

Plasma extracellular vesicle-packaged microRNAs as candidate diagnostic biomarkers for early-stage breast cancer

CHEN YAN^{1,2*}, JINTAO HU^{3*}, YIPENG YANG¹, HONG HU¹, DONGXIAN ZHOU¹, MIN MA⁴ and NAN XU¹

¹Department of Breast Surgery, The Second Clinical Medical College (Shenzhen People's Hospital) of Jinan University, Shenzhen, Guangdong 518020; ²Integrated Chinese and Western Medicine Postdoctoral Research Station, Jinan University, Guangzhou, Guangdong 510632; ³Department of Pathology, The Second Clinical Medical College (Shenzhen People's Hospital) of Jinan University, Shenzhen, Guangdong 518020; ⁴College of Traditional Chinese Medicine, Institute of Integrated Traditional Chinese and Western Medicine, Jinan University, Guangzhou, Guangdong 510632, P.R. China

Received December 18, 2018; Accepted June 12, 2019

DOI: 10.3892/mmr.2019.10669

Abstract. Extracellular vesicle-packaged microRNAs (miRNAs) are a class of circulating miRNAs located in the plasma that are packaged into extracellular vesicles. The present study examined the expression profiles of extracellular vesicles and tissue miRNAs with the aim of investigating the miRNA signatures in early-stage breast cancer. The present study identified and compared the extracellular vesicle-packaged miRNA expression signature and tissue miRNA expression signature from healthy individuals (n=10) and patients with early-stage breast cancer (n=12). A total of five miRNAs, including miRNA-375, miRNA-24-2-5p, miRNA-548b-5p, miRNA-655-3p and miRNA-376b-5p, were synchronized in extracellular vesicles and tissues of the breast cancer group when compared with the healthy group. The highest area under the curve (AUC) for a single miRNA was achieved with miRNA-548b-5p [AUC=0.785; 95% confidence interval (CI)=0.585-0.984; P=0.022]. The highest overall AUC was achieved by the combination of miRNA-375, miRNA-655-3p, miRNA-548b-5p and miRNA-24-2-5p (AUC=0.808; 95% CI=0.629-0.986; P=0.013). The Kaplan-Meier curves and log test analysis results of these five miRNAs, especially those

for miRNA-548b-5p, were partly consistent with the hypothesis. Two miRNAs (miRNA-548b-5p and miRNA-376b-5p) were positively associated with patient survival, while two miRNAs (miRNA-375 and miRNA-24-2-5p) were negatively associated with patient survival. The present study provided a set of plasma extracellular vesicle-packaged miRNA-based biomarkers for the diagnosis of early-stage breast cancer.

Introduction

Breast cancer (BC) is the most common malignancy in women, and one of the three most common cancer types worldwide (1). Early-stage breast cancer without detectable distant metastases is a potentially curable disease (1). The current *in vivo* diagnostic tools for the detection of early-stage BC include mammography and ultrasound (2). Due to the limited sensitivity of traditional diagnostic methods, certain micro-molecules such as microRNAs (miRNAs) have been considered as potential biomarkers for early-stage BC diagnosis (3,4).

miRNAs are a class of small, evolutionarily conserved, non-coding RNAs that are 18-25 nucleotides in length (5). miRNAs are capable of inducing translational repression or degradation of target mRNAs by binding to their 3' untranslated regions (3'UTRs), and participate in almost all key cellular processes (5). Extracellular vesicle-packaged miRNAs are a class of circulating miRNAs that are packaged into extracellular vesicle and can be detected in the serum (6,7). Extracellular vesicle-packaged miRNAs have systemic effects in primary breast cancer and contribute to processes within the blood circulation (8,9).

Extracellular vesicle-packaged miRNAs have the potential to serve as biomarkers for evaluating breast cancer, and numerous extracellular vesicle-packaged miRNAs indicative of breast cancer have been identified (10,11). For example, a systematic review and meta-analysis suggested that miRNA-21 was a potential biomarker for the early diagnosis of breast cancer, with high sensitivity and specificity (12). In addition, the combination of five serum miRNAs, including miRNA-1248, miRNA-1307-3p, miRNA-4634, miRNA-6861-5p and miRNA-6875-5p, has been reported to be able to detect early stage breast cancer with a sensitivity as high as 98% (13). In

Correspondence to: Professor Nan Xu, Department of Breast Surgery, The Second Clinical Medical College (Shenzhen People's Hospital) of Jinan University, 1017 Dongmen Road, Shenzhen, Guangdong 518020, P.R. China
E-mail: xunan2005@163.com

Professor Min Ma, College of Traditional Chinese Medicine, Institute of Integrated Traditional Chinese and Western Medicine, Jinan University, 601 Huangpu West Road, Guangzhou, Guangdong 510632, P.R. China
E-mail: tmamin@jnu.edu.cn

*Contributed equally

Key words: microRNA, extracellular vesicle, plasma, tissue, biomarker, breast cancer

addition, serum extracellular vesicle-packaged miRNA-373 was reported to be associated with more aggressive breast cancer (14).

As a potential biomarker for early-stage breast cancer diagnosis, plasma extracellular vesicle-packaged miRNA detection is a non-invasive procedure when compared with diagnostic procedures involving tissue biomarkers.

In order to screen potential extracellular vesicle-packaged miRNAs for early-stage breast cancer diagnosis, the present study attempted to analyze the extracellular vesicle-packaged miRNA expression profile in blood and tissue clinical samples. Specifically, the profiles of extracellular vesicle-packaged miRNAs extracted from the plasma of patients with early-stage breast cancer were compared with those of the control group. The profiles of miRNAs extracted from cancer tissues of patients with early-stage BC were also compared with the normal breast tissues of the control group. The four sets of data were analyzed in order to identify the extracellular vesicle-packaged miRNAs that can be used as diagnosis biomarkers of early-stage breast cancer.

Materials and methods

Patient cohorts. In the BC group, plasma samples were collected from patients with Stage I breast cancer (T1N0M0) (1), who did not have other systemic diseases or cancer at the time of their initial diagnosis, prior to receiving any treatment. BC tissues from these patients were collected at the time of surgery. All the patients involved were women. The age range of patients was 20–63 years.

In the control group, plasma samples were collected from patients with benign breast disease, including breast fibroadenoma and mammary adenosis; these patients also did not have other systemic diseases or cancer at the time of collection. The plasma samples were collected at the time of diagnosis prior to surgery, and normal breast tissue from the same patient was collected during surgery, which served as the benign lesion sample. The diagnoses of the patients were confirmed by clinical and pathological findings. A summary of the clinicopathological characteristics of control and BC groups is presented in Table I. All experiments were performed in accordance with the approved guidelines and regulations of the Shenzhen People's Hospital. All experimental protocols were approved by the Ethics Committees of Shenzhen People's Hospital. Written informed consent was obtained from all participants, in accordance with the Declaration of Helsinki.

Sample preparation. Venous blood samples (6 ml/patient) were collected in tubes containing 10.8 mg EDTA as the anticoagulant (BD Biosciences). Plasma was separated from blood cells by centrifugation at 1,000 × g for 10 min at 4°C. The plasma layer was further centrifuged at 16,000 × g for 10 min at 4°C prior to extracellular vesicle extraction. Tissue samples were identified by a clinical pathologist and cut into 1 cm³ sections for storage at -80°C until further processing.

Extraction of plasma extracellular vesicles. Extracellular vesicles were isolated from the plasma using ExoQuick precipitation (System Biosciences, LLC), according to the manufacturer's protocol. Briefly, the extracellular vesicles

were precipitated by incubation; the ExoQuick extracellular vesicle precipitation reagent was added at 4°C for 60 min. The extracellular vesicle pellet was collected by centrifugation at 1,500 × g for 10 min at 4°C and resuspended in 10 mM PBS to four times the original plasma volume.

Transmission electron microscopy. A copper mesh was placed on a clean wax plate and 100 μl of the extracellular vesicle suspension was added. After 4 min, the copper mesh was removed and placed in 2% phosphotungstic acid for 5 min at room temperature. The mesh was dried on a filter paper. The mesh was examined under a transmission electron microscope (HT7700; Hitachi, Ltd.) and images were captured.

