Circulating miR-148b and miR-133a as biomarkers for breast cancer detection

Jie Shen[1], Qiang Hu[2,5], Michael Schrauder[6], Li Yan[2], Dan Wang[2], Leonardo Medico[3], Yuqing Guo[3], Song Yao[3], Qianqian Zhu[2], Biao Liu[2], Maochun Qin[2], Matthias W. Beckmann[6], Peter A. Fasching[6], Reiner Strick[6], Candace S. Johnson[4], Christine B. Ambrosone[3], Hua Zhao[1], Song Liu[2]

[1] Department of Epidemiology, the University of Texas MD Anderson Cancer Center, Houston, TX, USA

[2] Department of Biostatistics & Bioinformatics, Roswell Park Cancer Institute, Buffalo, NY, USA

[3] Department of Cancer Prevention and Control, Roswell Park Cancer Institute, Buffalo, NY, USA

[4] Department of Pharmacology and Therapeutics, Roswell Park Cancer Institute, Buffalo, NY, USA

[5] Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences, School of Basic Medicine, Peking Union Medical College, Beijing, P.R.China

[6] Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany

Correspondence to: Hua Zhao, **email**: hzhao2@mdanderson.org

Song Liu, **email**: song.liu@roswellpark.org

ABSTRACT

Circulating microRNAs have drawn a great deal of attention as promising novel biomarkers for breast cancer. However, to date, the results are mixed. Here, we performed a three-stage microRNA analysis using plasma samples from breast cancer patients and healthy controls, with efforts taken to address several pitfalls in detection techniques and study design observed in previous studies. In the discovery phase with 122 Caucasian study subjects, we identified 43 microRNAs differentially expressed between breast cancer cases and healthy controls. When those microRNAs were compared with published data from other studies, we identified three microRNAs, including miR-148b, miR-133a and miR-409-3p, whose plasma levels were significantly higher in breast cancer cases than healthy controls and were also significant in previous independent studies. In the validation phase with 50 breast cancer cases and 50 healthy controls, we validated the associations with breast cancer detection for miR-148b and miR-133a (P = 1.5×10^{-6} and 1.3×10^{-10}, respectively). In the *in-vitro* study phase, we found that both miR-148b and miR-133a were secreted from breast cancer cell lines, showing their secretory potential and possible tumor origin. Thus, our data suggest that both miR-148b and miR-133a have potential use as biomarkers for breast cancer detection.

INTRODUCTION

microRNAs are a class of small noncoding RNAs that play a central role in the regulation of mRNA expression [1-5]. The discovery that microRNA expression is frequently dys-regulated in a cancer-specific manner provides an opportunity to develop these RNAs as biomarkers for cancer detection [6-12]. Most previous studies on microRNA expression have been performed on tissue specimens. However, because tumor-derived microRNAs can be present in blood and appear to be stable to certain degree and protected from endogenous ribonuclease activity in circulation, some studies have shown diagnostic and prognostic potential for circulating microRNAs [13-25].

The potential of circulating microRNAs as

biomarkers for cancer early detection is particularly relevant to breast cancer because it is the most common cancer in US women, regardless of race or ethnicity, despite improvement in cancer screening and treatment strategies. Mammography is the current gold standard, but can have false negative rates of up to 20% [26]. The diagnosis of breast cancer relies on the histological examination of tissue biopsies, or cytology of fine-needle aspirates, which are both invasive procedures. Known serum-based tumor markers, such as CA15.3 and BR27.29, cannot be used for breast cancer detection due to low sensitivity. There is thus a need for developing novel markers that are minimally invasive, for the improved detection of breast cancer.

Previously, our group, as well as others, have compared the profiles of circulating microRNAs between breast cancer patients and healthy controls, and attempted to identify circulating microRNA-based breast cancer detection biomarkers [16, 19, 20, 22, 23, 25, 27-52]. Although a number of circulating microRNAs have been identified, the results are widely inconsistent. To date, no consistent diagnostic signature for circulating microRNAs in breast cancer is available. Several issues, including patient heterogeneity, microRNA contamination from various blood components, microRNA quantification platforms, data normalization, and biological significance are likely contributors to the inconsistency [9, 32, 53]. Thus, a study to appropriately address those issues is highly desirable.

In light of these observations, in the first phase, we performed quantitative PCR-based plasma microRNA profiling analysis among invasive breast cancer patients, patients with ductal carcinoma in situ (DCIS), and healthy women. In the second phase, we compared the significant microRNAs identified from our study with other published studies and then validated those with confirmed associations in independent samples. In the third phase, we determined whether cultured breast cancer cell lines secreted those validated circulating microRNAs into the culture medium, attempting to understand the possible origins of those identified microRNAs.

RESULTS

In the discovery cohort, a total of 122 women, including 52 with invasive breast cancer, 35 with DCIS, and 35 healthy controls, were included in the analysis. All of the study subjects were Caucasians. The mean age of invasive breast cancer cases, DCIS cases, and controls were 52, 55, and 55 years, respectively. There was no statistically significant difference in age. All invasive breast cases had histologically confirmed early stage (I and II) invasive ductal carcinoma. Blood samples were drawn prior to surgery. The tumor size ranged between 0.2 to 2.5 cm. ER, PR and HER2/neu status data were available for 50 patients with invasive breast cancer: 39 ER+, 11 ER−;

41 PR+, 9 PR−; and 12 HER2/neu+, 38 HER2/neu−. Eight patients had triple negative breast cancer.

Unsupervised hierarchical clustering analysis based on the top 75% of the most variable microRNAs across 106 samples passing quality control indicates that the microRNA expression profiles of invasive samples were intrinsically different from those of DCIS and control groups, while the microRNA expression profiles of DCIS and control groups were not different from each other. As shown in Figure 1, three major clusters exist (from right to left): the first cluster with predominately DCIS and control samples, the second cluster with all invasive samples, and the third cluster with a mixture of DCIS, control and invasive samples. A similar clustering pattern was observed when all microRNAs in the PCR panel were used (data not shown).

Pair-wise comparisons between invasive, DCIS and control groups were performed in order to identify differentially expressed microRNAs with at least 2-fold expression change at the significance level of FDR <0.01. In consistent with the clustering analysis, no differentially expressed microRNAs were found between DCIS and control groups, and the expression levels of a number of microRNAs were found to be significantly elevated in the invasive samples compared with DCIS or controls. For the comparison between invasive and control groups, we identified 43 differentially expressed microRNAs, with 40 up-regulated in invasive samples and 3 down-regulated. For the comparison between invasive and DCIS groups, we identified 27 differentially expressed microRNAs, with 24 up-regulated in invasive samples and 3 down-regulated. The detailed list of differentially expressed microRNAs is summarized in Table 1. As shown in Figure 2, the majority (23 out of 27) of differentially expressed microRNAs identified from invasive versus DCIS comparison are included in the list of differentially expressed microRNAs identified from invasive versus control comparison.

