Subtype-specific micro-RNA expression signatures in breast cancer progression

Short title: miRNA invasiveness signatures of breast cancer

Keywords: miRNA, breast cancer invasion, biomarker, subtype, DCIS
Article type: Original article

Brief description: Markers of breast cancer progression may help tailoring treatment of patients with *in situ* carcinomas. We have identified miRNA-based subtype-specific breast signatures differentiating between DCIS and invasive carcinomas that deserve to be tested prospectively in a clinical setting. We also propose an externally validated DCIS signature to replace a previously unvalidated signature.

Authors: Vilde D Haakensen[1) 2)]* Vegard Nygaard[3)]*, Liliana Gregers[4)] *, Miriam R Aure[1) 2)], Bastian Fromm[3)], Ida R K Bukholm [5) 6)], Torben Lüders[6) 7)], Suet-Feung Chin[8)], Anna Git[8) 9)], Carlos Caldas[8) 9)], Vessela N Kristensen[1) 2) 7)], Alvis Brazma[4)], Anne-Lise Børresen-Dale[1) 2)], Eivind Hovig [3) 10) 11)], Åslaug Helland[1) 12)]. *These authors have contributed equally to the manuscript

Affiliations:
1) Department of Cancer Genetics, Institute for Cancer Research, Division of Cancer Medicine, Surgery and Transplantation, Oslo University Hospital Radiumhospitalet, Oslo, Norway
2) The K. G. Jebsen Center for Breast Cancer Research, Institute for Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway,
3) Department of Tumor Biology, Institute for Cancer Research, Division of Cancer Medicine, Surgery and Transplantation, Oslo University Hospital Radiumhospitalet, Oslo, Norway
4) European Molecular Biology Laboratory, European Bioinformatics Institute, EMBL-EBI, Wellcome Trust Genome Campus, Hinxton, UK
5) Department of Breast and Endocrine Surgery, Akershus University Hospital, Lorenskog, Norway
6) Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway
7) Department of Clinical Molecular Biology (EpiGen), Akershus University Hospital, Lørenskog, Norway
8) Cancer Research UK, Cambridge Institute, University of Cambridge, Cambridge CB2 0RE, UK
9) Department of Oncology, University of Cambridge, Cambridge CB2 2XZ, UK
10) Department of Informatics, University of Oslo, Oslo, Norway
11) Institute of cancer genetics and informatics, Division of Cancer Medicine, Surgery and Transplantation, Oslo University Hospital Radiumhospitalet, Oslo, Norway
12) Department of Oncology, Division of Cancer Medicine, Surgery and Transplantation, Oslo University Hospital Radiumhospitalet, Oslo, Norway

Corresponding author:
Vilde D Haakensen
Department of Cancer Genetics
Institute for Cancer Research
Oslo University Hospital Radiumhospitalet
0310 Oslo
Norway
Ph: +47 22 93 40 00
Fax: +47 22 78 13 95
E-mail: vilde.drageset.haakensen@rr-research.no

Field Code Changed

Electronic word count: ~~4375~~ 4511 (excluding abstract and references)
Abstract word count: 233

Formatted: Highlight
Formatted: Highlight

List of abbreviations
miRNA, microRNA; DCIS, ductal carcinoma *in situ;* ER, estrogen receptor; PGR, progesterone receptor; HER2, human epidermal growth factor 2, FDR, false discovery rate; KEGG, Kyoto

Encyclopedia of Genes and Genomes; PAM50, Prediction analysis of microarrays 50; IC, Integrated cluster.

Competing interests

The authors declare that they have no competing interest.

Financial support:

This work was performed as part of the EurocanPlatform which has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 260791.

Authors' contributions:

VDH was co-ordinating the project, interpreted data and wrote the manuscript. VN performed pre-processing of the data, mapped the miRNAs, made data-analyses, interpreted data and contributed to writing the paper. LG performed pre-processing of the data, mapped the miRNAs, made data-analyses and contributed to writing the paper. MRA mapped the miRNAs and identified predicted miRNAs ~~Ids~~. BF contributed with curation of miRNAs. IRKB, TL, AG, SFC, CC and VNK participated in acquisition of data. AB and EH contributed to data analysis. ALBD conceived and designed the study. ÅH conceived and designed the study and interpreted the data. All authors critically revised the manuscript for important intellectual content and approved the final version of the manuscript to be published, and agreed to be accountable for all aspects of the work, ensuring that the accuracy or integrity of the work are appropriately investigated and resolved.

Novelty and Impact Statements: The current debate about over-treatment of non-invasive and indolent breast carcinomas calls for diagnostic measures of malignant potential. Robust markers of progression may help reduce the overtreatment. MicroRNAs are stable molecules with impact on gene-expression. Breast cancer is a heterogeneous disease and subtype-specific markers of progression and invasiveness are needed in order to identify relevant alterations. In this study, a subtype-specific microRNA signature predicting malignant potential is proposed.

Abstract

Robust markers of invasiveness may help reduce the overtreatment of *in situ* carcinomas. Breast cancer is a heterogeneous disease and biological mechanisms for carcinogenesis vary between subtypes. Stratification by subtype is therefore necessary in order to identify relevant and robust signatures of invasive disease. We have identified miRNA alterations during breast cancer progression in two separate datasets and used stratification and external validation to strengthen the findings.

We analysed two separate datasets (METABRIC and AHUS) consisting of a total of 186 normal breast tissue samples, 18 ductal carcinoma *in situ* (DCIS) and 1338 invasive breast carcinomas. Validation in a separate dataset and stratification by molecular subtypes based on immunohistochemistry, PAM50 and integrated cluster classifications were performed.

We propose subtype-specific miRNA signatures of invasive carcinoma and a validated signature of DCIS. miRNAs included in the invasive signatures include down-regulation of miR-139-5p in aggressive subtypes and up-regulation of miR-29c-5p expression in the luminal subtypes. No miRNAs were differentially expressed in the transition from DCIS to invasive carcinomas on the whole, indicating the need for subtype stratification. A total of 27 miRNAs were included in our proposed DCIS-signature. Significant alterations of expression included up-regulation of miR-21-5p and the miR-200-family and down-regulation of let-7 family members in DCIS samples. The signatures proposed here can form the basis for studies exploring DCIS samples with increased invasive potential and serum biomarkers for *in situ* and invasive breast cancer.