Western blot analysis. The extracellular vesicle pellet was dissolved in RIPA protein lysis buffer (Sangon Biotech, Co., Ltd.), and the protein concentration was determined using the bicinchoninic acid protein assay kit (Thermo Fisher Scientific, Inc.). The proteins were separated on an SDS-PAGE gel (the concentrations of separating gel and stacking gel were 12 and 5% respectively). A total of 30 μg of protein was loaded per lane, while the protein ladder (Thermo Fisher Scientific, Inc.; 26616) was loaded 3 μl per lane. The proteins were electrotransferred onto a PVDF membrane. Membranes were blocked with 5% skimmed milk in TBST at room temperature for 1 h. The membrane was then incubated with the following primary antibodies at 4°C overnight: Tumor susceptibility gene 101 protein (TSG101; Abcam; ab83, dilution 1:1,000), CD9 (Abcam; ab92762, dilution 1:1,000) and heat shock cognate 70 protein (HSC70; Santa Cruz; sc7298, dilution 1:300). The membrane was then incubated with the HRP-conjugated corresponding secondary antibodies (Proteintech; SA00001-1, dilution 1:1,000 and CST; 7074P2, dilution 1:10,000) at room temperature for 1 h and treated with enhanced chemiluminescence detection reagents (4A Biotech; 4AW012-100). The specific protein bands were visualized and recorded on film using a chemiluminescence imaging system (2000; Tanon).

RNA preparation and miRNA sequencing. Total RNA was extracted from the extracellular vesicles and tissues using TRIzol[®] reagent (Life Technologies; Thermo Fisher Scientific, Inc.), according to the manufacturer's protocol. The purity of the isolated RNA was determined according to the ratio of radiation absorbance at 260 nm to that at 280 nm, using a Nanodrop ND-1000 system (Thermo Fisher Scientific, Inc.); 200 ng total RNA was used for the construction of the miRNA library using the NEBNext Multiplex Small RNA Library Prep Set for Illumina kit (New England BioLabs, Inc.). Each RNA sample underwent adaptor ligation, cDNA synthesis and polymerase chain reaction amplification. The quality of the RNA library was verified by Agilent 2100 Bioanalyzer (Agilent Technologies, Inc.). The library was sequenced on the Hiseq 2500 (Illumina Inc.). RNA-seq datasets were generated with 250 bp paired-end reads with a 10 million-read sequencing depth of the multiplexed samples.

miRNA data processing. Fastx_toolkit (http://hannonlab.cshl.edu/fastx_toolkit/) was used to remove low quality reads from the sequencing data. A quality statistic was performed using FastQC software (version 0.13.0) (<https://www>.

Table I. Characteristics of the study participants.

Factor	Breast cancer group, n=12	Control group, n=10
No. of blood samples	12	10
No. of tissue samples	5	7
Age, years	46.83±7.70	35.80±12.45
Tumor size (cm)	1.29±0.37	-
Estrogen receptor status		
Positive	10 (83%)	-
Negative	2 (17%)	-
Progesterone receptor status		
Positive	8 (67%)	-
Negative	4 (33%)	-
HER2 status		
Positive	3 (25%)	-
Negative	9 (75%)	-
Ki-67 index (%)		
<10%	2 (16%)	-
10-20%	5 (45%)	-
>20%	5 (42%)	-
Molecular subtype		
Luminal A	5 (42%)	-
HER2 positive	-	-
Triple-negative breast cancer	-	-
Luminal B	7 (58%)	-

bioinformatics.babraham.ac.uk/projects/fastqc/) as a quality control tool for the high-throughput sequencing data. All reads were compared with the reference sequence, human mature miRNA from miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw/php/index.php>), using the FastQC high precision sequence alignment algorithm (15). Gene expression was quantified in reads per kilobase of transcript per million mapped reads (RPKM) (16,17).

Bioinformatics analysis. Differential miRNA expression analysis was performed. EdgeR was used to determine the differentially expressed miRNAs (18). $P < 0.05$ and the absolute value of \log_2 (FoldChange) < 1 ($|\log_2$ (FoldChange)| > 1) were considered to indicate a statistically significant difference. Clustering analysis was performed using a hierarchical clustering method based on the expression level of differentially expressed miRNAs [$\log_{10}(\text{RPKM}+1)$]. miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw/php/index.php>) was used to identify potential human miRNA target genes. Gene Ontology (GO) analysis was performed using topGO (version 2.18.0) (<https://bioconductor.org/packages/release/bioc/html/topGO.html>) to investigate the biological processes, cellular components and specific molecular functions of the identified differentially expressed coding genes. Pathway analysis was performed using KOBAS (kobas2.0-20150126) (18) to determine the involvement of co-expressed genes in different biological pathways according to the Kyoto Encyclopedia of Genes and Genomes (KEGG).

The hypergeometric test was used for statistical analysis. miRNA-target gene-network analysis was performed using mirWalk (<http://mirwalk.umm.uni-heidelberg.de/>) (19). A receiver operating characteristic (ROC) curve was generated to calculate the relationship between sensitivity and specificity for the disease group compared with the healthy controls. Data analysis and ROC curve analysis was performed using SPSS (version 19.0; IBM Corp.) and GraphPad Prism (version 7; GraphPad Software, Inc.).

The Kaplan-Meier method was used to compute the survival analyses using OncoLnc (<http://oncolnc.org/>) and Oncomir (<http://www.oncomir.org>). OncoLnc is a tool for interactively exploring survival correlations, and for downloading clinical data coupled to expression data for mRNAs, miRNAs or long noncoding RNAs (lncRNAs). OncoLnc contains survival data for 8,647 patients from 21 cancer studies performed by The Cancer Genome Atlas (TCGA; <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>), along with miRNAs from TCGA. OncomiR is an online resource for exploring miRNA dysregulation in cancer. Using combined miRNA-seq, RNA-seq and clinical data from TCGA, it is able systematically perform statistical analyses to identify dysregulated miRNAs that are associated with tumor development and progression in the majority of cancer types (20).

Statistical analysis. The analysis of differentially expressed miRNAs was evaluated using R studio software (version 3.4.2) with the edgeR package (21,22). Differential expression was assessed for each gene using an exact test analogous to Fisher's exact test, but adapted for over-dispersed data (22). $P < 0.05$ was considered to indicate a statistically significant difference. Data are expressed as the median with range, and were analyzed using GraphPad Prism software (version 7.0). Pearson correlation analysis was performed to assess the correlation between age and miRNA expression profile using SPSS (version 19). The scatter plot was drawn using SPSS (version 19).

Results

Extracellular vesicles isolation and validation. The present study initially assessed the extracellular vesicles that were isolated from the plasma samples. Transmission electron microscopy revealed spherical vesicles of 30-100 nm in diameter in all samples (Fig. 1A), suggesting that the present study had successfully purified extracellular vesicles from the plasma. According to the standard curve, the mean protein concentration of extracellular vesicles in the breast cancer group was 24.16 $\mu\text{g}/\mu\text{l}$, and was 23.84 $\mu\text{g}/\mu\text{l}$ in the control group. Western blotting was performed to detect three conventional extracellular vesicle protein markers, namely TSG101, CD9 and HSC70 (15). With an equal amount of protein loaded in each lane, these three extracellular vesicle protein markers were all highly enriched in the isolated extracellular vesicles relative to the plasma (Fig. 1B). These results confirmed the successful purification of intact extracellular vesicles from all plasma samples.

Extracellular vesicle-packaged miRNA expression profiles in breast cancer and control groups. The optical density

