Mol Biol Rep (2013) 40:2175–2186
DOI 10.1007/s11033-012-2277-5

Genome-wide screen for aberrantly expressed miRNAs reveals miRNA profile signature in breast cancer

Li Guo · Yang Zhao · Sheng Yang · Min Cai · Qian Wu · Feng Chen

Received: 22 August 2012 / Accepted: 19 November 2012 / Published online: 30 November 2012

Abstract Dysregulation in the expression of miRNAs contributes to the occurrence and development of many human cancers. We herein attempted to obtain the potential association between miRNA expression profile and breast cancer by applying high-throughput sequencing technology. Small RNAs from seven paired tumor and adjacent normal tissue samples were sequenced. To determine the miRNA expression profiles in tissues and sera, another five equally pooled serum samples from 20 patients and 30 normal women were sequenced. Despite a similar number in abundantly expressed miRNAs across samples, we detected varying miRNA expression profiles. Some miRNAs showed inconsistent or opposite dysregulation trends across different tumor tissues, including some abundantly expressed miRNA gene clusters and gene families. Wilcoxon sign-rank test for paired samples analysis revealed that abnormal miRNAs showed a higher level of variation across the seven tumor samples. We also completely surveyed abnormal miRNAs expressed in tumor and serum tissues in the mixed datasets based on the relative expression levels. Most of these miRNAs were significantly down-regulated in tumor samples, but nine abnormal miRNAs (miR-18a, 19a, 20a, 30a, 103b, 126, 126*, 192, 1287) were consistently expressed in tumor tissues and serum samples. Based on experimentally validated target mRNAs, functional enrichment analysis indicated that these abnormal miRNAs and miRNA groups (miRNA gene clusters and gene families) have important roles in multiple biological processes. Dynamic miRNA expression profiles, various abnormal miRNA profiles and complexity of the miRNA regulatory network reveal that the miRNA expression profile is a potential biomarker for classifying or detecting human disease.

Keywords MicroRNA (miRNA) · Breast cancer · miRNA profile

Introduction

MicroRNAs (miRNAs), ~22 nt in length, negatively regulate gene expression at the post-transcriptional and/or translation levels [1]. The small ncRNAs are well-conserved phylogenetically and have similar roles in the development of diverse species [2–4]. They have critical roles in various biological processes, such as cell development, differentiation, proliferation, apoptosis and metabolism [5], and can even influence specific human diseases, including various cancers. Some miRNAs are located at fragile sites or genomic regions, and are prone to be aberrantly expressed in cancers due to frequent deletion or amplification [6]. The dysregulation is mainly from nucleotide deletion, mutation and methylation of miRNA genes in human disease [6]. A series of miRNAs have been characterized as potential tumor suppressors or oncomiRs

Electronic supplementary material The online version of this article (doi:10.1007/s11033-012-2277-5) contains supplementary material, which is available to authorized users.

L. Guo · Y. Zhao · S. Yang · M. Cai · F. Chen (✉)
Department of Epidemiology and Biostatistics, School of Public Health, Nanjing Medical University, Nanjing 210029, China
e-mail: fengchen@njmu.edu.cn

L. Guo
e-mail: gl8008@163.com; lguo@njmu.edu.cn

Q. Wu (✉)
State Key Laboratory of Reproductive Medicine, Department of Hygienic Analysis and Detection, School of Public Health, Nanjing Medical University, Nanjing 210029, China
e-mail: wuqian@njmu.edu.cn

[7, 8]. Many of them are differentially expressed between normal and tumor tissues, and have potential biological roles in the occurrence and development of cancer. The small RNAs are quite stable and are not degraded, and emerging evidence suggests that the small RNAs may be potential biomarkers, especially for circulating miRNAs. Some miRNAs were identified to be associated with human tumors, such as ovarian cancer and colon cancer [9–11]. Thus, miRNA expression profile can classify different human cancers [12].

In recent decades, miRNAs have been involved in the study of the human diseases, including breast cancer (BC) [13–24]. By applying miRNA microarrays and high-throughput sequencing technologies, several miRNAs are reportedly associated with BC. Dysregulation of miR-210 is correlated with tumor aggressiveness and poor prognosis [24], and miR-355 inhibits tumor reinitiation [25]. However, these studies have always implicated a single miRNA or few miRNAs in breast cancer. Here, we attempted to investigate miRNA expression profiles by applying high-throughput sequencing technology in paired tumor tissues and adjacent normal tissues from patients with breast cancer, and in pooled normal and tumor serum samples. Differentially expressed miRNA profiles based on each control were obtained and subjected to comparative analysis, especially between the seven paired tumor and adjacent normal tissues, and between tissue and serum samples. In addition, we also estimated differentially expressed miRNA species based on sequencing datasets of single paired samples and mixed datasets of samples.

Materials and methods

Sample collection and small RNA library preparation

7 paired tumor and adjacent normal tissues from seven patients with breast cancer were obtained. Serum samples from 20 patients and 30 normal women were collected. Due to less serum samples from patients, these rare samples were equally pooled as another two tumor samples and three normal samples (each pooled serum sample contained 10 samples) (Table S1). The institutional review board of Nanjing Medical University approved this study. Written informed consent was obtained from each participant or representative prior to tissue acquisition.

Total RNA was extracted from samples using TRIzol (Invitrogen). Small RNAs were then isolated using mirVana™ miRNA Isolation Kit (Ambion) and purified using Small RNA Expression Kit (Life Technologies). Purified small RNAs were then used to construct a small RNA library to be used for sequencing using the SOLiD™ sequencing platform (ABI, Life Technology).

SOLiD sequencing and small RNA analysis

SOLiD sequencing was conducted at the State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, China. Expressed miRNAs were comprehensively searched using SOLiD miRNA analysis pipeline (http://SOLiDsoftwaretools.com/gf/project/srna/). The pipeline included three steps: (1) filter other human non-coding RNAs (ncRNAs, including tRNAs, rRNAs, snoRNAs, etc.), (2) detect human known miRNAs based on known pre-miRNAs from the miRBase database (Release 16.0, http://www.mirbase.org/), and (3) discover novel candidate miRNAs based on the human genome sequence.

The final relative expression levels of miRNAs were determined by normalization with the total reads of all the expressed miRNAs in the sample. To comprehensively search for differentially expressed miRNA species and filter out those miRNAs with lower expression levels, fold change values were estimated by adding an additional and given low number (1 sequence counts), respectively. In order to estimate miRNA expression profiles more accurately, we only collected those miRNAs with at least 0.30 % relative expression levels in a sample.

Assessment of miRNA profiles and functional enrichment analysis

To characterize the differentially expressed miRNA profiles in BC, we also analyzed the miRNAs at the single dataset and mixed dataset levels based on the original sequence counts and relative expression levels, respectively. The mixed datasets were similar to equally pooled samples, based on relative expression levels. Different methods were used to assess miRNA expression profiles to help select and correct the estimation method. For example, for the breast cancer samples and their adjacent normal tissues, sequencing datasets of seven cancer samples and adjacent normal tissues were mixed as a larger tumor dataset and as an adjacent normal tissue dataset based on the relative expression levels, respectively.

miRNAs are prone to be clustered on chromosomes (miRNA gene cluster), and some miRNAs are composed of miRNA gene family with similar sequences. These special miRNAs may have close functional relationships by co-regulating biological processes. Therefore, in the study, we also surveyed some abundantly expressed miRNA gene clusters and gene families, and further investigated their expression patterns and potential functional relationship.

