

RESEARCH ARTICLE

Open Access

# miRNA expression profiling of 51 human breast cancer cell lines reveals subtype and driver mutation-specific miRNAs

Muhammad Riaz<sup>†</sup>, Marijn TM van Jaarsveld<sup>†</sup>, Antoinette Hollestelle, Wendy JC Prager-van der Smissen, Anouk AJ Heine, Antonius WM Boersma, Jingjing Liu, Jean Helmijr, Bahar Ozturk, Marcel Smid, Erik A Wiemer, John A Foekens and John WM Martens\*

## Abstract

**Introduction:** Breast cancer is a genetically and phenotypically complex disease. To understand the role of miRNAs in this molecular complexity, we performed miRNA expression analysis in a cohort of molecularly well-characterized human breast cancer cell lines to identify miRNAs associated with the most common molecular subtypes and the most frequent genetic aberrations.

**Methods:** Using a microarray carrying LNA<sup>TM</sup> modified oligonucleotide capture probes, expression levels of 725 human miRNAs were measured in 51 breast cancer cell lines. Differential miRNA expression was explored by unsupervised cluster analysis and was then associated with the molecular subtypes and genetic aberrations commonly present in breast cancer.

**Results:** Unsupervised cluster analysis using the most variably expressed miRNAs divided the 51 breast cancer cell lines into a major and a minor cluster predominantly mirroring the luminal and basal intrinsic subdivision of breast cancer cell lines. One hundred and thirteen miRNAs were differentially expressed between these two main clusters. Forty miRNAs were differentially expressed between basal-like and normal-like/claudin-low cell lines. Within the luminal-group, 39 miRNAs were associated with *ERBB2* overexpression and 24 with *E-cadherin* gene mutations, which are frequent in this subtype of breast cancer cell lines. In contrast, 31 miRNAs were associated with *E-cadherin* promoter hypermethylation, which, contrary to *E-cadherin* mutation, is exclusively observed in breast cancer cell lines that are not of luminal origin. Thirty miRNAs were associated with *p16<sup>INK4</sup>* status while only a few miRNAs were associated with *BRCA1*, *PIK3CA/PTEN* and *TP53* mutation status. Twelve miRNAs were associated with DNA copy number variation of the respective locus.

**Conclusion:** Luminal-basal and epithelial-mesenchymal associated miRNAs determine the subdivision of miRNA transcriptome of breast cancer cell lines. Specific sets of miRNAs were associated with *ERBB2* overexpression, *p16<sup>INK4</sup>* or *E-cadherin* mutation or *E-cadherin* methylation status, which implies that these miRNAs may contribute to the driver role of these genetic aberrations. Additionally, miRNAs, which are located in a genomic region showing recurrent genetic aberrations, may themselves play a driver role in breast carcinogenesis or contribute to a driver gene in their vicinity. In short, our study provides detailed molecular miRNA portraits of breast cancer cell lines, which can be exploited for functional studies of clinically important miRNAs.

\* Correspondence: j.martens@erasmusmc.nl

† Contributed equally

Erasmus University Medical Center, Daniel den Hoed Cancer Center,  
Department of Medical Oncology and Cancer Genomics Center, Dr.  
Molewaterplein 50, 3015 GE Rotterdam, the Netherlands

## Introduction

Numerous lines of evidence indicate that breast cancer is genetically and epigenetically not just one disease but a diverse group of diseases with diverse clinically relevant biological and phenotypical features. Recent technological advances in molecular profiling have led to the identification of an increasing number of molecular subtypes in breast cancer, each with distinct co-regulated and anti-regulated genes. However, the biology of these molecular subtypes and their underlying genetic drivers may be affected by numerous biological factors, including miRNAs.

miRNAs are a class of small nonprotein-coding genes that regulate the expression of genes post-transcriptionally via sequence-specific interaction with the 3' UTR of target mRNAs, resulting in inhibition of translation and/or mRNA degradation [1,2]. A large number of studies have established that miRNAs play essential roles in biological processes, such as development [3,4], cell proliferation [5], apoptosis [6], stress response, and tumorigenesis [7,8]. Aberrant expression levels of miRNAs have been observed in many solid cancers including breast cancer. In breast cancer, the expression levels of several miRNAs are significantly different between normal and cancerous tissues, between breast cancers of different molecular subtypes [9-11] with a different prognosis [12-14], and between breast cancers showing different responses to endocrine therapy [15,16]. Despite significant progress in the last few years on miRNA biology, the exact biological functions and the genetic factors driving their expression have been revealed for only a limited number of miRNAs in breast cancer.

Human breast cancer cell lines are excellent experimental models and renewable resources to investigate biological functions of clinically important miRNAs both in *in vitro* cultured conditions and *in vivo* when raised as xenografts [6,17-20]. Here, using microarrays we analyzed miRNA expression levels in 51 molecularly well-characterized human breast cancer cell lines. We explored the association of individual miRNA expression levels with intrinsic subtypes and the most common recurrent genetic aberrations. Through this analysis, we provide a catalog of miRNAs in human breast cancer cell lines, which can be used to understand underlying biology of clinically relevant miRNAs and to reveal the genetic factors that may be involved in their regulation.

## Materials and methods

### Breast cancer cell lines

The cohort of 51 human breast cancer cell lines used in this study is listed in Table 1. The origin of the cell lines has been described elsewhere [21]. Prior to expression profiling, all cell lines were established to be

genetically unique, monoclonal and of correct identity by performing STR profiling using the PowerPlex<sup>®</sup> 16 system (Promega, Madison, WI, USA). The PowerPlex<sup>®</sup> 16 system included 15 STRs and one gender-discriminating locus (Penta E, D18S51, D21S11, TH01, D3S1358, FGA, TPOX, D8S1179, vWA, Amelogenin, Penta D, CSF1PO, D16S539, D7S820, D13S317, and D5S818).

The method involved isolation of the genomic DNA from each breast cancer cell line using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany) and 10 ng of the isolated DNA was used as the input for the multiplex PCR. The following PCR conditions were used: 95°C for 11 minutes, 96°C for 1 minute, 10×(94°C for 30 seconds, 60°C for 30 seconds, 70°C for 45 seconds), 22×(90°C for 30 seconds, 60°C for 30 seconds, 70°C for 45 seconds), and then 60°C for 30 minutes. The PCR was carried out using primers linked with fluorescent dyes (6-carboxy-4',5'-dichloro-2',7'-dimethoxy-fluorescein, fluorescein, and carboxy-tetramethylrhodamine). The labeled amplicons were detected using the 3130xl Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) and data were analyzed using the Genemarker 1.91 software from Softgenetics (State College, PA, USA).

The end result for each cell line was an electropherogram with each STR allele represented as one or more peaks of an appropriate fluorophore. The authenticity of all cell lines, except SUM cell lines, were assessed by comparing the generated STR profiles with the source STR profiles present in the American Type Culture Collection and the Deutsche Sammlung von Mikroorganismen und Zellkulturen. As no reference is available, the STR profiles of the SUM cell lines were matched with the profiles generated from the earliest passage of these cell lines stored in the in-house culture collection.

For experiments, each cell line was cultured in triplicate on collagen-coated petri dishes in RPMI 1640 medium supplemented with 10% heat-inactivated fetal bovine serum and antibiotic agents (100 µg/ml penicillin G and 80 µg/ml streptomycin). The petri dishes were placed in a humidified atmosphere of 5% CO<sub>2</sub> and 95% air at 37°C until cultures were 70 to 80% confluent. Ethical approval of our study was not necessary as our experiment involved only *in vitro* propagated human breast cancer cell lines.

### Total RNA isolation

Total RNA from all samples was isolated using RNeasy Lysis Reagent (Qiagen, Crawley, UK) according to the manufacturer's manual. Briefly, a biological sample was lysed in RNeasy Lysis Reagent and the lysate was separated into an aqueous phase and an organic phase after the addition of chloroform. DNA

**Table 1 Molecular and biochemical characterization of 51 human breast cancer cell lines**

Cell line	Molecular classification Intrinsic subtype <sup>a</sup>	Clinically relevant protein expression							
		ER	PGR	ERBB2	EGFR	CK5	CK8-18	CK19	CK14
SUM185PE	Luminal	-	-	-	-	-	+	+	-
MDA-MB-175VII	Luminal	+	-	-	-	-	+	+	-
BT483	Luminal	+	-	-	-	-	+	+	-
T47D	Luminal	+	+	-	+	-	+	+	-
MDA-MB-415	Luminal	+	+	-	-	-	+	+	-
ZR75-1	Luminal	+	+	-	-	-	+	+	-
MCF-7	Luminal	+	+	-	-	-	+	ND	-
SUM52PE	Luminal	+	-	-	-	-	+	ND	-
UACC812	Luminal	+	+	++	-	-	+	+	-
MDA-MB-361	Luminal	+	+	++	-	-	+	+	-
ZR75-30	Luminal	+	-	++	-	-	+	+	-
SK-BR-5	Luminal	-	-	++	-	-	+	+	-
OCUB-F	Luminal	-	-	++	-	+	+	+	-
MPE600	Luminal	+	-	++	-	-	+	+	-
MDA-MB-134VI	Luminal	+	-	-	-	-	+	+	-
SUM44PE	Luminal	+	+	-	-	-	+	+	-
CAMA-1	Luminal	+	-	-	-	-	+	+	-
BT474	Luminal	+	+	++	-	-	+	+	-
MDA-MB-330	Luminal-ERBB2+	+	-	++	-	-	+	+	-
HCC1419	Luminal-ERBB2+	+	-	++	-	-	+	+	-
HCC202	Luminal-ERBB2+	-	-	++	-	-	+	+	-
SUM190PT	Luminal-ERBB2+	-	-	++	-	-	+	+	-
SUM225CWN	Luminal-ERBB2+	-	-	++	-	-	+	+	-
UACC893	Luminal-ERBB2+	-	-	++	+	-	+	+	-
SK-BR-3	Luminal-ERBB2+	-	-	++	-	-	+	+	-
EVSA-T	Luminal-ERBB2+	-	+	++	-	-	+	+	-
MDA-MB-453	Luminal-ERBB2+	-	-	++	-	-	+	+	-
HCC1569	ER-negative-ERBB2+	-	-	++	+	-	-	-	-
HCC1954	ER-negative-ERBB2+	-	-	++	+	+	+	-	-
HCC1500	ER-negative-ERBB2+	+	+	-	-	-	+	+	-
DU4475	Other	-	-	-	-	+	-	-	-
SUM229PE	Basal-like	-	-	-	+	+	+	+	-
HCC1937	Basal-like	-	-	-	+	+	+	-	-
MDA-MB-468	Basal-like	-	-	-	+	+	+	+	-
HCC1806	Basal-like	-	-	-	+	+	+	+	+
HCC70	Basal-like	-	-	-	+	+	+	+	-
HCC1143	Basal-like	-	-	-	+	+	+	+	-
BT20	Basal-like	-	-	-	+	+	+	+	-
SUM149PT	Basal-like	-	-	-	+	+	+	+	-
HCC1395	Basal-like	-	-	-	+	+	-	-	-
SK-BR-7	Normal-like/claudin-low	-	-	-	+	-	+	-	-
Hs578T	Normal-like/claudin-low	-	-	-	+	-	-	-	-
MDA-MB-231	Normal-like/claudin-low	-	-	-	+	-	-	-	-
SUM1315M02	Normal-like/claudin-low	-	-	-	+	-	-	-	-
MDA-MB-436	Normal-like/claudin-low	-	-	-	+	-	-	-	-
BT549	Normal-like/claudin-low	-	-	-	+	-	-	-	-
MDA-MB-157	Normal-like/claudin-low	-	-	-	+	-	-	-	-
SUM159PT	Normal-like/claudin-low	-	-	-	+	-	-	-	-
MDA-MB-435s	Normal-like/claudin-low	-	-	-	-	-	-	-	-
SUM102PT	Normal-like/claudin-low	ND	ND	ND	ND	ND	ND	ND	ND
HCC38	Normal-like/claudin-low	ND	ND	ND	ND	ND	ND	ND	-

CK, cytokeratin; ER, estrogen receptor; EGFR, epidermal growth factor receptor; PGR, progesterone receptor. <sup>a</sup>Cell lines are classified based on expression of the intrinsic gene set defined by Perou and Colleagues [23]. For protein expression: +, expression; ++, overexpression; -, no expression; ND, not determined. Other subtype includes DU4475 since it is not classified to any subtype of breast cancer by the intrinsic gene set.

and protein were subsequently removed by carefully transferring the aqueous phase containing RNA to a fresh Eppendorf tube. The RNA was obtained from the aqueous phase by an isopropanol precipitation, washed with ethanol and air dried for subsequent procedures. The purity of the isolated RNA was checked using NanoDrop<sup>®</sup> ND-1000 (Isogen Life Science, De Meern, the Netherlands) ensuring spectrophotometric ratios of  $A_{260\text{ nm}}/A_{280\text{ nm}} \sim 2$  and  $A_{260\text{ nm}}/A_{230\text{ nm}} \geq 2$ , and the quality control checks were performed according to the previously described methodology [22].