Introduction

Breast cancer is a life-threatening disease where early diagnosis reduces the risk of systemic involvement and thus improves prognosis. Identification of molecular alterations during the carcinogenic process from normal tissue to ductal carcinoma in situ (DCIS) and invasive cancer is important in order to understand the underlying biological processes, to identify biological tumour markers for early detection and to discover possible preventive and predictive measures. We propose microRNA (miRNA) expression signatures characteristic of the transition from normal breast tissue to DCIS and of DCIS to invasive cancer based on breast tissue whole genome miRNA expression data collected from two separate studies, METABRIC and AHUS.

The molecular mechanisms responsible for progression are poorly understood, although alterations between the normal and benign breast tissue as well as of malignant breast tumours have been published previously [1,2]. Due to power insufficiency, these studies did not investigate progression in a subtype-specific manner, a gap we are proposing to fill with the present dataset and analysis. Women with high mammographic density or benign breast lesions are at higher risk of developing malignant breast disease, but having difficult to interpret mammograms may mask the early stages of malignancy [3]. Therefore, having molecular markers, particularly if available in the circulation, may aid the diagnostic process.

DCIS is a disease that is not in itself life-threatening. Most lesions will not progress to invasive disease [4], but since we lack knowledge of which lesions will progress, all patients are treated radically to prevent progression. Biological markers of progression may enable reduction of the over-treatment of this group of patients. Since most patients with DCIS are treated with surgery and/or radiotherapy and very few women choose watchful surveillance there is a lack of information of which DCIS lesions will progress and therefore surrogate end-points must be used. One commonly employed strategy is to identify characteristics of the women with DCIS that will relapse or be diagnosed with invasive breast cancer later. Such a cohort captures women who are insufficiently treated for DCIS, but will mostly consist of women with second primary malignant breast disease and may hence lead to identification of factors associated with increased risk rather than progression [5]. An alternative strategy is to look for characteristics of invasive carcinoma in DCIS lesions [1]. Signatures of miRNAs differentially expressed between invasive carcinomas and DCIS samples, as developed in this study, may be used to create single sample predictors that look for invasiveness-like expression pattern in DCIS samples, first in a retrospective and later in a diagnostic setting.

The driving mechanisms of breast carcinogenesis and the resulting intra-tumour heterogeneity have been explained by clonal evolution [6]. This theory suggests that cancer arises in a normal cell through the accumulation of carcinogenic genetic aberrations in a multistep process. Later, cancer cells will acquire new genetic aberrations giving rise to new clones. This multistep process transforms the clone from normal cells to hyperplastic/benign cells and further to carcinoma cells that are confined by the basement membrane (DCIS) and, finally to carcinoma cells that have the capability of invading surrounding tissues and metastasizing. According to this model, it should be possible to detect lesions at pre-malignant stages to prevent invasive disease.

miRNAs are a class of small RNA molecules that regulate their target mRNAs by translational repression, or by aiding increased rates of mRNA degradation. miRNAs are associated with several cellular processes such as proliferation, apoptosis, stress response and carcinogenesis [7]. Enerly and colleagues identified miRNAs differentially expressed between different molecular subtypes of breast cancer and miRNAs related to molecular characteristics of the tumour (such as *TP53* mutations and estrogen receptor (ER) status) [8]. Aure et al identified miRNAs whose expression was consistently deregulated by DNA copy number alterations and methylation [9] and identified clusters of miRNAs that are collectively associated with the expression of key proteins (growth factors, their receptors and serine-tyrosine kinases) in breast cancer [10]. Dvinge and colleagues published an in-depth analysis of the miRNA profiles in the METABRIC dataset, which is also included in our analyses [11].

They focused on prognostic signatures, as well as an exploration of functional associations of miRNA networks, but they did not report progression signatures.

Several different approaches are used to subtype breast cancer. In the clinic, immunohistochemistry and fluorescent imaging classify tumours according to expression of ER and progesterone receptor (PGR), human epidermal growth factor 2 (HER2) receptor status and partly by the proliferation marker Ki67. Based on gene expression microarray status, five intrinsic subtypes were identified [12]. These subtypes have proven to be robust and clinically relevant. Refinement of the gene lists used for subtyping has resulted in the PAM50 classification [13]. Combining gene expression with copy number alterations, integrated clusters (IC) were identified carrying additional clinical and prognostic relevance [14].

Field Code Changed

Here, we report miRNA profiles characterizing the different transitions that occur during breast carcinogenesis, from normal breast tissue to DCIS and from DCIS to subtype-specific invasive carcinomas based on material from two different studies. The overlap between the DCIS-signature described here and that previously published by Volinia et al [1] is reported as validated miRNA signature of progression.

Material and methods

miRNA gene expression datasets have been collected from two different studies. Data from a total of 1542 breast tissue samples were included in this study: 186 normal breast tissue samples, 18 DCIS lesions and 1338 invasive carcinomas (~~Table~~ Additional file 1A). The cohorts are described below.

AHUS (Akershus University Hospital)

The Akershus University Hospital has sequentially collected breast tissue specimens from breast cancer patients and from women undergoing surgery for breast reduction. The patients were included from 2003 to 2009. miRNA expression profiling was obtained for 55 invasive carcinomas, 8 DCIS samples and 70 normal breast tissue samples (including 29 tumour-adjacent normal tissue samples and 41 breast reduction samples). The samples were hybridized on Agilent 8x15K arrays (Agilent Technologies, Santa Clara, CA, USA), catalogue number 4470B (v2) and 4470C (v3), and the features were extracted using Agilent Feature Extraction. Parts of the dataset has been published previously [2].The complete dataset is available in the ArrayExpress database (www.ebi.ac.uk/arrayexpress) under accession number E-MTAB-3759. Ethical approval was obtained by the Regional ethical committee South-East by reference numbers 429-04148 and 2014/895.