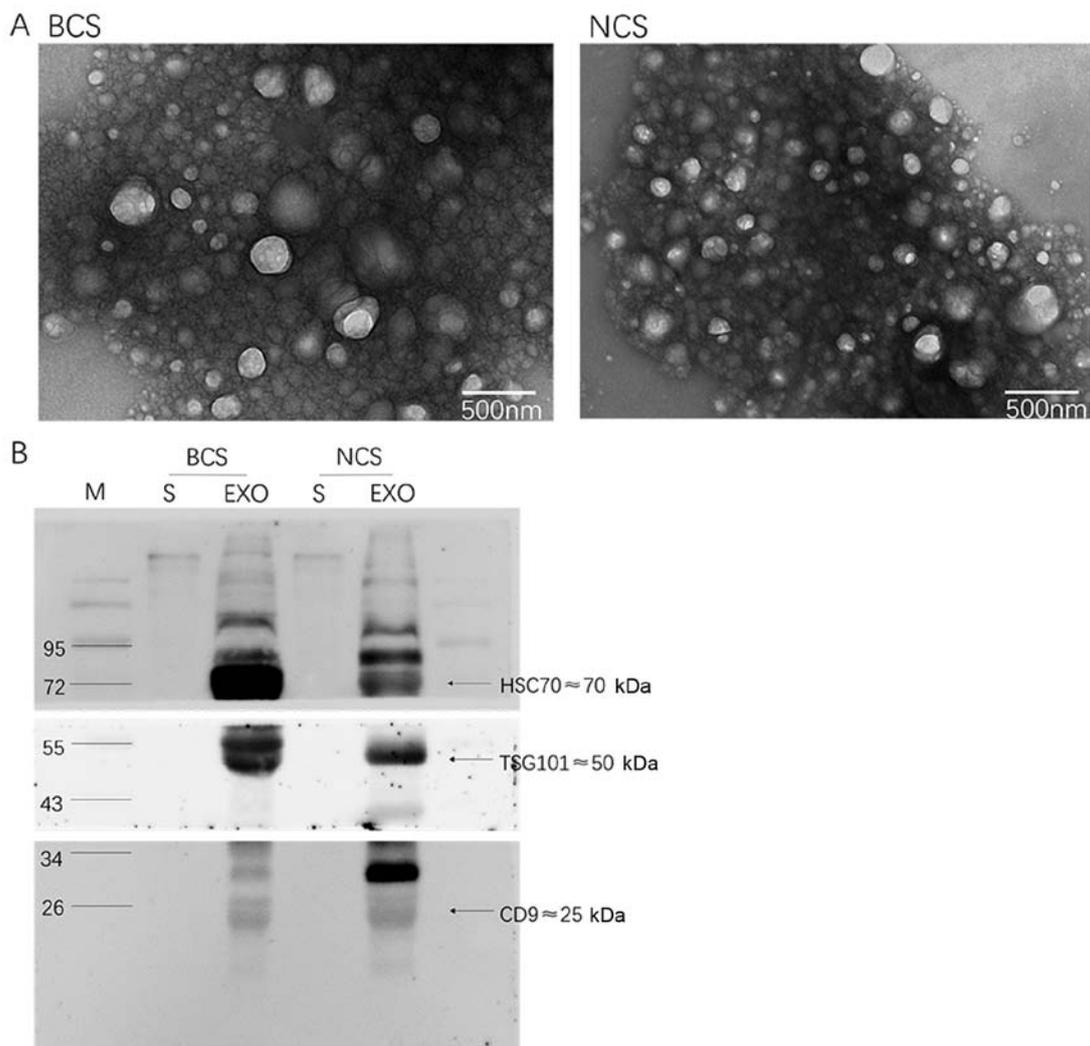


Figure 1. Characterization of plasma extracellular vesicles. (A) Electron microscopy image of extracellular vesicles isolated from healthy control and breast cancer patients. (B) Western blot analysis of the extracellular vesicle protein markers HSC70, TSG101 and CD9 in total plasma and isolated extracellular vesicles. BCS, breast cancer serum; NCS, normal control serum; HSC70, heat shock cognate 71 kDa protein; TSG101, tumor susceptibility gene 101 protein; M, marker; S, serum; EXO, exosome.

(OD)260/OD280 value range of total RNA extracted from extracellular vesicles in each samples ranged between 1.70 and 2.00, which was indicative of high RNA purity. Differential miRNA expression analysis was performed to identify the candidate extracellular vesicle-packaged miRNAs. A series of bioinformatics analyses were performed to reveal the relationship between extracellular vesicle-packaged miRNAs and the pathogenesis and progression of early-stage breast cancer. The edgeR method, based on negative binomial distribution, was adopted. When compared with the control group, there were 41 differentially expressed miRNAs out of 1,111 miRNAs in the blood samples of the BC group, of which 21 were significantly upregulated, while 20 were significantly downregulated miRNAs ($P < 0.05$). In addition, when compared with the control group, there were 181 differentially expressed miRNAs in the tissue samples of the BC group, of which 102 were significantly upregulated, while 79 were significantly downregulated ($P < 0.05$) (Table II). A volcano plot illustrating the distribution of differential miRNA between breast cancer and control group is presented in Fig. 2A.

Hierarchical clustering was performed using R software based on the RPKM values of differentially expressed miRNAs. Hierarchical clustering analysis revealed differences in the extracellular vesicle-packaged miRNA expression in the blood samples between the breast cancer and control groups (Fig. 2B).

GO and pathway analysis. miRTarBase was used to identify potential human miRNA target genes of differentially expressed miRNAs. GO was used to classify the functions of the target genes from three structured networks, including biological processes, cellular components and molecular function.

In the present study, GO enrichment was performed in blood and tissue samples comparing the early-stage breast cancer group with the control group (Fig. 3). In the biological process analysis, the 20 most prominent terms in the blood and tissue samples were compared, and 9 out of 20 terms were identical, including 'cellular macromolecule metabolic process', 'regulation of gene expression' and 'RNA metabolic

Table II. List of differentially expressed extracellular particle miRNAs between the two groups.

Differential expression	Blood samples, early-stage BC vs. control	Tissue samples, early-stage BC vs. control
Significantly upregulated miRNA (P<0.05)	21	102
Significantly downregulated miRNA (P<0.05)	20	79

miRNA, microRNA; BC, breast cancer.

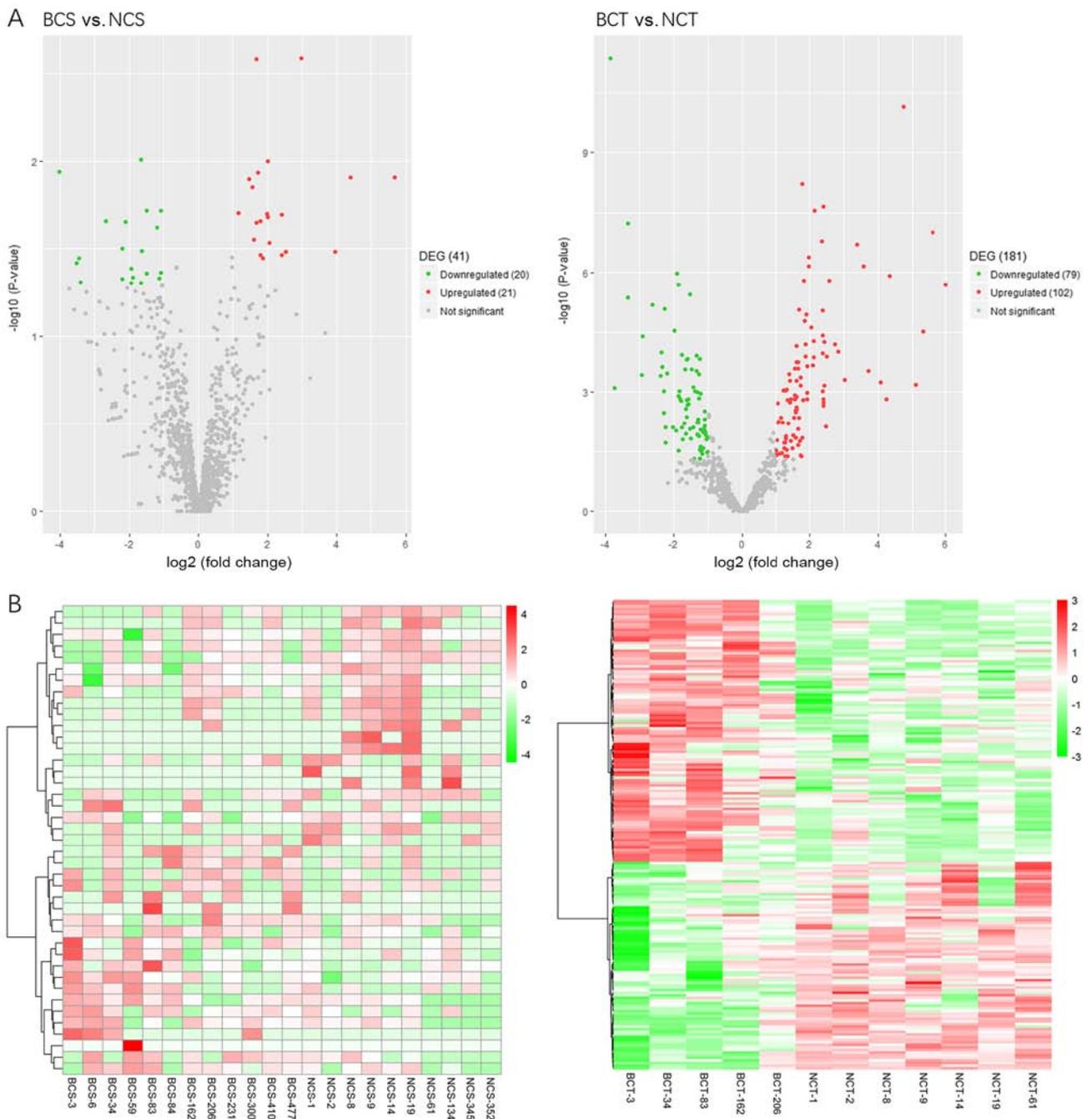


Figure 2. Differences in extracellular vesicle-packaged miRNA expression. (A) Volcano plot of extracellular vesicle-packaged miRNAs; the red spots represent the significantly upregulated miRNAs in the breast cancer group [$\log_2(\text{FoldChange}) > 1$ and $P < 0.05$], while the green spots represent the significantly downregulated miRNAs in the breast cancer group [$\log_2(\text{FoldChange}) < -1$ and $P < 0.05$]. edgeR based on the negative binomial distribution was used for the statistical analysis. (B) Heatmap of extracellular vesicle differential miRNAs profiles in the breast cancer and control groups. Cluster analysis was performed using the $\log_{10}(\text{RPKM}+1)$ values. The upregulated miRNAs are marked in red, while the downregulated miRNAs are marked in green. miRNA, microRNA; RPKM, reads per kilobase of transcript per million mapped reads; DEG, differentially expressed gene; BCS, breast cancer serum; NCS, normal control serum; BCT, breast cancer tissue; NCT, normal control tissue.

