



HAL
open science

The proteins FABP7 and OATP2 are associated with the basal phenotype and patient outcome in human breast cancer

H. Zhang, E. A. Rakha, G. R. Ball, I. Spiteri, M. Aleskandarany, E. C. Paish, D. G. Powe, R. D. Macmillan, C. Caldas, I. O. Ellis, et al.

► **To cite this version:**

H. Zhang, E. A. Rakha, G. R. Ball, I. Spiteri, M. Aleskandarany, et al.. The proteins FABP7 and OATP2 are associated with the basal phenotype and patient outcome in human breast cancer. *Breast Cancer Research and Treatment*, 2009, 121 (1), pp.41-51. 10.1007/s10549-009-0450-x . hal-00485060

HAL Id: hal-00485060

<https://hal.science/hal-00485060>

Submitted on 20 May 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The proteins FABP7 and OATP2 are associated with the basal phenotype and patient outcome in human breast cancer

H Zhang¹, EA Rakha¹, GR Ball², I Spiteri⁴, M Aleskandarany¹, EC Paish¹, DG Powe¹, RD Macmillan³, C Caldas⁴, IO Ellis¹, AR Green¹

¹Division of Pathology, School of Molecular Medical Sciences, ³The Breast Institute, Nottingham University Hospitals and University of Nottingham

²School of Biomedical and Natural Sciences, Nottingham Trent University

⁴Cancer Research UK, Cambridge Research Institute, Li Ka-Shing Centre, Cambridge

Address correspondence to: Andrew Green, Division of Pathology, School of Molecular Medical Sciences, University of Nottingham, Queen's Medical Centre, Nottingham, UK NG7 2UH. e-mail: andrew.green@nottingham.ac.uk Tel: +44 1158230786 Fax: +44 1159709479

Running head: FABP7 expression in basal-like breast cancer

Keywords: basal phenotype, breast cancer, FABP7, OATP2, immunohistochemistry, tissue microarray

Abstract

Purpose

The basal-like or basal phenotype class of breast cancers (BP) have recently attracted attention as a poor prognostic form of breast cancer. However, BP appear to encompass biologically and clinically heterogeneous tumours; resulting in a lack of consensus definition of BP.

Methods

We analysed 48,000 gene transcripts in 132 invasive breast carcinomas to identify two novel genes (OATP2 and FABP7) significantly associated with BP (defined by cytokeratin (CK)5/6 and/or CK14 positivity). Using a series of invasive breast carcinoma cases (n=899), prepared as tissue microarrays, we assessed OATP2 and FABP7 protein expression using immunocytochemistry to investigate associations with clinicopathological variables, patients' outcome and ability to refine BP classification.

Results

A total of 7.9% and 15.6% cases were OATP2 and FABP7 positive respectively. OATP2 was associated with tumours of high histological grade ($p < 0.01$), ER and PgR negativity ($p < 0.01$) and shorter breast cancer specific survival (BCSS; $p = 0.04$). FABP7 expression was associated with lower lymph node stage ($p < 0.01$), ER and PgR negativity ($p < 0.01$). BP tumours which were FABP7 positive had a significantly longer BCSS ($p = 0.05$) and disease-free survival (DFS; $p = 0.01$) compared with FABP7 negative basal tumours ($p < 0.01$). OATP2 positive tumours were associated with adverse survival and increased risk of early recurrence.

Conclusions

This study confirms the biological and clinical heterogeneity of the BP in breast cancer. We have identified a novel subgroup of basal tumours showing FABP7 expression that have significantly better clinical outcome. Further studies analysing the role of FABP7 are therefore warranted.

Introduction

Advances in the understanding of breast cancer molecular and genetic mechanisms have led to the realisation of the heterogeneity of the disease and the promise of a new era of individualised management for breast cancer patients. Microarray-based class discovery studies pioneered by Perou and colleagues [1] have demonstrated that breast cancer could be classified into molecularly distinct groups based upon global gene expression profiles and their similarity to those of normal cell counterparts. Their results have subsequently been confirmed and expanded by multiple independent studies [2-5]. One of these biologically and clinically distinct groups, which has attracted attention in recent years as a molecularly distinct subtype of breast cancer with poor prognosis, is basal-like cancer [1, 3, 4, 6]. Basal-like cancers also show negativity for markers that are targeted by endocrine therapy (estrogen receptor (ER) and progesterone receptor (PgR)) and Herceptin (HER2). In breast cancer, the class of tumours showing basal differentiation has been described more than three decades ago [7]. Furthermore the expression of basal CKs in breast cancer and their association with poor prognosis has confirmed by several studies [8-11]. However, basal-like class of tumours have recently received much attention after their identification in gene expression profiling (GEP) studies as a molecularly distinct subtype of breast cancer