When comparing between the invasive breast cancer and the healthy control group, the most significantly differentially expressed microRNA was miR-199a-5p. The plasma levels of miR-199a-5p were more than 14-fold higher in patients with invasive breast cancer than in healthy controls. After adjusting for multiple comparisons, the P value is 4.1×10^{-12}. We also performed receiver operating characteristic curve (ROC) analysis for each differentially expressed microRNA. The areas under the curve (AUC) for miR-199a-5p was 0.91, indicating clear separation between the invasive breast cancer group and the healthy control group from our discovery cohort (Table 1). When comparing between invasive breast cancer and DCIS groups, the most significantly differentially expressed microRNA was miR-136. The plasma levels of miR-136 were more than 18-fold higher in patients with invasive breast cancer than in healthy controls. After adjusting for multiple comparisons, the P value reached 7.7×10^{-8}. The AUC for miR-136 was 0.89, indicating

Table 1: Differentially expressed microRNAs identified from Invasive versus Control/DCIS comparisons

Invasive versus Control				Invasive versus DCIS			
ID	FC	FDR	AUC	ID	FC	FDR	AUC
hsa-miR-136	19.63386	1.27×10^{-08}	0.867218	hsa-miR-136	18.59352	7.70×10^{-08}	0.88628
hsa-miR-199a-5p	14.33637	4.15×10^{-12}	0.910559	hsa-miR-33a	9.180245	3.91×10^{-05}	0.827586
hsa-miR-485-3p	9.954621	1.10×10^{-08}	0.874704	hsa-miR-199a-5p	6.217108	5.35×10^{-07}	0.892883
hsa-miR-33a	9.615268	2.47×10^{-05}	0.789204	hsa-miR-326	5.678471	7.92×10^{-06}	0.870139
hsa-miR-495	9.239414	2.72×10^{-08}	0.853034	hsa-miR-130a	4.440197	1.48×10^{-06}	0.882979
hsa-miR-543	8.382473	1.10×10^{-08}	0.86643	hsa-miR-495	4.092802	2.79×10^{-04}	0.799707
hsa-miR-326	6.789743	7.23×10^{-07}	0.878251	hsa-miR-485-3p	3.689859	5.47×10^{-04}	0.7854
hsa-let-7d	5.873482	1.25×10^{-08}	0.888495	hsa-let-7d	3.640281	1.88×10^{-05}	0.872707
hsa-miR-130a	4.630065	4.36×10^{-07}	0.86643	hsa-miR-142-5p	3.54654	2.13×10^{-05}	0.845194
hsa-miR-221	4.575009	9.21×10^{-08}	0.841608	hsa-miR-543	3.374637	4.91×10^{-04}	0.768525
hsa-miR-382	4.410125	3.24×10^{-04}	0.720646	hsa-miR-1974	3.29879	2.03×10^{-03}	0.812546
hsa-miR-199a-3p	4.346987	8.51×10^{-05}	0.836091	hsa-miR-22	3.17841	7.24×10^{-06}	0.827586
hsa-miR-22*	4.335605	7.43×10^{-06}	0.839638	hsa-miR-22*	3.095461	5.04×10^{-04}	0.781731
hsa-miR-766	4.213962	2.20×10^{-08}	0.855004	hsa-miR-107	2.732845	7.24×10^{-06}	0.855833
hsa-miR-339-5p	4.143662	1.10×10^{-08}	0.866036	hsa-miR-221	2.720919	3.05×10^{-04}	0.790536
hsa-miR-1974	3.999563	3.08×10^{-04}	0.849094	hsa-miR-339-5p	2.67183	3.65×10^{-05}	0.825385
hsa-miR-127-3p	3.912698	5.63×10^{-04}	0.753743	hsa-miR-766	2.32762	5.74×10^{-04}	0.787968
hsa-miR-409-3p	3.884377	7.75×10^{-05}	0.778566	hsa-miR-423-3p	2.250449	3.91×10^{-05}	0.810712
hsa-miR-151-5p	3.730705	1.11×10^{-09}	0.877463	hsa-miR-151-5p	2.241882	5.21×10^{-05}	0.80741
hsa-miR-133a	3.652671	2.82×10^{-05}	0.795508	hsa-miR-331-3p	2.241426	8.03×10^{-03}	0.717168
hsa-miR-99b	3.399696	4.13×10^{-06}	0.808905	hsa-miR-32	2.237895	7.25×10^{-03}	0.809978