Field Code Changed

METABRIC (University of Cambridge)

The University of Cambridge has provided a dataset consisting of miRNA expression from 1283 breast cancer patients with 116 matched tumour-adjacent normal breast tissues as part of the METABRIC dataset previously published [11] and available in the European Genome-phenome Archive (EGAS00000000122). From the 1302 tumour samples, those described as phyllodes (n=5), benign (n=3) or null (n=1) were removed. 10 additional DCIS samples were included. The samples were hybridized on custom Agilent-miRNA-60K arrays (Agilent Technologies, Santa Clara, CA, USA), as previously described [11]. Ethical approval for the METABRIC study protocol was obtained from the ethics committees in Cambridge and Vancouver.

All patients in both studies have signed informed consent. Sample annotation for both cohorts is available in Additional file 2~~1~~, including subtype classification. ~~.~~

Validation cohort

Volinia and colleagues [1] published lists of differentially expressed genes based on a dataset published by Farazi and colleagues [15] consisting of sequenced miRNA profiles from 11 normal breast tissues, 17 DCIS samples and 151 invasive breast carcinomas, in addition to 6 cell lines. Of these, Volinia and colleagues used the profiles of 6 normal breast tissues, 8 DCIS samples and 80

invasive breast carcinomas to create transition signatures from normal to DCIS and from DCIS to invasive breast cancer, containing 66 and 9 miRNAs respectively [1].

Data processing

The METABRIC dataset was provided in a processed format, as described in Dvinge et al. The AHUS samples were provided as Agilent Feature Extraction result files. These were read in using the the R library AgiMicroRna [16] and then normalized using rma [17] without background subtraction [18]. MicroRNAs detected in fewer than 75% of the samples in at least one clinical category were omitted. To account for discrepancies in miRNA nomenclature due to different miRBase releases, the feature names were converted to MIMAT ids using a mapping file downloaded from miRBase [19]. Only miRNA features common to both datasets were analysed.

Statistical analyses

A clear batch effect was observed between the two data sets. To mitigate this, two analytical approaches were used, both using the limma package to detect differentially expressed genes between the clinical groups [20]. In the "meta" approach, the data sets were analyzed individually as well as combined by using the p-values in a meta-analysis. In the "merged" approach, the two data sets were combined and then analyzed with the data set source included as a blocking factor. A miRNA was reported as differentially expressed if its false discovery rate (FDR) was <0.05 in both analyses [21]. Information about the genomic location of the miRNAs was collected from the UCSC genome browser, hg19 assembly [22].

Differential expression (limma [20]) was used in an attempt to identify miRNAs accompanying the progression from DCIS to carcinoma. Due to the relatively small number of DCIS samples, they were included in the contrasts as a single group. Tumours were classified in alternative ways, by using routine immunohistochemical parameters (status of ER, PGR and HER2 receptor), gene expression (PAM50 [13]) or integrative genomic analysis [14]. Since data on copy number alterations was only available for the METABRIC samples, stratification by the IC-subgroups was only performed on this dataset. PAM50 subtype is given for DCIS samples where mRNA data was available (Additional file 1A and 2). Hierarchical clustering with Euclidean distance and complete linkage was performed by the function heatmap.2 within the R-package gplots.

miRNAs curation

miRNA probes were classified according to their likelihood to represent *bona fide* miRNA genes as described by Fromm et al [23]. Briefly, a consistent set of criteria for annotation of miRNAs is used that includes (1) proven genomic origin of (2) expressed hairpin precursor that (3) show at least 16 nt complementarity of both strands; (4) expression of 20-26 nt long reads with (5) highly conserved start positions (5' homogeneity); (6) reads originate from both arms of the precursor and (7) show a 2 nt offset created by the two sequential RNAse III cuts (Drosha/Dicer). Probes that clearly violated at least one of the criteria were omitted from proposed signatures, as were miRNAs no longer listed in miRBASE. miRNA target search is explained in Additional file 1B.

Reproducible research

In an effort to make our results reproducible [24], most of the scripts used in this work are accessible through a github repository at https://github.com/ous-uio-bioinfo-core/Eurocan_microRNA. Lists of all miRNAs differentially expressed are available in the additional material with explanation given in Additional file 1C~~2~~. The repository contains r-scripts, r-markdown, QC-plots and additional descriptions, but not the data which is accessible from https://www.ebi.ac.uk/arrayexpress accession number E-MTAB-3759 (AHUS dataset) *[Accesson granted reviewer with Username: Reviewer_E-MTAB-3759 and Password: hjynmtmx]* and https://www.ebi.ac.uk/ega/ accession number EGAS00000000122 (METABRIC dataset, approval needed). Validation was performed using an external dataset, Volinia et al [1].

Field Code Changed

Field Code Changed

Results and discussion

Lack of replicability is a major challenge of gene expression profiling studies. To improve robustness of our signatures, two separate cohorts were used where only miRNAs significant in two separate statistical approaches were included. Validation was performed using results from a third, external dataset was performed. Stratification by subtypes helps identify robust signatures by reducing heterogeneity.

Transition from normal breast tissue to DCIS

A total of 70 miRNAs were identified as differentially expressed between normal breast tissue and DCIS in both datasets and with both analytical approaches (meta-analysis and merged approach) (Additional file 3). Of these, 53 were analysed by Volinia et al [1] and 27 were found differentially expressed in the same direction in their study and are proposed as a DCIS signature (Table 1~~2~~ and Figure 1). miRNA-210-3p was significantly up-regulated in both our analyses (Additional file 3), but was down-regulated in the same transition in Volinia et al and is hence excluded from our proposed signature.

We found that the major alterations in miRNAs expression occur early in breast carcinogenesis and are detected in the transition from normal breast tissue to DCIS, as previous studies have shown [1,2]. In a previous analysis published by Tahiri et al, we studied the deregulation of miRNA expression in normal-benign-malignant specimens [2]. Of the 53 miRNAs analysed in the transition from normal to DCIS by this study and Volinia et al, 24 were also found deregulated in benign tumours (Additional file 1D ~~4 and 5~~). Of the 27 validated miRNAs differentially expressed between normal and DCIS, 14 are also found as differentially expressed in benign tumours in the same direction (Additional file 1D~~4~~). These include miR-21, let-7 and the miR-200 family. These results suggest that miRNA deregulation appears very early in ~~proliferation~~ progression and cannot be easily separated from markers of malignancy. Despite a limited number of DCIS samples, 70 miRNAs were identified as significantly altered using both statistical approaches (Additional file 3). Contrasting small numbers of DCIS and normal samples inherently makes all statistical approaches used to measure miRNA levels and the different bioinformatic processing of the raw data bias-prone. Nevertheless, the proposed signature of 27 miRNAs robustly emerges in all three independent cohorts, including an external dataset [1]. These miRNAs are proposed as a transition signature from normal breast tissue to DCIS (Table 1~~2~~). Hierarchical clustering of all normal and DCIS samples in each dataset based on these 27 miRNAs is illustrated in Figure 2, while clustering on the 256 expressed miRNAs is shown in Additional file 1E.