For those aberrantly expressed miRNAs in tumor tissues, further functional enrichment analysis was performed using CapitalBio Molecule Annotation System V4.0 (MAS, http://bioinfo.capitalbio.com/mas3/) based on

their known experimentally validated target mRNAs in the miRTarBase database[26].

Statistical analysis

Hierarchical clustering analysis was performed using Cluster 3.0 program, and the results were visualized with the TreeView 1.60 program [27, 28]. Both the two softwares were obtained from http://rana.lbl.gov/eisen. Wilcoxon sign-rank test for paired samples was used to identify differentially expressed miRNA profiles between paired tumor and adjacent normal tissues. The association between miRNA species with aberrant expression and breast cancer was evaluated by principal component analysis (PCA). The scatter-plot matrix was drawn by using the "car" package in R. We used the top principal components (PCs) explaining at least 80 % of the total variation to discriminate tumor and normal controls. Receiver Operating Characteristic (ROC) curves were applied to evaluate the discriminant ability of PCA. All the analyses were performed using Statistical Analysis System software (Version 9.1.3, SAS Institute, NC) and R. A P value <0.05 was considered statistically significant, and all tests of significance were two-tailed.

Results

Overview of miRNA expression profiles

There were 34–56 (45.90 ± 6.40) abundantly expressed miRNA species in all samples, and significant differences were detected between tissue and serum samples (tissue samples: 43.43 ± 5.29, serum samples: 52.80 ± 3.56, $P < 0.05$) (Table S1). No significant difference was detected between tumor and adjacent normal tissues ($P = 0.75$). Generally, several miRNA species were quite abundantly expressed in vivo, while others were always expressed at lower levels (Figure S1). However, although similar expression patterns were detected across samples, the relative expression profiles vary. For example, in the tumor tissue of patient 3, let-7b was the most abundantly expressed miRNA species with a 22.45 % relative expression level, while in the tumor tissue of patient 5, miR-23a was the most abundant miRNA with a 12.41 % relative expression level. Despite the detected miRNAs showing large divergence across different samples, they always shared nearly half or more common miRNA species between each paired tumor and adjacent normal samples (Figure S2). Furthermore, we also collected those miRNAs with higher relative expression levels (more than 1.00 % in a sample) to characterize their distribution patterns across different samples. 10–26 miRNA species were detected in a specific sample, but a total of 74 miRNAs were detected in all samples (Fig. 1). These samples showed varying expression profiles of abundantly expressed miRNAs. A larger level of divergence was detected between seven tumor tissues and seven adjacent normal tissues (Fig. 1). No significant difference was detected between the total numbers of abundantly expressed miRNAs in tumor and adjacent normal samples ($P = 0.72$).

Moreover, we also mixed samples of the same origin based on their original sequence counts and relative expression levels. Thus, the four experimental samples include tumor tissue sample, adjacent normal tissue sample, serum tumor, and normal serum samples. Compared to the miRNA profile of single samples, mixed samples indicated similar abundant miRNA populations despite the involvement of different miRNA species. Different methods of combining datasets also led to similar profile of abundantly expressed miRNAs. Similar to what was observed using single samples, a highly variant miRNA profiles across the mixed four samples were also detected based on the mixed datasets of original sequence counts (Figure S3A). Most of these miRNAs were "private miRNAs" or shared by 1–3 mixed samples. For the miRNA species that were commonly expressed, their expression pattern in tissue and serum tumor was variable (Figure S3B). For example, let-7a was down-regulated in tissue and serum tumor samples. On the other hand, miR-191 was significantly up-regulated in tumor tissue, whereas it was significantly down-regulated in serum tumor.

Differentially expressed miRNAs in breast cancer tissues

Among the set of abundant miRNAs, aberrantly expressed miRNAs were then determined based on adjacent normal tissue in each patient. A miRNA was considered to be significantly dysregulated if the fold change (log2) was more than 2.0 or less than -2.0. Based on this criterion, various differentially expressed miRNA populations were obtained in the seven tumor tissues, and the expression profiles of these miRNAs across the tissue samples were inconsistent in spite of all seven patients being diagnosed with the same pathology (breast cancer) (Table S2). The significantly dysregulated miRNAs (from 8 to 58) exhibited up-regulation or down-regulation patterns. Of these seven paired samples, rare dysregulated miRNAs were found, and most of them were private miRNAs (Figure S4; Table S2). Only three commonly up-regulated miRNAs (miR-192, 223 and 451) were identified in three tumor tissue samples, whereas others were always shared by two tumor tissue samples, including 12 up-regulated and 10 down-regulated miRNA species (Figure S4). Strikingly, miR-150 showed an inconsistent trend of dysregulation