Although patients with basal-like tumours are generally characterised by shorter disease free survival, high mortality rate and experience a unique pattern of distant metastasis [12, 13], it must be noted that this is not always the case, as some authors have reported a heterogeneity within basal like tumours and other studies have failed to show an association with poorer outcome [14-16]. Identification of the basal subset can be achieved by GEP or an immunophenotype (IHC). Despite the claimed greater accuracy of identification using GEP [17],

the more practical and cheaper method of IHC is more suited to a routine clinical setting. However, the inherent heterogeneity of basal tumours [18-20] has resulted in a lack of robust definition of the basal phenotype in clinical diagnosis. In clinical practice the 'triple-negative' designation based on negative expression of ER, PgR and HER2 has been adopted and promoted in some centres as a simple and pragmatic definition [21, 22]. In spite of this, recent studies have raised the issue that some basal CK expressing (CK5/6, 14, 17) tumours are positive for ER and HER2, questioning general validity of the triple-negative method [23] and its reliance on negative results without inclusion of positive entry criteria. Several studies have attempted to refine the classification of the group by using a broader range of markers including the basal cytokeratins 5/6 and 14 and other basal associated markers such as epidermal growth factor receptor (EGFR), vimentin and P-cadherin [14, 24].

Most IHC based definitions use positive expression of basal CKs as the basis to define the basal group [12, 13, 15, 25-27] as additional markers have so far failed to refine further the identification of poor survival cases.. Other authors argue that basal CKs are not always expressed by the basal-like group of tumours identified by gene analyses, and hence novel markers must be incorporated into its definition [28].

The aim of this study was to determine and characterise the protein expression of two novel markers associated with basal CKs expression, identified from gene profiling analysis, in large well-characterised cohort of invasive breast tumours and to investigate their association with other clinicopathological variables, patients' outcome and their ability to refine the classification of the basal phenotype.

Methods

Patient samples

The patient cohort comprised 899 primary operable invasive breast carcinoma cases from women presenting between 1986-1993, selected from the well characterised Nottingham-Tenovus Primary Breast Carcinoma Series. All patients had been assessed in a standardised manner for clinical history and tumour characteristics including age, menopause status, tumour type [29], histological grade [30], tumour size, tumour recurrence, presence of distant metastasis, lymph node status, vascular invasion and Nottingham Prognostic Index (NPI) [31]. Data relating to survival was collated in a prospective manner including breast cancer specific survival (BCSS), defined as the interval (in months) from the date of the primary treatment to the time of death due to breast cancer, and disease free interval (DFI), defined as the interval (in months) from the date of the primary treatment to the first loco-regional recurrence or distant metastasis. Patient management was based on tumour characteristics by NPI and hormone receptor status. Patients with an NPI score ≤ 3.4 received no adjuvant therapy, those with a NPI score >3.4 received Tamoxifen if oestrogen receptor (ER) positive (\pm Zoladex if pre-menopausal) or classical cyclophosphamide, methotrexate and 5-fluorouracil if ER negative and fit enough to tolerate chemotherapy [32].

The cohort of patients (n=899) consisted of 621 patients aged over or equal to 50 years, and 278 under 50 years old. Disease characteristics varied amongst the cohort; at primary diagnosis 196 (21.8%) were grade 1, 301 (33.5%) grade 2, 392 (43.6%) grade 3. A total of 561 (62.4%) patients had tumours larger than 1.5cm and 246 (27.4%) had definite vascular invasion. Follow-up data was collected initially at 3 months, followed by intervals of 6 months and then 12 months (median 160 months; range 1 – 224 months). During this time, recurrence occurred in 394

(43.8%) cases, distant metastases in 115 (12.8%) cases, and 278 (30.9%) patients died from breast cancer. The immunoreactivity, scoring and categorisation of oestrogen receptor (ER), progesterone receptor (PgR), cytokeratin (CK)5/6, CK14, CK18, CK19, HER2, p53, BRCA1, basal phenotype (BP) and triple-negative phenotype (TN) were defined in this series as previously described [33-36].

This research was approved by Nottingham Research Ethics Committee 2 under the title of “Development of a molecular genetic classification of breast cancer”.

Transcript expression profiling

Total RNA was isolated from 128 frozen tumours retrieved from Nottingham Hospitals NHS Trust Tumour Bank between 1986 and 1992 as described [19, 37]. RNA integrity and genomic DNA contamination were analysed using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA). Total RNA was biotin-labelled using the Illumina TotalPrep RNA Amplification kit (Ambion) following manufacturers instructions. Biotin-labelled cRNA (1.5 µg) was used for each hybridisation on Sentrix Human-6 BeadChips (Illumina, San Diego, CA, USA) in accordance with manufacturer's protocol.[also described in [38].] Illumina gene expression data containing 47,293 features was processed and summarized in the Illumina BeadStudio software. Analyses of the probe level data were performed using the *beadarray* Bioconductor package [39] . The expression data is deposited in ArrayExpress at the EBI with accession number E-TABM-576 (<http://www.ebi.ac.uk/miamexpress/>).