hsa-miR-107	3.28739	1.11×10^{-07}	0.865642	hsa-let-7f	2.195287	3.22×10^{-03}	0.761555
hsa-miR-142-5p	3.286851	4.51×10^{-05}	0.816785	hsa-miR-28-5p	2.025899	8.11×10^{-04}	0.768158
hsa-let-7f	3.278603	9.59×10^{-06}	0.810481	hsa-miR-342-3p	-2.00785	3.24×10^{-04}	0.787601
hsa-miR-139-5p	3.274487	4.18×10^{-06}	0.789992	hsa-miR-140-3p	-2.05119	8.04×10^{-03}	0.862436
hsa-miR-28-5p	3.137736	1.47×10^{-07}	0.845942	hsa-miR-720	-2.46154	2.33×10^{-04}	0.780998
hsa-miR-22	2.922818	1.91×10^{-05}	0.845548	hsa-miR-144*	-3.72644	7.05×10^{-07}	0.936537
hsa-miR-423-3p	2.545573	2.70×10^{-06}	0.816391				
hsa-miR-335	2.519429	5.94×10^{-03}	0.732467				
hsa-miR-328	2.446737	7.08×10^{-06}	0.803783				
hsa-miR-23b	2.367462	9.18×10^{-03}	0.763593				
hsa-miR-30c	2.33765	4.90×10^{-04}	0.767139				
hsa-miR-484	2.312884	9.74×10^{-05}	0.841608				
hsa-miR-331-3p	2.267972	5.81×10^{-03}	0.71119				
hsa-miR-148b	2.166974	1.96×10^{-05}	0.810087				
hsa-miR-339-3p	2.131754	6.73×10^{-05}	0.760441				
hsa-miR-152	2.073642	2.40×10^{-03}	0.746257				
hsa-miR-126	2.052121	2.78×10^{-05}	0.774232				
hsa-miR-374a	2.04563	4.35×10^{-03}	0.751379				
hsa-miR-30b	2.015793	3.24×10^{-04}	0.764381				
hsa-miR-194	-2.12943	2.27×10^{-03}	0.806147				
hsa-miR-375	-2.77381	3.40×10^{-03}	0.739559				
hsa-miR-144*	-3.74132	3.48×10^{-07}	0.94011				
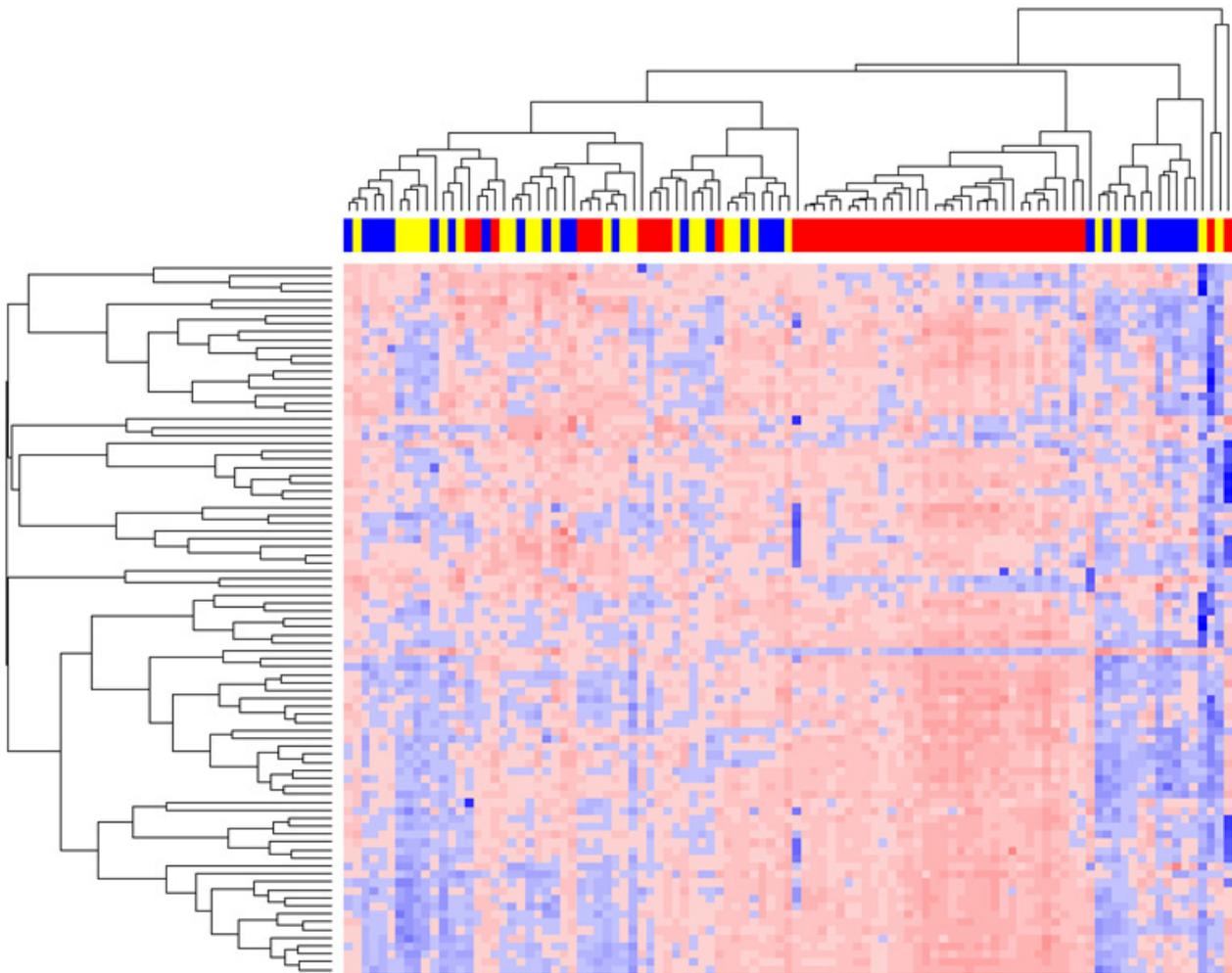
Figure 1: Unsupervised hierarchical clustering analysis of serum microRNA expression profiles. In the clustering heat map, red indicates up-regulated while blue indicates down regulated. In the sample clustering dendrogram, red indicates invasive samples, blue indicates DCIS samples, while yellow indicates control samples.
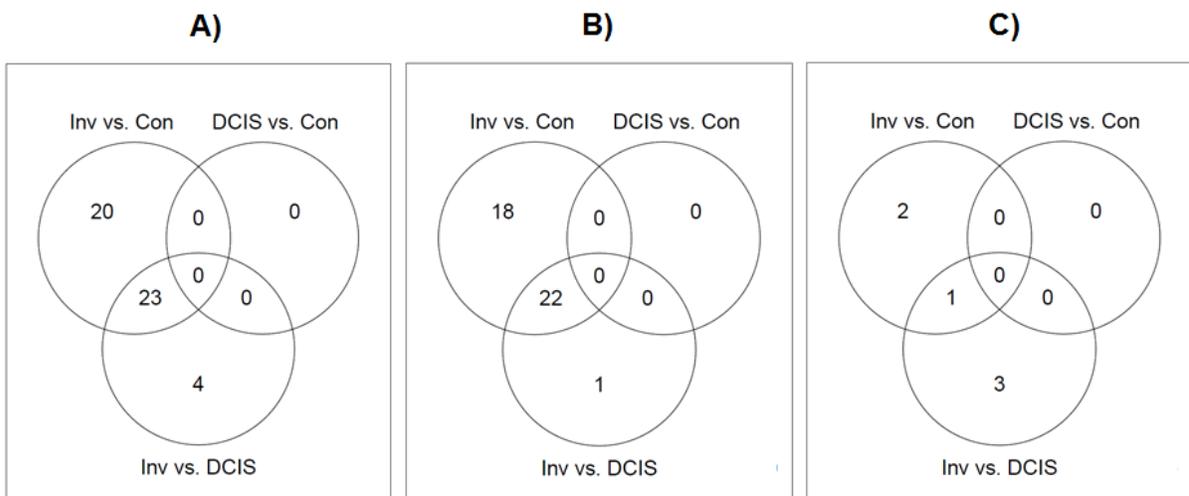
Figure 2: Venn diagrams showing the overlap of differentially expressed microRNAs identified from each of the three comparisons: invasive group versus control group, invasive group versus DCIS group, and DCIS group versus control group, respectively. A) The overlap of all differentially expressed microRNAs; B) The overlap of up-regulated differentially expressed microRNAs; C) The overlap of down-regulated differentially expressed microRNAs.

clear separation between invasive breast cancer and DCIS groups in our discovery cohort (Table 1).

We are mindful of the wide inconsistency among the previous circulating microRNA studies. Thus, we compared our results with others extracted from published studies [16, 19, 20, 22, 23, 25, 27-52], to determine if the significant microRNAs found in our studies were independently reproduced in any other studies. To ensure the quality, we restricted review to studies with at least one validation population. After an extensive literature search, none of our top significant microRNAs were reported previously in other studies. However, we found three microRNAs, namely miR-148b (fold change=2.17, adjusted $P=1.96\times10^{-5}$), miR-133a (fold change = 3.65, adjusted $P=2.82\times10^{-5}$) and miR-409-3p (fold change=3.88, adjusted $P=7.75\times10^{-5}$), which have also been reported to be up-regulated in circulation from breast cancer patients compared to healthy controls in other studies [38, 54].