#### Gene expression profiling

Total RNA (200 ng) was reverse transcribed, copied into double-strand cDNA and labeled to yield biotin-labeled cRNA using the 3' IVT express Kit (Affymetrix, Santa Clara, CA, USA) according to the manufacturer's instructions. Biotin-labeled cRNA was subsequently fragmented and loaded onto an Affymetrix GeneTitan instrument. The hybridization cocktail was applied to Human Genome HT\_HG-U133\_Plus\_PM GeneChip 96-well arrays. All steps including hybridization, washing and scanning were carried out automatically inside the instrument. The raw data (.CEL files) were normalized by the RMA method using the default settings of the Affymetrix Expression Console<sup>™</sup> software and were used for statistical analysis. The microarray data were deposited in the Gene Expression Omnibus data repository [GEO:GSE41313].

For subtype classification of the cell lines, Perou and colleagues' intrinsic gene set of 496 genes [23] were matched to the Affymetrix probe sets using Unigene cluster numbers. Some of the genes have multiple probe sets present. To ensure analysis of only the informative probe sets, the probe sets that did not vary across all samples were removed, leaving the most variable ones for analysis (66% of the probe sets). These genes were then used to cluster the breast cancer cell lines. The intrinsic molecular subtypes were assigned as follows: luminal-type cell lines that exhibited higher expression of *ESR1*, *GATA3*, *TFF3*, and *FOXA1*; ERBB2-positive cell lines that showed higher expression of *ERBB2*, *GRB7*, and *STARD3*; basal-like cell lines, which were characterized by higher expression of *KRT5*, *KRT17*, *BST2*, and *FABP7*; and normal-like/ Claudin-low cell lines that did not show *KRT5* and *KRT17* expression and have low expression of claudin 3, claudin 4, and claudin 7 genes. The ERBB2-positive cell lines were further designated as luminal-ERBB2-positive and estrogen receptor (ER)-negative/ERBB2-positive cell lines, because the former in addition to *ERBB2* overexpression also show *ESR1* gene expression on microarray while the latter do not.

#### miRNA expression analysis

miRNA expression profiling was performed using miRNA microarrays according to a previously published method [24]. In brief, 1  $\mu\text{g}$  total RNA was labeled with Cy3 using the ULS aRNA labeling kit (Kreatech, Amsterdam, the Netherlands). The LNA<sup>™</sup> modified oligonucleotide capture probe set (miRBase version 10.0, annotation version 13; Exiqon, Vedbaek, Denmark) was spotted in duplicate on Nexterion E glass slides in Nexterion Spot buffer (Schott, Elmsford, NY, USA) using a Virtek Chipwriter Pro (Bio-Rad, Hercules, CA, USA). The RNA sample with a labeling efficiency  $> 15$  pmol Cy3/ $\mu\text{g}$  RNA was used for hybridization in a salt-based hybridization buffer (Ocimum Biosolutions, Hyderabad, India) overnight at 60° C using a Tecan HS4800 pro hybridization station. Hybridized slides were scanned in a Tecan LS Reloaded scanner. Data were extracted using Imagene software (6.0 standard edition; TCAN, Chapel Hill, NC, USA). The raw data were normalized using quantile normalization and used for statistical analysis. The normalized expression data are provided in Table S1 in Additional file 1. To validate our findings, subsets of the differentially expressed miRNAs between two miRNA-driven clusters were quantified using the Taqman Human MicroRNA Assay Set from Applied Biosystems (Nieuwerkerk aan den IJssel, the Netherlands) as described previously [12].

#### SNP arrays

Genomic DNA from all cell lines was extracted using the QIAamp DNA Mini Kit (Qiagen). Genomic DNA (500 ng) was used as the starting material to capture genome-wide chromosomal aberrations with the aid of the Affymetrix Genome-Wide Human SNP 6.0 array technology. The steps were performed according to Affymetrix's recommended protocols. In summary, after digestion of the genomic DNA using either restriction enzymes NspI or StyI, adaptors were ligated to the obtained DNA fragments. These fragments were subsequently amplified using PCR, fragmented, end-labeled with biotin and hybridized onto GeneChip SNP 6.0 arrays. After hybridization the arrays were washed and scanned to generate the raw data (.CEL files) using the Affymetrix Genotyping Console<sup>™</sup> software. The chromosomal gains and losses were calculated using SNP copy number variation (CNV) on the same chromosome. The chromosomal regions containing gains or losses were correlated with the expression level of the miRNAs located on the same genomic regions. The SNP data were deposited in the GEO data repository [GEO:GSE41313].

#### Protein expression and mutational analysis

Protein expression data of ER, progesterone receptor, ERBB2, epidermal growth factor receptor, and cytokeratin

(CK) 5, CK8-18, CK19, and CK14 were used from previously published work [25], except for 10 HCC cell lines that were characterized by immunohistochemistry using the same protocols as described before [25]. Mutational analysis of *p16<sup>INK4a</sup>*, *BRCA1*, *E-cadherin*, *PIK3CA*, and *PTEN* and promoter hypermethylation analysis of *E-cadherin* were previously reported for all cell lines except for HCC cell lines. These were separately analyzed according to the previously published methods [25] (data presented in Table S2 in Additional file 1).

#### Hierarchical clustering and statistical analyses

Hierarchical clustering analyses of significant miRNAs were performed using Cluster 3 software [26] and the expression patterns of miRNAs and mRNAs in the heat maps were visualized using Treeview 1.1.6 R2. Average linkage clustering was carried out on both samples and mRNA and miRNA expression data, respectively, using Pearson correlation as a distance measure. Differential expression of miRNAs between two groups was determined using the univariate *t* test in BRB-array tools 3.7 developed by Dr. Richard Simon and the BRB-Array-Tools Development Team). A permutation value of  $P < 0.05$  was considered statistically significant and used to select differentially expressed miRNAs for supervised hierarchical cluster analysis. The associations between continuous variables were tested using Spearman Rank correlation (*R*s values). The Kruskal-Wallis test was used to evaluate differences among groups and  $P < 0.05$  was considered statistically significant.

## Results

### Molecular features of human breast cancer cell lines

The 51 human breast cancer cell lines that were used in this study are listed in Table 1, together with the protein expression results of ER, progesterone receptor, ERBB2, epidermal growth factor receptor and CK5, CK8-18, CK19, and CK14. All cell lines were first profiled for global mRNA expression using microarrays, and the cell lines in Table 1 are grouped into molecular subtypes based on the expression of the intrinsic genes originally described by Perou and colleagues [23].

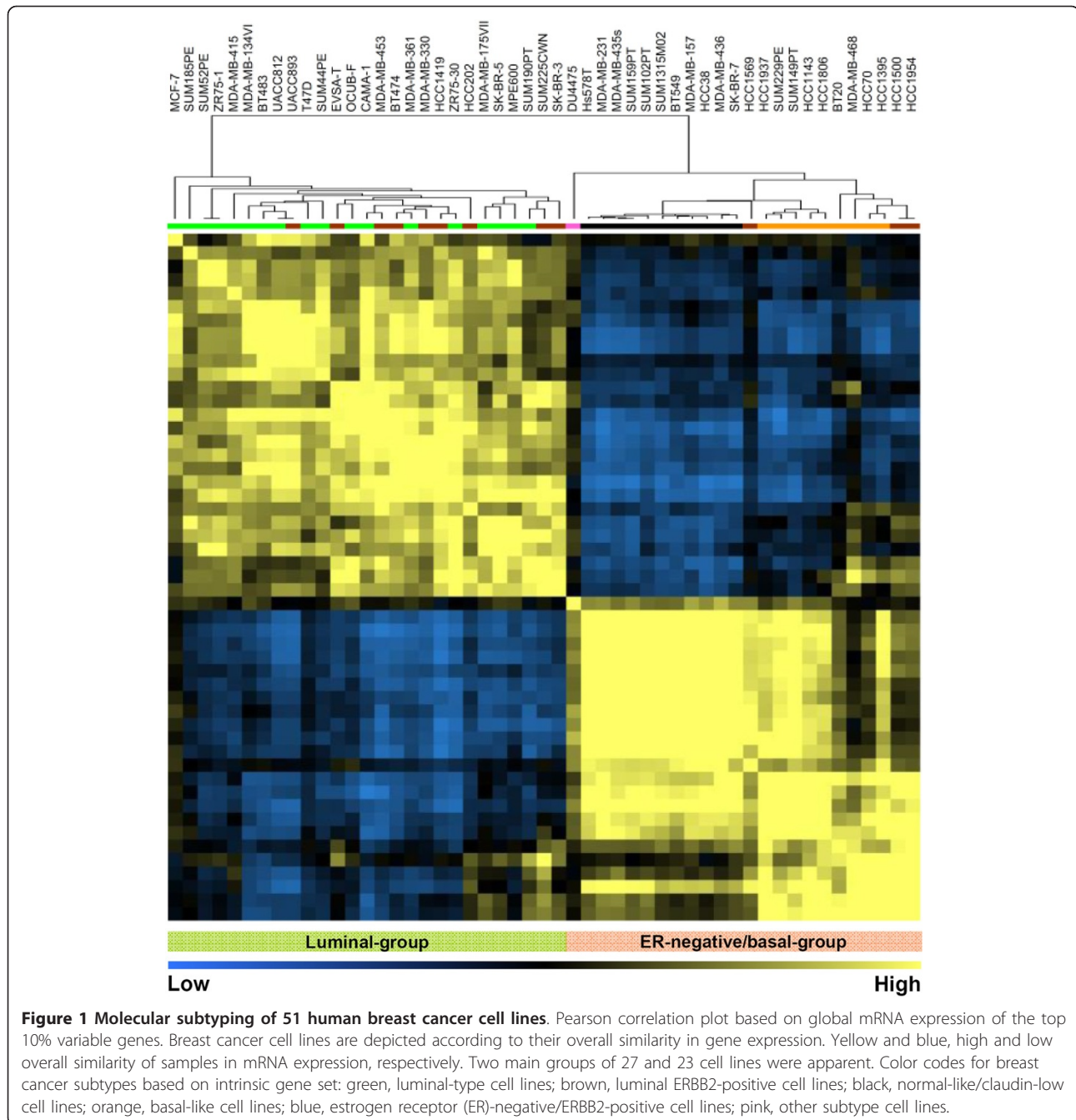
A Pearson correlation based on the top 10% of variably expressed genes classified the cell line cohort into two obvious groups (Figure 1). The first major group included 27 cell lines, of which 17 expressed ER protein. All cell lines in this group showed higher expression of luminal intrinsic genes as defined by Perou and colleagues [23]. We therefore defined this group of breast cancer cell lines as the luminal-group (Figure 1, left block). On the other hand, the second group included 23 breast cancer cell lines all of which were ER-negative and showed a predominant expression of basal intrinsic genes. We defined this group of breast cancer cell lines

as “theER-negative/basal-group (Figure 1, right block). Moreover, according to classification using the intrinsic gene expressions [23], the luminal-group included nine ERBB2-overexpressing breast cancer cell lines, which did not cluster as a separate group. This result is in line with clustering of clinical specimens where the majority of luminal, ERBB2-positive tumors cluster with the luminal B tumors [27]. Within the ER-negative/basal-group, the basal-like and normal-like/claudin-low cell lines clustered distinctively into two subgroups (Figure 1). The DU4475 cell line could not be assigned to any subtype using the intrinsic gene set and was therefore designated as “the other” subtype.