Identification of novel genes associated with basal phenotype

Artificial Neural Networks (ANNs) were used to identify genes significantly associated with BP. The ANN modelling used a supervised learning approach, multi-layer perceptron architecture with a sigmoidal transfer function, where weights were updated by a back propagation algorithm [40, 41]. Learning rate and momentum were set at 0.1 and 0.5 respectively. The ANN architecture utilised five hidden nodes in the hidden layer and randomised initial weights. The output node was coded as 0 if a case was BP negative (n=93), and 1 if BP positive (n=35). Data consisted of 128 samples each with 47,293 corresponding variables specifying the Log₁₀ expression ratio of each transcript. Prior to ANN training, the data was randomly divided into three subsets; 60% for training, 20% for testing (to assess model performance during the training process) and 20% for validation (to independently test the model on data completely blind to the model). This process of random sample cross validation enabled the generation of confidence intervals for the predictions on a separate blind data set, and therefore avoided over-fitting of the data.

The intensity of each gene was used as an individual input in an ANN model, creating n individual models, where n was the number of transcripts on the array (47,293). These n models were then split into three subsets (described above) and trained. This random resampling and training process was repeated 50 times to generate predictions and associated error values for each sample with respect to the validation (blind) data. Inputs were ranked in ascending order based on predictive error and the gene that performed with the lowest error was selected for further training. Next, each of the remaining genes were sequentially added to the previous best gene, and were used in combination in a model, creating n-1 models each containing two genes as inputs. Training was repeated and performance evaluated. The model with the highest modelling performance was again selected and the process repeated creating n-2 models each

containing three inputs. This process was repeated until no significant gain was evident from the addition of further inputs. This resulted in a final model containing those transcripts that most accurately classified the patients according to BP. Receiver Operating Characteristic (ROC) curves was generated to provide statistics regarding model sensitivity and specificity.

Immunohistochemistry

Breast cancer tissue microarrays (TMA) and whole breast cancer tissue sections were prepared and immunohistochemistry performed employing the standard streptavidin-biotin complex method as previously described [32, 42]. Primary antibodies used were Fatty acid binding protein-7 (FABP7; 1:50; Abcam Ltd, Cambridge, UK) and Organic Anion Transporting Polypeptide 2 (OATP2; clone MDQ; 1:20, Abcam) for 45 minutes incubation. 3,3'-Diaminobenzidine tetrahydrochloride (Dako liquid DAB plus, K3468, Dako, Denmark) was used as a chromagen. The TMA sections were counterstained with Mayer's haematoxylin. Positive controls for FABP7 (brain tissue) and OATP2 (liver tissue) were used according to the supplier's data sheet.

Immunohistochemical Scoring

Immunoreactivity of OATP2 and FABP7 in the TMA cores was scored as percentage of cells stained. Only staining of the invasive malignant cells within the tissue cores was considered. TMAs were scored using high resolution digital images (NanoZoomer; Hamamatsu Photonics, Welwyn Garden City, UK), at x20 magnification, using a web-based interface (Distiller; Slidepath Ltd, Dublin, Ireland). All samples were scored by one observer (HZ) and a proportion of these counter-scored by a further observer (ARG) to ensure reproducibility.

Statistical Analysis

Statistical analysis was performed using SPSS v15.0 statistical software. Frequency histograms were examined to group the expression of the biomarkers into expression groups, which were used throughout the analyses. Chi-squared analyses were used for inter-relationships between steroid receptor co-regulators and for comparison with patho-clinical parameters, steroid receptors and biomarkers. The Kaplan-Meier method was used for the calculation of survival curves and disease free interval analyses. Cox proportional hazards model was used to test the statistical independence and significance of predictors on disease-free interval (DFI) and the breast cancer specific survival (BCSS). A p-value of <0.05 was considered significant.

Results

Novel transcripts associated with basal phenotype

In order to identify novel transcript associated with BP, analysis of the expression of 47,293 transcripts in 128 invasive breast carcinomas from gene microarray experiments [38] were screened using an Artificial Neural Network (ANN) predictive algorithm. This approach stratified the transcripts on their ability to classify the BP (as defined by CK5/6 and 14 positivity [43], n = 35) compared with non-basal tumours (n = 93). Figure 1 illustrates the top 21 transcripts according to BP status. Two of these transcripts, FABP7 and OATP2, showed significantly higher expression in tumours with BP compared with non-basal tumours and were therefore chosen for subsequent translational analysis in TMAs due to the availability of commercial antibodies suitable for immunohistochemistry.

