Acids Res.* **48**, W521–W528, https://doi.org/10.1093/nar/gkaa309

48 Ji, B.Y., You, Z.H., Cheng, L., Zhou, J.R., Alghazzawi, D. and Li, L.P. (2020) Predicting miRNA-disease association from heterogeneous information network with GraRep embedding model. *Sci. Rep.* **10**, 6658, https://doi.org/10.1038/s41598-020-63735-9

49 Hayes, J., Peruzzi, P.P. and Lawler, S. (2014) MicroRNAs in cancer: biomarkers, functions and therapy. *Trends Mol. Med.* **20**, 460–469, https://doi.org/10.1016/j.molmed.2014.06.005

50 Bourdakou, M.M., Athanasiadis, E.I. and Spyrou, G.M. (2016) Discovering gene re-ranking efficiency and conserved gene-gene relationships derived from gene co-expression network analysis on breast cancer data. *Sci. Rep.* **6**, 20518, https://doi.org/10.1038/srep20518

51 Zhou, X., Menche, J., Barabasi, A.L. and Sharma, A. (2014) Human symptoms-disease network. *Nat. Commun.* **5**, 4212, https://doi.org/10.1038/ncomms5212

52 Ryu, S., McDonnell, K., Choi, H., Gao, D., Hahn, M., Joshi, N. et al. (2013) Suppression of miRNA-708 by polycomb group promotes metastases by calcium-induced cell migration. *Cancer Cell* **23**, 63–76, https://doi.org/10.1016/j.ccr.2012.11.019

53 Eyre, R., Alferez, D.G., Santiago-Gomez, A., Spence, K., McConnell, J.C., Hart, C. et al. (2019) Microenvironmental IL1β promotes breast cancer metastatic colonisation in the bone via activation of Wnt signalling. *Nat. Commun.* **10**, 5016, https://doi.org/10.1038/s41467-019-12807-0

54 Yu, H., Chen, X. and Lu, L. (2017) Large-scale prediction of microRNA-disease associations by combinatorial prioritization algorithm. *Sci. Rep.* **7**, 43792, https://doi.org/10.1038/srep43792

55 Li, X.Y., Lin, Y.P. and Gu, C.L. (2017) A network similarity integration method for predicting microRNA-disease associations. *RSC Adv.* **7**, 32216–32224, https://doi.org/10.1039/C7RA05348G

56 Chen, H., Guo, R., Li, G., Zhang, W. and Zhang, Z. (2020) Comparative analysis of similarity measurements in miRNAs with applications to miRNA-disease association predictions. *BMC Bioinformatics* **21**, 176, https://doi.org/10.1186/s12859-020-3515-9

57 Cui, X., He, H., He, F., Wang, S., Li, F. and Bo, X. (2015) Network fingerprint: a knowledge-based characterization of biomedical networks. *Sci. Rep.* **5**, 13286, https://doi.org/10.1038/srep13286

58 Cheng, L., Li, J., Ju, P., Peng, J. and Wang, Y. (2014) SemFunSim: a new method for measuring disease similarity by integrating semantic and gene functional association. *PLoS ONE* **9**, e99415, https://doi.org/10.1371/journal.pone.0099415

59 Lakiotaki, K., Georgakopoulos, G., Castanas, E., Roe, O.D., Borboudakis, G. and Tsamardinos, I. (2019) A data driven approach reveals disease similarity on a molecular level. *NPJ Syst. Biol. Appl.* **5**, 39, https://doi.org/10.1038/s41540-019-0117-0

60 Jin, S., Zeng, X., Fang, J., Lin, J., Chan, S.Y., Erzurum, S.C. et al. (2019) A network-based approach to uncover microRNA-mediated disease comorbidities and potential pathobiological implications. *NPJ Syst. Biol. Appl.* **5**, 41, https://doi.org/10.1038/s41540-019-0115-2

61 Zhang, Y., Xue, Z., Guo, F., Yu, F., Xu, L. and Chen, H. (2020) Nc2Eye: a curated ncRNAomics knowledgebase for bridging basic and clinical research in eye diseases. *Front. Cell Dev. Biol.* **8**, 75, https://doi.org/10.3389/fcell.2020.00075