Those three microRNAs were included in the validation stage analysis. Our validation cohort included a total of 50 patients with stage I and II breast cancer and 50 healthy controls. All study subjects were Caucasians. In this independent cohort, we found the plasma levels of miR-148b and miR-133a were also significantly higher in breast cancer patients than healthy controls (Figure 3). For miR-148b, the fold change was 5.1 ($P = 1.5\times10^{-6}$). For miR-133a, the fold change was 8.3 ($P = 1.3\times10^{-10}$). The AUC for the combination of these two microRNAs is 0.86. No significant association was found for miR-409-3p in the validation cohort.

Lastly, we determined whether miR-148b and miR-133a act as secretory microRNAs and are excreted into the culture media by MCF-7 and MDA-231 cell lines. We compared levels of miR-148 and miR133a in culture media among baseline (0 hr), 24 and 48 hrs. We observed both microRNA expression increased, suggesting that both
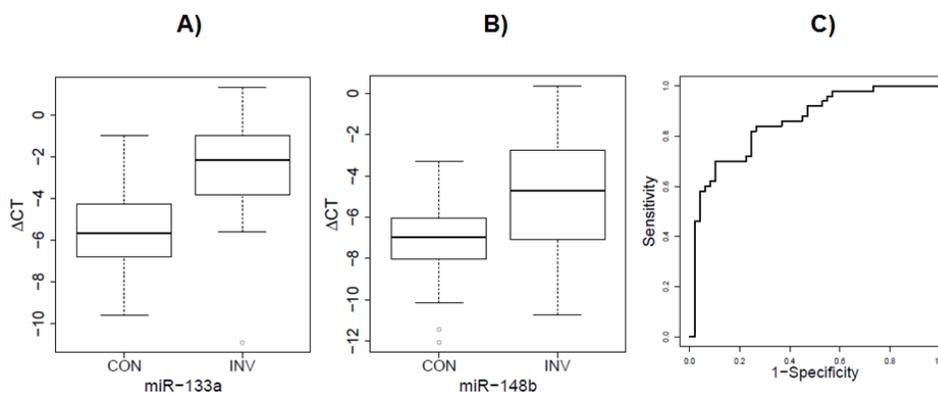
Figure 3: Elevated serum level of of miR-133a and mir-148b in invasive group versus control group of validation cohort. A) The fold change of miR-133a was 8.3 ($P = 1.3\times10^{-10}$) in the validation cohort. B) The fold change of miR-148b was 5.1 ($P = 1.5\times10^{-6}$) in the validation cohort. C) The ROC curve derived from the combinations of these two miRNAs (AUC = 0.86) in the validation cohort.
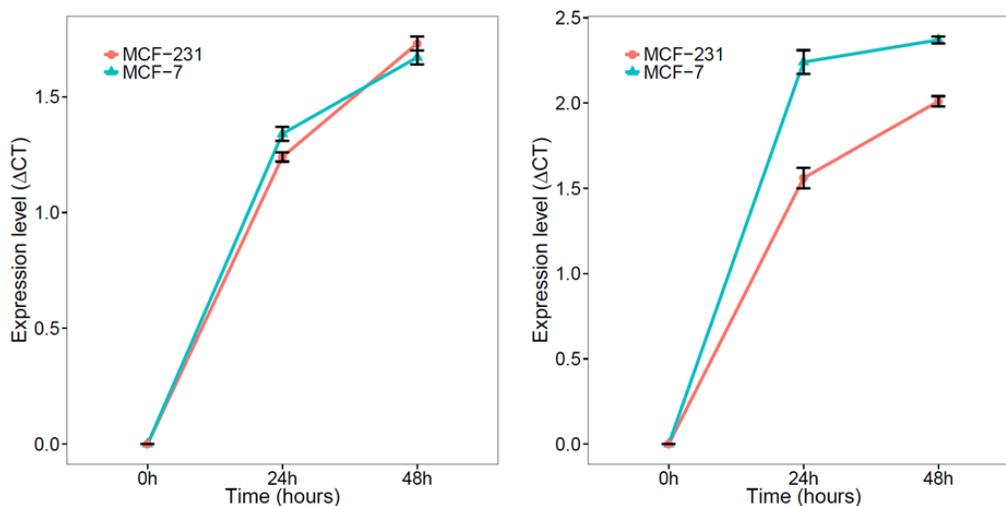
Figure 4: Expression of miR-148b and miR-133a in culture media of breast cancer cell lines (MCT-7 and MD-231). A) miR-148b levels in media of both MCF-7 and MD-231 increased with longer incubation intervals; B) miR-133a levels in media of both MCF-7 and MD-231 increased with longer incubation intervals.

microRNAs are possible secretory microRNAs (Figure 4).

DISCUSSION

Published studies on circulating microRNAs as cancer detection biomarkers have identified a wide variety of microRNAs. Unfortunately, few of them have been validated in other studies. Recently, Leidner et al compared the circulating microRNAs results from several breast cancer studies [32] and were unable to demonstrate reproducibility by various measures across the different datasets. The widespread inconsistency may reflect technical difficulties, such as lack of standardization of biospecimens, quantification platforms, and internal controls. For example, the biospecimen type varies among studies. Some use plasma and serum, and some use whole blood. Serum and plasma are considered equivalent, although microRNA concentration appeared to be higher in serum [55]. The use of whole blood will lead to the isolation of microRNAs from many cell types including those within blood cells, but not just circulating microRNAs, warranting caution when comparing microRNA profiles derived from blood with those from serum or plasma [56]. As demonstrated by Pritchard et al, even serum and plasma samples may be contaminated with microRNAs from a variety of blood cells [56]. Such microRNAs may include mir-16, mir-150, mir-486-5p, let-7a, mir-574-3p, mir-223, mir-197, mir-451, and mir-92a. To be cautious, we did not include any of them in the analysis. The results may also vary by experimental approaches. So far, microarray profiling, next generation sequencing, quantitative RT-PCR profiling, or targeted assays of specific microRNAs have been employed in previous studies. Even with the same samples, the results from different platforms could be inconsistent, and RT-PCR, generally considered as the "gold standard" assay, is required to validate the discovery. In our current study, the RT-PCR approach was used in both discovery and validation cohort. There is also a problem about internal controls. To date, there is no consistent internal control microRNA available for normalizing circulating microRNA expression. The use of spike-in or a small RNA for data normalization in similar studies has sometimes been considered to be problematic due to their suspected instability. In the current study, because of using a larger profiling cohort, we had the capacity to select internal reference microRNAs empirically.