Expression of OATP2 and FABP7 in breast tumours

Evaluation of OATP2 and FABP7 in whole breast cancer tissue sections (n=20) demonstrated that OATP2 immunoreactivity was localised to the cytoplasm of invasive cells, malignant *in situ* components where present, vascular endothelial cells and fibroblasts within the tumour stroma (Fig 2a, b). FABP7 immunoreactivity was present in normal adipocytes of the interlobular breast stroma delineating the plasma membrane of the cells and staining the nuclei strongly. Normal ductal/acinar epithelial cells also showed positive nuclear and cytoplasmic FABP7 immunoreactivity with some vascular endothelial cells of small vessels and capillaries also positive (Fig 2c). Both the cytoplasm and nuclei of invasive breast carcinoma cells showed positive granular immunoreactivity for FABP7 ranging from weak to strong with some cases showing no immunoreactivity (Fig 2d).

In the TMA breast tumours series, both proportion and intensity of FABP7 and OATP2 immunoreactivity varied in the cohort. The number of informative cases for FABP7 and OATP2 was 578 and 584 respectively. FABP7 and OATP2 expression were each categorised into two groups, based on the observations of the frequency histogram distributions (FABP7, negative/low cytoplasmic expression $\leq 60\%$, positive cytoplasmic expression $> 60\%$; OATP2, negative $\leq 15\%$, positive expression $> 15\%$). Using these cut-off points, a total of 90/578 (15.6%) cases showed strong FABP7 cytoplasmic immunoreactivity, whilst 46/584 (7.9%) cases showed positive cytoplasmic expression of OATP2. We also noted a small number of tumours ($n = 34$) showed nuclear immunoreactivity for FABP7 but no correlation was found between FABP7 nuclear expression and any clinico-pathological parameters (data not shown).

Association of OATP2 and FABP7 with clinico-pathological parameters

Table 1 summarises the associations between OATP2 and FABP7 expression and clinico-pathological variables. Positive OATP2 expression was associated with higher histological grade ($p=0.009$) and probable or definite vascular invasion ($p=0.035$). Strong FABP7 cytoplasmic expression showed a negative association with lymph node stage ($p=0.004$), a positive association with the good and moderate Nottingham Prognostic Index groups (NPI; $p=0.010$) and with histological tumour type (Table 2, $p=0.035$), where negative FABP7 staining was associated with classical lobular and lobular mixed tumours. OATP2 and FABP7 expression were not associated with either tumour size, age at diagnosis, or development of distant metastases.

Association of OATP2 and FABP7 with biomarkers

OATP2 and FABP7 expression was compared with the expression of previously determined biomarkers, summarised in Table 3. A highly significant inverse association between OATP2

and FABP7 was observed where 68/84 (81%) cases with negative/low expression for FABP7 were OATP2 positive ($p < 0.001$).

Positive OATP2 protein expression was also significantly associated with the positive expression of basal CKs [CK5/6 ($p = 0.001$), CK14 ($p < 0.001$) and CK17 ($p = 0.002$)], p53 ($p = 0.037$), p-cadherin ($p = 0.012$) and EGFR ($p = 0.020$). Positive OATP2 expression was further significantly associated with lack of ER ($p < 0.001$) and PgR ($p < 0.001$) expression but was not associated with HER2 ($p = 0.565$).

Strong FABP7 immunoreactivity also showed a significantly positive association with the expression of basal CKs [CK5/6 ($p = 0.025$), CK14 ($p = 0.001$) and CK17 ($p = 0.006$)], positive EGFR ($p = 0.035$), P-cadherin ($p = 0.002$) and luminal cytokeratin CK18 ($p = 0.026$) expression. Similar to OATP2, strong FABP7 immunoreactivity was correlated with negative expression of steroid receptors ER ($p = 0.002$), PgR ($p = 0.006$) but likewise was not associated with HER2 ($p = 0.647$). There was also no association between OATP2 or FABP7 expression with CK7/8, CK19, caveolin 1, caveolin 2, smooth muscle actin or MIB1 (data not shown).

There was also a highly significant association observed between OATP2 and FABP7 with BP either defined as CK5/6 and/or CK14 positivity ($p < 0.001$, Table 4) or with the inclusion of CK17 positivity ($p < 0.001$, data not shown). Within the basal-like tumours only, OATP2 expression was significantly associated with ER positive cases ($p = 0.010$) whilst FABP7 expression was not associated with ER status ($p = 0.657$; Table 4). However there was no association with either OATP2 or FABP7 with the triple negative phenotype (ER, PgR and HER2 negative; Table 4).

Association of OATP2 and FABP7 with patient outcome

Univariate survival analyses showed those tumours positive for OATP2 had a significantly poorer BCSS compared with OATP2 negative tumours ($p=0.042$, Figure 3A). Although positive expression of OATP2 was not significantly correlated with DFI over the full follow up period ($p=0.266$, Figure 3B), there is clearly a dramatic difference in the probability of recurrence during the first five years where positive OATP2 expression is associated with a shorter DFI ($p=0.014$). In contrast, strong immunoreactivity for FABP7 in breast cancer was associated with a trend for a longer BCSS, although this did not reach statistical significance ($p=0.052$, Figure 3C). Furthermore, FABP7 was significantly correlated with DFI where a longer DFI was observed for those patients with tumours showing strong immunoreactivity for FABP7 ($p=0.019$, Figure 3D).