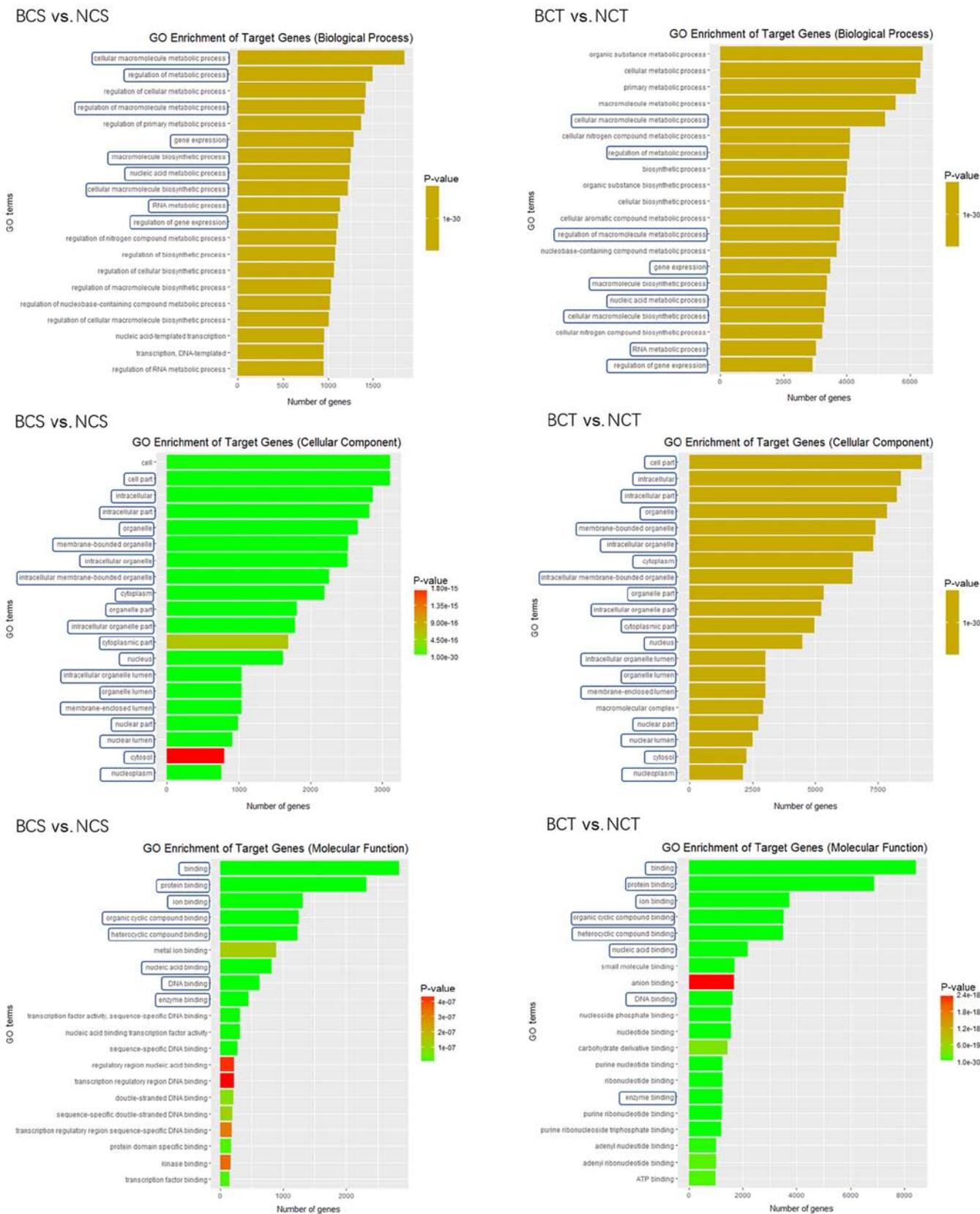


Figure 3. GO analysis. The top 20 most significantly changed GO terms of the predicted target genes are illustrated. The GO analysis included biological processes, cellular components and molecular function. The method of statistical analysis used was analogous to Fisher's exact test. Terms tagged with box were the same terms in both analyses. GO, gene ontology; BCS, breast cancer serum; NCS, normal control serum; BCT, breast cancer tissue; NCT, normal control tissue.

process'. In the cellular components analysis, the present study contrasted the 20 most prominent terms in the blood and tissue

samples, and 19 out of 20 terms were identical, including 'nucleoplasm', 'cytosol' and 'nuclear lumen'. In the molecular

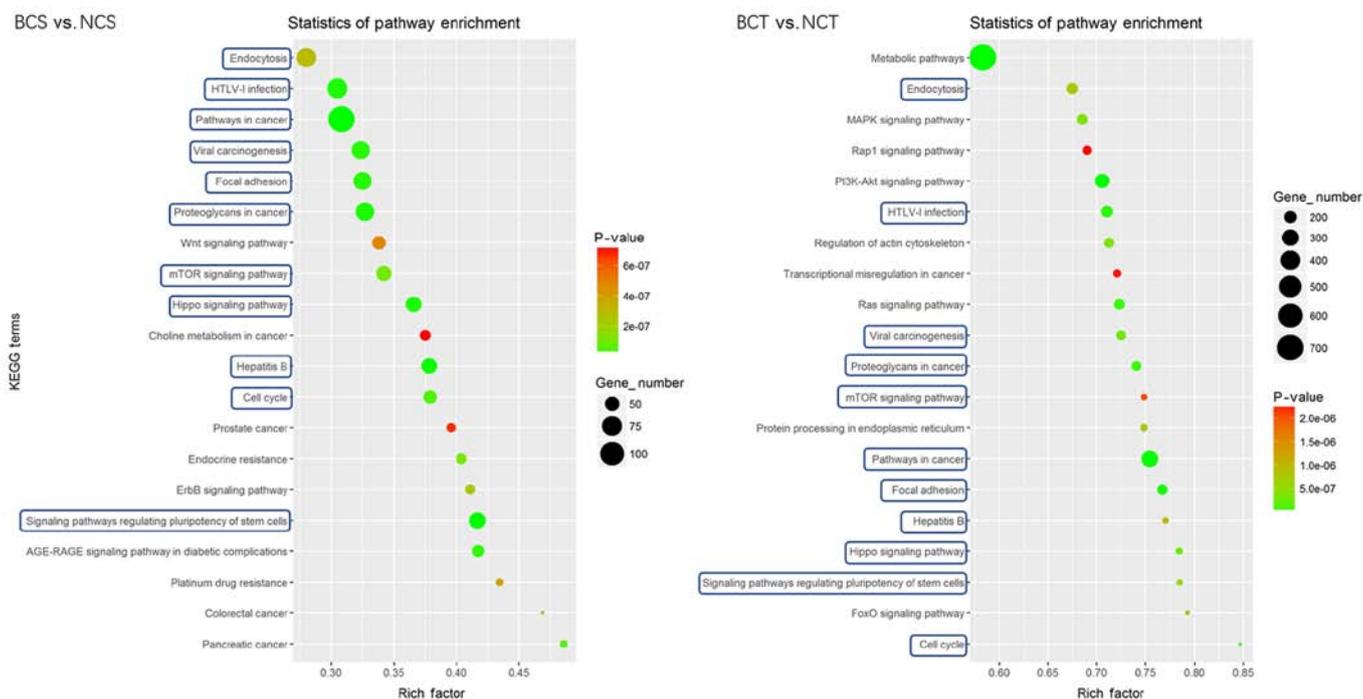


Figure 4. KEGG analysis. The 20 most significantly changed pathways of the predicted target genes are illustrated. The method of statistical analysis was the hypergeometric test. Terms indicated with boxes appear in both analyses. BCS, breast cancer serum; NCS, normal control serum; BCT, breast cancer tissue; NCT, normal control tissue; KEGG, Kyoto Encyclopedia of Genes and Genomes.

function analysis, the present study contrasted the 20 most prominent terms in the blood and tissue samples, and 8 out of 20 terms were observed to be identical, including 'enzyme binding', 'DNA binding' and 'nucleic acid binding'.