Notably, the entire luminal-group of 27 cell lines expressed luminal cytokeratins (such as CK8/18 and CK19). The basal-like cell lines within the ER-negative/basal-group expressed basal and luminal cytokeratins (CK5, CK19, CK8-18), whereas all normal-like/claudin-low cell lines lacked the expression of these cytokeratins except for SK-BR-7 that exhibited CK8-18 expression (Table 1). Furthermore, except for MDA-MB-435s, the entire ER-negative/basal-group of 23 cell lines also expressed epidermal growth factor receptor protein (Table 1), which is a known marker for basal-type breast cancer.

### miRNA expression profiling revealed two groups of breast cancer cell lines

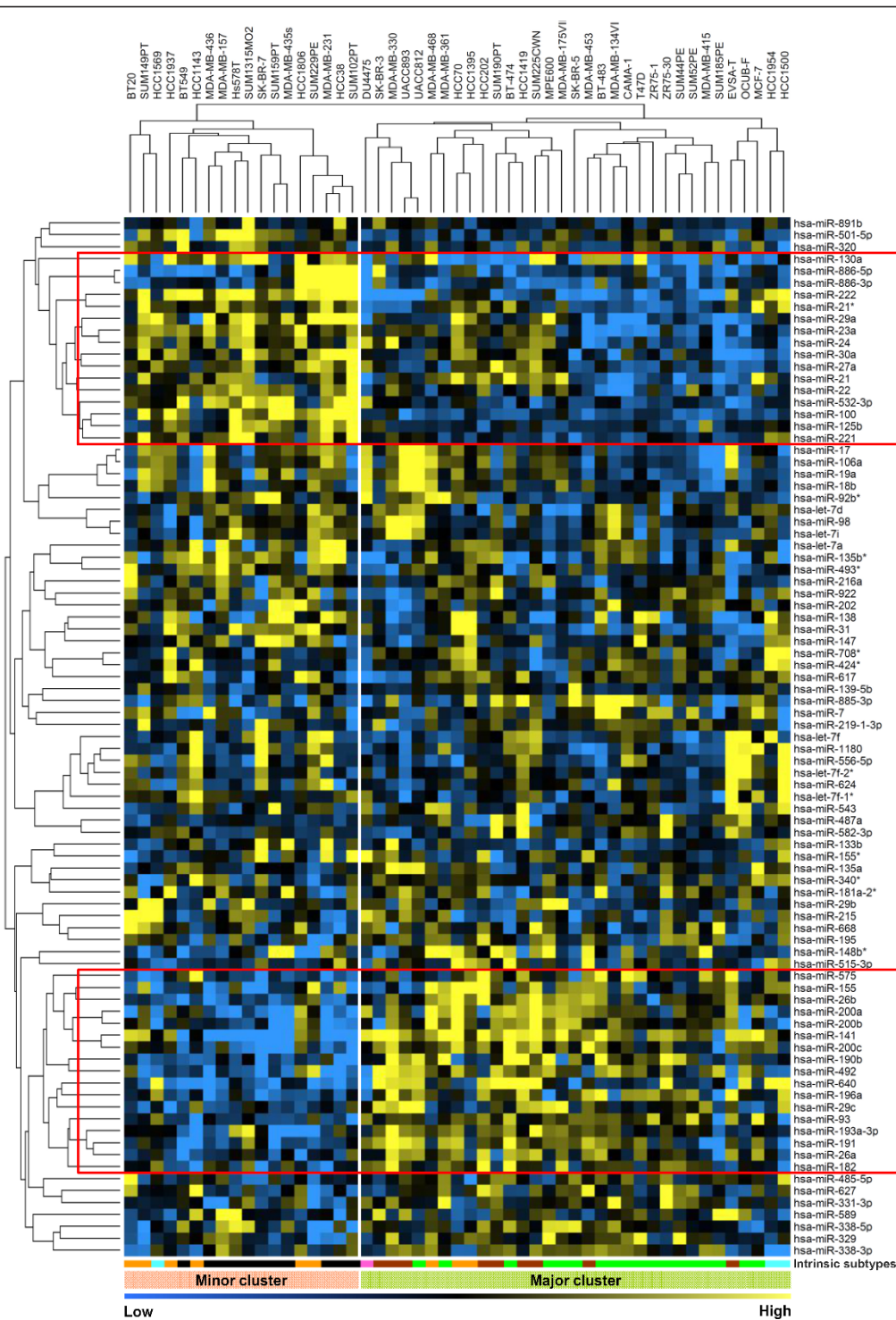
miRNA expression profiling was carried out using microarrays containing LNA™ capture probe sets against 725 human miRNAs. After filtering out miRNAs that were undetectable in most of the cell lines (57%), 413 miRNAs remained and were selected for further analysis. Unsupervised hierarchical clustering based on the most variable 87 miRNAs (with standard deviation  $\geq 0.75$ ) revealed two prominent cell line clusters: a minor cluster that included 18 cell lines and a major cluster that included 33 cell lines (Figure 2). Based on intrinsic subtyping, 17 of 18 cell lines present in the minor cluster belonged to either the basal-like or the normal-like/claudin-low subtypes, and interestingly 16 of them were deficient of ER, progesterone receptor and ERBB2 protein expression (triple-negative). The other two cell lines (HCC38 and SUN102PT) remained undetermined for protein analysis due to technical reasons. On the other hand, 29 of 33 breast cancer cell lines present in the major cluster were of the luminal-type, and all cell lines except SUM185 exhibited either one of the hormone receptor proteins or showed ERBB2 protein expression. The DU4475 cell line, which was previously proposed to have a stem cell-like phenotype due to activation of the Wnt signaling pathway [25,28], clustered at the edge of the major cluster. As a whole, similar to mRNA expression, miRNA expression profiling discriminated cell lines into two major subtypes of breast cancer.



Among the 87 most-variable expressed miRNAs across the entire panel, a group of 15 miRNAs (*hsa-miR-130a*, *hsa-miR-886-5p*, *hsa-miR-886-3p*, *hsa-miR-222*, *hsa-miR-21\**, *hsa-miR-29a*, *hsa-miR-23a*, *hsa-miR-24*, *hsa-miR-30a*, *hsa-miR-27a*, *hsa-miR-22*, *hsa-miR-532-3p*, *hsa-miR-100*, *hsa-miR-125b*, *hsa-miR-221*) was significantly higher expressed in the minor cluster as opposed to other miRNAs (Figure 2, top red box). Another group of 17 miRNAs (*hsa-miR-575*, *hsa-miR-155*, *hsa-miR-26b*, *hsa-miR-200a*, *hsa-miR-200b*, *hsa-miR-141*, *hsa-miR-*

*200c*, *hsa-miR-190b*, *hsa-miR-492*, *hsa-miR-640*, *hsa-miR-196a*, *hsa-miR-29c*, *hsa-miR-93*, *hsa-miR-193a-3p*, *hsa-miR-191*, *hsa-miR-26a*, *hsa-miR-182*) showed significantly higher expression in the major cluster compared with the other miRNAs (fold change  $\geq 1.5$ ) (Figure 2, bottom red box).

To confirm the validity of miRNA array data, we had previously measured miRNA expression levels by real-time RT-PCR in our laboratory available from 20 miRNAs differentially expressed between the major cluster



**Figure 2 Global miRNA expression analysis of 51 human breast cancer cell lines.** Unsupervised clustering across the breast cancer cell lines using the 87 most variably expressed miRNAs. Yellow and blue, high expression and low expression of the particular miRNA in the particular sample, respectively. Cell lines in the dendrogram of the hierarchical clustering based on miRNA expression are coded according to their intrinsic subtypes and depicted at the bottom of the figure (for color codes see Figure 1 caption). Red boxes enlist the groups of miRNAs that show the most robust cluster-specific differential expression in the cell lines. miRNA *hsa-miR-21* in the red box of minor cluster failed to reach our criteria of fold change  $\geq 1.5$ .

and the minor cluster for four ER-positive cell lines and four ER-negative cell lines. On average, the expression levels of these 20 miRNAs in these eight cell lines measured on two different platforms showed a positive correlation ( $R_s = 0.53$  and  $R_s = 0.35$  for miRNAs that were positively and negatively associated with ER status, respectively) (see Table S3 in Additional file 1). These data indicate that the differential expression of these miRNAs measured by the LNA<sup>TM</sup>-based microarray in most of the cases could be validated by an independent technique.

As the miRNA-driven major and minor clusters included ER-positive (18 of 33) and ER-negative (all) cell lines, respectively, we determined whether differential expression of miRNAs between the two clusters is driven by ER status of the cell lines. For this analysis, we grouped cell lines into ER-positive and ER-negative based on mRNA expression of ER and performed supervised cluster analysis on these groups. We identified 79 differentially expressed miRNAs between ER-positive and ER-negative cell lines ( $P < 0.05$ ) (see Figure S1 in Additional file 2 and Table S4A in Additional file 1), of which 54 miRNAs were common to the 113 miRNAs that were differentially expressed between the two clusters identified through unsupervised analysis (see Table S5 in Additional file 1). Interestingly, however, 59 miRNAs were uniquely differentially expressed between the two miRNA-driven clusters whereas 25 miRNAs were uniquely related to ER status of the cell lines (see Table S6 in Additional file 1). To reveal whether ER-related miRNAs in cell lines could be associated with the ER status in clinical tumors, we used a publicly available dataset from Cimino and colleagues comprising 77 primary breast tumors (53 ER-positive tumors and 24 ER-negative tumors) [29] and determined differentially expressed miRNAs between ER-positive and ER-negative tumors. Among 37 miRNAs commonly detectable in our data and in Cimino and colleagues' dataset, the differential expression of six miRNAs showed an overlap between cell lines and primary tumors. Notably, however, five of these six miRNAs (*hsa-miR-141*, *hsa-miRNA-26a*, *hsa-miR-29c*, *hsa-miR-148b*, *hsa-miR-193a-3p*) showed significantly higher expression in both ER-positive cell lines and primary tumors, and one miRNA (*hsa-miR-532-3p*) showed significantly lower expression in both ER-positive cell lines and ER-positive tumors (see Table S4B in Additional file 1).