Within the BP (defined as CK5/6 and CK14 positive), FABP7 expression significantly affected patient outcome. Basal-like tumours which showed strong expression for FABP7 had a significantly longer BCSS than FABP7 negative tumours showing BP ($p=0.006$, Figure 4A). Similarly, FABP7 immunoreactivity was associated with a significantly longer DFI in breast tumours showing BP ($p=0.002$, Figure 4B). There was no difference in BCSS or DFI in non-basal tumours with respect FABP7 expression. OATP2 expression was not associated with either BCSS or DFI within the basal and non-basal tumours (data not shown).

In multivariate Cox-proportional hazards analysis, FABP7 expression was a significant predictor of longer DFI ($p=0.033$) independent of tumour grade, size, lymph node stage, ER status, BP status (Table 5). There was also a strong trend for FABP7 expression to independently predict a longer BCSS ($p=0.052$)

Discussion

The basal-like type or BP of breast cancer has been subject to extensive investigation in recent years, mainly due to its association with poor patient survival [13, 15, 22, 23, 43, 44]. Despite this interest, there is currently no robust definition for designation of ‘basal-like’ or BP breast cancer in routine clinical practice, furthermore tumours with a BP are also now recognised to exhibit heterogeneity [14, 23]. In this study, the protein expression of two novel markers associated with basal phenotype, identified by a gene profiling analysis, OATP2 and FABP7, were investigated in a large series of breast tumours.

The expression of OATP2 was significantly associated with poor patient outcome. OATP2 has been found to be co-expressed at high levels in the placenta and the liver with breast cancer resistance protein (BCRP), an ATP-binding cassette (ABC) protein [45]. BCRP, when over-expressed *in vitro*, has the ability to induce high levels of resistance to anti-cancer chemotherapy drugs such as anthracyclines, mitoxantrone, and bisantrene [46]. It could be inferred therefore that over-expression of OATP2 in breast tumours may give the tumour some chemotherapy-resistant properties. Unfortunately, the number of patients used in this study was not sufficient to investigate whether or not OATP2 had an effect on chemotherapy response and therefore further studies determining the expression of the marker in all patients who received chemotherapy and/or hormone therapy should be conducted.

FABP7 has previously been linked to BP of breast tumours in gene expression studies [3], and the findings in this study confirm this association. However protein expression of FABP7 in human breast cancer has not previously been investigated. FABP7 belongs to a group of proteins called Lipid Binding Proteins (LBP) which are involved in the transportation and storage of fatty

acids for metabolism [47]. Studies on FABP7 expression in the brain have demonstrated that it is linked with poorer survival in patients with glioblastomas [48, 49], but the findings of this study showed that FABP7 positive breast tumours had longer BCSS and DFI which appears independent of other known prognostic factors in breast cancer. Various *in vitro* studies of the FABP7 gene have concluded that it acts as an inhibitor of breast cancer cell growth, with inhibition of growth being proportional to the expression of the gene, thereby explaining its alternative name, Mammary Derived Growth Inhibitor-Related gene (MRG) [50]. We have observed that FABP7 is preferentially expressed by breast tumours exhibiting a basal phenotype and significantly associated with better outcome in this form of breast cancer. This incongruity could suggest that FABP7 positive tumours behave less aggressively particularly as FABP7 is known to inhibit cell growth [50]. It has been demonstrated that EGFR activation up-regulates FABP7 and induces its nuclear translocation *in vitro*. Inhibition of FABP7 expression in turn inhibits EGF-induced migration in glial cells, suggesting that EGFR induced invasion properties is partly mediated by FABP7 [48]. The link between FABP7 and EGFR is of interest in breast carcinomas, especially the basal-like group, as EGFR has been identified by several publications to be one of the discriminating factors of this particular group [51-53].

In conclusion, the study has demonstrated the biological, pathological and clinical features of FABP7 and OATP2 protein expression in human breast carcinoma and with respect to FABP7 particularly its relationship and significance in BP breast cancer [12, 13, 15, 18, 22, 23, 25, 28, 43, 54]. These novel biomarkers have significant associations, both independently and within the BP, on tumour behaviour, expression of other biomarkers, and clinical outcome. Thus, this also further study confirms the clinical importance of the BP, and that screening for these basal markers may be relevant to integral to determining the outcome of individual patients with this

form of breast cancer. This study highlights an important niche in the literature where only a number of immune response markers in basal-like cancers have been previously identified that have an association with improved patient survival[55]. The identification of FABP7 through gene analyses and the analysis of its protein expression may suggest that novel markers may be able to differentiate between basal tumours associated with poor or favourable patient outcome. As a result, such markers need to be considered for incorporation into the definition of the basal phenotype, in the search for improved and robust ways to differentiate this group clinically.