KEGG pathway analysis was performed to identify significantly affected pathways. The results indicated that in the blood samples, the target genes of differentially expressed miRNAs were enriched in 'ErbB signaling pathway', 'cell cycle', 'signaling pathways regulating pluripotency of stem cells' and 'pathways in cancer'. In the tissue samples, the target genes of differentially expressed miRNAs were enriched in 'PI3K-Akt signaling pathway', 'MAPK signaling pathway', 'mTOR signaling pathway' and 'pathways in cancer'. Comparison of the 20 most prominent terms of the KEGG pathway analysis in the blood and tissue samples revealed that 11 out of 20 terms were the same, including 'signaling pathways regulating pluripotency of stem cells', 'pathways in cancer', 'cell cycle', 'proteoglycans in cancer' and the 'mTOR signaling pathway' (Fig. 4).

Cross-referencing of the blood and tissue analyses. Comparison of the early BC group with the control group revealed that six miRNAs were significantly altered in the blood and tissue analyses (Table III and Fig. 5). miRNA-205-5p was significantly upregulated in the blood samples ($P=0.011650$), yet was significantly downregulated in the tissue samples ($P=0.000304$). Excluding miRNA-205-5p, a total of five miRNAs exhibited identical changing trends in the blood and tissue analyses. These miRNAs included miRNA-375, miRNA-24-2-5p, miRNA-548b-5p, miRNA-655-3p and miRNA-376b-5p. Network analysis of miRNAs and target genes was performed based on the five miRNAs, and it was revealed that these five miRNAs were associated with the same target genes and

regulated each other, thereby affecting biological processes (Fig. 6). At the same time, the statistical analysis of patient's age showed that the mean age between the breast cancer group and control group was significantly different ($P=0.019$). However, bivariate correlation analysis revealed that there was no significant correlation between these 5 miRNAs (miR-375, miR-24-2-5p, miR-548b-5p, miR-655-3p and miR-376b-5p) and age in the blood groups (Table SI and Fig. S1).

Evaluation of extracellular vesicle-packaged miRNAs as early-stage breast cancer diagnostic biomarkers. To evaluate the utility of extracellular vesicle-packaged miRNA levels for discriminating cases of early-stage BC from controls, ROC curve analysis was performed. A total of five miRNAs identified in the primary analyses were highly differentially expressed between patients with early-stage BC and controls (Fig. 7). The highest area under the curve (AUC) for a single miRNA was achieved with miRNA-548b-5p [AUC=0.785; 95% confidence interval (CI)=0.585-0.984; $P=0.022$].

In addition, the different combinations of these miRNAs were investigated (these results are presented in Table SI), and the highest AUC was achieved via the combination of miRNA-375, miRNA-24-2-5p, miRNA-548b-5p and miRNA-655-3p (AUC=0.808; 95% CI=0.629-0.986; $P=0.013$). An increase in the AUC was observed when miRNA-375 was combined with other miRNAs (Fig. 7), such as miRNA-375 + miRNA-655-3p (AUC=0.762; 95% CI=0.557-0.966; $P=0.035$), miRNA-375 + miRNA-655-3p + miRNA-548b-5p (AUC=0.792; 95% CI=0.605-0.979; $P=0.018$).

Because the highest AUC was achieved via the combination of miRNA-375, miRNA-24-2-5p, miRNA-548b-5p and miRNA-655-3p, KEGG pathway analysis of these four

Table III. Overexpressed and downregulated miRNAs in blood and tissues from BC patients as compared to the control group.

miRNA	P-value	FC (log ₂)	Upregulated/downregulated
miRNA-375			
BCS vs. NCS	0.019806	1.169516	Upregulated
BCT vs. NCT	8.92x10 ⁻⁶	2.378307	Upregulated
miRNA-24-2-5p			
BCS vs. NCS	0.020269	2.415220	Upregulated
BCT vs. NCT	0.016187	1.026315	Upregulated
miRNA-548b-5p			
BCS vs. NCS	0.023904	-1.197655	Downregulated
BCT vs. NCT	0.007955	-1.418292	Downregulated
miRNA-655-3p			
BCS vs. NCS	0.041027	-1.954611	Downregulated
BCT vs. NCT	0.008961	-1.211824	Downregulated
miRNA-376b-5p			
BCS vs. NCS	0.049708	-1.948349	Downregulated
BCT vs. NCT	0.024284	-1.217745	Downregulated
miRNA-205-5p			
BCS vs. NCS	0.011650	1.726517	Upregulated
BCT vs. NCT	0.000304	-1.333712	Downregulated

BCS, breast cancer serum; NCS, normal control serum; BCT, breast cancer tissue; NCT, normal control tissue; miRNA, microRNA; FC, fold change.

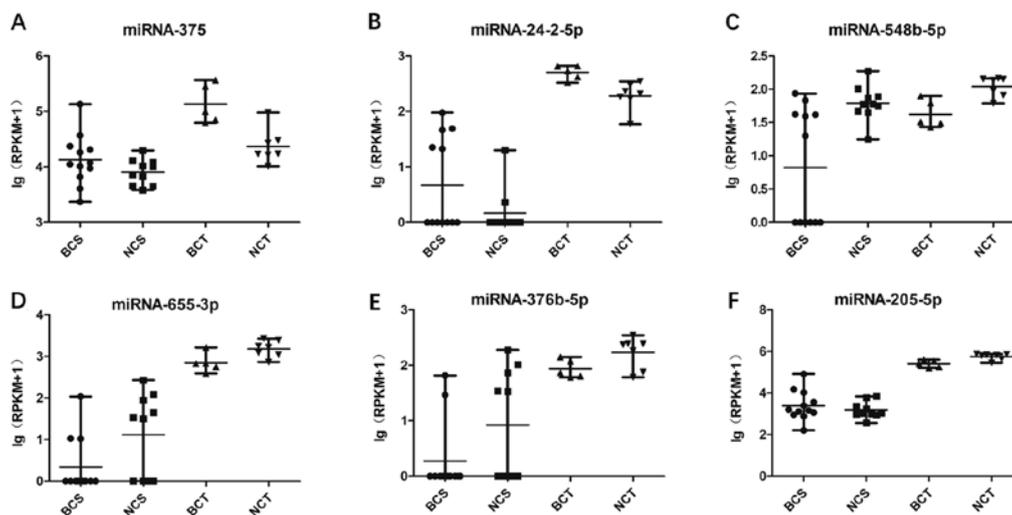


Figure 5. Distributions of miRNAs expression in tissue and blood samples between the two groups. (A) miRNA-375. (B) miRNA-24-2-5p. (C) miRNA-548b-5p. (D) miRNA-655-3p. (E) miRNA-376b-5p. (F) miRNA-205-5p. miRNA, microRNA; RPKM, reads per kilobase of transcript per million mapped reads; BCS, breast cancer serum; NCS, normal control serum; BCT, breast cancer tissue; NCT, normal control tissue.

miRNAs was subsequently performed. As illustrated in Fig. 8, the Venn diagram had little overlap, indicating that the four miRNAs were involved in independent pathways. Taken together, these data demonstrated that extracellular vesicle-packaged miRNAs exhibit potential as diagnostic markers for early-stage breast cancer.