#### miRNA expression of cell lines associated with molecular intrinsic subtypes of breast cancer

We next aimed to identify miRNAs that are differentially expressed between the intrinsic subtypes of breast cancer. As expected, we observed a substantial overlap of differentially expressed miRNAs between ER-positive

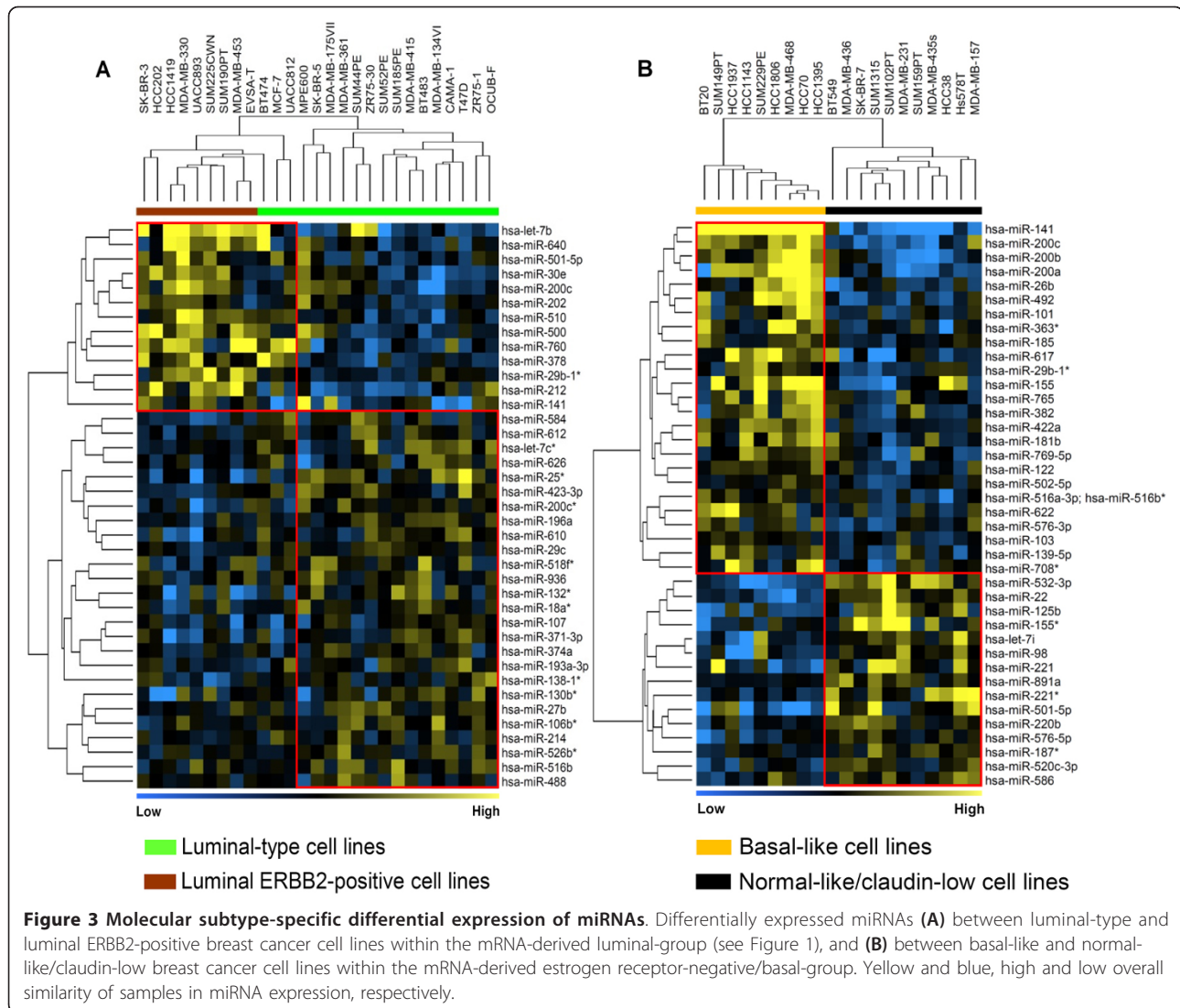
and ER-negative cell lines and the luminal-group and ER-negative/basal-group of cell lines (see Figure S1 in Additional file 2 and Tables S4A and S7 in Additional file 1). To identify additional miRNAs that are related to intrinsic subtypes but to avoid the confounding effect of ER, we compared cell lines with and without ERBB2 overexpression within the luminal-group. Likewise we compared the intrinsic basal-like cell lines with normal-like/claudin-low cell lines within the ER-negative/basal-group. These comparisons revealed 39 differentially expressed miRNAs in luminal cell lines with or without overexpressed ERBB2; 40 miRNAs were found differentially expressed between basal-like and normal-like/claudin-low cell lines ( $P < 0.05$ ) (Figure 3A, B; see Tables S8 and S9 in Additional file 1).

The top most highly expressed miRNAs in ERBB2 overexpressing cell lines included *hsa-let-7b*, *hsa-miR-640*, *hsa-miR-200c*, *hsa-miR-378*, *hsa-miR-141*, *hsa-miR-196a*, *hsa-miR-29c*, and *hsa-miR-18a\**, whereas *hsa-miR-501-5p*, *hsa-miR-202*, *hsa-miR-760*, and *hsa-miR-626* were more highly expressed in luminal cell lines lacking ERBB2 overexpression (fold change  $\geq 1.5$ ) (see Table S8 in Additional file 1). Similarly, the most highly expressed miRNAs in normal-like/claudin-low cell lines were *hsa-miR-22*, *hsa-miR-532-3p*, *hsa-miR-125b*, *hsa-miR-501-5p*, and *hsa-miR-155\**, whereas in basal-like cell lines miRNAs of the miR-200 family (*hsa-miR-492*, *hsa-miR-26b*, *hsa-miR-617*, *hsa-miR-155*) were highly expressed (fold change  $\geq 2$ ) (see Table S9 in Additional file 1). Notably, *hsa-miR-155\** and *hsa-miR-155*, which derived from the same precursor, showed an opposing expression pattern in basal-like compared with normal-like/claudin-low breast cancer cell lines.

#### Differential expression of miRNAs and E-cadherin loss in breast cancer cell lines

E-cadherin elicits a growth suppressive effect in mammary epithelial cells and its expression is frequently lost in human breast tumors [30]. Loss of E-cadherin expression occurs either by gene mutation or by gene promoter hypermethylation. In this cohort of cell lines and in clinical specimens, *E-cadherin* is preferentially mutated in luminal/ER-positive cell lines while it is frequently inactivated by promoter hypermethylation in basal/ER-negative cell lines [21]. To determine whether loss of *E-cadherin* expression in human breast cancer cell lines is associated with differential expression of miRNAs, we compared 10 *E-cadherin* mutant cell lines with 17 wild-type cell lines within the luminal-group of cell lines (Figure 4A) and compared nine *E-cadherin* promoter hypermethylated cell lines with eight nonmethylated cell lines within the ER-negative/basal-group (Figure 4B). Our comparison revealed signatures of 24 and 31 differentially expressed miRNAs associated with *E-cadherin*

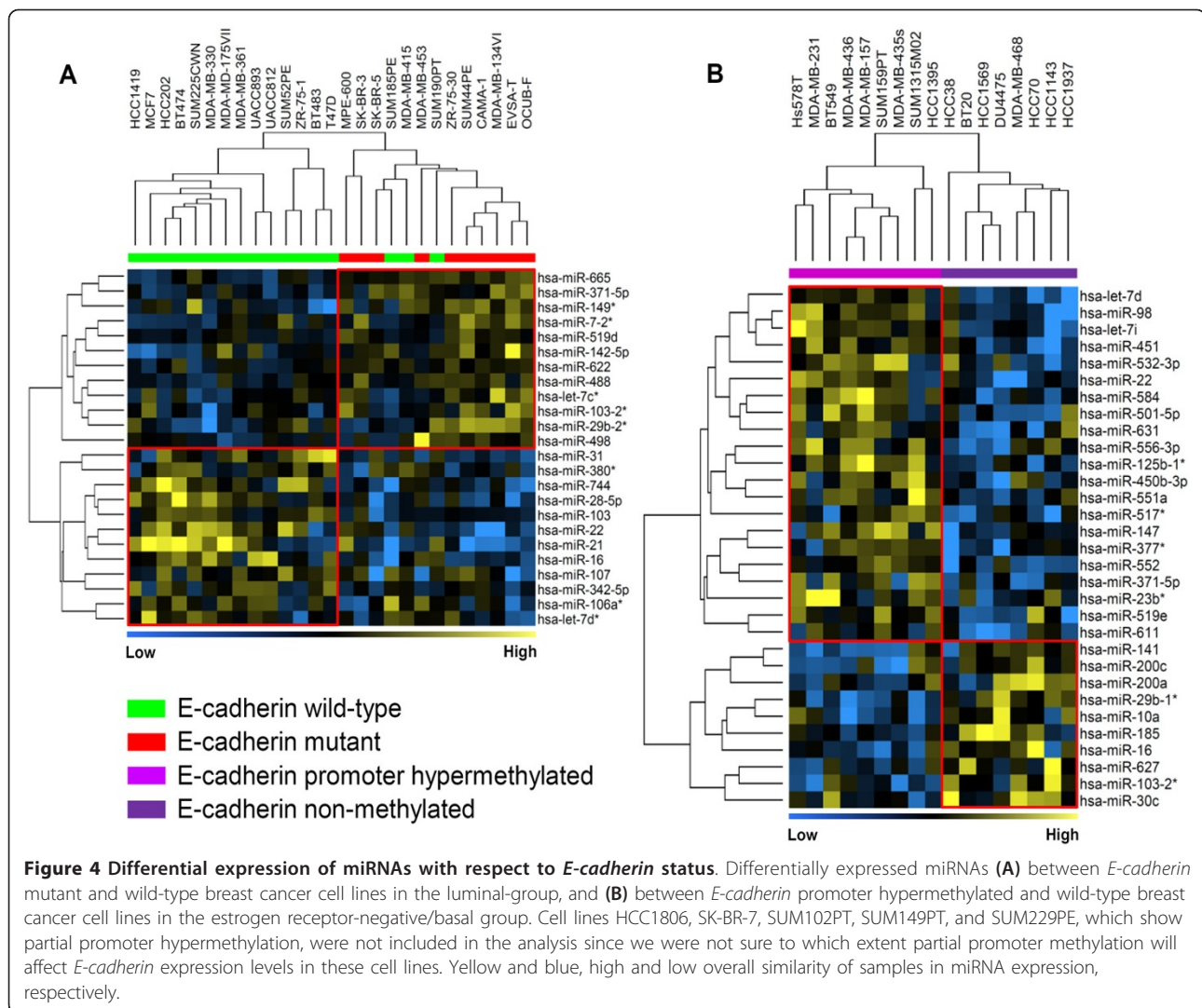




mutant and wild type cell lines and with cell lines having *E-cadherin* promoter hypermethylation or not, respectively ( $P < 0.05$ ) (Figure 4A, B; in see Tables S10 and S11 in Additional file 1). Interestingly, both *E-cadherin* signatures had no miRNAs in common showing similar behavior.