62 Farazi, T.A., Ten Hoeve, J.J., Brown, M., Mihailovic, A., Horlings, H.M., van de Vijver, M.J. et al. (2014) Identification of distinct miRNA target regulation between breast cancer molecular subtypes using AGO2-PAR-CLIP and patient datasets. *Genome Biol.* **15**, R9, https://doi.org/10.1186/gb-2014-15-1-r9

63 Yerukala Sathipati, S. and Ho, S.Y. (2018) Identifying a miRNA signature for predicting the stage of breast cancer. *Sci. Rep.* **8**, 16138, https://doi.org/10.1038/s41598-018-34604-3

64 Bhattacharyya, M., Nath, J. and Bandyopadhyay, S. (2015) MicroRNA signatures highlight new breast cancer subtypes. *Gene* **556**, 192–198, https://doi.org/10.1016/j.gene.2014.11.053

65 Kurozumi, S., Yamaguchi, Y., Kurosumi, M., Ohira, M., Matsumoto, H. and Horiguchi, J. (2017) Recent trends in microRNA research into breast cancer with particular focus on the associations between microRNAs and intrinsic subtypes. *J. Hum. Genet.* **62**, 15–24, https://doi.org/10.1038/jhg.2016.89

66 Tran, N., Abhyankar, V., Nguyen, K., Weidanz, J. and Gao, J. (2018) MicroRNA dysregulational synergistic network: discovering microRNA dysregulatory modules across subtypes in non-small cell lung cancers. *BMC Bioinformatics* **19**, 504, https://doi.org/10.1186/s12859-018-2536-0

67 Lorente-Cebrian, S., Herrera, K., Milagro, F.I., Sanchez, J., de la Garza, A.L. and Castro, H. (2019) miRNAs and novel food compounds related to the browning process. *Int. J. Mol. Sci.* **20**, 5998, https://doi.org/10.3390/ijms20235998

Supplementary Data

Human microRNA similarity in breast cancer

Ying Jing[1], Donghai Li[1*]

[1]State Key Laboratory of Pharmaceutical Biotechnology, Jiangsu Engineering Research Center for MicroRNA Biology and Biotechnology, Nanjing Advanced Institute for Life Sciences (NAILS), School of Life Sciences, Nanjing University, Jiangsu 210023, P.R. China.

* Correspondence:

Donghai Li

State Key Laboratory of Pharmaceutical Biotechnology, Jiangsu Engineering Research Center for MicroRNA Biology and Biotechnology, Nanjing Advanced Institute for Life Sciences (NAILS), School of Life Sciences, Nanjing University, Jiangsu 210023, P.R. China

E-mail: donghaili@nju.edu.cn
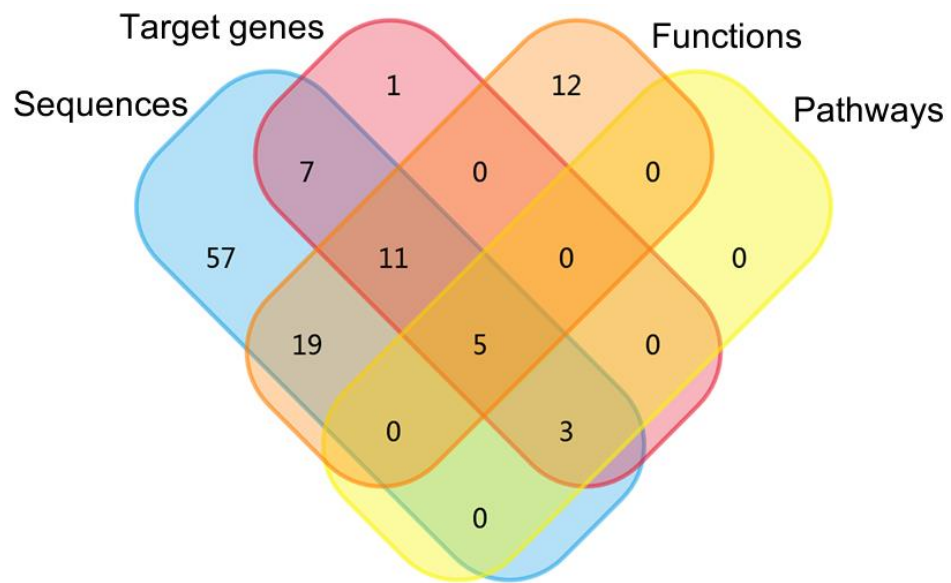
Supplementary Figure S1. Venn diagram comparing the miRNAs related to sequences, target genes, pathways and functions in similarity networks.