The inconsistency may also reflect the limitations stemming from study design. For example, most of the published studies do not have a validation population and the sample size for most of the studies is modest, thus patient heterogeneity becomes an ever bigger issue. To overcome those issues, the current study included both discovery and validation populations, both of which have adequate sample size. To further foster the consistency, we performed a literature review on circulating microRNAs and breast cancer, and compared our data with others [16, 19, 20, 22, 23, 25, 27-52]. Three microRNAs which were significantly different between breast cancer cases and controls in the discovery cohort of our study, namely miR-148b, miR-133a and miR-409-3p, were independently reproduced in other studies. miR-148b and miR-409-3p were reported previously by Cuk et al [38]. Using a two-step approach (discovery and validation), they found that 4 microRNAs (miR-148b, miR-376c, miR-409-3p and miR-801) were shown to be significantly up-regulated in the plasma of breast cancer patients in the validation cohort (N=127). Using three microRNAs (miR-148b, miR-409-3p and miR-801) in combination in ROC curve analysis, the discriminatory power reaches to 0.69 between breast cancer cases and healthy controls. miR-133a was reported in a recent study by Chan et al [54]. Using paired breast cancer tumor and normal tissues and serum samples, they found miR-133a, miR-133b, miR-1 and miR-92a were the most important diagnostic markers in the discovery cohort (N=32), which were then successfully validated (N=132). Results from our validation cohort confirmed the association with breast cancer detection for miR-148b and miR-133a.

One of the major problems about circulating microRNA research is the lack of an understanding of how those identified significant microRNAs enter the circulation. It has been hypothesized that most of the circulating microRNAs are scavengers of tumor cell apoptosis or necrosis. If that proves to be true, we expect to observe correlations between paired tumor tissues and serum/plasma. However, several recent studies have shown that the microRNA profiles between sera and the corresponding matched tumor are largely dissimilar, and circulating microRNAs do not reflect their abundance in the malignant cells [38, 54]. Those observations raise the possibility of the existence of alternative mechanism, through which tumor cells may specifically secrete circulating microRNAs. Such action may help modify the surroundings and create a favorable environment for tumor progression. In the current study, we observed that both miR-133a and miR-148b can be secreted by proliferating breast cancer cell lines. Although we do not know whether this observation is only limited to breast cancer cell lines or can also be observed in tumor tissues, our findings underscore the need to explore the underlying molecular mechanisms of circulating microRNAs.

Recently, The Cancer Genome Atlas (TCGA) data have been analyzed and released to the public. The breast cancer data in TCGA include a large number of primary breast tumors characterized by genomic DNA copy number arrays, DNA methylation, exome sequencing, mRNA arrays, and microRNA sequencing. Those data create a valuable resource for us to explore the biological functions of specific microRNAs. Although the available clinical information is still limited, we found that the expression of miR-148b was significant differed by

histological types ($P = 7.588 \times 10^{-12}$ and $q=3.29 \times 10^{-9}$). In addition, in a recent analysis using TCGA RNA and small RNA data, Volinia et al found that miR-148b was among a panel of mRNAs/microRNAs which were associated with overall survival [57]. Although we did not observe similar associations for miR-133a, miR-133a has been reported to be a tumor suppressor and had prognostic value in esophageal squamous cell carcinoma [58, 59] and other cancers [60].

In addition to cancers, circulating microRNAs, especially inflammatory-related circulating microRNAs, may also be used as biomarkers for aging and other aging-related diseases [61, 62]. In a recent study by Noren Hooten et al, they found serum levels of miR-151a-5p, miR-181a-5p and miR-1248 were significantly decreased in 20 older individuals compared to 20 younger individuals [61]. In humans, miR-1248 was found to regulate the expression of mRNAs involved in inflammatory pathways and miR-181a was found to correlate negatively with the pro-inflammatory cytokines IL-6 and TNFα and to correlate positively with the anti-inflammatory cytokines TGFβ and IL-10. In addition, a number of circulating miRNAs, which are functionally related to proinflammatory, seem to be promising biomarkers for the major age-related diseases such as cardiovascular disease (CVD), type 2 diabetes mellitus (T2DM), Alzheimer Disease (AD), and rheumatoid arthritis (RA) [62].

In summary, we discovered and validated two circulating microRNAs, miR-148b and miR-133a, associated with early stage breast cancer. Since both microRNAs can be secreted from breast cancer cell lines, their existence in circulation may potentially serve as noninvasive screening and prognostic biomarkers for breast cancer. A large-scale prospective trial is needed to confirm their clinical usage.

MATERIALS AND METHODS

Study population

The study was approved by Institutional Research Board (IRB) of Roswell Park Cancer Institute (RPCI). Biosepcimens from RPCI were used as an initial discovery cohort. Anonymized biospecimens and questionnaire data used in this study were made available through the Data Bank and BioRepository (DBBR) of RPCI [63]. Patients are enrolled through site-specific clinics prior to surgery and chemotherapy, and controls are cancer-free individuals who are visitors or family members of patients, or are enrolled through community events. Relationships between patients and controls are carefully annotated, so to avoid overmatching patients to their own family or friends. Written consent is obtained prior to enrollment into the DBBR, which allows DBBR to provide anonymized

biospecimens and questionnaire data for research (such as this study) without further consent. Patients and controls are consented to provide a non-fasting blood sample and to complete an epidemiological questionnaire. Blood samples are drawn in phlebotomy and transferred to the DBBR laboratory. Following DBBR standard operating procedure (SOP), samples are processed and blood components stored within one hour of collection to minimize degradation. Ten milliliters of whole blood was obtained from each study subject. EDTA-plasma was extracted by centrifuging whole blood at 3,000 rpm for 10 minutes at room temperature. All extracted plasma samples are stored in phased liquid nitrogen. To minimize the effect of freeze-thaw on circulating microRNAs, we only used plasma samples which had not been previously thawed. In this study, a total of 52 women with early stage invasive breast cancer (stage I and II), 35 women with DCIS, and 35 cancer-free women were included in the microRNA profiling analysis.

The validation cohort consisted of 50 early stage (stage I and II) breast cancer patients and 50 healthy controls who participated in a prospective case-control study for the molecular detection of breast cancer (MODE-B Study), conducted at the University Hospital Erlangen, Erlangen, Germany. Patients were newly diagnosed with breast cancer and pre-treatment blood samples were collected. For each patient, five tubes of peripheral blood (serum, plasma, PAXgene®, CPDA, EDTA), urine samples, fresh frozen tissue samples from the core needle biopsy at diagnosis and paraffin embedded tumor samples were available. An epidemiological questionnaire was completed via an in-person interview by cases and controls included in this study.

Cell cultures

To determine the secretory potential of the significant microRNAs identified from microRNA profiling, two breast cancer cell lines, MCF-7 and MDA-231, were cultured and a fraction of the culture medium was collected at 0, 24 and 48 hours after the initial seeding of cells in 10-cm dish. 0.5×10^6 cells were seeded in the initial dish.