Kaplan-Meier analysis. Based on online data from TCGA, the Kaplan-Meier curves were constructed, and log-rank test

analysis was performed. Low expression of miRNA-24-2-5p was associated with a high survival rate over the period between 3,000 and 6,000 days. A similar trend was also observed in miRNA-375; low expression of miRNA-375 was associated with high survival over the period between 2,000 and 6,000 days. For miRNA-655-3p, a decrease in this miRNA in breast cancer was identified, although its lower expression was also associated with a higher survival rate. miRNA-548b-5p was decreased in the breast cancer group, and the high

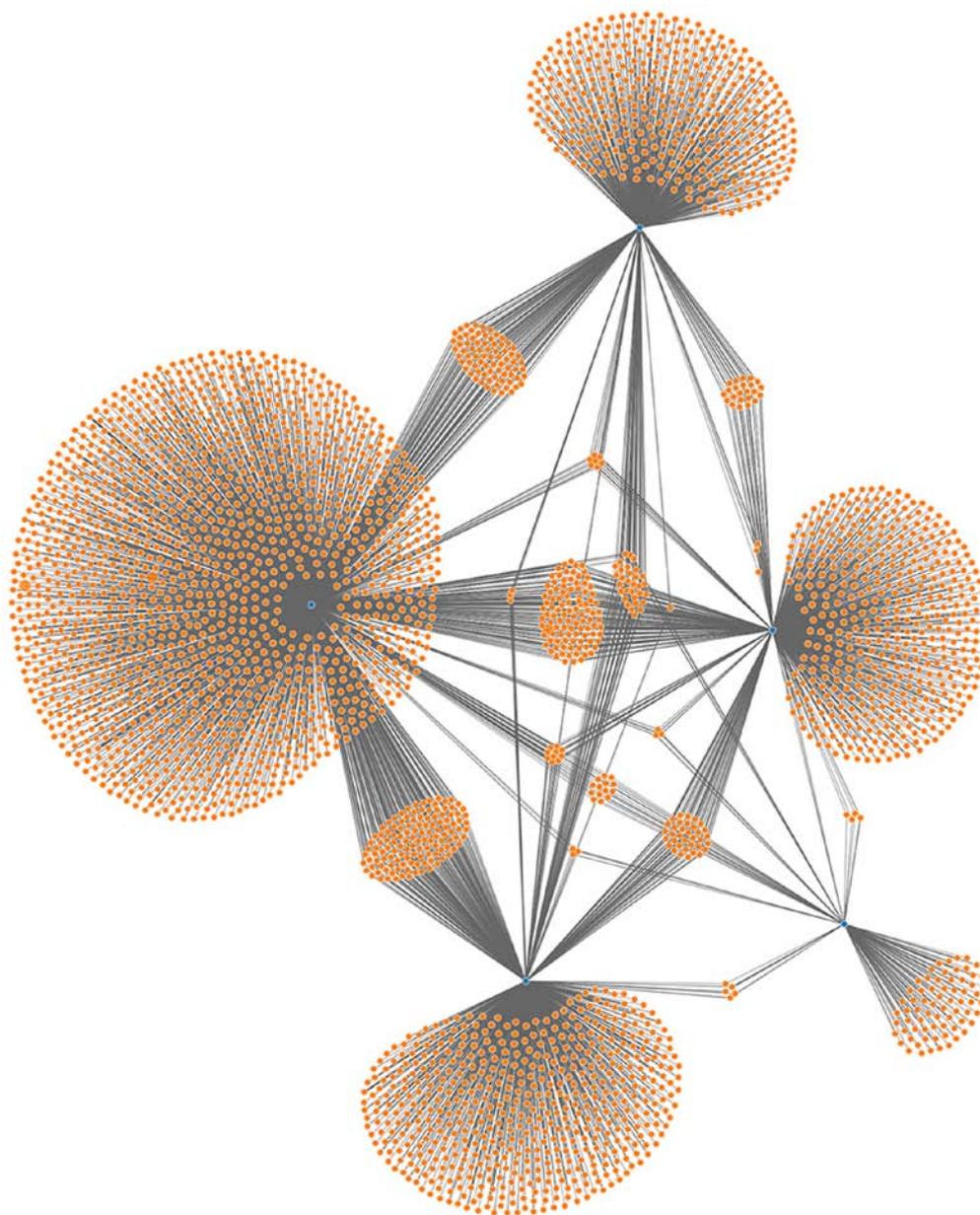


Figure 6. Extracellular vesicle-packaged miRNA-gene network for the five miRNAs. Blue spots indicate the miRNAs, while orange spots indicate the target genes. miRNA, microRNA.

expression of miRNA-548b-5p was associated with a high survival rate (Fig. 9). However, it was not possible to include miRNA-376b-5p in the Kaplan-Meier analysis, as there were no expression data for miRNA-376b-5p in breast cancer within TCGA database, but the Kaplan-Meier curve and log-rank test analysis of miRNA-376b-5p was available based on cancer of the bladder, colon, pancreas, rectum and stomach. Apart from bladder cancer, the results demonstrated that the increased expression of miRNA-376b-5p was associated with a higher survival rate in the other cancer types examined (Fig. S2).

Discussion

The aim of the present study was to identify the plasma extracellular vesicle-packaged miRNAs that had the potential to be biomarkers of early-stage BC, and that served an important

role in the tumorigenesis of BC. The present study identified 35 differentially expressed extracellular vesicle-packaged miRNAs in the plasma and the 175 differentially expressed miRNAs in the tissues, and the differentially expressed miRNAs in the plasma and tissue groups were compared. A total of 1,111 miRNAs in all blood and tissue samples were sequenced. miRNA-375 and miRNA-24-2-5p were upregulated in both blood and tissue samples from the breast cancer group, when compared with the control. However, the remaining miRNAs (miRNA-548b-5p, miRNA-655-3p and miRNA-376-5p) were downregulated in both blood and tissue sample of the breast cancer group, when compared with the control. The cause of this phenomenon may be associated with the selective loading of extracellular vesicles. Taken together, the results suggested that these five miRNAs may exert pivotal functions in the tumorigenesis of breast cancer.

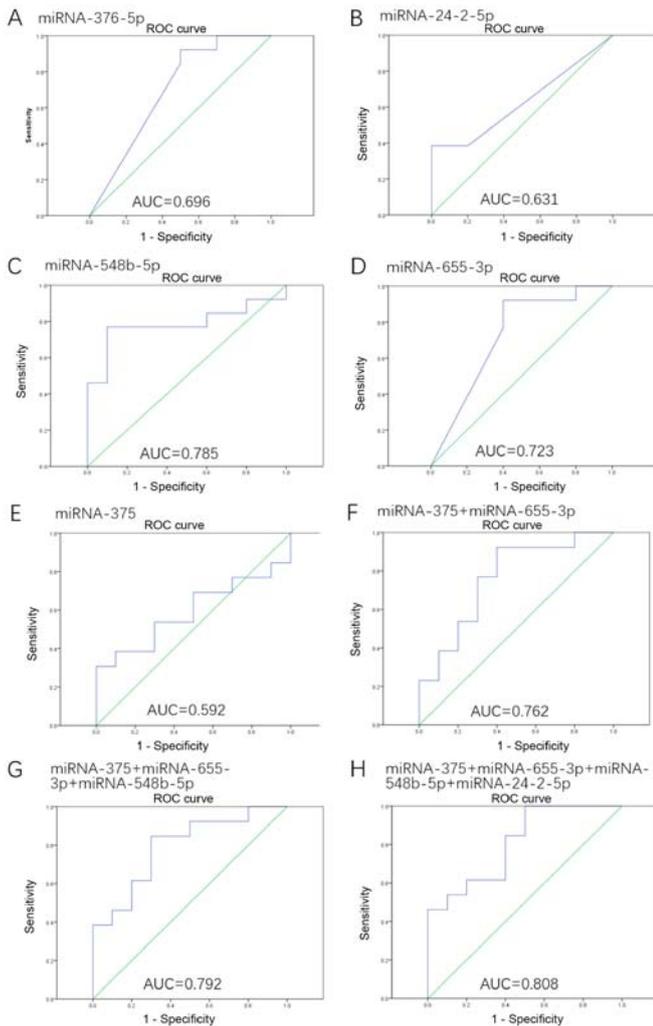


Figure 7. ROC curves for the miRNAs that were significantly different in patients with early-stage BC when compared with controls. ROC curve with AUC for (A) miRNA-376b-5p, (B) miRNA-24-2-5p, (C) miRNA-548b-5p, (D) miRNA-655-3p, (E) miRNA-375, (F) the combination of miRNA-375 and miRNA-655-3p, (G) the combination of miRNA-375, miRNA-655-3p and miRNA-548b-5p, and (H) the combination of miRNA-375, miRNA-655-3p, miRNA-548b-5p and miRNA-24-2-5p. miRNA, microRNA; ROC, receiver operating characteristic; AUC, area under the curve.