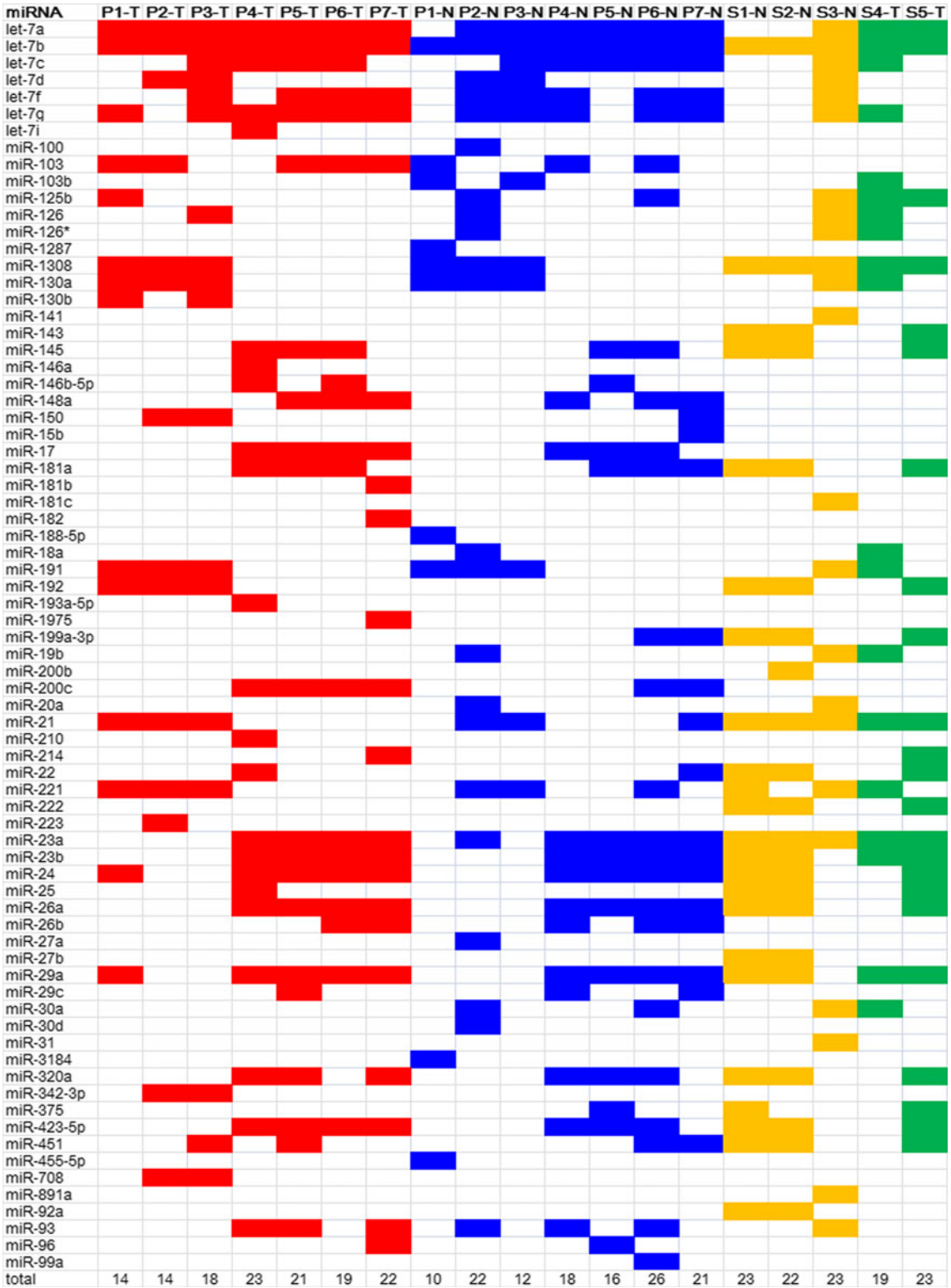

Fig. 1 Distribution patterns across different samples based on abundantly expressed miRNAs. These miRNAs have higher relative expression levels (more than 1.00 % at least in a sample). A total of 74 miRNAs are shown in the figure. These miRNAs exhibited a dynamic expression distribution across different samples. Pn-T: indicates tumor sample from patient 1 to 7 (in *red* background); Pn-N: indicates adjacent normal sample from patient 1 to 7 (in *blue* background); Sn-T: indicates tumor serum samples (in *yellow* background); Sn-N: indicates normal serum samples (in *green* background). (Color figure online)

across samples. It was up-regulated in two tumor tissues, whereas it was down-regulated in another two tumor tissues (Figure S4). Moreover, we also found that another 18 miRNAs were shared by up-regulated and down-regulated miRNA populations (Figure S4; Table S2).

A hierarchical clustering analysis was performed on all of the listed miRNAs in Figure S4. The result revealed various expression patterns across different samples (Fig. 2), despite these samples originating from patients with the same pathology (breast cancer). Although these miRNAs might be dysregulated between some paired samples, the average fold changes (log2) were moderate and variable (Fig. 2). Most of these miRNAs that we identified to be dysregulated did not show a significant dysregulated trend. The average fold change values of 4 miRNAs (miR-145, 130b, 192 and 223) were more than 1.5, while only one miRNA (miR-103b) was less than −1.5 (Fig. 2).

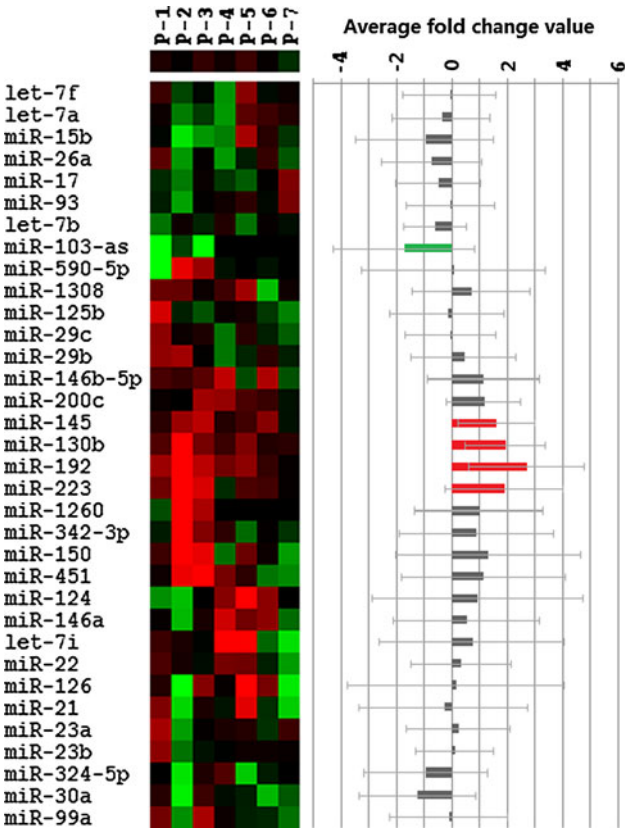

Fig. 2 Hierarchical clustering of miRNA expression. All of these miRNAs were listed in Fig. 1, and were detected to be abundantly expressed in at least two patients. miRNA expression levels are shown in a spectrum with *red* indicating up-regulation and *green* indicating down-regulation. P1-7 indicates tumor and normal adjacent samples from the seven patients. Each paired samples showed differentially expression miRNA profiles. The average fold changes (log2) are presented on the *right*. Dysregulated miRNAs are highlighted in *red* (up-regulation) and *green* (down-regulation) background. (Color figure online)

For the seven paired tumor and adjacent normal samples, the aberrantly expressed miRNAs were identified from abundantly expressed miRNAs using the Wilcoxon sign-rank test for paired samples. 23 miRNAs were characterized as aberrantly expressed miRNA species, but 15 of them were down-regulated in breast cancer samples. These abnormal miRNAs showed larger levels of variation (Table 1). If the samples with similar origins were mixed based on their original sequence counts and relative expression levels, aberrantly expressed miRNA profiles showed high inconsistency in expression (Tables 1, S3). The different estimation methods resulted in variable differentially expressed miRNA profiles.