Acknowledgements

This work was funded by the Breast Cancer Campaign.

References

1. Perou CM, Sorlie 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, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D (2000) Molecular portraits of human breast tumours. *Nature* 406:747-752.
2. Zhao H, Langerod A, Ji Y, Nowels KW, Nesland JM, Tibshirani R, Bukholm IK, Karesen R, Botstein D, Borresen-Dale AL, Jeffrey SS (2004) Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. *Mol Biol Cell* 15:2523-2536.
3. Sorlie 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, Eystein Lonning P, Borresen-Dale AL (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98:10869-10874.
4. Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, Botstein D (2003) Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A* 100:8418-8423.
5. Hu Z, Fan C, Oh DS, Marron JS, He X, Qaqish BF, Livasy C, Carey LA, Reynolds E, Dressler L, Nobel A, Parker J, Ewend MG, Sawyer LR, Wu J, Liu Y, Nanda R, Tretiakova M, Ruiz Orrico A, Dreher D, Palazzo JP, Perreard L, Nelson E, Mone M, Hansen H, Mullins M, Quackenbush JF, Ellis MJ, Olopade OI, Bernard PS, Perou CM (2006) The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* 7:96.

6. Sotiriou C, Neo SY, McShane LM, Korn EL, Long PM, Jazaeri A, Martiat P, Fox SB, Harris AL, Liu ET (2003) Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 100:10393-10398.
7. Murad TM (1971) A proposed histochemical and electron microscopic classification of human breast cancer according to cell of origin. *Cancer* 27:288-299.
8. Moll R, Franke WW, Schiller DL, Geiger B, Krepler R (1982) The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. *Cell* 31:11-24.
9. Nagle RB, Bocker W, Davis JR, Heid HW, Kaufmann M, Lucas DO, Jarasch ED (1986) Characterization of breast carcinomas by two monoclonal antibodies distinguishing myoepithelial from luminal epithelial cells. *J Histochem Cytochem* 34:869-881.
10. Altmannsberger M, Dirk T, Droese M, Weber K, Osborn M (1986) Keratin polypeptide distribution in benign and malignant breast tumors: subdivision of ductal carcinomas using monoclonal antibodies. *Virchows Arch B Cell Pathol Incl Mol Pathol* 51:265-275.
11. Dairkee SH, Ljung BM, Smith H, Hackett A (1987) Immunolocalization of a human basal epithelium specific keratin in benign and malignant breast disease. *Breast Cancer Res Treat* 10:11-20.
12. Rakha E, Putti, TC, Abd El-Rehim, DM, et al (2006) Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. *Journal of Pathology* 208:495-506.
13. van de Rijn M, Perou, CM, Tibshirani, R, et al (2002) Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome. *American Journal of Pathology* 161:1991-1996.
14. Rakha EA, Reis-Filho JS, Ellis IO (2008) Basal-like breast cancer: a critical review. *J Clin Oncol* 26:2568-2581.
15. Jones C, Nonni AV, Fulford L, Merrett S, Chaggar R, Eusebi V, Lakhani SR (2001) CGH analysis of ductal carcinoma of the breast with basaloid/myoepithelial cell differentiation. *Br J Cancer* 85:422-427.
16. Makretsov NA, Huntsman DG, Nielsen TO, Yorlida E, Peacock M, Cheang MC, Dunn SE, Hayes M, van de Rijn M, Bajdik C, Gilks CB (2004) Hierarchical clustering analysis of tissue microarray immunostaining data identifies prognostically significant groups of breast carcinoma. *Clin Cancer Res* 10:6143-6151.
17. C. M. Perou TS, M. B. Eisen, et al. (2000) Molecular portraits of human breast tumours. *Nature* 406:747-752.
18. Kreike B vKM, Horlings H, Weigelt B, Bartelink H, Van de Vijver MJ (2007) Gene expression profiling and histopathological characterization of triple negative/basal-like breast carcinomas. *Breast Cancer Research* 9:R65.
19. Chin SF, Teschendorff AE, Marioni JC, Wang Y, Barbosa-Morais NL, Thorne NP, Costa JL, Pinder SE, van de Wiel MA, Green AR, Ellis IO, Porter PL, Tavaré S, Brenton JD, Ylstra B, Caldas C (2007) High-resolution array-CGH and expression profiling identifies a novel genomic subtype of ER negative breast cancer. *Genome Biol* 8:R215.
20. Teschendorff AE, Caldas C (2008) A robust classifier of high predictive value to identify good prognosis patients in ER-negative breast cancer. *Breast Cancer Res* 10:R73.
21. Bauer KR BM, Cress RD et al (2007) Descriptive analysis of estrogen receptor (ER)-negative, progesterone receptor (PR)-negative, and HER2-negative invasive breast cancer, the so-called triple-negative phenotype: a population-based study from the California cancer Registry. *Cancer* 109:1721-1728.