**Association of miRNAs expression with common genetic aberrations (*p16<sup>INK4a</sup>*, *TP53*, *BRCA1*, *PIK3CA*, *PTEN*)**  
 The *BRCA1*, *p16<sup>INK4a</sup>*, *TP53*, and *PTEN* genes play a growth suppressive role in mammary epithelial cells and are frequently inactivated in human breast tumor tissue. *PIK3CA* is also frequently mutated in breast cancer but has an oncogenic role. To reveal differentially expressed miRNAs associated with these mutations, we performed supervised analyses between mutant and wild-type cell lines for these genes. To avoid the confounding effects

of ER, we performed these analyses only in ER-positive or ER-negative cell lines.  
 In our cohort, 13 cell lines had lost the *p16<sup>INK4a</sup>* locus, of which the majority are ER-negative. The comparison revealed that 30 miRNAs were differentially expressed between nine *p16<sup>INK4a</sup>* mutant cell lines and seven wild-type cell lines in the ER-negative/basal-group (Figure 5; see Table S12 in Additional file 1). The most highly expressed miRNAs in *p16<sup>INK4a</sup>* mutant cell lines were *hsa-miR-29a* and *hsa-miR-100* (fold change  $\geq 2$ ). Notably, the *p16<sup>INK4a</sup>* miRNA signature in the ER-negative/basal-group of cell lines appeared to be pronounced, and when used for clustering could almost completely separate the cell lines mutant for the *p16<sup>INK4a</sup>* gene from those that were wild-type. One should, however, emphasize that the tumor suppressor genes *p16<sup>INK4a</sup>* and *p14<sup>ARF</sup>* are co-localized and are often simultaneously

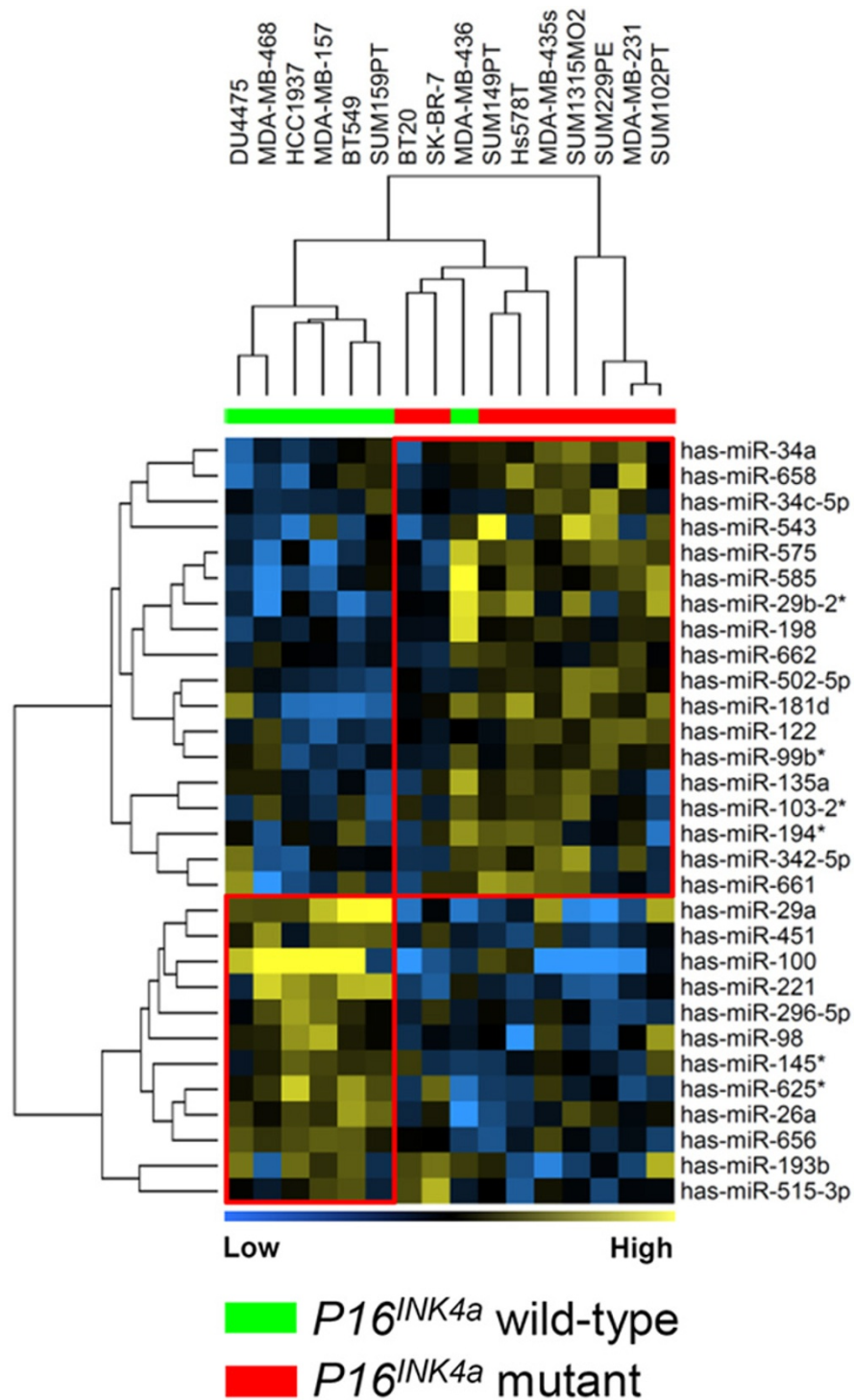


lost amongst cell lines. We therefore cannot determine which of these genes might drive the differential expression of these miRNAs.

Furthermore, the current cohort of breast cancer cell lines contained 15 *PIK3CA* mutant cell lines (four ER-negative, eleven ER-positive) and eight *PTEN* mutant cell lines (four ER-negative, four ER-positive). Both *PIK3CA* and *PTEN* are components of the PI3K signaling pathway and mutated mutually exclusively in clinical specimens [31] and also in this cohort of cell lines [32]. Separate analyses between *PIK3CA* or *PTEN* mutant versus wild-type cell lines for these genes did not yield a considerable number of differentially expressed miRNAs. We therefore combined all cell lines harboring mutations either in *PIK3CA* or *PTEN* and compared them with wild-type cell lines. As *PIK3CA/PTEN* mutations are present in both the luminal-group and the ER-negative/basal-group of cell lines, we therefore performed

multiple analyses. First, we analyzed all cell lines and found 49 miRNAs, which were differentially expressed between mutant and wild-type cell lines (see Figure S2A in Additional file 2 and Table S13 in Additional file 1). With successive analyses, we found eight and 42 differentially expressed miRNAs among the ER-negative/basal-group and the luminal group, respectively (see Figure S2A in Additional file 2 and Tables S14 and S15 in Additional file 1).

With respect to *BRCA1*, which is mutated in four of the ER-negative cell lines, we identified 13 miRNAs being differentially expressed between four mutant cell lines and 12 wild-type breast cancer cell lines (see Figure S2B in Additional file 2 and Table S16 in Additional file 1). The two most highly expressed miRNAs in *BRCA1* mutant cell lines were *hsa-miR-29b* and *hsa-miR-891b* (fold change  $\geq 2$ ). Finally, we analyzed differentially expressed miRNAs between *TP53* mutant cell



**Figure 5** Differentially expressed miRNAs between *p16<sup>INK4a</sup>* mutant and wild-type cell lines in the estrogen receptor-negative/basal-group. The analysis was restricted to only the estrogen receptor (ER)-negative/basal-group of cell lines since the majority of *p16<sup>INK4a</sup>* mutant cell lines were ER-negative. Yellow and blue, high and low overall similarity of samples in miRNA expression, respectively.

lines ( $n = 39$ ) and wild-type cell lines ( $n = 9$ ) and found 18 to be associated with the *TP53* mutation status (see Figure S3 in Additional file 2 and Table S17 in Additional file 1). Important to note, however, is that the differentially expressed miRNA signatures associated with *BRCA1*, *PIK3CA/PTEN*, and *TP53* mutations did not strongly discriminate between mutant and wild-type cell lines, and therefore their significance may be of limited value.

#### miRNA genes show genomic aberrations in breast cancer cell lines

Previous studies show that DNA CNVs in breast tumor tissues can lead to differential expression of genes and miRNAs [33-35]. To investigate this in our cohort of breast cancer cell lines, we first determined the DNA CNVs of the cell lines by performing whole-genome SNP profiling. These CNVs were then correlated with the expression levels of the 87 most variably expressed miRNAs in the cell lines. The correlation revealed 12 miRNAs, which were significantly associated with DNA CNVs (Kruskal-Wallis test,  $P < 0.05$ ) (see Table S18 in Additional file 1). The top four most significantly associated miRNAs - *hsa-miR-130a* (11q12.1), *hsa-miR-22* (17p13.1), *hsa-miR-93* (7q22.1) and *hsa-miR-383* (8p22) - with DNA CNVs in breast cancer cell lines are shown in Figure 6.

#### Discussion

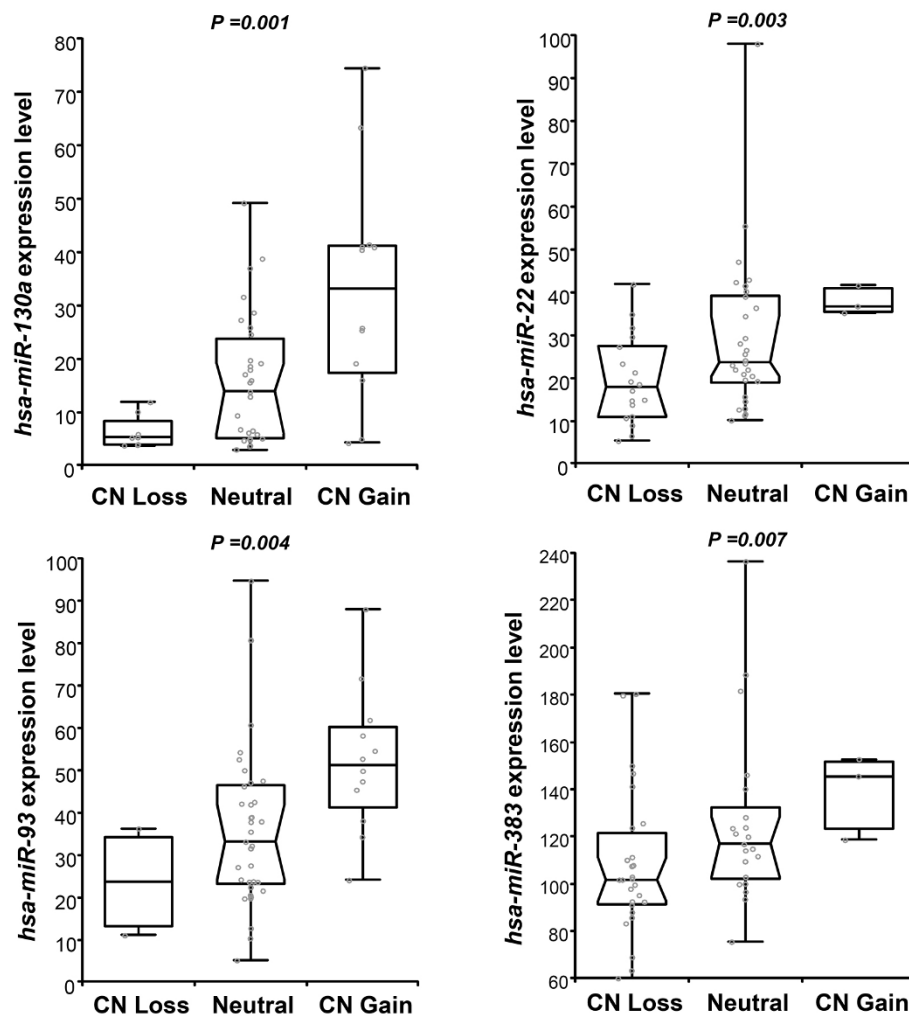
Human breast cancer cell lines are renewable resources that are extensively utilized as reliable workhorses to explore biological functions of clinically relevant molecules in breast cancer. Extensive molecular characterization and gene mutation analysis by us and other researchers have suggested that breast cancer cell lines have retained significant molecular features that are commonly observed in clinical breast tumors [36,37]. This suggestion prompted us to use 51 human breast cancer cell lines as a discovery cohort to identify differentially expressed miRNAs between known intrinsic subtypes of breast cancer as well as those associated with common genetic aberrations present in the cell lines.