RNA isolation

Plasma and culture medium microRNAs were isolated using the miRNeasy kit (Qiagen) with minor modifications. In brief, 700 µl of QIAzol reagent was added to 400 µl of plasma sample or 1ml culture medium. The sample was mixed in a tube, followed by the addition of 3 µl of miSPIKE, spiked-in microRNA, at a concentration of 0.1 µM (IDT) and 140 µl of chloroform. After mixing vigorously for 15 s, the sample was then centrifuged at 12,000 g for 15 minutes. The upper aqueous

phase was carefully transferred to a new collection tube, and 1.5 volume of ethanol was added. The sample was then applied directly to a silica membrane-containing column and the RNA was bound and cleaned using buffers provided by the manufacturer to remove impurities. The immobilized RNA was then collected from the membrane with a low salt elution buffer. Similar method was used to extract microRNAs from the cell culture medium. The quality and quantity of the RNA was evaluated by 260/280 ratio using NanoDrop spectrophotometry (NanoDrop ND-1000 Technologies Inc.) and Agilent 2100 Bioanalyzer (Agilent Technologies). The efficiency of small RNA isolation was monitored by the amount of spiked-in microRNA recovered by using PCR with sequence specific primers (IDT).

microRNA profiling

In the discovery cohort, microRNA expression in the plasma samples was profiled using Exiqon MiRCURY LNA Universal RT microRNA PCR Technology, following the manufacturer's recommended protocol (Exiqon). The Serum/Plasma Focus microRNA PCR Panel is a microRNA-specific, LNA™-based system designed for sensitive and accurate detection of serum/plasma microRNA by quantitative real-time PCR using SYBR® Green. A total of 168 human microRNAs commonly found in human serum/plasma and 7 reference microRNAs were included in this panel. Only 20 ng total RNAs is needed in each analysis. Briefly, 20 ng total RNAs were reverse transcribed using the RT enzyme. The RT mixture was incubated for 60 min at 42˚C followed by heat-inactivation of the reverse transcriptase for 5 min at 95˚C. Multiplex RT reactions were diluted 62.5-fold with water, and 55µl of each diluted product was combined with 55µl of 2X Universal SYBR® Green master mix. One-hundred µl of the sample/master mix from each Multiplex pool was loaded into the array. Then, the array was centrifuged and mechanically sealed with the Applied Biosystems sealer device. qRT-PCR was carried out on an Applied BioSystems 7900HT real-time PCR instrument using the manufacturer's recommended cycling conditions. The experiment was performed at RPCI Genomic Core Shared Resource.

For the validation cohort, the expression levels of microRNAs were confirmed with a Taqman-based real-time quantitative PCR (RT-qPCR) using individual microRNA-specific primers and probes as described by the manufacturer (Applied Biosystems). The first-strand microRNA-cDNA PCR template was generated from 50 ng of total RNA according to the manufacturer's instructions. Approximately 2.5 ng of cDNA was then used in the PCR on a CFX96 Touch Real-Time PCR Detection System from Bio-Rad. Triplicate samples, validated endogenous controls, and inter-assay controls were used throughout. The qRT-PCR results were analyzed by SDS

2.2.2. To be consistent with the discovery cohort, we chose miR-93 as the endogenous control. RT-qPCR data were the normalized expression values in which the endogenous control miR-93 was used as the reference gene. For each assay, the Ct (Cycle threshold) of microRNA of interest in the TaqMan qPCR assay was subtracted from the average miR-93 Ct value to obtain a ΔCt value (miR-93 - microRNA of interest). A higher delta Ct value indicates a higher expression level of the microRNA of interest. For cell culture medium, similar qPCR was applied at baseline, 24 and 48 hours. For each assay at 24 and 48 hours, the Ct (Cycle threshold) of microRNA of interest in the TaqMan qPCR assay was subtracted from the baseline Ct value to obtain a ΔCt value.

Data Analysis

For data quality control, we excluded samples with less than half of the profiled microRNAs with a Ct value less than 32. A total of 19 (out of 122) samples were removed, which left a total of 103 study subjects, including 47 patients with invasive breast cancer, 29 patients with DCIS, and 27 controls, for further analysis. We also excluded the list of blood cells derived microRNAs (including miR-16, miR-150, miR-486-5p, let-7a, miR-574-3p, miR-223, miR-197, miR-451, and miR-92a) from downstream analysis. The Exiqon's Serum/Plasma Focus microRNA PCR Panel supplies five microRNAs (miR-93, miR-103, miR-191, miR-423-3p, and miR-425) as candidate reference microRNAs for normalization. In our samples, miR-93 was the most stably expressed (Figure S1), so miR-93 was chosen as reference microRNA for normalization of the RT-PCT data.

Unsupervised hierarchical clustering algorithm based on the average linkage and Pearson correlation metric was performed based on the normalized expression profiles from the top 75% of the most variable (i.e., largest variance) microRNAs across 103 samples, as well as from all microRNAs [64]. We used the Limma program in the R-based Bioconductor package to calculate the statistical significance for the level of differential expression [65]. Briefly, a linear model was fit to the data, with cell means corresponding to the different conditions and a random effect for array. The Benjamini & Hochberg method was used to control the false discovery rate (FDR) for multiple testing [66].

REFERENCES

1. Wu W, Sun M, Zou GM and Chen J. MicroRNA and cancer: Current status and prospective. International journal of cancer Journal international du cancer. 2007; 120(5):953-960.

2. Wang Y, Stricker HM, Gou D and Liu L. MicroRNA: past and present. Frontiers in bioscience : a journal and virtual library. 2007; 12:2316-2329.

3. Chen PY and Meister G. microRNA-guided posttranscriptional gene regulation. Biological chemistry. 2005; 386(12):1205-1218.

4. Alvarez-Garcia I and Miska EA. MicroRNA functions in animal development and human disease. Development. 2005; 132(21):4653-4662.

5. Gregory RI and Shiekhattar R. MicroRNA biogenesis and cancer. Cancer research. 2005; 65(9):3509-3512.

6. Lehmann U. Aberrant DNA methylation of microRNA genes in human breast cancer - a critical appraisal. Cell and tissue research. 2014.

7. Xiao YF, Yong X, Fan YH, Lu MH, Yang SM and Hu CJ. microRNA detection in feces, sputum, pleural effusion and urine: novel tools for cancer screening (Review). Oncology reports. 2013; 30(2):535-544.

8. de Planell-Saguer M and Rodicio MC. Analytical aspects of microRNA in diagnostics: a review. Analytica chimica acta. 2011; 699(2):134-152.

9. Andorfer CA, Necela BM, Thompson EA and Perez EA. MicroRNA signatures: clinical biomarkers for the diagnosis and treatment of breast cancer. Trends in molecular medicine. 2011; 17(6):313-319.

10. Zoon CK, Starker EQ, Wilson AM, Emmert-Buck MR, Libutti SK and Tangrea MA. Current molecular diagnostics of breast cancer and the potential incorporation of microRNA. Expert review of molecular diagnostics. 2009; 9(5):455-467.

11. Deng S, Calin GA, Croce CM, Coukos G and Zhang L. Mechanisms of microRNA deregulation in human cancer. Cell cycle. 2008; 7(17):2643-2646.