Some of the miRNAs have previously been found similarly altered. miR-21 is one of the most studied cancer-related miRNAs and several studies have found miR-21 up-regulated in DCIS samples compared with non-malignant breast tissue [25]. The expression of this miRNA was found at elevated levels with increasing stages of breast cancer, and a corresponding increase could be measured in the serum [26], suggesting it as a likely serum biomarker. This was also the miRNA with the highest fold change in our study. miR-21 is induced by inflammatory stimuli and is involved in the regulation of inflammation. Particularly, miR-21 acts as a mediator of the anti-inflammatory response in macrophages, but it has also been found to promote inflammatory mediators in carcinogenesis [27]. An increased expression of miR-21 in DCIS may thus reflect both a carcinogenic drive and the inflammatory host response to malignant alterations. The *MIR21* gene is located at 17q23.2, in a region that harbours several oncogenes and that has been found amplified in breast carcinomas [28]. Residing close to known oncogenes, its expression levels ~~will be a compound~~may be both a driver and passenger effect. The expression of this miRNA correlates positively with copy number alterations and negatively with methylation [9].

All five members of the miRNA-200 family (residing on two different chromosomes) are up-regulated in DCIS compared to normal breast tissue in our data; (miR-200a-3p, miR-200b-3p, miR-200c-3p, miR-141-3p and miR-429). Three of the miRNAs are validated in Volinia et al (miR-200b-3p, miR-200c-3p

and miR-429). These miRNAs are enriched in epithelial tissues [29] and are induced by TP53 [30]. The tumour suppressive function of these miRNAs is demonstrated by high expression inhibiting the epithelial to mesenchymal transition through repressed transcription of *ZEB1* and *ZEB2*, two E-cadherin transcriptional repressors [30]. At the same time, miR-200-expression has been associated with increased risk of metastasis in breast cancer and in mouse models [31]. An up-regulation of these miRNAs in the DCIS samples could indicate repression of epithelial mesenchymal transition and increase in E-cadherin transcription which is in line with the non-invasive nature of this disease. A relative increase in the epithelial compartment of the DCIS samples cannot be excluded. We do not find members of the miR-200 family up-regulated in any of the subtypes of invasive primary tumours, and have no data from metastases.

DCIS is heterogeneous, and molecular subtypes have been defined [32]. Due to a limited sample size, stratification by subtypes was not possible in our analyses and lack of stratification may increase noise so that we lose some information about the biology involved in the early stages of the carcinogenic process.

Target search of the 27 miRNAs was performed separately for miRNAs up- and down-regulated in the transition to DCIS as described in Additional file 1B. For miRNAs up-regulated in DCIS, 5616 predicted and validated targets were identified. Of these, 284 were cancer-related. For miRNAs down-regulated in DCIS, 7178 predicted and validated targets were identified, of which 272 were cancer-associated. Complete lists are given in Additional file 4.

Transition from DCIS to invasive breast cancer

No miRNAs were consistently identified as differentially expressed when DCIS and invasive samples *per se* were compared in both analytical approaches. Volinia et al have published an invasiveness signature of miRNAs altered in the transition from DCIS to invasive carcinoma [1]. Of their nine miRNAs, only one was borderline significant in our analysis: Let-7d with an FDR 0.06 and a log fold change of 0.65 (Additional file 5~~6~~).

The inconsistency of miRNA signatures in the transition from DCIS to invasive carcinoma may partly be due to low statistical power as the number of DCIS samples is low but it is, in our view, also due to the heterogeneity of breast cancer.

Overall, we show a higher rate of inter-study concordance in differentially-expressed miRNAs observed for normal-to-DCIS transition compared to DCIS-to-carcinoma transition. This may reflect the more fundamental changes taking place at the initiation of tumourigenesis, statistically reinforced by the relative uniformity of the normal samples compared to carcinomas. Thus, a subtype-stratified approach for carcinomas is better suited to identify biological processes~~d~~ in early DCIS-to-carcinoma transitions. Still, a subtype-stratified approach may give more representative information about the biological processes in the early transitional stages as well (from normal tissue to DCIS).

Subtype-specific signatures of invasive breast carcinomas

We have identified subtype-specific signatures of invasive breast carcinomas (Tables ~~3~~2 and ~~4~~3, Figure 1). Fold change and FDR-values of the miRNAs differentially expressed from DCIS to each subtype are listed in Additional files 6~~7~~-2~~45~~. Subtype stratification based on molecular signatures (PAM50 or IC) results in more significant miRNAs than stratification based on ER, PgR and HER2 alone, indicating a better representation of the intrinsic biology of the samples. For IHC-based subtypes (Additional files 2~~2~~1-2~~5~~4), no miRNAs were differentially expressed between DCIS and the ER+ subtypes. Six miRNAs were down-regulated in ER-/HER2+ invasive samples compared with DCIS of which five belong to the miR-30-family (miR-30a-5p, miR-30a-3p, miR-30c-2-3p, miR-30c-5p-miR-30e-3p, miR-139-5p and miR-221-3p). miR-139-5p was down-regulated in both ER- subtypes, while miR-887-3p was down-regulated in triple negative breast cancer only.

Ten miRNAs were deregulated in several subtypes. These may serve as more general markers, but they were not identified without stratification by the invasive subtypes. miR-125b2-3p, miR-139-5p, miR-140-3p, miR-378a-3p, miR-221-3p, miR-31p and miR-433-3p were all down-regulated in the invasive subtypes and have been reported as tumour suppressors in the literature [33-36]. Three miRNAs (miR-505, miR-29c-5p and miR-30a-3p) have a conflicting direction of regulation in the transition from DCIS to invasive subtypes. The two latter are up-regulation in the transition from DCIS to luminal subtypes, but down-regulated in the more aggressive HER2 subtype.