In this study, we demonstrated that global miRNA expression profiling can assign our cell line collection into two clusters. These two clusters predominantly mirror the ER-based dichotomy present in human breast cancer cell lines, which may point to the fact that, like mRNA, miRNA expression profiling allows the discrimination between ER-positive and ER-negative cell lines. This suggests that a significant number of miRNAs may be under the control of ER regulation. In line with this, we observed a trend of finding ER binding sites closer to the significantly differentially expressed miRNAs between ER-positive and ER-negative cell lines than nonsignificant miRNAs (see Table S19 in Additional file

1 and Figure S4 in Additional file 2). Interestingly, not all miRNAs discriminating ER-positive and ER-negative cell lines are associated with ER status and therefore the division of the cell lines into two clusters may not be purely dictated by the expression of ER-related miRNAs, but these miRNAs are probably also related to luminal versus basal cellular differentiation.

Our supervised analysis revealed a signature of 79 differentially expressed miRNAs between ER-positive and ER-negative cell lines. Six of these miRNAs were also found similarly differentially expressed between ER-positive and ER-negative primary tumor specimens as they were in cell lines [29] (see Table S4B in Additional file 1). This overlap may seem small but we have observed discrepancies between miRNA profiling platforms (unpublished observations), and our study used LNA<sup>TM</sup> technology-based microarrays to quantify miRNA expression in the cell lines whereas Cimino and colleagues used Agilent microarrays to measure miRNAs in tumors. Also not all miRNAs are present on both platforms, which already explains one-half of the discrepancies. Furthermore, tumor heterogeneity and stromal contribution (such as connective tissues, blood vessels and immune infiltrates) and a likely selection bias in the cell lines - which are more ER-negative, mesenchymal, and *TP53* mutant than primary tumors - could explain the observed discrepancies. A perfect correlation would thus not have been anticipated. Reassuringly, however, we found miRNAs previously reported to be ER-regulated miRNAs, such as a cluster of *hsa-miR-221/222* [12,38]. In line with this, we also found a significant inverse correlation between the expression levels of these ER-regulated miRNAs and mRNA expression levels of their functionally validated target genes - for instance, *hsa-miR-221* and *hsa-miR-222* have been shown to target the *CDKN1B* and *ESR1* genes (see Table S20 in Additional file 1). Such miRNAs have been shown to interact directly with ER and cause a phenotypic shift from ER-positive to ER-negative tumor cells [38]. Our miRNA analyses in this cohort of cell lines therefore confirms previous findings but also revealed new miRNAs (discussed below), which may have potentially interesting biological roles in ER-driven cancer.

Among these potentially novel ER-related miRNAs, four miRNAs (*hsa-miR-26a*, *hsa-miR-92b*, *hsa-miR-191*, *hsa-miR-492*) appear to show consistently higher expression across all the ER-positive cell lines (fold change  $\geq 1.5$ ), and as yet only *hsa-miR-26a* has been implicated in breast carcinogenesis whereas *hsa-miR-26a* and *hsa-miR-92b* have also been implicated in brain tumors [39,40]. In breast cancer, high expression of the *hsa-miR-26a* miRNA downregulates *EZH2* and is therefore related to a favorable outcome on tamoxifen in metastatic breast cancer [41], and also interacts with *CDK4*



**Figure 6** Association of miRNA expression with genomic copy number variation in breast cancer cell lines. The top four most significant miRNAs are represented (see Table S16 in Additional file 1 for a complete list). The Kruskal-Wallis test was used to reveal significant associations of miRNAs with genome copy number (CN) variation. y axis, expression levels of miRNA; x axis, number of samples with CN loss/gain or neutral.

and *CENTG1* oncogenes and forms an integrated onco-mir/oncogene DNA cluster, which promotes glioblastoma tumor growth via RB1, PI3K/AKT, and JNK pathways [39]. On the other hand, *hsa-miR-92b* has been found to be exclusively overexpressed in primary brain tumors but serves as a biomarker to discriminate brain primary cancer from metastasis [40]. The other two miRNAs (*hsa-miR-191*, *hsa-miR-492*) have been linked to hepatic cancer. The increased expression of *hsa-miR-191* stimulates proliferation in hepatocellular carcinoma cell lines and its therapeutic targeting suppresses tumor masses *in vivo* [42]. The *hsa-miR-492* miRNA has been shown to be processed from the *keratin 19* gene and upregulated in metastatic hepatoblastoma [43]. How these miRNAs play their biological roles in the context of breast cancer remains unknown and demands clinical and functional validation studies.

Another major contribution to the overall miRNA profiling of this cell line cohort is the identification of a signature of 42 differentially expressed miRNAs, which discriminates between basal-like and normal-like/claudin-low breast cancer cell lines. These miRNAs together with the ER-associated miRNAs are the major determinants of the overall clustering of the cell lines. Importantly, this signature includes all four members of the *hsa-miR-200* family (*hsa-miR-200a*, *hsa-miR-200b*, *hsa-miR-200c*, *hsa-miR-141*), *hsa-miR-155*, and *hsa-miR-622* miRNAs. Several studies have implicated these miRNAs to be involved in epithelial-mesenchymal transition (EMT), to be related to the stem-cell-like phenotype, and to be associated with a switch in paclitaxel responsiveness [44-47]. In breast cancer, these miRNAs are known to regulate the EMT process by targeting the ZEB family of the transcription factors through an active

negative feedback loop [48,49]. Importantly, the ZEB family of transcription factors was reported to be a repressor of E-cadherin expression in several epithelial carcinomas including breast carcinoma.

Interestingly, we observed a significant inverse correlation between the expression levels of all miRNAs of the *hsa-miR-200* family and the mRNA expression level of the ZEB1 transcription factor, which is a well-known target of this miRNA family (see Table S20 in Additional file 1). Additionally, our finding that the *hsa-miR-200* cluster showed lower expression in normal-like/claudin-low breast cancer cell lines fits with the fundamental literature that these miRNAs are indeed involved in EMT, because the majority of the normal-like/claudin-low cell lines lack E-cadherin protein expression, exhibit low claudin expression, and generally display EMT-like features as well as the breast cancer stem cell phenotype (CD44<sup>high</sup>/CD24<sup>low</sup>) [50-53]. Other known significant miRNAs of this signature are *hsa-miR-155* and *hsa-miR-622*, which were also linked to enhanced tumorigenesis in various cancer types besides breast cancer [54,55]. In concordance with this, we recently found that the normal-like/claudin-low cell lines with low expression of these EMT-related miRNAs show highly aggressive growth characteristics *in vivo* when raised as xenografts in nude mice (M Riaz and colleagues, unpublished data). Importantly, this signature includes three previously unknown miRNAs (*hsa-miR-492*, *hsa-miR-26b*, *hsa-miR-617*; fold change  $\geq 1.5$ ) found to be associated with cell lines frequently showing EMT-like characteristics.

We observed that the miRNA signatures associated with *E-cadherin* mutation and promoter hypermethylation include distinct miRNAs. This may point to an existence of unique biology at the miRNA level in tumors that show *E-cadherin* inactivation due to gene mutation rather than those that lose *E-cadherin* expression due to promoter hypermethylation. Important to note is that *E-cadherin* promoter hypermethylated cell lines include all normal-like/claudin-low cell lines, and most EMT-related miRNAs were also associated with *E-cadherin* promoter hypermethylation.

With respect to common genetic aberrations present in breast cancer, our study reveals differentially expressed miRNA signatures associated with commonly mutated tumor suppressors and oncogenes. Most significantly, a signature of 30 miRNAs associated with *p16<sup>INK4a</sup>* mutant cell lines can strongly discriminate between mutant and wild-type ER-negative cell lines. *P16<sup>INK4a</sup>* is a tumor suppressor gene located on chromosomal band 9p21 that has been frequently altered in many human cancers [56]. *P16<sup>INK4a</sup>* regulates cell cycle progression by targeting CDK4/6 through the pRb signaling pathway. Interestingly, this signature includes some miRNAs that are involved in

cell cycle regulation. For instance, *hsa-miR-100*, which is highly expressed in the *p16<sup>INK4a</sup>* wild-type cell lines, targets the *RBSP3* gene that in acute myeloid leukemia regulates the cell cycle through partial modulation of pRB/E2F1 [57]. *hsa-miR-34a* has been reported as a suppressor of cell proliferation and migration in colon cancer [58] and its mechanism of growth inhibition also involves cell cycle arrest followed by apoptosis [59]. Besides this, we found a few miRNAs to be only associated with mutation in *BRCA1*, *TP53* and *PIK3CA/PTEN*; these observed associations require further independent validation in breast cancer specimens.

Our finding that the differential expression of miRNAs is associated with ERBB2-overexpressing luminal-type cell lines is also intriguing since the miRNAs showing elevated expression in ERBB2-positive cell lines (fold change  $\geq 1.5$ ) are not located on the ERBB2 amplicon (17q12) and thus may be regulated indirectly by genes in the ERBB2 amplicon. It is also important to mention that a majority of these miRNAs have already been implicated in various types of epithelial carcinoma including breast cancer [60-64]. One should, however, note that our study is associative and does not necessarily reveal causality. We therefore propose functional studies on these miRNAs to reveal more biological insights into their role regarding ERBB2 overexpression in breast cancer.

Finally, we also identified 12 miRNAs to be associated with CNVs in breast cancer cell lines (see Table S17 in Additional file 1). The majority of these miRNAs (*hsa-miR-130a*, *hsa-miR-93*, *hsa-miR-383*, *hsa-miR29c*, *hsa-miR-382*, *hsa-miR-31*) were already found to be located in regions that exhibited DNA copy number abnormalities in breast cancer tumors [65]. Importantly, this repertoire of miRNAs includes *hsa-miR-22* previously shown to be regulated by ER [66] and we provide evidence that it can also be regulated by the loss of the locus containing this miRNA. Six miRNAs of this repertoire are located in the genomic regions containing known protein coding genes. All miRNAs of this repertoire have also been implicated in various cancers but a thorough validation of these miRNAs with respect to DNA copy number changes in clinical specimen is imperative.