Principal component analyses (PCA) suggested that the top principle components of several miRNAs (miR-30a, 145, 223, 130b and 192) could distinguish tumor samples from normal samples based on seven paired tumor and adjacent normal samples (Fig. 3a). Further, we also analyzed the nine

Table 1 Aberrantly expressed miRNAs in breast cancer based on the adjacent normal tissues by Wilcoxon sign-rank test for paired samples (n = 7)

miRNAs	Mean ± SD $\log_2$ (fold change value)	*P* value
let-7b	−0.59 ± 1.13	0.03
let-7g	0.02 ± 1.09	0.00
let-7i	0.59 ± 3.20	0.00
miR-100	−0.83 ± 1.39	0.00
miR-141	−0.32 ± 1.95	0.00
miR-143	−0.02 ± 1.65	0.00
miR-145	1.45 ± 1.29	0.00
miR-146a	0.54 ± 2.62	0.00
miR-148a	−1.09 ± 1.32	0.00
miR-17	−0.46 ± 1.53	0.00
miR-181b	−0.14 ± 0.89	0.00
miR-181c	−1.45 ± 2.31	0.00
miR-1975	0.34 ± 0.66	0.00
miR-19b	−0.58 ± 2.34	0.00
miR-200c	0.74 ± 1.52	0.00
miR-20a	−1.18 ± 1.86	0.00
miR-22	0.38 ± 1.80	0.00
miR-23a	0.17 ± 1.81	0.02
miR-23b	0.11 ± 1.40	0.00
miR-26a	−0.73 ± 1.81	0.00
miR-30d	0.02 ± 1.99	0.00
miR-320a	−0.43 ± 1.59	0.02
miR-342-3p	0.88 ± 2.62	0.02
miR-423-5p	−0.81 ± 2.28	0.00
miR-451	0.86 ± 2.56	0.02
miR-99a	−0.08 ± 2.13	0.02

These miRNAs are abundantly expressed (>1.00 %)

miRNAs (miR-18a, 19a, 20a, 30a, 103b, 126, 126*, 192, 1287) in the 19 samples (including serum samples). Although Tumor and normal samples were not separated completely as well as in the 14 samples, it is worth noting that most were separated (Fig. 3b). ROC curves indicated that these miRNA species can discriminate breast cancer samples from normal samples (Fig. 3c).

Abnormal miRNAs in tissue and serum samples

Here, seven paired tumor and adjacent normal tissue samples, five pooled serum tumor and control samples (each sample was obtained from 10 patients or 10 normal women) were sequenced. miRNA profiles showed various patterns due to individual differences, especially for differentially expressed miRNA profiles (Figure S2; Table S2). To avoid the influence of inconsistent miRNA profiles, we mixed samples with similar origins based on their relative expression levels to survey aberrantly expressed miRNAs between tumor tissue and serum samples. In the set of abundantly expressed miRNAs, 28 miRNAs were dysregulated in tumor tissues or serum samples (Table 2). They always showed a consistent trend in expression levels and most of them (20 miRNAs) were consistently down-regulated. Only nine miRNAs (miR-18a, 19a, 20a, 30a, 103b, 126, 126*, 192, 1287) showed consistent dysregulation patterns between tumor tissues and tumor sera samples (Table 2). Interestingly, of these consistently dysregulated miRNAs, only miR-192 was up-regulated in tumor samples, while others were down-regulated. Except for miR-103b and miR-192, other aberrantly expressed miRNAs were registered in the miRWalk database (http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/) [29] as miRNAs associated with human breast neoplasms. However, miR-103 was correlated with mouse breast neoplasms.

Aberrantly expressed miRNAs were subjected further to pathway analysis using the miRTarBase database. mRNAs were included in the analysis if they were regulated by at least two abnormal miRNAs. The analysis revealed that the putative target mRNAs may contribute to multiple biological processes including MAPK signaling pathway, TGF-beta signaling pathway, ErbB signaling pathway, p53 signaling pathway, and occurrence and development of some cancers [e.g. Chronic myeloid leukemia, Bladder cancer, Pancreatic cancer, etc. (Table S4)]. These results suggest that these abnormally regulated miRNAs have important biological roles in multiple signaling pathways.

miRNA gene clusters/gene families associated with breast cancer

According to the viewpoint that several miRNAs can co-regulate or co-contribute to complex biological processes, we also collected some abundantly expressed miRNA gene clusters and gene families to further characterize their potential expression and functional relationships. These clustered and homologous miRNAs with close functional relationships were analyzed as a special miRNA group. We selected let-7, miR-23, 30, 221 gene families, and miR-221, 23a, 30b and 143 gene clusters for further analysis to be able to understand the potential functional implication of miRNA groups (miR-221 and miR-222 were clustered in a gene cluster, and they both belong in the miR-221 gene family). Certain miRNAs are included in multiple miRNA gene clusters or gene families. For example, miR-23a is a member of the miR-23a gene cluster and the miR-23 gene family. These miRNA groups showed a dynamic expression pattern across the 19 samples based on their relative expression levels, especially for the let-7 gene family (Fig. 4a). The let-7 group exhibited more than 50 % expression in adjacent normal sample of patient 3, while it was rarely expressed in other samples (less than 5 %).

Differential expression analysis based on each adjacent normal sample also showed various miRNA expression patterns in the seven patients, showing an inconsistent dysregulation trend. For example, with the miR-23 and miR-30 gene families, the miR-30b gene cluster exhibited a variable trend in dysregulation (Fig. 4b). According to the average fold change (log2), although most of these miRNAs showed consistent expression patterns, the opposite trends were also detected between tissue and serum tumor samples, especially for let-7 and miR-30 gene families, and miR-30b gene cluster (Fig. 4c). Similar to the single abnormally expressed miRNAs, these abnormally expressed miRNA groups may also contribute to various biological processes, including p53 signaling pathway, Wnt signaling pathway, MAPK signaling pathway and some cancers, based on their experimentally validated target mRNAs (Table S5).

Discussion

Dynamic miRNA expression profiles across different samples

Analysis of miRNA profiles revealed dynamic expression patterns across different samples, including tissue and serum samples (Figs. 1, S1). This may be derived from the spatial/temporal expression of miRNAs and individual differences. Based on the profile of abundantly expressed miRNAs, the number of miRNAs expressed was similar and no significant difference was detected between samples, despite of the serum samples containing a higher proportion of the abundant miRNA species (Table S1). Indeed, expression analysis suggested an expression bias based on the relative expression level. Several miRNA species had higher expression