One miRNA, miR-139-5p, was down-regulated in the poor prognosis subtypes; HER2-enriched, luminal B, basal-like and in the HER2-enriched IC5 tumours as well as in triple negative tumours, but not in the good prognosis subtypes, luminal A and IC3 (in the latter, no miRNAs are significantly deregulated). This miRNA is a well known tumour suppressor and is a putative biomarker for several cancer types and has been found to bedown-regulated in triple negative breast carcinomas and with increasing involvement of lymph nodes [34].

Down-regulation of several members of the miR-30 family was observed in HER2-associated tumours (Tables ~~3~~2 and Additional file 2~~4~~3). Expression of miRNAs in this family has been reported to confer favourable prognosis [37]. miR-30a-3p was up-regulated in luminal A tumours and in IC8 which is dominated by luminal A samples. In the literature, this miRNA is described as a tumour suppressor and low expression has been associated with early recurrence [38]. Expression of miR-30c-2-3p, which was down-regulated in HER2-enriched and HER2-positive tumours in our data, has previously been found up-regulated in luminal A tumours and down-regulated in basal-like and HER2-enriched tumours [37]. This miRNA is an independent positive prognostic factor and has been found to reduce resistance to doxorubicin and paclitaxel [37] and to regulate breast cancer cell invasion [39]. A low expression in the poor prognosis HER2-enriched subtype is in line with studies that have shown that high expression of HER2 is associated with risk of progression [40]. Both miR-30a-3p and 30c-2-3p are down-regulated in the transition from DCIS to IC5 which is dominated by the HER2 subtype ($p<0.02$), but the significance does not hold after correction for multiple testing. miR-31-5p was also down-regulated in several subtypes (luminal B, IC1, IC2 and IC8) and is also known to be a tumour suppressor [41].

miR-29c-5p was up-regulated in luminal A and B tumours and down-regulated in basal-like tumours, suggesting that it might be a marker of luminal invasiveness. The same miRNA is up-regulated from DCIS to IC1, 2, 6, 7 and 8 which are all dominated by luminal samples, but down-regulated in the transition to IC10 which is dominated by basal-like samples. Dvinge and colleagues found the miR--29 family involved in the modulatory network of IC4 related to inflammatory response, but it was not significantly regulated in the transition from DCIS to IC4. Our findings correspond well with previous findings that miR-29c suppress cell migration and invasion (in colorectal cancer)[42] and that high expression of miR-29c is associated with favourable prognosis [43].

Chromosomal location

An overview of miRNA expression related to chromosomal location is presented in Figure ~~2~~3. Down-regulation of miRNA transcripts of genes located on chromosome 14 is dominating. Chromosomal location for deregulated miRNAs was determined and cytobands with more than one deregulated miRNA during breast cancer progression were identified (Additional file 1F~~2~~6). Two miRNA-clusters located on 14q32 were deregulated in the transition from DCIS to various invasive subtypes, primarily luminal B and IC10 (good outcome basal-like samples). One miRNA was up-regulated (miR-342-3p) whereas the others were all down-regulated. ~~14q32 is a very well studied imprinted genomic region, where the miRNA clusters are often internally co-regulated and have been found down-regulated in many cancers, including breast cancer [9]. Deletions in this region have been associated with BRCA2-mutations [42]which in turn often share a similar phenotype with luminal B tumours [43] which matches our results well. Loss in 14q32.33 has been associated with poor prognosis [44], which is another feature of luminal B and IC10 tumours. There are however many miRNAs located in~~

~~14q32 and of the 31 miRNAs from this region expressed in our data, the majority (n=17) of the expressed miRNAs (n=31) were not significantly altered.~~

Generally, there is a down-regulation of deregulated miRNAs during progression, but chromosome 1q (particularly 1q32.2) has overrepresentation of up-regulated miRNAs (Figure 3~~2~~). This chromosomal region is amplified in many breast carcinomas, both of indolent and aggressive subtypes [44]. Chromosome 14 (particularly 14q32.2) is characterized by a large proportion of down-regulated miRNAs.

Four miRNAs with genes located at 5q32 were down-regulated in the transition from normal breast tissue to DCIS (Table ~~2~~1); miR-143-5p, miR-145-3p, miR-145-5p and miR-378a-3p, the latter being down-regulated in several invasive sub-types as well (luminal A, luminal B, IC2, IC6 and IC8) (Additional file 1F). ~~. miR-143 and miR-145 have been suggested to have tumour suppressor function [46] and have been found down-regulated in triple negative breast cancer [47]. Chivukala and colleagues found that these miRNAs are not expressed in non-malignant colonic epithelial cells and the apparent down-regulation in tumours may be due to a smaller fraction of stromal cells expressing these miRNAs [48,49]. This is in line with our findings that all breast tumour subtypes have low expression of these miRNAs compared with non-malignant tissue which usually contains more stromal cells.~~

Step-wise progression from normal to DCIS to invasive cancer

Fourteen of the 27 miRNAs validated in the transition from normal breast tissue to DCIS, were also found significantly altered in the progression from DCIS to a specific breast cancer subtype. Of these, seven were consistently down-regulated (let-7c-5p, miR-125b-5p, miR-140-3p, miR-145-3p, miR-145-5p, miR-193a-5p and miR-378a-3p) consistent with a subtype-specific tumour suppressive function~~.~~. Four miRNA were consistently up-regulated, indicating an oncogenic potential (miR-106b-5p, miR-142, miR-342-3p, miR-425-5p).