## Conclusion

In summary, our analyses show that the dichotomy in breast cancer cell lines is, in line with the general consensus in the field, related to ER status. However, part of this dichotomy may be related to ER-independent luminal versus basal differentiation status of cell lines. Secondly, another prominent dichotomy observed among breast cancer cell lines discriminating basal-like cell lines from normal-like/claudin-low cell lines

involves miRNAs related to EMT and/or stemness. Furthermore, we reveal sets of miRNAs associated with genes frequently amplified (ERBB2) or mutated (*p16<sup>INK4a</sup>*, *PIK3CA* and/or *PTEN*, *E-cadherin*, *BRCA1*) in breast cancer cell lines and thus these miRNAs may contribute to the function of these oncogenes/tumor suppressors. Finally, certain miRNAs themselves are located in genomic regions, which frequently show genetic aberrations, implying that these miRNAs themselves may have a potential driver role or contribute to known drivers in these genomic regions. Our current findings call for further validation of these signatures in clinical specimens. Importantly, our study provides a unique molecular miRNA portrait of human breast cancer cell lines, which can be exploited for functional studies of clinically important miRNAs.

## Additional material

**Additional file 1: an Excel file of Tables S1 to S20 containing lists of differentially expressed miRNAs associated with breast cancer molecular subtypes and recurrent genetic aberrations.**

**Additional file 2: Figure S1 showing differential expression of miRNAs between ER-positive and ER-negative human breast cancer cell lines.** Cell lines have been grouped into ER-positive and ER-negative groups based on their mRNA expression levels measured by microarray. Yellow and blue, high and low overall similarity of samples in miRNA expression, respectively. Figure S2 showing differential expression of miRNAs between cell lines mutant and wild-type for *BRCA1* and *PIK3CA/PTEN* genes. (A) miRNA differential expression associated with *PIK3CA/PTEN* mutations in all, the luminal-group, and the ER-negative/basal-group of breast cancer cell lines. (B) miRNA differential expression associated with *BRCA1* mutation. Yellow and blue, high and low overall similarity of samples in miRNA expression, respectively. Figure S3 showing differential expression miRNAs between cell lines mutant and wild-type for the *TP53* gene. Cell lines with *TP53* heterozygous mutation (OCUB-F) and with undetermined *TP53* mutation status (HCC202, HCC1395, HCC1500) were excluded from the analysis. Yellow and blue, high and low overall similarity of samples in miRNA expression, respectively. Figure S4 showing comparison of median genomic distances of ER binding sites and significantly and nonsignificantly differentially associated miRNAs with ER status of the cell lines. The Mann-Whitney test was used to compare the median genomic distances of ER binding sites and miRNAs.

## Abbreviations

BRCA1: breast cancer 1; BST2: bone marrow stromal cell antigen 2; CDK: cyclin-dependent kinase; CK: cytokeratin; CNV: copy number variation; EMT: epithelial-mesenchymal transition; ER: estrogen receptor; ERBB2: erythroblastic leukemia viral oncogene homolog 2; ESRT: estrogen receptor; FABP7: fatty acid binding protein 7; FOXA1: forkhead box protein A1; GRB7: growth factor receptor-bound protein 7; JNK: c-Jun N-terminal kinase; KRT: keratin; miRNA: microRNA; PI3K: phosphoinositide 3-kinase; PCR: polymerase chain reaction; PTEN: phosphatase and tensin homolog; RB: retinoblastoma protein; SNP: single nucleotide polymorphism; STR: short tandem repeat; TFF3: trefoil factor 3; TP53: tumor protein 53; UTR: untranslated region.

## Authors' contributions

JAF, JWMM, EAW, and MR designed the study. MR, JAF and JWMM wrote the manuscript. MTMvJ and AWMB performed the miRNA expression analysis. JL and AH performed gene mutation analysis. JH performed STR analysis. BO and WJCP-vdS carried out RNA isolation and quantification.

AAJH, MS, and MR performed the statistical data analyses. All authors approved the final version of the manuscript.

## Competing interests

The authors declare that they have no competing interests.

## Acknowledgements

Janssen Pharmaceutical, a Johnson and Johnson Company, Beerse, Belgium, sponsored a part of the molecular profiling performed within this study. This study was supported financially by the Netherlands Genomics Initiative/ Netherlands Organization for Scientific Research (NWO) and by the Dutch Cancer Society (EMCR 2007-3794).

Received: 23 November 2012 Revised: 21 February 2013

Accepted: 14 March 2013 Published: 19 April 2013

## References

1. Pasquinelli AE: **MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship.** *Nat Rev Genet* 2012, **13**:271-282.
2. Bartel DP: **MicroRNAs: genomics, biogenesis, mechanism, and function.** *Cell* 2004, **116**:281-297.
3. Alvarez-Garcia I, Miska EA: **MicroRNA functions in animal development and human disease.** *Development* 2005, **132**:4653-4662.
4. Zhao Y, Srivastava D: **A developmental view of microRNA function.** *Trends Biochem Sci* 2007, **32**:189-197.
5. Enerly E, Steinfeld I, Kleivi K, Leivonen SK, Aure MR, Russnes HG, Rønneberg JA, Johnsen H, Navon R, Rødland E, Mäkelä R, Naume B, Perälä M, Kallioniemi O, Kristensen VN, Yakhini Z, Børresen-Dale AL: **miRNA-mRNA integrated analysis reveals roles for miRNAs in primary breast tumors.** *PLoS One* 2011, **6**:e16915.
6. Hwang HW, Mendell JT: **MicroRNAs in cell proliferation, cell death, and tumorigenesis.** *Br J Cancer* 2006, **94**:776-780.
7. Leung AK, Sharp PA: **MicroRNA functions in stress responses.** *Mol Cell* 2010, **40**:205-215.
8. Ventura A, Jacks T: **MicroRNAs and cancer: short RNAs go a long way.** *Cell* 2009, **136**:586-591.
9. Blenkiron C, Goldstein LD, Thorne NP, Spiteri I, Chin SF, Dunning MJ, Barbosa-Morais NL, Teschendorff AE, Green AR, Ellis IO, Tavaré S, Caldas C, Miska EA: **MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype.** *Genome Biol* 2007, **8**:R214.
10. Iorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M, Ménard S, Palazzo JP, Rosenberg A, Musiani P, Volinia S, Nenci I, Calin GA, Querzoli P, Negrini M, Croce CM: **MicroRNA gene expression deregulation in human breast cancer.** *Cancer Res* 2005, **65**:7065-7070.
11. Van der Auwera I, Limame R, van Dam P, Vermeulen PB, Dirix LY, Van Laere SJ: **Integrated miRNA and mRNA expression profiling of the inflammatory breast cancer subtype.** *Br J Cancer* 2010, **103**:532-541.
12. Foekens JA, Sieuwerts AM, Smid M, Look MP, de Weerd V, Boersma AW, Klijn JG, Wiemer EA, Martens JW: **Four miRNAs associated with aggressiveness of lymph node-negative, estrogen receptor-positive human breast cancer.** *Proc Natl Acad Sci USA* 2008, **105**:13021-13026.
13. Camps C, Buffa FM, Colella S, Moore J, Sotiropoulos C, Sheldon H, Harris AL, Gleadow JM, Ragoussis J: **hsa-miR-210 is induced by hypoxia and is an independent prognostic factor in breast cancer.** *Clin Cancer Res* 2008, **14**:1340-1348.
14. Tavazoie SF, Alarcon C, Oskarsson T, Padua D, Wang Q, Bos PD, Gerald WL, Massague J: **Endogenous human microRNAs that suppress breast cancer metastasis.** *Nature* 2008, **451**:147-152.
15. Rodriguez-Gonzalez FG, Sieuwerts AM, Smid M, Look MP, Meijer-van Gelder ME, de Weerd V, Sleijfer S, Martens JW, Foekens JA: **MicroRNA-30c expression level is an independent predictor of clinical benefit of endocrine therapy in advanced estrogen receptor positive breast cancer.** *Breast Cancer Res Treat* 2011, **127**:43-51.
16. Jansen MP, Reijm EA, Sieuwerts AM, Ruigrok-Ritstier K, Look MP, Rodriguez-Gonzalez FG, Heine AA, Martens JW, Sleijfer S, Foekens JA, Berns EM: **High miR-26a and low CDC2 levels associate with decreased EZH2 expression and with favorable outcome on tamoxifen in metastatic breast cancer.** *Breast Cancer Res Treat* 2012, **133**:937-947.
17. Ryu S, Joshi N, McDonnell K, Woo J, Choi H, Gao D, McCombie WR, Mittal V: **Discovery of novel human breast cancer microRNAs from deep**