22. Rakha EA E-SM, Green AR et al (2007) Prognostic markers in triple-negative breast cancer. *Cancer* 109:25-32.
23. Emad Rakha IEaJR-F (2007) Triple-negative and basal-like breast cancer: synonyms? *Clinical Cancer Research*.
24. Marc Tischkowitz J-SB, Louis R Bégin,, David G Huntsman MCC, Lars A Akslen,, Foulkes TONaWD (2007) Use of immunohistochemical markers can refine prognosis in triple negative breast cancer. *BMC Cancer* 7.
25. EA Rakha TP, DM Abd El-Rehim, C Paish, AR Green, DG Powe, AH Lee, JF Robertson, IO Ellis (2006) Morphological and immunophenotypic analysis of breast carcinomas with basal and myoepithelial differentiation. *The Journal of Pathology* 208:495-506.
26. Mervi Laakso NL, Åke Borg, Jorma Isola (2005) Cytokeratin 5/14-positive breast cancer: true basal phenotype confined to BRCA1 tumors. *Modern Pathology* 18:1321-1328.
27. Potemski P KR, Watala C, Pluciennik E, Bednarek AK, Kordek R (2005) Prognostic relevance of basal cytokeratin expression in operable breast cancer. *Oncology* 69:478-485.
28. Jumppanen M, Gruvberger-Saal, S., Kauraniemi, P., Tanner, M. et al. (2007) Basal-like phenotype is not associated with patient survival in estrogen-receptor-negative breast cancers. *Breast Cancer Res* 9:R16.
29. Ellis IO, Galea M, Broughton N, Locker A, Blamey RW, Elston CW (1992) Pathological prognostic factors in breast cancer. II. Histological type. Relationship with survival in a large study with long-term follow-up. *Histopathology* 20:479-489.
30. Elston CW, Ellis IO (1991) Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 19:403-410.
31. Galea MH, Blamey RW, Elston CE, Ellis IO (1992) The Nottingham Prognostic Index in primary breast cancer. *Breast Cancer Res Treat* 22:207-219.
32. Madjd Z, Pinder SE, Paish C, Ellis IO, Carmichael J, Durrant LG (2003) Loss of CD59 expression in breast tumours correlates with poor survival. *J Pathol* 200:633-639.
33. Abd El-Rehim DM, Ball G, Pinder SE, Rakha EA, Paish C, Robertson JFR, MacMillan D, Blamey RW, Ellis IO (2005) High throughput protein expression analysis using Tissue Microarray Technology of a large well characterised series identifies biologically distinct classes of breast cancer confirming recent cDNA expression analyses. *Int J Cancer* 116:340-350.
34. Abd El-Rehim DM, Pinder SE, Paish CE, Bell JA, Rampaul RS, Blamey RW, Robertson JF, Nicholson RI, Ellis IO (2004) Expression and co-expression of the members of the epidermal growth factor receptor (EGFR) family in invasive breast carcinoma. *Br J Cancer* 91:1532-1542.
35. Rakha EA, El-Rehim DA, Paish C, Green AR, Lee AH, Robertson JF, Blamey RW, Macmillan D, Ellis IO (2006) Basal phenotype identifies a poor prognostic subgroup of breast cancer of clinical importance. *Eur J Cancer* 42:3149-3156.
36. Rakha EA, El-Sayed ME, Green AR, Lee AH, Robertson JF, Ellis IO (2007) Prognostic markers in triple-negative breast cancer. *Cancer* 109:25-32.
37. Naderi A, Teschendorff AE, Barbosa-Morais NL, Pinder SE, Green AR, Powe DG, Robertson JF, Aparicio S, Ellis IO, Brenton JD, Caldas C (2006) A gene-expression signature to predict survival in breast cancer across independent data sets. *Oncogene*.