One miRNA was down-regulated in the transition to DCIS, but had conflicting deregulation in the invasive transition (miR-99a-5p was down-regulated in luminal B and up-regulated in normal-like). Two miRNAs were down-regulated in DCIS and up-regulated in invasive subtypes (luminal B and IC1 tumours for miR-193b-3p and normal-like for miR-497-5p)~~.~~. The stepwise e~~E~~xpression of these miRNAs is illustrated in Additional file 2~~5~~2~~7~~).

miR-210 has previously been identified as a marker of poor prognosis in breast cancer and other carcinomas [1]. Volinia et al found miR-210 to be down-regulated in DCIS compared with normal breast tissue, but up-regulated in invasive carcinomas compared with DCIS [1]. In our data, miR-210 was up-regulated in DCIS compared with normal tissue and was not detected as significantly altered in any invasive subtype. The role of this miRNA in breast carcinogenesis is still unclear.

~~**Possible clinical application**~~

~~These miRNA signatures provide candidates for miRNA serum markers of breast carcinogenesis. Most of the miRNAs identified are detectable in serum (unpublished data from our lab). Testing these miRNAs in serum and their association with breast cancer disease may enable diagnostic tests that may add information to mammography/ultrasound in the diagnosis/follow-up of women with high risk of developing breast malignancy.~~

The main purpose of identifying miRNA signatures of progression is the study of the carcinogenic process at a molecular level. However, miRNAs associated with breast cancer progression may also serve as markers of the disease in surgically removed tissue and possibly even in the serum. There is reason to believe that miRNAs are selectively secreted into circulation and many miRNAs present in the tissue will not be present at significant levels in the serum. Separate studies are therefore needed in plasma/serum to explore these possibilities. Of the 12 miRNAs in our DCIS signature that was up-regulated in DCIS, 6 are identified at a significant level (median of all samples had a higher than minimum abundance in serum): miR-21-5p, miR-106b-5p, 142-3p, miR-342-3p, miR-425-5p, let-

7b-5p). Of the remaining up-regulated miRNAs, the literature supports a possible serum biomarker role for miR-155 [45] and miR-200c [46].

Replicability of miRNA signatures

Several studies have shown limited overlap of differentially expressed miRNAs from microarrays and sequencing [47]. Although sequencing of miRNAs is becoming more of a gold standard for measuring the expression levels of miRNAs, there are challenges to both library preparation and data analysis. The most varying miRNAs should, however, be identified by both methods. The Agilent microarray platform has been found to be of the best performers in identifying true differentially expressed miRNAs [47].

Dvinge and colleagues studied miRNA-expression in the large METABRIC dataset, which was also included in this study. They found that individual miRNAs are not robust prognostic markers [11]. They also found that miRNA signatures were not highly correlated with clinical or histopathological traits, but to a larger extent with molecular traits such as PAM50 and IC-subtype. This corresponds with our finding that miRNA signatures were identified in the transition to PAM50 subtypes, but not to immunohistologically defined subtypes.

Using two different statistical approaches, we aimed at removing statistically obtained false positive miRNAs. The number of miRNAs significant in each approach is given in Additional files 3 and 5~~6~~-2~~45~~. Generally, the merged approach identified more miRNAs than the meta-analysis, while most of the miRNAs identified in the meta-analysis were also present in the merged approach.

Conclusions

We propose miRNA signatures characterizing the transition from normal breast to DCIS and the invasive transformation from DCIS to several PAM50 and IC subtype of breast cancer. These signatures are candidate biomarkers of breast carcinogenesis (DCIS and invasive subtypes of breast cancer) and their clinical value deserves to be tested in prospective setting. Most of the miRNAs are detectable in the serum and could possibly add clinical value to mammography in a diagnostic setting. We have also demonstrated that the heterogeneous nature of breast cancer necessitates stratification by subtype for identification of robust signatures and that stratification by PAM50 or IC subtype gives more significant miRNAs than stratification based on IHC scores.

Acknowledgements

We would like to thank Stefano Volinia for communication about his work. This work was performed as part of the EurocanPlatform which has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement No. 260791. Portions of this research (Venn diagram creator) were supported by the W.R. Wiley Environmental Molecular Science Laboratory, a national scientific user facility sponsored by the U.S. Department of Energy's Office of Biological and Environmental Research and located at PNNL. PNNL is operated by Battelle Memorial Institute for the U.S. Department of Energy under contract DE-AC05-76RL0 1830

Reference List

1. Volinia S, Galasso M, Sana ME, Wise TF, Palatini J, Huebner K *et al.*: **Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA.** *Proc Natl Acad Sci U S A* 2012, **109:** 3024-3029.

2. Tahiri A, Leivonen SK, Luders T, Steinfeld I, Ragle AM, Geisler J *et al*.: **Deregulation of cancer-related miRNAs is a common event in both benign and malignant human breast tumors.** *Carcinogenesis* 2013.

3. Ursin G, Ma H, Wu AH, Bernstein L, Salane M, Parisky YR *et al*.: **Mammographic density and breast cancer in three ethnic groups.** *Cancer Epidemiol Biomarkers Prev* 2003, **12:** 332-338.

4. Rosen PP, Braun DW, Jr., Kinne DE: **The clinical significance of pre-invasive breast carcinoma.** *Cancer* 1980, **46:** 919-925.

5. Kerlikowske K, Molinaro A, Cha I, Ljung BM, Ernster VL, Stewart K *et al*.: **Characteristics associated with recurrence among women with ductal carcinoma in situ treated by lumpectomy.** *J Natl Cancer Inst* 2003, **95:** 1692-1702.

6. Nik-Zainal S, Van LP, Wedge DC, Alexandrov LB, Greenman CD, Lau KW *et al*.: **The life history of 21 breast cancers.** *Cell* 2012, **149:** 994-1007.

7. Calin GA, Croce CM: **MicroRNA signatures in human cancers.** *Nat Rev Cancer* 2006, **6:** 857-866.

8. Enerly E, Steinfeld I, Kleivi K, Leivonen SK, Aure MR, Russnes HG *et al*.: **miRNA-mRNA integrated analysis reveals roles for miRNAs in primary breast tumors.** *PLoS One* 2011, **6:** e16915.

9. Aure MR, Leivonen SK, Fleischer T, Zhu Q, Overgaard J, Alsner J *et al*.: **Individual and combined effects of DNA methylation and copy number alterations on miRNA expression in breast tumors.** *Genome Biol* 2013, **14:** R126.