There are conflicting reports about miRNA-375. It has previously been reported to be a key driver of the proliferation of ER α -positive breast cells (23); conversely, it has also been reported that microRNA-375 may inhibit the viability, migration and invasion of human BC cells by targeting paired box 6 (24). An additional report stated that miRNA-375 could inhibit the cancer stem cell phenotype and tamoxifen resistance by degrading homeobox B3 in human estrogen receptor (ER)-positive BC (25). However, the present study suggested that miRNA-375 functions as a tumor promoter, as it was upregulated in both plasma and tissue samples from the BC group. In contrast with the present findings, another study demonstrated that miRNA-24 was downregulated in patients with BC; however, that study focused on circulating miRNAs, as opposed to extracellular vesicle-packaged miRNAs in the blood (26). In addition, miRNA-24 was reported to participate in the acquisition of drug-resistance in BC cells (27). Coincidentally, it was reported that miR-24-2 could control the

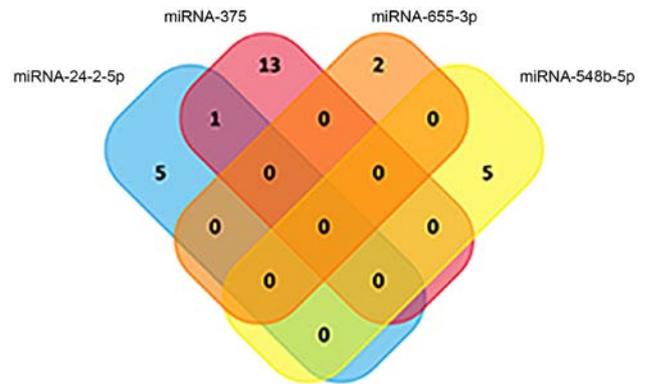


Figure 8. Venn diagram of the pathways that miRNA-375, miRNA-655-3p, miRNA-548b-5p and miRNA-24-2-5p are involved in. miRNA, microRNA.

expression of H2A histone family member X and induce apoptosis by targeting the anti-apoptotic gene BCL-2 in BC (28). miR-548 has been reported to be upregulated in ER-positive BC cells when compared with ER-negative BC cells (29). miRNA-548 decreases nuclear paraspeckle assembly transcript 1 expression and promotes the apoptosis of BC cells (30), in accordance with the present findings. In the present study, miRNA-548-5p was downregulated in both plasma and tissue samples from the BC group, and it has been reported that miRNA-548-3p inhibits the proliferation of BC cells by regulating the expression of enoyl-CoA hydratase, short chain 1 (31). miRNA-655 suppresses epithelial-to-mesenchymal transition by targeting paired related homeobox 1 in triple-negative BC (32). miRNA-655 can inhibit pituitary tumor cell tumorigenesis and is involved in a p53/PTTG1 regulator of sister chromatid separation, securin regulation feedback loop (33). In the present study, miRNA-655-3p was downregulated in the plasma and tissue samples of the BC group, thereby functioning as a tumor suppressor, similar to the above reports. miRNA-376b has been reported to promote BC metastasis by directly targeting homeobox D10 (34). miRNA-376b-5p represses angiogenesis both *in vivo* and *in vitro* (35), and thus miRNA-376b-5p was suggested to be a potential tumor suppressor. In the present study, miRNA-376b-5p was downregulated in the plasma and tissue samples from the BC group. Overall, the current study observed the same changes in these five miRNAs between early BC and control groups in clinical samples. The bioinformatics analysis revealed that there is a common target between the miRNAs, and they may form a network of mutual regulation.

The present study performed pathway analysis and GO analysis based on the differentially expressed extracellular vesicle-packaged miRNAs in the plasma and tissue samples. There was a large difference in the differentially expressed miRNAs between the plasma and tissue samples. However, a comparison of the 20 most prominent terms of KEGG pathway and GO analysis revealed that some of the terms were identical: 45% (9/20) in biological process, 95% (19/20) in cellular component, 40% (8/20) in molecular function and 55% (11/20) of the KEGG pathways. This indicated that the differentially expressed miRNAs had similar effects in the blood and tissue microenvironments. The ROC curve analysis revealed that miRNA-548b-5p had

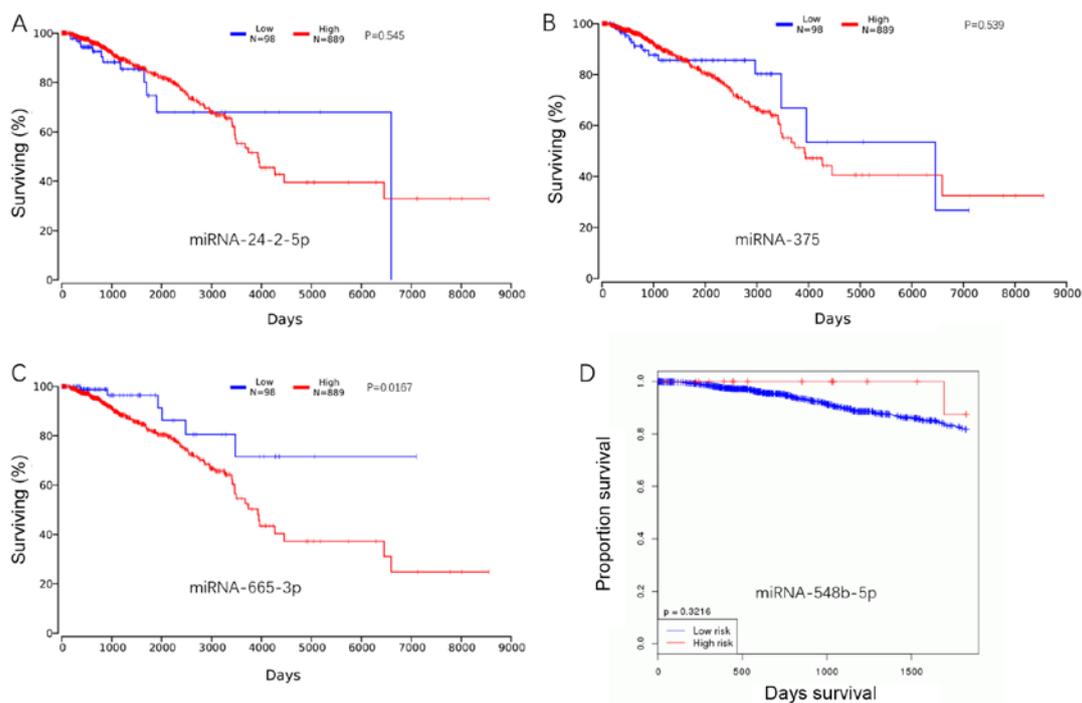


Figure 9. Prognostic value of miRNAs in breast cancer based on The Cancer Genome Atlas database (Kaplan-Meier plots). (A) miRNA-24-2-5p. (B) miRNA-375. (C) miRNA-665-3p. (D) miRNA-548b-5p. miRNA, microRNA.

a greater distinguishing ability among the five miRNAs. However, the combination of miRNA-375, miRNA-655-3p, miRNA-548b-5p and miRNA-24-2-5p had the highest distinguishing ability. Furthermore, this analysis revealed that there was a gradual increase in the AUC when miRNA-375 was combined with other miRNAs. The KEGG pathways of these four miRNAs was subsequently analyzed, and there was almost no overlap. Therefore, it can be assumed that these four miRNAs have independent functions.

Based on online data from TCGA, Kaplan-Meier curve and log-rank test analysis were performed. Consistent with the previous results, miRNA-24-2-5p and miRNA-375, which were increased in the breast cancer group, were partially negatively associated with patient survival. However, the Kaplan-Meier curve and log-rank test analysis result for miRNA-655-3p was the opposite, as a decrease in the expression of this miRNA in breast cancer was associated with a higher survival rate. It was hypothesized that the reason for this discrepancy was the limited sample size in the study cohort. However, miRNA-548b-5p, an miRNA that was decreased in the breast cancer group, was positively associated with patient survival. It is notable that no expression data for miRNA-376b-5p in breast cancer were identified in the TCGA database. The Kaplan-Meier curve and log-rank test analysis of miRNA-376b-5p was available based on cancer of the bladder, colon, pancreas, rectum and stomach. Apart from bladder cancer, the results demonstrated that the increased expression of miRNA-376b-5p was associated with a higher survival rate in the other examined cancer types.

In conclusion, the present study provided a set of plasma extracellular vesicle-packaged miRNA-based biomarkers for the diagnosis of early-stage BC and indicated the complexity

of miRNA regulation in the microenvironments of the blood and tissue. These findings may enrich the presently available knowledge on early-stage breast cancer diagnosis biomarkers, thus improving the curability of breast cancer in the future. However, the expression of the identified miRNAs was not verified by quantitative PCR analysis; there are plans to use a larger cohort of patients and compare the expression of the identified miRNAs among breast cancer patients with different disease stages in a future study.

Acknowledgements

Not applicable.

Funding

The present study was supported by Sanming Project of Medicine in Shenzhen (grant no. SZSM201512015).

Availability of data and materials

The datasets generated and analyzed during the current study are not publicly available due to the 'Interim measures for the management of human genetic resources' promulgated by the Chinese government; however, they are available from corresponding author on reasonable request.

Authors' contributions

NX and MM conceived and designed the study. CY and JH performed the experiments and wrote the manuscript. MM and DZ performed the data analysis and revised the manuscript critically. YY and HH collected the patient's samples. All

authors contributed to the preparation of the final manuscript and lent shape to the final paper.