12. Zhang L, Yang N and Coukos G. MicroRNA in human cancer: one step forward in diagnosis and treatment. Advances in experimental medicine and biology. 2008; 622:69-78.

13. Kosaka N, Iguchi H and Ochiya T. Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis. Cancer science. 2010; 101(10):2087-2092.

14. Wang J, Zhang KY, Liu SM and Sen S. Tumor-Associated Circulating MicroRNAs as Biomarkers of Cancer. Molecules. 2014; 19(2):1912-1938.

15. Farina NH, Wood ME, Perrapato SD, Francklyn CS, Stein GS, Stein JL and Lian JB. Standardizing Analysis of Circulating MicroRNA: Clinical and Biological Relevance. Journal of cellular biochemistry. 2013.

16. Schwarzenbach H. Circulating nucleic acids as biomarkers in breast cancer. Breast cancer research : BCR. 2013; 15(5):211.

17. Ma R, Jiang T and Kang X. Circulating microRNAs in cancer: origin, function and application. Journal of experimental & clinical cancer research : CR. 2012; 31:38.

18. Yu DC, Li QG, Ding XW and Ding YT. Circulating MicroRNAs: Potential Biomarkers for Cancer. International journal of molecular sciences. 2011; 12(3):2055-2063.

19. Mostert B, Sieuwerts AM, Martens JW and Sleijfer S. Diagnostic applications of cell-free and circulating tumor cell-associated miRNAs in cancer patients. Expert review of molecular diagnostics. 2011; 11(3):259-275.

20. Heneghan HM, Miller N, Kelly R, Newell J and Kerin MJ. Systemic miRNA-195 differentiates breast cancer from other malignancies and is a potential biomarker for detecting noninvasive and early stage disease. The oncologist. 2010; 15(7):673-682.

21. Vlassov VV, Laktionov PP and Rykova EY. Circulating nucleic acids as a potential source for cancer biomarkers. Current molecular medicine. 2010; 10(2):142-165.

22. Heneghan HM, Miller N, Lowery AJ, Sweeney KJ, Newell J and Kerin MJ. Circulating microRNAs as novel minimally invasive biomarkers for breast cancer. Annals of surgery. 2010; 251(3):499-505.

23. Zhu W, Qin W, Atasoy U and Sauter ER. Circulating microRNAs in breast cancer and healthy subjects. BMC research notes. 2009; 2:89.

24. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, Knudsen BS, Stirewalt DL, Gentleman R, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105(30):10513-10518.

25. Zhao H, Shen J, Medico L, Wang D, Ambrosone CB and Liu S. A pilot study of circulating miRNAs as potential biomarkers of early stage breast cancer. PloS one. 2010; 5(10):e13735.

26. Blanchard K, Colbert JA, Kopans DB, Moore R, Halpern EF, Hughes KS, Smith BL, Tanabe KK and Michaelson JS. Long-term risk of false-positive screening results and subsequent biopsy as a function of mammography use. Radiology. 2006; 240(2):335-342.

27. Waters PS, McDermott AM, Wall D, Heneghan HM, Miller N, Newell J, Kerin MJ and Dwyer RM. Relationship between circulating and tissue microRNAs in a murine model of breast cancer. PloS one. 2012; 7(11):e50459.

28. Healy NA, Heneghan HM, Miller N, Osborne CK, Schiff R and Kerin MJ. Systemic mirnas as potential biomarkers

for malignancy. International journal of cancer Journal international du cancer. 2012; 131(10):2215-2222.

29. Wu X, Somlo G, Yu Y, Palomares MR, Li AX, Zhou W, Chow A, Yen Y, Rossi JJ, Gao H, Wang J, Yuan YC, Frankel P, Li S, Ashing-Giwa KT, Sun G, et al. De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer. Journal of translational medicine. 2012; 10:42.

30. Kumar S, Keerthana R, Pazhanimuthu A and Perumal P. Overexpression of circulating miRNA-21 and miRNA-146a in plasma samples of breast cancer patients. Indian journal of biochemistry & biophysics. 2013; 50(3):210-214.

31. Eichelser C, Flesch-Janys D, Chang-Claude J, Pantel K and Schwarzenbach H. Deregulated serum concentrations of circulating cell-free microRNAs miR-17, miR-34a, miR-155, and miR-373 in human breast cancer development and progression. Clinical chemistry. 2013; 59(10):1489-1496.

32. Leidner RS, Li L and Thompson CL. Dampening enthusiasm for circulating microRNA in breast cancer. PloS one. 2013; 8(3):e57841.

33. Zeng RC, Zhang W, Yan XQ, Ye ZQ, Chen ED, Huang DP, Zhang XH and Huang GL. Down-regulation of miRNA-30a in human plasma is a novel marker for breast cancer. Medical oncology. 2013; 30(1):477.

34. Mar-Aguilar F, Mendoza-Ramirez JA, Malagon-Santiago I, Espino-Silva PK, Santuario-Facio SK, Ruiz-Flores P, Rodriguez-Padilla C and Resendez-Perez D. Serum circulating microRNA profiling for identification of potential breast cancer biomarkers. Disease markers. 2013; 34(3):163-169.

35. Ng EK, Li R, Shin VY, Jin HC, Leung CP, Ma ES, Pang R, Chua D, Chu KM, Law WL, Law SY, Poon RT and Kwong A. Circulating microRNAs as specific biomarkers for breast cancer detection. PloS one. 2013; 8(1):e53141.

36. Chen W, Cai F, Zhang B, Barekati Z and Zhong XY. The level of circulating miRNA-10b and miRNA-373 in detecting lymph node metastasis of breast cancer: potential biomarkers. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine. 2013; 34(1):455-462.

37. Si H, Sun X, Chen Y, Cao Y, Chen S, Wang H and Hu C. Circulating microRNA-92a and microRNA-21 as novel minimally invasive biomarkers for primary breast cancer. Journal of cancer research and clinical oncology. 2013; 139(2):223-229.

38. Cuk K, Zucknick M, Heil J, Madhavan D, Schott S, Turchinovich A, Arlt D, Rath M, Sohn C, Benner A, Junkermann H, Schneeweiss A and Burwinkel B. Circulating microRNAs in plasma as early detection markers for breast cancer. International journal of cancer Journal international du cancer. 2013; 132(7):1602-1612.

39. Sun Y, Wang M, Lin G, Sun S, Li X, Qi J and Li J. Serum microRNA-155 as a potential biomarker to track disease in breast cancer. PloS one. 2012; 7(10):e47003.

40. Madhavan D, Zucknick M, Wallwiener M, Cuk K, Modugno C, Scharpff M, Schott S, Heil J, Turchinovich A, Yang R, Benner A, Riethdorf S, Trumpp A, Sohn C, Pantel K, Schneeweiss A, et al. Circulating miRNAs as surrogate markers for circulating tumor cells and prognostic markers in metastatic breast cancer. Clinical cancer research : an official journal of the American Association for Cancer Research. 2012; 18(21):5972-5982.