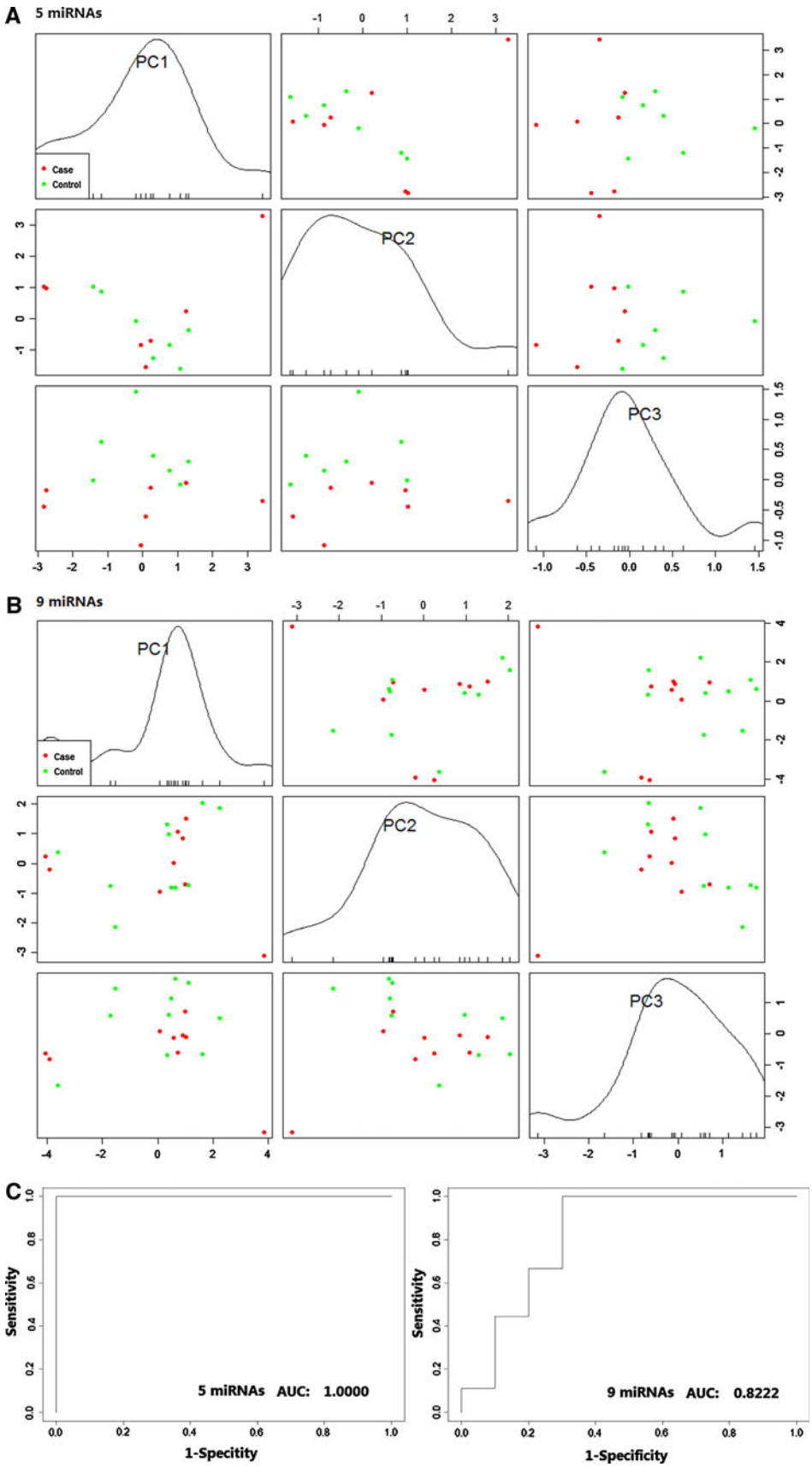

Fig. 3 PCA and ROC curves based on the selected miRNAs. **a** Scatter-plot matrix of the top three principal components (PCs) using the five miRNAs (miR-30a, 145, 223, 130b, 192) in 14 samples from seven patients (the seven paired tumor and adjacent normal samples). **b** Scatter-plot matrix of the top three PCs using the nine miRNAs (miR-18a, 19a, 20a, 30a, 103b, 126, 126*, 192, 1287) in 19 samples, including another 5 mixed serum samples. **c** The AUC of discriminate models use the three top PCs can explain 100 % (5 miRNAs in 14 samples) and 82.22 % (9 miRNAs in 19 samples) of the total variation

Table 2 miRNA profiling in tissue and serum samples based on mixed datasets of relative expression levels

miRNAs	log2 (fold change value) (tissue samples)	log2 (fold change value) (serum samples)
let-7a	−0.01	−2.16
let-7d	−0.24	−3.68
let-7e	−0.32	−3.08
miR-18a	−1.48	−4.84
miR-19b	−1.58	−3.94
miR-20a	−1.46	−3.92
miR-22	0.08	2.04
miR-27b	−0.41	2.00
miR-30a	−1.47	−3.05
miR-30d	−0.87	−3.44
miR-93	0.45	−3.05
miR-103b	−4.53	−7.58
miR-124	2.40	0.94
miR-126	−1.68	−4.21
miR-126*	−2.47	−3.64
miR-130a	1.00	−3.81
miR-181c	−0.79	−3.43
miR-188-5p	−7.36	−0.74
miR-192	3.48	1.50
miR-210	2.45	0.67
miR-223	2.51	1.18
miR-342-3p	2.92	−0.27
miR-375	−0.22	2.00
miR-376c	−0.84	−3.86
miR-708	1.31	−3.88
miR-891a	0.72	−5.67
miR-1287	−7.72	−1.88
miR-4303	1.38	−5.36

levels, while others were rarely detected (Figure S1). These abundant miRNAs showed inconsistent expression across different samples. The large variation in expression might be derived from difference in original sequence counts, influenced by polymerase chain reaction (PCR) amplification through sample preparation. The higher sensitivity and high-throughput sequencing datasets from deep sequencing technology also contributed to the large variation, but the larger expression difference still implicated expression bias between different miRNA species (Figure S1). Except for a series of abundantly expressed miRNAs, the rest were expressed at lower expression levels. This difference in the level of expression possibly indicates essential biological roles of the abundant miRNAs in organisms.

We also further characterized the distribution of abundantly expressed miRNA populations (more than 1 %) in all samples. Variable expression patterns of several miRNAs were observed despite most of them being present at similar numbers (Fig. 1). Except for let-7b which was abundantly expressed in all samples, others showed variable expression patterns across different samples. For example, let-7a was highly expressed in most of the samples, while miR-100 was only highly expressed in the adjacent normal sample of patient 2. About 20 miRNAs (19.30 ± 4.07) were characterized as abundant miRNAs in a single sample, but a total of 74 miRNAs were abundantly expressed in all samples. The results revealed various miRNA profiles with less common miRNAs, especially those miRNAs that were detected in all samples (only let-7b was predominantly expressed in all samples). Most of these miRNAs were expressed in several samples, or as private miRNAs expressed in a specific sample (Fig. 1). Although abundant miRNAs were defined arbitrarily, the results still revealed that rare miRNAs were shared by different samples or individuals, and a specific individual or sample had a private profile of abundant miRNAs. However, miRNA profiles between tumor samples and the adjacent normal tissue from the same patient exhibited high similarity. For example, each paired tumor and adjacent normal samples shared nearly half or more abundant miRNAs, although some miRNAs were abnormally expressed (Figure S2). When these samples were grouped into 4 experimental groups according to their origin, the same dynamic and variable distribution patterns were detected across samples (Figs. 1, S2, S3).