Supplementary Table S1. Pathway analysis of hsa-miR-20b-5p, hsa-miR-106b-5p, hsa-miR-141-3p, hsa-miR-195-5p and hsa-miR-429 (miRPathDB, *p*<0.05).

MiRNAs	Pathway	*p*-value	Targets
hsa-miR-20b-5p	Pathways in cancer	5.00E-04	*AKT3, CDKN1A, FZD6, HIF1A, MAPK9, PPARG, PTEN, STAT3, VEGFA*
hsa-miR-106b-5p	Cell cycle	4.38E-05	*CCND1, CCND2, CDKN1A, E2F1, E2F5, RB1, RBL1, RBL2, WEE1*
hsa-miR-141-3p	MiRNAs in cancer	2.82E-06	*CDC25A, CDC25C, CYP1B1, E2F3, MCL1, MDM4, PTEN, SHC1, TGFB2, ZEB1, ZEB2, ZFPM2*
hsa-miR-195-5p	Pathways in cancer	3.99E-07	*ABL1, AKT2, APC2, AR, ARAF, ARNT, ARNT2, AXIN2, BCL2, BCL2L1, BCR, BID, BIRC5, BMP2, CASP3, CASP9, CBL, CBLB, CBLC, CCDC6, CCND1, CCNE1, CCNE2, CDC42, CDH1, CDK4, CDK6, CDKN1A, CDKN2B, CHUK, CKS1B, COL4A1, COL4A3, COL4A4, CRK, CRKL, CTBP2, CTNNA1, CTNNA3, CXCL8, CYCS, DAPK3, DCC, DVL1, DVL2, DVL3, E2F1, E2F2, E2F3, EGFR, EGLN1, EGLN3, ELOB, ELOC, FGF1, FGF10, FGF11, FGF14, FGF18, FGF19, FGF2, FGF20, FGF23, FGF5, FGF8, FGF9, FGFR1, FGFR2, FGFR3, FLT3, FOS, FOXO1, FZD10, FZD3, FZD4, FZD6, FZD7, FZD8, FZD9, GLI2, GLI3, GSK3B, HDAC2, HHIP, HIF1A, HRAS, HSP90AA1, HSP90B1, IGF1, ITGA2, ITGB1, JAK1, JUP, KIT, KLK3, KRAS, LAMA4, LAMC1, LAMC3, LEF1, MAP2K1, MAP2K2, MAPK1, MAPK10, MAPK3, MAPK8, MAPK9, MAX, MECOM, MITF, MSH2, MSH3, MSH6, NFKB1, NKX3-1, NOS2, PAX8, PDGFA, PDGFRA, PDGFRB, PIAS2, PIK3CA, PIK3CB, PIK3CG, PIK3R1, PIK3R2, PIK3R3, PIK3R5, PLCG1, PLD1, PML, PPARD, PPARG, PRKCA, PRKCG, PTCH1, PTCH2, PTEN, PTGS2, RAC2, RAF1, RALB, RALBP1, RALGDS, RARB, RASSF1, RBX1, RET, RHOA, RUNX1, RUNX1T1, RXRA, SHH, SMAD2, SMAD3, SMO, SOS2, STAT1, STAT3, STAT5A, STAT5B, STK36, SUFU, TCF7, TCF7L2, TFG, TGFA, TGFB2, TGFBR1, TGFBR2, TP53, TPM3, TRAF1, TRAF2, TRAF3, TRAF4, TRAF5, TRAF6, VEGFA, VEGFB, VEGFD, WNT2B, WNT3, WNT3A, WNT4, WNT5B, WNT7A, WNT7B, WNT8A, WNT8B, WNT9A, WNT9B, XIAP, ZBTB16*
hsa-miR-429	MiRNAs in cancer	8.22E-15	*BCL2, BCL2L11, CDKN1B, CRKL, DDIT4, DNMT1, EP300, EZH2, FSCN1, IRS1, KRAS, MAPK7, MYC, PLCG1, PTEN, SHC1, VEGFA, ZEB1, ZEB2, ZFPM2*