Ethics approval and consent to participate

The present study was approved by Ethics Committee of Shenzhen People's Hospital (no. LL-KT-201801099) and written informed consent was obtained from all patients.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

- Harbeck N and Gnant M: Breast cancer. *Lancet* 389: 1134-1150, 2017.
- Bertoli G, Cava C and Castiglioni I: MicroRNAs: New biomarkers for diagnosis, prognosis, therapy prediction and therapeutic tools for breast cancer. *Theranostics* 5: 1122-1143, 2015.
- Opstal-van Winden AW, Rodenburg W, Pennings JL, van Oostrom CT, Beijnen JH, Peeters PH, van Gils CH and de Vries A: A bead-based multiplexed immunoassay to evaluate breast cancer biomarkers for early detection in pre-diagnostic serum. *Int J Mol Sci* 13: 13587-13604, 2012.
- Sinn P, Aulmann S, Wirtz R, Schott S, Marmé F, Varga Z, Lebeau A, Kreipe H and Schneeweiss A: Multigene assays for classification, prognosis, and prediction in breast cancer: A critical review on the background and clinical utility. *Geburtshilfe Frauenheilkd* 73: 932-940, 2013.
- Herranz H and Cohen SM: MicroRNAs and gene regulatory networks: Managing the impact of noise in biological systems. *Genes Dev* 24: 1339-1344, 2010.
- Ashby J, Flack K, Jimenez LA, Duan Y, Khatib AK, Somlo G, Wang SE, Cui X and Zhong W: Distribution profiling of circulating microRNAs in serum. *Anal Chem* 86: 9343-9349, 2014.
- Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF, Mitchell PS, Bennett CF, Pogosova-Agadjanyan EL, Stirewalt DL, *et al*: Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc Natl Acad Sci USA* 108: 5003-5008, 2011.
- Desrochers LM, Antonyak MA and Cerione RA: Extracellular vesicles: Satellites of information transfer in cancer and stem cell biology. *Dev Cell* 37: 301-309, 2016.
- Tomasetti M, Lee W, Santarelli L and Neuzil J: Exosome-derived microRNAs in cancer metabolism: Possible implications in cancer diagnostics and therapy. *Exp Mol Med* 49: e285, 2017.
- Graveel CR, Calderone HM, Westerhuis JJ, Winn ME and Sempere LF: Critical analysis of the potential for microRNA biomarkers in breast cancer management. *Breast Cancer (Dove Med Press)* 7: 59-79, 2015.
- Hannafon BN, Trigos YD, Calloway CL, Zhao YD, Lum DH, Welm AL, Zhao ZJ, Blick KE, Dooley WC and Ding WQ: Plasma exosome microRNAs are indicative of breast cancer. *Breast Cancer Res* 18: 90, 2016.
- Gao Y, Cai Q, Huang Y, Li S, Yang H, Sun L, Chen K and Wang Y: MicroRNA-21 as a potential diagnostic biomarker for breast cancer patients: A pooled analysis of individual studies. *Oncotarget* 7: 34498-34506, 2016.
- Shimomura A, Shino S, Kawachi J, Takizawa S, Sakamoto H, Matsuzaki J, Ono M, Takeshita F, Niida S, Shimizu C, *et al*: Novel combination of serum microRNA for detecting breast cancer in the early stage. *Cancer Sci* 107: 326-334, 2016.
- Eichelscher C, Stuckrath I, Muller V, Milde-Langosch K, Wikman H, Pantel K and Schwarzenbach H: Increased serum levels of circulating exosomal microRNA-373 in receptor-negative breast cancer patients. *Oncotarget* 5: 9650-9663, 2014.
- Xiao CL, Mai ZB, Lian XL, Zhong JY, Jin JJ, He QY and Zhang G: FANSe2: A robust and cost-efficient alignment tool for quantitative next-generation sequencing applications. *PLoS One* 9: e94250, 2014.
- Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B: Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5: 621-628, 2008.
- Bloom JS, Khan Z, Kruglyak L, Singh M and Caudy AA: Measuring differential gene expression by short read sequencing: Quantitative comparison to 2-channel gene expression microarrays. *BMC Genomics* 10: 221, 2009.
- Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, Kong L, Gao G, Li CY and Wei L: KOBAS 2.0: A web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* 39 (Web Server Issue): W316-W322, 2011.
- Herrera-Perez Z, Gretz N and Dweep H: A comprehensive review on the genetic regulation of cisplatin-induced nephrotoxicity. *Curr Genomics* 17: 279-293, 2016.
- Wong NW, Chen Y, Chen S and Wang X: OncomiR: An online resource for exploring pan-cancer microRNA dysregulation. *Bioinformatics* 34: 713-715, 2018.
- Rajkumar AP, Qvist P, Lazarus R, Lescai F, Ju J, Nyegaard M, Mors O, Børghlum AD, Li Q and Christensen JH: Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. *BMC Genomics* 16: 548, 2015.
- Robinson MD, McCarthy DJ and Smyth GK: edgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139-140, 2010.
- de Souza Rocha Simonini P, Breiling A, Gupta N, Malekpour M, Youns M, Omranipour R, Malekpour F, Volinia S, Croce CM, Najmabadi H, *et al*: Epigenetically deregulated microRNA-375 is involved in a positive feedback loop with estrogen receptor alpha in breast cancer cells. *Cancer Res* 70: 9175-9184, 2010.
- Zou Q, Yi W, Huang J, Fu F, Chen G and Zhong D: MicroRNA-375 targets PAX6 and inhibits the viability, migration and invasion of human breast cancer MCF-7 cells. *Exp Ther Med* 14: 1198-1204, 2017.
- Fu H, Fu L, Xie C, Zuo WS, Liu YS, Zheng MZ and Yu JM: miR-375 inhibits cancer stem cell phenotype and tamoxifen resistance by degrading HOXB3 in human ER-positive breast cancer. *Oncol Rep* 37: 1093-1099, 2017.
- Schrauder MG, Strick R, Schulz-Wendtland R, Strissel PL, Kahmann L, Loehberg CR, Lux MP, Jud SM, Hartmann A, Hein A, *et al*: Circulating micro-RNAs as potential blood-based markers for early stage breast cancer detection. *PLoS One* 7: e29770, 2012.
- Yamamoto Y, Yoshioka Y, Minoura K, Takahashi RU, Takeshita F, Taya T, Horii R, Fukuoka Y, Kato T, Kosaka N and Ochiya T: An integrative genomic analysis revealed the relevance of microRNA and gene expression for drug-resistance in human breast cancer cells. *Mol Cancer* 10: 135, 2011.
- Srivastava N, Manvati S, Srivastava A, Pal R, Kalaiarasan P, Chattopadhyay S, Gochhait S, Dua R and Bamezai RN: miR-24-2 controls H2AFX expression regardless of gene copy number alteration and induces apoptosis by targeting antiapoptotic gene BCL-2: A potential for therapeutic intervention. *Breast Cancer Res* 13: R39, 2011.
- Lee YM, Lee JY, Ho CC, Hong QS, Yu SL, Tzeng CR, Yang PC and Chen HW: miRNA-34b as a tumor suppressor in estrogen-dependent growth of breast cancer cells. *Breast Cancer Res* 13: R116, 2011.
- Ke H, Zhao L, Feng X, Xu H, Zou L, Yang Q, Su X, Peng L and Jiao B: NEAT1 is required for survival of breast cancer cells through fus and mir-548. *Gene Regul Syst Bio* 10: 11-17, 2016.
- Shi Y, Qiu M, Wu Y and Hai L: miR-548-3p functions as an anti-oncogenic regulator in breast cancer. *Biomed Pharmacother* 75: 111-116, 2015.
- Lv ZD, Kong B, Liu XP, Jin LY, Dong Q, Li FN and Wang HB: miR-655 suppresses epithelial-to-mesenchymal transition by targeting Prrx1 in triple-negative breast cancer. *J Cell Mol Med* 20: 864-873, 2016.
- Liang HQ, Wang RJ, Diao CF, Li JW, Su JL and Zhang S: The PTTG1-targeting miRNAs miR-329, miR-300, miR-381, and miR-655 inhibit pituitary tumor cell tumorigenesis and are involved in a p53/PTTG1 regulation feedback loop. *Oncotarget* 6: 29413-29427, 2015.
- An N, Luo X, Zhang M and Yu R: MicroRNA-376b promotes breast cancer metastasis by targeting Hoxd10 directly. *Exp Ther Med* 13: 79-84, 2017.
- Li LJ, Huang Q, Zhang N, Wang GB and Liu YH: miR-376b-5p regulates angiogenesis in cerebral ischemia. *Mol Med Rep* 10: 527-535, 2014.