Tumor and normal tissue samples were separated by the PCA analysis using selected several miRNAs (Fig. 3). However, PCA based on the abundantly expressed miRNAs could not completely separate the tumor and normal samples (Figure S5). These results indicate that it may be difficult to separate tumor samples from normal samples based on miRNA expression profiles, but that a few specific miRNAs could clearly distinguish tumor from normal samples. Although PCA analysis based on miRNA expression profiles did not separate tumor and normal samples, the complexity of interaction and functional relationships between miRNAs may be more revealing. It is not enough to assess tumor samples using PCA analysis based on the complexity of miRNAs. Indeed, many studies have reported that a single miRNA or some special miRNAs could serve as potential biomarkers to identify or classify human diseases. However, the interaction of miRNA and its target mRNA is quite complex. Generally, a biological pathway will involve a set of specific miRNAs, and a specific miRNA will target multiple mRNAs. More importantly, the interaction between miRNAs should be not ignored such as endogenous sense/antisense miRNAs [30–34]. Therefore, a specific miRNA or a set of few miRNAs might not be suitable biomarkers to classify human diseases, given the dynamic and highly variable miRNA expression profiles. Based on the complexity of interaction and potential functional relationships between miRNAs, comprehensive miRNA profiles, especially for abundantly expressed and

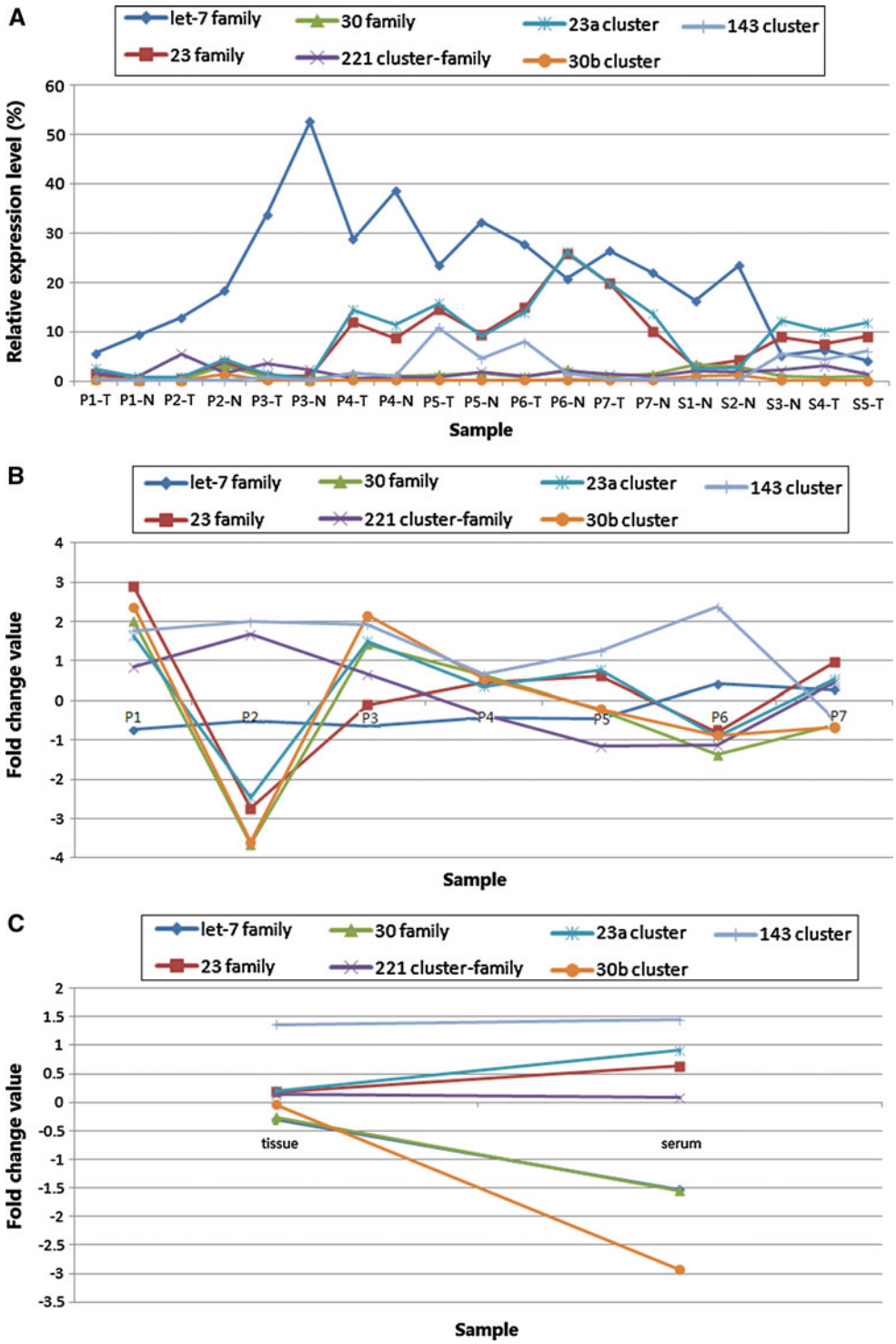

Fig. 4 Dynamic expression patterns of several miRNA gene clusters and gene families. **a** Expression patterns across different samples based on relative expression levels. P1-7 indicates tumor (T) and normal adjacent samples (N) from the seven patients. S1-5 indicates tumor (T) and normal serum samples (N). These miRNA groups exhibited various expression patterns. miR-221 and miR-222 are clustered in a gene cluster, and both comprise the miR-221 gene family. **b** Fold change values of each paired tumor and adjacent normal samples were estimated after taking the logarithm in seven patients. Some miRNA gene clusters and gene families showed inconsistent dysregulation trends. **c** Dysregulation trends were estimated according to the average fold change values in tissue and serum samples. Many of these miRNAs showed consistent expression trends

functional miRNAs, may serve as a potential biomarker (discussed below).

Diversity of differentially expressed miRNAs

Similar to the variation observed across miRNA profiles, abnormally expressed miRNAs also showed dynamic expression patterns (Fig. 2; Tables 1, 2, S2, S3). For example, seven tumor tissues showed inconsistent expression of abnormal miRNA populations, in spite of these seven samples originating from the same disease diagnosis (breast cancer). Only three abnormal miRNAs were shared by three patients, and about 10 were shared by two patients (Figure S4). Some abnormal miRNAs also showed inconsistent

up-regulation or down-regulation (Figure S4). Hierarchical clustering analysis also indicated that aberrantly expressed miRNAs showed larger differences across different samples (Fig. 2a). Most of these miRNAs (at least in a sample) characterized to be abnormal were classified as moderate miRNAs with stable expression (Fig. 2b). Wilcoxon sign-rank test for paired samples results indicated the widespread larger standard deviation (Table 1), which demonstrated the larger divergence among the seven patients. This was probably due to the dynamic or inconsistent dysregulation trends, especially for those exhibiting an opposite pattern of dysregulation (Fig. 2; Tables 1, 2, S2, S3). Individual differences across samples and higher sensitivity of deep sequencing technology also largely contributed to the expression divergence. Moreover, miRNA expression is time-specific and/or tissue-specific which also largely contributes to the dynamic miRNA expression profiles. For these reasons, it is difficult to obtain a consistent trend of dysregulation in miRNAs of patients with breast cancer. In fact, some miRNAs exhibited an opposite dysregulation pattern across different patients (discussed below).