- sequencing data by analysis of pri-microRNA secondary structures. *PLoS One* 2011, **6**:e16403.
18. Farazi TA, Spitzer JI, Morozov P, Tuschl T: **miRNAs in human cancer.** *J Pathol* 2011, **223**:102-115.
  19. Miller TE, Ghoshal K, Ramaswamy B, Roy S, Datta J, Shapiro CL, Jacob S, Majumder S: **MicroRNA-221/222 confers tamoxifen resistance in breast cancer by targeting p27Kip1.** *J Biol Chem* 2008, **283**:29897-29903.
  20. Voorhoeve PM, le Sage C, Schrier M, Gillis AJ, Stoop H, Nagel R, Liu YP, van Duijse J, Drost J, Griekspoor A, Zlotorynski E, Yabuta N, De Vita G, Nojima H, Looijenga LH, Agami R: **A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors.** *Cell* 2006, **124**:1169-1181.
  21. Hollestelle A, Elstrodt F, Timmermans M, Sieuwerts AM, Klijn JG, Foekens JA, den Bakker MA, Schutte M: **Four human breast cancer cell lines with biallelic inactivating alpha-catenin gene mutations.** *Breast Cancer Res Treat* 2010, **122**:125-133.
  22. Sieuwerts AM, Meijer-van Gelder ME, Timmermans M, Trapman AM, Garcia RR, Arnold M, Goedheer AJ, Portengen H, Klijn JG, Foekens JA: **How ADAM-9 and ADAM-11 differentially from estrogen receptor predict response to tamoxifen treatment in patients with recurrent breast cancer: a retrospective study.** *Clin Cancer Res* 2005, **11**:7311-7321.
  23. Perou CM, Sørlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lønning PE, Børresen-Dale AL, Brown PO, Botstein D: **Molecular portraits of human breast tumors.** *Nature* 2000, **406**:747-752.
  24. Pothof J, Verkaik NS, van IW, Wiemer EA, Ta VT, van der Horst GT, Jaspers NG, van Gent DC, Hoeijmakers JH, Persengiev SP: **MicroRNA-mediated gene silencing modulates the UV-induced DNA-damage response.** *EMBO J* 2009, **28**:2090-2099.
  25. Hollestelle A, Nagel JH, Smid M, Lam S, Elstrodt F, Wasielewski M, Ng SS, French PJ, Peeters JK, Rozendaal MJ, Riaz M, Koopman DG, Ten Hagen TL, de Leeuw BH, Zwarthoff EC, Teunisse A, van der Spek PJ, Klijn JG, Dinjens WN, Ethier SP, Clevers H, Jochemsen AG, den Bakker MA, Foekens JA, Martens JW, Schutte M: **Distinct gene mutation profiles among luminal-type and basal-type breast cancer cell lines.** *Breast Cancer Res Treat* 2010, **121**:53-64.
  26. de Hoon MJ, Imoto S, Nolan J, Miyano S: **Open source clustering software.** *Bioinformatics* 2004, **20**:1453-1454.
  27. Sørlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lønning PE, Børresen-Dale AL: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci USA* 2001, **98**:10869-10874.
  28. Zeng YA, Nusse R: **Wnt proteins are self-renewal factors for mammary stem cells and promote their long-term expansion in culture.** *Cell Stem Cell* 2010, **6**:568-577.
  29. Cimino D, De Pittà C, Orso F, Zampini M, Casara S, Penna E, Quagliano E, Forni M, Damasco C, Pinatel E, Ponzoni R, Romualdi C, Brisken C, De Bortoli M, Biglia N, Provero P, Lanfranchi G, Taverna D: **miR148b is a major coordinator of breast cancer progression in a relapse-associated microRNA signature by targeting ITGA5, ROCK1, PIK3CA, NRAS, and CSF1.** *FASEB J* 2012, **27**:1223-1235.
  30. Vos CB, Cleton-Jansen AM, Bex G, de Leeuw WJ, ter Haar NT, van Roy F, Cornelisse CJ, Peterse JL, van de Vijver MJ: **E-cadherin inactivation in lobular carcinoma in situ of the breast: an early event in tumorigenesis.** *Br J Cancer* 1997, **76**:1131-1133.
  31. Saal LH, Holm K, Maurer M, Memeo L, Su T, Wang X, Yu JS, Malmström PO, Mansukhani M, Enoksson J, Hibshoosh H, Borg A, Parsons R: **PIK3CA mutations correlate with hormone receptors, node metastasis, and ERBB2, and are mutually exclusive with PTEN loss in human breast carcinoma.** *Cancer Res* 2005, **65**:2554-2559.
  32. Hollestelle A, Elstrodt F, Nagel JH, Kallemeijn WW, Schutte M: **Phosphatidylinositol-3-OH kinase or RAS pathway mutations in human breast cancer cell lines.** *Mol Cancer Res* 2007, **5**:195-201.
  33. Pollack JR, Sørlie T, Perou CM, Rees CA, Jeffrey SS, Lønning PE, Tibshirani R, Botstein D, Børresen-Dale AL, Brown PO: **Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors.** *Proc Natl Acad Sci USA* 2002, **99**:12963-12968.
  34. Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousses S, Rozenblum E, Ringnér M, Sauter G, Monni O, Elkahlon A, Kallioniemi OP, Kallioniemi A: **Impact of DNA amplification on gene expression patterns in breast cancer.** *Cancer Res* 2002, **62**:6240-6245.
  35. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM: **Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers.** *Proc Natl Acad Sci USA* 2004, **101**:2999-3004.
  36. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T, Clark L, Bayani N, Coppe JP, Tong F, Speed T, Spellman PT, DeVries S, Lapuk A, Wang NJ, Kuo WL, Stilwell JL, Pinkel D, Albertson DG, Waldman FM, McCormick F, Dickson RB, Johnson MD, Lippman M, Ethier S, Gazdar A, Gray JW: **A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes.** *Cancer Cell* 2006, **10**:515-527.
  37. Kao J, Salari K, Bocanegra M, Choi YL, Girard L, Gandhi J, Kwei KA, Hernandez-Boussard T, Wang P, Gazdar AF, Minna JD, Pollack JR: **Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery.** *PLoS One* 2009, **4**:e6146.
  38. Di Leva G, Gasparini P, Piovon C, Ngankeu A, Garofalo M, Taccioli C, Iorio MV, Li M, Volinia S, Alder H, Nakamura T, Nuovo G, Liu Y, Nephew KP, Croce CM: **MicroRNA cluster 221-222 and estrogen receptor alpha interactions in breast cancer.** *J Natl Cancer Inst* 2010, **102**:706-721.
  39. Kim H, Huang W, Jiang X, Pennicooke B, Park PJ, Johnson MD: **Integrative genome analysis reveals an oncomir/oncogene cluster regulating glioblastoma survivorship.** *Proc Natl Acad Sci USA* 2010, **107**:2183-2188.
  40. Nass D, Rosenwald S, Meiri E, Gilad S, Tabibian-Keissar H, Schlosberg A, Kuker H, Sion-Vardy N, Tobar A, Kharenko O, Sitbon E, Lithwick Yanai G, Elyakim E, Cholak H, Gibori H, Spector Y, Bentwich Z, Barshack I, Rosenfeld N: **MiR-92b and miR-9/9\* are specifically expressed in brain primary tumors and can be used to differentiate primary from metastatic brain tumors.** *Brain Pathol* 2009, **19**:375-383.
  41. Jansen MP, Reijm EA, Sieuwerts AM, Ruigrok-Ritstier K, Look MP, Rodríguez-González FG, Heine AA, Martens JW, Sleijfer S, Foekens JA, Berns EM: **High miR-26a and low CDC2 levels associate with decreased EZH2 expression and with favorable outcome on tamoxifen in metastatic breast cancer.** *Breast Cancer Res Treat* 2012, **133**:937-947.
  42. Elyakim E, Sitbon E, Faerman A, Tabak S, Montia E, Belanis L, Dov A, Marcusson EG, Bennett CF, Chajut A, Cohen D, Yerushalmi N: **hsa-miR-191 is a candidate oncogene target for hepatocellular carcinoma therapy.** *Cancer Res* 2010, **70**:8077-8087.
  43. von Frowein J, Pagel P, Kappler R, von Schweinitz D, Roscher A, Schmid I: **MicroRNA-492 is processed from the keratin 19 gene and up-regulated in metastatic hepatoblastoma.** *Hepatology* 2011, **53**:833-842.
  44. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, Vadas MA, Khew-Goodall Y, Goodall GJ: **The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1.** *Nat Cell Biol* 2008, **10**:593-601.
  45. Korpala M, Lee ES, Hu G, Kang Y: **The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2.** *J Biol Chem* 2008, **283**:14910-14914.
  46. Adam L, Zhong M, Choi W, Qi W, Nicoloso M, Arora A, Calin G, Wang H, Siefker-Radtke A, McConkey D, Bar-Eli M, Dinney C: **miR-200 expression regulates epithelial-to-mesenchymal transition in bladder cancer cells and reverses resistance to epidermal growth factor receptor therapy.** *Clin Cancer Res* 2009, **15**:5060-5072.
  47. Gupta PB, Onder TT, Jiang G, Tao K, Kuperwasser C, Weinberg RA, Lander ES: **Identification of selective inhibitors of cancer stem cells by high-throughput screening.** *Cell* 2009, **138**:645-659.
  48. Hurteau GJ, Carlson JA, Spivack SD, Brock GJ: **Overexpression of the microRNA hsa-miR-200c leads to reduced expression of transcription factor 8 and increased expression of E-cadherin.** *Cancer Res* 2007, **67**:7972-7976.
  49. Christoffersen NR, Silaharoglu A, Orom UA, Kauppinen S, Lund AH: **miR-200b mediates post-transcriptional repression of ZFH1B.** *RNA* 2007, **13**:1172-1178.
  50. Park SM, Gaur AB, Lengyel E, Peter ME: **The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2.** *Genes Dev* 2008, **22**:894-907.
  51. Prat A, Parker JS, Karginova O, Fan C, Livasy C, Herschkowitz JI, He X, Perou CM: **Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.** *Breast Cancer Res* 2010, **12**:R68.



52. Sarrjo D, Franklin CK, Mackay A, Reis-Filho JS, Isacke CM: **Epithelial and mesenchymal subpopulations within normal basal breast cell lines exhibit distinct stem cell/progenitor properties.** *Stem Cells* 2012, **30**:292-303.
53. Hollestelle A, Peeters JK, Smid M, Timmermans M, Verhoog LC, Westenend PJ, Heine AA, Chan A, Sieuwerts AM, Wiemer EA, Klijn JG, van der Spek PJ, Foekens JA, Schutte M, den Bakker MA, Martens JW: **Loss of E-cadherin is not a necessity for epithelial to mesenchymal transition in human breast cancer.** *Breast Cancer Res Treat* 2013, **138**:47-57.
54. Guo XB, Jing CQ, Li LP, Zhang L, Shi YL, Wang JS, Liu JL, Li CS: **Down-regulation of miR-622 in gastric cancer promotes cellular invasion and tumor metastasis by targeting ING1 gene.** *World J Gastroenterol* 2011, **17**:1895-1902.
55. Donnem T, Eklo K, Berg T, Sorbye SW, Lonvik K, Al-Saad S, Al-Shibli K, Andersen S, Stenvold H, Bremnes RM, Busund LT: **Prognostic impact of MiR-155 in non-small cell lung cancer evaluated by in situ hybridization.** *J Transl Med* 2011, **9**:6.
56. Sherr CJ: **Cancer cell cycles.** *Science* 1996, **274**:1672-1677.
57. Zheng YS, Zhang H, Zhang XJ, Feng DD, Luo XQ, Zeng CW, Lin KY, Zhou H, Qu LH, Zhang P, Chen YQ: **MiR-100 regulates cell differentiation and survival by targeting RBSP3, a phosphatase-like tumor suppressor in acute myeloid leukemia.** *Oncogene* 2012, **31**:80-92.
58. Tazawa H, Tsuchiya N, Izumiya M, Nakagama H: **Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells.** *Proc Natl Acad Sci USA* 2007, **104**:15472-15477.
59. Cole KA, Attiyeh EF, Mosse YP, Laquaglia MJ, Diskin SJ, Brodeur GM, Maris JM: **A functional screen identifies miR-34a as a candidate neuroblastoma tumor suppressor gene.** *Mol Cancer Res* 2008, **6**:735-742.
60. Redova M, Poprach A, Nekvindova J, Iliev R, Radova L, Lakomy R, Svoboda M, Vyzula R, Slaby O: **Circulating miR-378 and miR-451 in serum are potential biomarkers for renal cell carcinoma.** *J Transl Med* 2012, **10**:55.
61. Zhao JJ, Lin J, Lwin T, Yang H, Guo J, Kong W, Dessureault S, Moscinski LC, Reznia D, Dalton WS, Sotomayor E, Tao J, Cheng JQ: **microRNA expression profile and identification of miR-29 as a prognostic marker and pathogenetic factor by targeting CDK6 in mantle cell lymphoma.** *Blood* 2010, **115**:2630-2639.
62. Kriegel AJ, Liu Y, Fang Y, Ding X, Liang M: **The miR-29 family: genomics, cell biology, and relevance to renal and cardiovascular injury.** *Physiol Genomics* 2012, **44**:237-244.
63. Jedlinski DJ, Gabrovská PN, Weinstein SR, Smith RA, Griffiths LR: **Single nucleotide polymorphism in hsa-mir-196a-2 and breast cancer risk: a case control study.** *Twin Res Hum Genet* 2011, **14**:417-421.
64. Vrba L, Jensen TJ, Garbe JC, Heimark RL, Cress AE, Dickinson S, Stampfer MR, Futscher BW: **Role for DNA methylation in the regulation of miR-200c and miR-141 expression in normal and cancer cells.** *PLoS One* 2010, **5**:e8697.
65. Zhang L, Huang J, Yang N, Greshock J, Megraw MS, Giannakakis A, Liang S, Naylor TL, Barchetti A, Ward MR, Yao G, Medina A, O'Brien-Jenkins A, Katsaros D, Hatzigeorgiou A, Gimotty PA, Weber BL, Coukos G: **microRNAs exhibit high frequency genomic alterations in human cancer.** *Proc Natl Acad Sci USA* 2006, **103**:9136-9141.
66. Pandey DP, Picard D: **miR-22 inhibits estrogen signaling by directly targeting the estrogen receptor alpha mRNA.** *Mol Cell Biol* 2009, **29**:3783-3790.

doi:10.1186/bcr3415

Cite this article as: Riaz et al.: miRNA expression profiling of 51 human breast cancer cell lines reveals subtype and driver mutation-specific miRNAs. *Breast Cancer Research* 2013 **15**:R33.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
www.biomedcentral.com/submit