10. Aure MR, Jernstrom S, Krohn M, Vollan HK, Due EU, Rodland E *et al*.: **Integrated analysis reveals microRNA networks coordinately expressed with key proteins in breast cancer.** *Genome Med* 2015, **7:** 21.

11. Dvinge H, Git A, Graf S, Salmon-Divon M, Curtis C, Sottoriva A *et al*.: **The shaping and functional consequences of the microRNA landscape in breast cancer.** *Nature* 2013, **497:** 378-382.

12. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H *et al*.: **Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.** *Proc Natl Acad Sci U S A* 2001, **98:** 10869-10874.

13. Parker JS, Mullins M, Cheang MC, Leung S, Voduc D, Vickery T *et al*.: **Supervised risk predictor of breast cancer based on intrinsic subtypes.** *J Clin Oncol* 2009, **27:** 1160-1167.

14. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ *et al*.: **The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups.** *Nature* 2012, **486:** 346-352.

15. Farazi TA, Horlings HM, Ten Hoeve JJ, Mihailovic A, Halfwerk H, Morozov P *et al*.: **MicroRNA sequence and expression analysis in breast tumors by deep sequencing.** *Cancer Res* 2011, **71:** 4443-4453.

16. Lopez-Romero P. AgiMicroRna: Processing and Differential Expression Analysis of Agilent microRNA chips. R package version 2.18.0. 2015.
Ref Type: Online Source

17. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U *et al*.: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Biostatistics* 2003, **4:** 249-264.

18. Lopez-Romero P, Gonzalez MA, Callejas S, Dopazo A, Irizarry RA: **Processing of Agilent microRNA array data.** *BMC Res Notes* 2010, **3:** 18.

19. Ambros V, Bartel B, Bartel DP, Burge CB, Carrington JC, Chen X *et al*.: **A uniform system for microRNA annotation.** *RNA* 2003, **9:** 277-279.

20. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W *et al*.: **limma powers differential expression analyses for RNA-sequencing and microarray studies.** *Nucleic Acids Res* 2015.

21. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J Royal Stat Soc Ser B* 1995, **57:** 289-300.

22. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM *et al*.: **The human genome browser at UCSC.** *Genome Res* 2002, **12:** 996-1006.

23. Fromm B, Billipp T, Johansen M, Peck LE, Tarver JE, King B *et al*.: **A Uniform System For The Annotation Of Human microRNA Genes And The Evolution Of The Human microRNAome.** *Ann Rev Genetics* 2015, **49**.

24. Peng RD: **Reproducible research in computational science.** *Science* 2011, **334:** 1226-1227.

25. Hannafon BN, Sebastiani P, de las MA, Lu J, Rosenberg CL: **Expression of microRNA and their gene targets are dysregulated in preinvasive breast cancer.** *Breast Cancer Res* 2011, **13:** R24.

26. Si H, Sun X, Chen Y, Cao Y, Chen S, Wang H *et al*.: **Circulating microRNA-92a and microRNA-21 as novel minimally invasive biomarkers for primary breast cancer.** *J Cancer Res Clin Oncol* 2013, **139:** 223-229.

27. Sheedy FJ: **Turning 21: Induction of miR-21 as a Key Switch in the Inflammatory Response.** *Front Immunol* 2015, **6:** 19.

28. Ueno T, Emi M, Sato H, Ito N, Muta M, Kuroi K *et al*.: **Genome-wide copy number analysis in primary breast cancer.** *Expert Opin Ther Targets* 2012, **16 Suppl 1:** S31-S35.

29. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D *et al*.: **MicroRNA expression profiles classify human cancers.** *Nature* 2005, **435:** 834-838.

30. Kim T, Veronese A, Pichiorri F, Lee TJ, Jeon YJ, Volinia S *et al*.: **p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2.** *J Exp Med* 2011, **208:** 875-883.

31. Korpal M, Ell BJ, Buffa FM, Ibrahim T, Blanco MA, Celia-Terrassa T *et al*.: **Direct targeting of Sec23a by miR-200s influences cancer cell secretome and promotes metastatic colonization.** *Nat Med* 2011, **17:** 1101-1108.

32. Muggerud AA, Hallett M, Johnsen H, Kleivi K, Zhou W, Tahmasebpoor S *et al*.: **Molecular diversity in ductal carcinoma in situ (DCIS) and early invasive breast cancer.** *Mol Oncol* 2010, **4:** 357-368.

33. Sun YM, Lin KY, Chen YQ: **Diverse functions of miR-125 family in different cell contexts.** *J Hematol Oncol* 2013, **6:** 6.

34. Zhang HD, Jiang LH, Sun DW, Li J, Tang JH: **MiR-139-5p: promising biomarker for cancer.** *Tumour Biol* 2015, **36:** 1355-1365.

35. Yuan Y, Shen Y, Xue L, Fan H: **miR-140 suppresses tumor growth and metastasis of non-small cell lung cancer by targeting insulin-like growth factor 1 receptor.** *PLoS One* 2013, **8:** e73604.

36. Weiner-Gorzel K, Dempsey E, Milewska M, McGoldrick A, Toh V, Walsh A *et al*.: **Overexpression of the microRNA miR-433 promotes resistance to paclitaxel through the induction of cellular senescence in ovarian cancer cells.** *Cancer Med* 2015, **4:** 745-758.

37. Bockhorn J, Dalton R, Nwachukwu C, Huang S, Prat A, Yee K *et al*.: **MicroRNA-30c inhibits human breast tumour chemotherapy resistance by regulating TWF1 and IL-11.** *Nat Commun* 2013, **4:** 1393.

38. Zhang N, Wang X, Huo Q, Sun M, Cai C, Liu Z *et al*.: **MicroRNA-30a suppresses breast tumor growth and metastasis by targeting metadherin.** *Oncogene* 2014, **33:** 3119-3128.

39. Bockhorn J, Yee K, Chang YF, Prat A, Huo D, Nwachukwu C *et al*.: **MicroRNA-30c targets cytoskeleton genes involved in breast cancer cell invasion.** *Breast Cancer Res Treat* 2013, **137:** 373-382.

40. Lu J, Guo H, Treekitkarnmongkol W, Li P, Zhang J, Shi B *et al*.: **14-3-3zeta Cooperates with ErbB2 to promote ductal carcinoma in situ progression to invasive breast cancer by inducing epithelial-mesenchymal transition.** *Cancer Cell* 2009, **16:** 195-207.