Specific circulating miRNAs have attracted attention for their specific characteristics as potential non-invasive biomarkers in various human diseases. In the study, some equally pooled serum samples were also used to further understand the potential relationship between tissue and serum samples. Due to the small amount of serum samples available, we sequenced small RNAs using equally pooled serum samples from 10 patients and 10 normal women (Table S1). Compared to the single sample analysis, using mixed samples might mask the dynamic expression of miRNAs to some extent, especially for those miRNAs with opposite expression trends. Here, we also surveyed abnormal miRNAs in tumor tissues using mixed datasets based on relative expression levels, which was similar to equally pooled samples. Inconsistent expression patterns led to various aberrantly expressed miRNA populations based on analysis of mixed datasets (Figure S3; Table S3). We also obtained abnormal miRNAs in tissues and serum samples with BC based on mixed datasets. Interestingly, compared to normal samples, most of these miRNAs were down-regulated in tumor samples (Table 2). 9 miRNAs showed consistent dysregulation patterns in tissue and serum samples, and only miR-192 was up-regulated (Table 2). Although these samples were obtained from different patients, the abnormal expression of some miRNAs was consistent between tumor tissues and serum samples, suggesting that a serum-based biomarker may be feasible in detecting human disease. Indeed, abundantly expressed miRNA profiles were similar between tumor and adjacent normal samples from the same individual (Figure S2). The close relationship between tumor and serum samples supported that serum miRNAs can be potential biomarker for diagnosing human disease. Based on experimentally validated target mRNAs, all of these abnormal miRNAs contribute to multiple biological processes and to the occurrence and development of human diseases (Table S4). Almost all the abnormal miRNAs have been registered in the miRWalk database as target miRNAs in breast cancer, and they may contribute to multiple biological processes, including some human diseases (Table S4). However, the regulatory network of miRNAs are more complex than we thought and therefore, we further analyzed some abundantly expressed miRNA gene clusters and gene families.

Based on the expression profiles of the abundantly expressed miRNA gene clusters and gene families, these miRNA groups showed higher levels of expression. They also exhibited various expression and dysregulation patterns with possible biological implications (Fig. 4; Table S5). Although a series of miRNAs are co-transcribed in a gene cluster or homologous members in a gene family, the members showed inconsistent expression patterns in each sample (data not shown) [35, 36]. The dynamic expression pattern is mainly derived from the complexity of the degradation mechanism, such as variations in the half-life. With co-regulation of or co-contribution in biological processes, the special expression pattern may reveal functional implication. Indeed, most studies often survey and validate differentially expressed miRNA gene clusters and gene families in various human diseases, especially for human cancers, e.g. miR-17-92 gene cluster in several cancers [37–39]. In the study, as expected, some miRNA groups showed dysregulation patterns in tumor samples (Fig. 4). Similar to the findings from analysis of single abnormal miRNAs, these miRNA groups also have multiple biological roles (Table S5). This study not only revealed close functional relationships across miRNAs, but also the complexity of its regulatory network based on the various expression patterns.

miRNA profiles: more reasonable potential biomarker?

Generally, miRNAs may have multiple target mRNAs and could play crucial roles in various biological processes. Moreover, most mRNAs are regulated by more than one miRNA [40]. For example, miR-124 contributes to the development of nervous system [41], and has been identified to have 150 experimentally validated human target mRNAs according to the miRTarBase database [26]. Clustered miRNAs on special genomic regions, possibly co-transcribed from genomic DNA sequence, co-regulate biological processes. The coordinated regulation patterns can be extended to members in a miRNA gene family, such as the miR-17-92 gene cluster and the miR-17 gene family [38, 39]. Although many miRNAs have been identified as aberrantly expressed in cancers, it is difficult to characterize tumor-specific miRNAs because most miRNAs have crucial

and multiple roles in essential biological processes, such as cell development, differentiation, proliferation, apoptosis and metabolism [5]. Indeed, accumulating reports suggest that a series of special miRNAs based on miRNA gene cluster and gene family contribute to tumourigenesis. However, given the complexity of regulatory network of miRNAs, analysis at the single miRNA level or through a series of special miRNAs is not enough to elucidate pathological processes in a complex organism. It is essential that we should consider the potential interaction between different miRNAs, including their co-regulation. Although a specific miRNA is dysregulated, other miRNAs might perform similar biological roles. This redundancy and complexity in miRNA regulation makes it difficult to draw conclusions from studies based on a specific miRNA or a series of miRNAs. Indeed, in this study, a dynamic expression pattern was observed across different samples. Some miRNAs might show inconsistent dysregulation patterns due to individual differences (Figs. 2, S3, S4).

Collectively, the miRNA expression profiles vary across different individuals. Individual difference largely contributes to the diversity of miRNA expression profiles. Equally pooled samples, especially limited serum or plasma samples from patients, have been used to detect differentially expressed miRNA profiles and miRNAs associated with human diseases. The method can identify abnormal miRNAs, but cannot reveal individual differences. For example, those miRNAs with opposite dysregulation trends would be regarded as stable miRNAs. However, generally, each biological process which involves multiple and complex regulatory networks employs many miRNAs. Some miRNAs, including miRNA gene clusters and gene families, may co-regulate or co-contribute to biological processes. Many abnormal miRNAs had been identified in patients with breast cancer: down-regulated miRNAs include miR-31, 125b, 145, 200c and 342; up-regulated miRNAs include miR-10b, 21, 27a, 221 and 222 [13–20] (Table S6). Indeed, more miRNAs associated with breast cancer have been identified using experimental validation. Here, we also obtained abnormal miRNA profiles, including some novel aberrantly expressed miRNA species in BC. The main reason is derived from high sensitivity of deep sequencing, and especially the study is simultaneously involved in the tissue and serum samples. The aberrant expression of miRNAs might be a potential biomarker combined with pathological data. According to the results of the study, we also validated that miR-222 was significantly increased in the serum through further validation using larger samples [23]. However, as discussed above, these validated abnormal miRNAs should not singly contribute to the occurrence and development of breast cancer, whereas they may co-contribute or regulate multiple biological processes underlying breast cancer. For example, a series of miRNAs, including some miRNA gene clusters and gene families (e.g. miR-221 gene cluster and gene family), have been reported as potential biomarkers to classify human diseases. Serum miRNA expression profile (7 miRNAs) has high sensitivity and specificity for distinguishing various stages of pancreatic cancer [42].

Taken together, miRNA profiles, especially for those abundantly expressed miRNA profiles, should be a better potential biomarker than a single miRNA or several related miRNAs. Based on the close relationships of miRNAs in tumor tissues and serum samples, it may be possible to diagnose breast cancer through the miRNA expression profile using chip inspection in the future. The diagnosis method can be further extended to other human diseases.

Acknowledgments We appreciate all the patients and healthy controls who participated in this research. We thank Juncheng Dai, Yongyue Wei and Jianling Bai for their help in statistic analysis. The work was supported by National Natural Science Foundation of China (Nos. 30901232, 81072389 and 81102182), China Postdoctoral Science Foundation funded project (No. 2012M521100), University Science Research Project of Jiangsu Province (No. 12KJB310003), Jiangsu Planned Projects for Postdoctoral Research Funds (No. 1201022B), Science and Technology Development Fund Key Project of Nanjing Medical University (No. 2012NJMU001), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD).

Conflict of interest The authors declare no potential conflict of interests with respect to the authorship and/or publication of this paper.