41. Cortez MA, Welsh JW and Calin GA. Circulating microRNAs as noninvasive biomarkers in breast cancer. Recent results in cancer research Fortschritte der Krebsforschung Progres dans les recherches sur le cancer. 2012; 195:151-161.

42. Wang H, Tan G, Dong L, Cheng L, Li K, Wang Z and Luo H. Circulating MiR-125b as a marker predicting chemoresistance in breast cancer. PloS one. 2012; 7(4):e34210.

43. Jung EJ, Santarpia L, Kim J, Esteva FJ, Moretti E, Buzdar AU, Di Leo A, Le XF, Bast RC, Jr., Park ST, Pusztai L and Calin GA. Plasma microRNA 210 levels correlate with sensitivity to trastuzumab and tumor presence in breast cancer patients. Cancer. 2012; 118(10):2603-2614.

44. van Schooneveld E, Wouters MC, Van der Auwera I, Peeters DJ, Wildiers H, Van Dam PA, Vergote I, Vermeulen PB, Dirix LY and Van Laere SJ. Expression profiling of cancerous and normal breast tissues identifies microRNAs that are differentially expressed in serum from patients with (metastatic) breast cancer and healthy volunteers. Breast cancer research : BCR. 2012; 14(1):R34.

45. Schwarzenbach H, Milde-Langosch K, Steinbach B, Muller V and Pantel K. Diagnostic potential of PTEN-targeting miR-214 in the blood of breast cancer patients. Breast cancer research and treatment. 2012; 134(3):933-941.

46. Schrauder MG, Strick R, Schulz-Wendtland R, Strissel PL, Kahmann L, Loehberg CR, Lux MP, Jud SM, Hartmann A, Hein A, Bayer CM, Bani MR, Richter S, Adamietz BR, Wenkel E, Rauh C, et al. Circulating micro-RNAs as potential blood-based markers for early stage breast cancer detection. PloS one. 2012; 7(1):e29770.

47. Zhao R, Wu J, Jia W, Gong C, Yu F, Ren Z, Chen K, He J and Su F. Plasma miR-221 as a predictive biomarker for chemoresistance in breast cancer patients who previously received neoadjuvant chemotherapy. Onkologie. 2011; 34(12):675-680.

48. Yang M, Chen J, Su F, Yu B, Su F, Lin L, Liu Y, Huang JD and Song E. Microvesicles secreted by macrophages shuttle invasion-potentiating microRNAs into breast cancer cells. Molecular cancer. 2011; 10:117.

49. Sieuwerts AM, Mostert B, Bolt-de Vries J, Peeters D, de Jongh FE, Stouthard JM, Dirix LY, van Dam PA, Van Galen A, de Weerd V, Kraan J, van der Spoel P, Ramirez-Moreno R, van Deurzen CH, Smid M, Yu JX, et al. mRNA and microRNA expression profiles in circulating tumor cells and primary tumors of metastatic breast cancer patients. Clinical cancer research : an official journal of the American

Association for Cancer Research. 2011; 17(11):3600-3618.

50. Asaga S, Kuo C, Nguyen T, Terpenning M, Giuliano AE and Hoon DS. Direct serum assay for microRNA-21 concentrations in early and advanced breast cancer. Clinical chemistry. 2011; 57(1):84-91.

51. Roth C, Rack B, Muller V, Janni W, Pantel K and Schwarzenbach H. Circulating microRNAs as blood-based markers for patients with primary and metastatic breast cancer. Breast cancer research : BCR. 2010; 12(6):R90.

52. Ross JS. Measuring circulating miRNAs: the new "PSA" for Breast Cancer? The oncologist. 2010; 15(7):656.

53. Cheng HH, Yi HS, Kim Y, Kroh EM, Chien JW, Eaton KD, Goodman MT, Tait JF, Tewari M and Pritchard CC. Plasma processing conditions substantially influence circulating microRNA biomarker levels. PloS one. 2013; 8(6):e64795.

54. Chan M, Liaw CS, Ji SM, Tan HH, Wong CY, Thike AA, Tan PH, Ho GH and Lee AS. Identification of circulating microRNA signatures for breast cancer detection. Clinical cancer research : an official journal of the American Association for Cancer Research. 2013; 19(16):4477-4487.

55. Wang K, Yuan Y, Cho JH, McClarty S, Baxter D and Galas DJ. Comparing the MicroRNA spectrum between serum and plasma. PloS one. 2012; 7(7):e41561.

56. Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyaji MM, Tait JF and Tewari M. Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. Cancer prevention research. 2012; 5(3):492-497.

57. Volinia S and Croce CM. Prognostic microRNA/mRNA signature from the integrated analysis of patients with invasive breast cancer. Proceedings of the National Academy of Sciences of the United States of America. 2013; 110(18):7413-7417.

58. Akanuma N, Hoshino I, Akutsu Y, Murakami K, Isozaki Y, Maruyama T, Yusup G, Qin W, Toyozumi T, Takahashi M, Suito H, Hu X, Sekino N and Matsubara H. MicroRNA-133a regulates the mRNAs of two invadopodia-related proteins, FSCN1 and MMP14, in esophageal cancer. British journal of cancer. 2014; 110(1):189-198.

59. Kano M, Seki N, Kikkawa N, Fujimura L, Hoshino I, Akutsu Y, Chiyomaru T, Enokida H, Nakagawa M and Matsubara H. miR-145, miR-133a and miR-133b: Tumor-suppressive miRNAs target FSCN1 in esophageal squamous cell carcinoma. International journal of cancer Journal international du cancer. 2010; 127(12):2804-2814.

60. Nohata N, Hanazawa T, Enokida H and Seki N. microRNA-1/133a and microRNA-206/133b clusters: dysregulation and functional roles in human cancers. Oncotarget. 2012; 3(1):9-21.

61. Noren Hooten N, Fitzpatrick M, Wood WH, 3rd, De S, Ejiogu N, Zhang Y, Mattison JA, Becker KG, Zonderman AB and Evans MK. Age-related changes in microRNA levels in serum. Aging. 2013; 5(10):725-740.

62. Olivieri F, Rippo MR, Procopio AD and Fazioli F. Circulating inflamma-miRs in aging and age-related diseases. Frontiers in genetics. 2013; 4:121.

63. Ambrosone CB, Nesline MK and Davis W. Establishing a cancer center data bank and biorepository for multidisciplinary research. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology. 2006; 15(9):1575-1577.

64. Eisen MB, Spellman PT, Brown PO and Botstein D. Cluster analysis and display of genome-wide expression patterns. Proceedings of the National Academy of Sciences of the United States of America. 1998; 95(25):14863-14868.

65. Smyth GK. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical applications in genetics and molecular biology. 2004; 3:Article3.

66. Hochberg Y and Benjamini Y. More powerful procedures for multiple significance testing. Statistics in medicine. 1990; 9(7):811-818.