41. O'Day E, Lal A: **MicroRNAs and their target gene networks in breast cancer.** *Breast Cancer Res* 2010, **12:** 201.

42. Zhang JX, Mai SJ, Huang XX, Wang FW, Liao YJ, Lin MC *et al*.: **MiR-29c mediates epithelial-to-mesenchymal transition in human colorectal carcinoma metastasis via PTP4A and GNA13 regulation of beta-catenin signaling.** *Ann Oncol* 2014, **25:** 2196-2204.

43. Nygren MK, Tekle C, Ingebrigtsen VA, Makela R, Krohn M, Aure MR *et al*.: **Identifying microRNAs regulating B7-H3 in breast cancer: the clinical impact of microRNA-29c.** *Br J Cancer* 2014, **110:** 2072-2080.

44. Bergamaschi A, Kim YH, Wang P, Sorlie T, Hernandez-Boussard T, Lonning PE *et al*.: **Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer.** *Genes Chromosomes Cancer* 2006, **45:** 1033-1040.

45. Roth C, Rack B, Muller V, Janni W, Pantel K, Schwarzenbach H: **Circulating microRNAs as blood-based markers for patients with primary and metastatic breast cancer.** *Breast Cancer Res* 2010, **12:** R90.

46. Le MT, Hamar P, Guo C, Basar E, Perdigao-Henriques R, Balaj L *et al*.: **miR-200-containing extracellular vesicles promote breast cancer cell metastasis.** *J Clin Invest* 2014, **124:** 5109-5128.

47. Git A, Dvinge H, Salmon-Divon M, Osborne M, Kutter C, Hadfield J *et al.*: **Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression.** *RNA* 2010, **16:** 991-1006.

Additional files

Additional file 1: Text-file. Additional information. A) Sample overview, B) miRNA target search, C) Explanation for lists of differentially expressed miRNAs, D) Comparison to Tahiri et al E) Hierarchical clustering, F) Chromosomal location. (txt)

Additional file 2: Table. Sample annotation. Annotation of all samples from METABRIC and AHUS cohorts with tissue type (normal/DCIS/invasive) and subtypes (PAM50, IC and IHC). (csv)

Additional file 3: Table. Differentially expressed miRNAs between normal breast tissue and DCIS. Explanation given in Additional file 1C). (csv)

Additional file 4: Table. Lists of predicted and validated targets of the 27 miRNAs in the DCIS signature stratified into up- and down-regulated signatures. Identification of cancer-related targets and overlap between targets of up- and down-regulated miRNAs. (xlsx)

Additional file 5: Table. Differentially expressed miRNAs between DCIS and invasive samples. Explanation given in Additional file 2. (csv)

Additional file 6: Table. Differentially expressed miRNAs between DCIS and luminal A invasive samples. Explanation given in Additional file 2. (csv)

Additional file 7: Table. Differentially expressed miRNAs between DCIS and luminal B invasive samples. Explanation given in Additional file 2. (csv)

Additional file 8: Table. Differentially expressed miRNAs between DCIS and HER2 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 9: Table. Differentially expressed miRNAs between DCIS and basal-like invasive samples. Explanation given in Additional file 2. (csv)

Additional file 10: Table. Differentially expressed miRNAs between DCIS and normal-like invasive samples. Explanation given in Additional file 2. (csv)

Additional file 11: Table. Differentially expressed miRNAs between DCIS and IC1 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 12: Table. Differentially expressed miRNAs between DCIS and IC2 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 13: Table. Differentially expressed miRNAs between DCIS and IC3 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 14: Table. Differentially expressed miRNAs between DCIS and IC4 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 15: Table. Differentially expressed miRNAs between DCIS and IC5 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 16: Table. Differentially expressed miRNAs between DCIS and IC6 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 17: Table. Differentially expressed miRNAs between DCIS and IC7 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 18: Table. Differentially expressed miRNAs between DCIS and IC8 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 19: Table. Differentially expressed miRNAs between DCIS and IC9 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 20: Table. Differentially expressed miRNAs between DCIS and IC10 invasive samples. Explanation given in Additional file 2. (csv)

Additional file 21: Table. Differentially expressed miRNAs between DCIS and ER+/HER2+ invasive samples. Explanation given in Additional file 2. (csv)

Additional file 22: Table. Differentially expressed miRNAs between DCIS and ER+/HER2- invasive samples. Explanation given in Additional file 2. (csv)

Additional file 23: Table. Differentially expressed miRNAs between DCIS and ER-/HER2+ invasive samples. Explanation given in Additional file 2. (csv)

Additional file 24: Table. Differentially expressed miRNAs between DCIS and triple negative invasive samples. Explanation given in Additional file 2. (csv)

Additional file 25: Figure S2. Boxplots of the expression of miRNAs differentially expressed in both transitions. (pdf)

Figure legends and tables

Figure 1. Summary of differentially expressed miRNAs. miRNAs differentially expressed between normal breast tissue and DCIS are validated in an external dataset. miRNAs differentially expressed between DCIS and the PAM50 and IC subtypes of invasive breast cancer are proposed as signatures of invasive breast carcinomas. All significant miRNAs are displayed for most subtypes, but selected miRNAs are displayed for the transition from normal to DCIS and from DCIS to subtypes luminal B, IC1, IC8 and IC10 due to large number of significant miRNAs (n>15). See Tables 1~~2~~, ~~3~~2, and ~~4~~3 for complete lists. Bold miRNAs are consistently differentially expressed from normal breast tissue to DCIS and from DCIS to an invasive subtype. Red = up-regulated. Blue = down-regulated.

Figure 2. Hierarchical clustering of normal and DCIS samples based on the proposed DCIS miRNA signature. The clustering was performed for the two datasets AHUS and METABRIC separately.

Figure ~~2~~3. Overview of genomic localisation of miRNAs deregulated during breast cancer progression. Alterations from normal tissue to DCIS and from DCIS to PAM50 and IC subtypes are shown. No miRNAs are deregulated from DCIS to IC3 or IC4. Red = up-regulated. Blue = down-regulated.