References

1. Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297
2. Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T (2001) Identification of novel genes coding for small expressed RNAs. Science 294:853–858
3. Lau NC, Lim LP, Weinstein EG, Bartel DP (2001) An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science 294:858–862
4. Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet 8:93–103
5. Perera RJ, Ray A (2007) MicroRNAs in the search for understanding human diseases. BioDrugs 21:97–104
6. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E et al (2004) Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci USA 101:2999–3004
7. Calin GA (2006) A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia (vol 353, pg 1793, 2005). New Engl J Med 355:533
8. Caldas C, Brenton JD (2005) Sizing up miRNAs as cancer genes. Nat Med 11:712–714
9. Cha YH, Kim NH, Park C, Lee I, Kim HS et al (2012) miRNA-34 intrinsically links p53 tumor suppressor and Wnt signaling. Cell Cycle 11:1273–1281
10. Cubillos-Ruiz JR, Baird JR, Tesone AJ, Rutkowski MR, Scarlett UK et al (2012) Reprogramming tumor-associated dendritic cells in vivo using miRNA mimetics triggers protective immunity against ovarian cancer. Cancer Res 72:1683–1693

11. Zhu R, Yang Y, Tian Y, Bai J, Zhang X et al (2012) Ascl2 knockdown results in tumor growth arrest by miRNA-302b-related inhibition of colon cancer progenitor cells. PLoS ONE 7:e32170
12. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J et al (2005) MicroRNA expression profiles classify human cancers. Nature 435:834–838
13. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R et al (2005) MicroRNA gene expression deregulation in human breast cancer. Cancer Res 65:7065–7070
14. Ma L, Teruya-Feldstein J, Weinberg RA (2007) Tumour invasion and metastasis initiated by microRNA 10b in breast cancer. Nature 449:682–688
15. Valastyan S, Reinhardt F, Benaich N, Calogrias D, Szasz AM et al (2009) A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. Cell 137:1032–1046
16. Miller TE, Ghoshal K, Ramaswamy B, Roy S, Datta J et al (2008) MicroRNA-221/222 confers tamoxifen resistance in breast cancer by targeting p27Kip1. J Biol Chem 283:29897–29903
17. Shimono Y, Zabala M, Cho RW, Lobo N, Dalerba P et al (2009) Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. Cell 138:592–603
18. Mertens-Talcott SU, Chintharlapalli S, Li MR, Safe S (2007) The oncogenic microRNA-27a targets genes that regulate specificity protein transcription factors and the G(2)-M checkpoint in MDA-MB-231 breast cancer cells. Cancer Res 67:11001–11011
19. Adams BD, Cowee DM, White BA (2009) The role of miR-206 in the epidermal growth factor (EGF) induced repression of estrogen receptor-alpha (ER alpha) signaling and a luminal phenotype in MCF-7 breast cancer cells. Mol Endocrinol 23:1215–1230
20. Bhaumik D, Scott GK, Schokrpur S, Patil CK, Campisi J et al (2008) Expression of microRNA-146 suppresses NF-kappaB activity with reduction of metastatic potential in breast cancer cells. Oncogene 27:5643–5647
21. Kong W, He L, Coppola M, Guo J, Esposito NN et al (2010) MicroRNA-155 regulates cell survival, growth, and chemosensitivity by targeting FOXO3a in breast cancer. J Biol Chem 285:17869–17879
22. Castaneda CA, Agullo-Ortuno MT, Fresno Vara JA, Cortes-Funes H, Gomez HL et al (2011) Implication of miRNA in the diagnosis and treatment of breast cancer. Expert Rev Anticancer Ther 11:1265–1275
23. Wu Q, Wang C, Lu Z, Guo L, Ge Q (2012) Analysis of serum genome-wide microRNAs for breast cancer detection. Clin Chim Acta 413:1058–1065
24. Volinia S, Galasso M, Sana ME, Wise TF, Palatini J et al (2012) Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. Proc Natl Acad Sci USA 109:3024–3029
25. Png KJ, Yoshida M, Zhang XH, Shu W, Lee H et al (2011) MicroRNA-335 inhibits tumor reinitiation and is silenced through genetic and epigenetic mechanisms in human breast cancer. Genes Dev 25:226–231
26. Hsu SD, Lin FM, Wu WY, Liang C, Huang WC et al (2011) miRTarBase: a database curates experimentally validated microRNA–target interactions. Nucleic Acids Res 39:D163–D169
27. Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95:14863–14868
28. Chiang DY, Brown PO, Eisen MB (2001) Visualizing associations between genome sequences and gene expression data using genome-mean expression profiles. Bioinformatics 17:S49–S55
29. Dweep H, Sticht C, Pandey P, Gretz N (2011) miRWalk–database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform 44:839–847
30. Guo L, Liang T, Gu W, Xu Y, Bai Y et al (2011) Cross-mapping events in miRNAs reveal potential miRNA-mimics and evolutionary implications. PLoS ONE 6:e20517
31. Shearwin KE, Callen BP, Egan JB (2005) Transcriptional interference—a crash course. Trends Genet 21:339–345
32. Hongay CF, Grisafi PL, Galitski T, Fink GR (2006) Antisense transcription controls cell fate in *Saccharomyces cerevisiae*. Cell 127:735–745
33. Stark A, Bushati N, Jan CH, Kheradpour P, Hodges E et al (2008) A single Hox locus in Drosophila produces functional microRNAs from opposite DNA strands. Gene Dev 22:8–13
34. Lai EC, Wiel C, Rubin GM (2004) Complementary miRNA pairs suggest a regulatory role for miRNA:miRNA duplexes. RNA 10:171–175
35. Guo L, Lu Z (2010) Global expression analysis of miRNA gene cluster and family based on isomiRs from deep sequencing data. Comput Biol Chem 34:165–171
36. Yu J, Wang F, Yang GH, Wang FL, Ma YN et al (2006) Human microRNA clusters: genomic organization and expression profile in leukemia cell lines. Biochem Bioph Res Co 349:59–68
37. Bomben R, Gobessi S, Dal Bo M, Volinia S, Marconi D et al (2012) The miR-17 approximately 92 family regulates the response to Toll-like receptor 9 triggering of CLL cells with unmutated IGHV genes. Leukemia 26:1584–1593
38. Tong MH, Mitchell DA, McGowan SD, Evanoff R, Griswold MD (2012) Two miRNA clusters, Mir-17-92 (Mirc1) and Mir-106b-25 (Mirc3), are involved in the regulation of spermatogonial differentiation in mice. Biol Reprod 86:72
39. Feuermann Y, Robinson GW, Zhu BM, Kang K, Raviv N et al (2012) The miR-17/92 cluster is targeted by STAT5 but dispensable for mammary development. Genesis 50:665–671
40. Volinia S, Galasso M, Costinean S, Tagliavini L, Gamberoni G et al (2010) Reprogramming of miRNA networks in cancer and leukemia. Genome Res 20:589–599
41. Cheng LC, Pastrana E, Tavazoie M, Doetsch F (2009) miR-124 regulates adult neurogenesis in the subventricular zone stem cell niche. Nature Neurosci 12:399–408
42. Liu R, Chen X, Du Y, Yao W, Shen L et al (2011) Serum microRNA expression profile as a biomarker in the diagnosis and prognosis of pancreatic cancer. Clin Chem 58:610–618