38. Git A, Spiteri I, Blenkiron C, Dunning MJ, Pole JC, Chin SF, Wang Y, Smith J, Livesey FJ, Caldas C (2008) PMC42, a breast progenitor cancer cell line, has normal-like mRNA and microRNA transcriptomes. *Breast Cancer Res* 10:R54.
39. Dunning MJ, Smith ML, Ritchie ME, Tavare S (2007) beadarray: R classes and methods for Illumina bead-based data. *Bioinformatics* 23:2183-2184.
40. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. *Nature* 323:533-536.
41. Lancashire LJ, Powe DG, Reis-Filho JS, Rakha E, Lemetre C, Weigelt B, Abdel-Fatah TM, Green AR, Mukta R, Blamey R, Paish EC, Rees RC, Ellis IO, Ball GR (2009) A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. *Breast Cancer Res Treat*.
42. Green AR, Burney C, Granger CJ, Paish EC, El-Sheikh S, Rakha EA, Powe DG, Macmillan RD, Ellis IO, Stylianou E (2007) The prognostic significance of steroid receptor co-regulators in breast cancer: co-repressor NCOR2/SMRT is an independent indicator of poor outcome. *Breast Cancer Res Treat*.
43. Rakha E, El-Rehim, DA, Paish, C, et al (2006) Basal phenotype identifies a poor prognostic subgroup of breast cancer of clinical importance. *European Journal of Cancer* 42:3149-3156.
44. T. Sorlie CMP, R. Tibshirani, T. Aas, S. Geisler, H. Johnsen, T. Hastie, M. B. Eisen, M. van de Rijn, S. S. Jeffrey et al (2001) Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 98:10869-10874.
45. Grube M RS, Meyer zu Schwabedissen H, Köck K, Draber K, et al. (2007) Organic Anion Transporting Polypeptide 2B1 and Breast Cancer Resistance Protein Interact in the Transepithelial Transport of Steroid Sulfates in Human Placenta Drug Metabolism and Disposition 35:30-34.
46. Ejendal KF HC (2002) Multidrug resistance and cancer: the role of the human ABC transporter ABCG2. *Curr Protein Pept Sci* 3:503-511.
47. Chmurzyńska A (2006) The multigene family of fatty acid-binding proteins (FABPs): function, structure and polymorphism. *Journal of Applied Genetics* 47:39-48.
48. Kaloshi G MK, Carpentier C, Taillibert S et al. (2007) FABP7 expression in glioblastomas: relation to prognosis, invasion and EGFR status. *J Neurooncology* 84:245-248.
49. Liang Y BA, Nicholas MK, Gupta N (2005) Id4 and FABP7 are preferentially expressed in cells with astrocytic features in oligodendrogliomas and oligoastrocytomas. *BMC Clinical Pathology* 5.
50. Wang M LY, Ni J, Aygun B, Goldberg ID, Shi EY (2000) Induction of Mammary Differentiation by Mammary-derived Growth Inhibitor-related Gene That Interacts with an -3 Fatty Acid on Growth Inhibition of Breast Cancer Cells. *Cancer Research* 60:6482-6487.
51. Nielsen TO JK, et al. (2004) Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma. *Clin Cancer Res* 10:5367-5374.
52. Kim MJ AS, et al. (2006) Clinicopathologic significance of the basal-like subtype of breast cancer: a comparison with hormone receptor and HER2/neu-overexpressing phenotypes. *Hum Pathol* 37:117-126.
53. Rodriguez-Pinilla SM HE, et al. (006) Prognostic significance of basal-like phenotype and fascin expression in node-negative invasive breast carcinomas. *Clin Cancer Res* 12:1533-1539.

54. Kreike B, van de Vijver, M. (2007) Are triple-negative tumours and basal-like breast cancer synonymous? Authors' response.
55. Teschendorff AE, Miremadi A, Pinder SE, Ellis IO, Caldas C (2007) An immune response gene expression module identifies a good prognosis subtype in estrogen receptor negative breast cancer. *Genome Biol* 8:R157.

Table 1 OATP2 and FABP7 expression and clinico-pathological parameters

	OATP2			FABP7		
	Negative	Positive	p value	Negative	Positive	p value
Grade						
1	107	4	0.009	94	13	0.180
2	172	9		160	25	
3	253	32		228	52	
Total	532	45		482	90	
LN Stage						
1	342	33	0.358	302	72	0.004
2	133	7		124	15	
3	57	5		56	3	
Total	532	45		482	90	
Tumour size						
≤1.5cm	180	16	0.466	168	314	0.338
>1.5cm	352	29		34	56	
Total	532	45		202	370	
Distant metastases						
No	370	26	0.183	311	67	0.147
Definite	168	20		170	22	
Total	538	46		481	89	
Nottingham Prognostic Index						
Good	158	9	0.489	140	24	0.010
Moderate	295	30		264	62	
Poor	77	6		76	4	
Total	530	45		480	90	
Vascular Invasion						
Negative	306	30	0.035	285	63	0.130
Probable	58	2		44	8	
Definite	166	11		151	17	
Total	530	43		480	88	

Table 2 OATP2 and FABP7 expression and histological tumour type

	OATP2*		FABP7†	
	Negative	Positive	Negative	Positive
Invasive Ductal/No Special Type	292	25	271	47
Tubular Mixed	110	8	93	19
Atypical Medullary	14	4	9	8
Typical Medullary	3	0	3	0
Classical Lobular	30	2	35	2
Solid Lobular	2	0	2	0
Tubulo-lobular	3	0	3	0
Lobular Mixed	19	1	18	3
Tubular	13	2	10	2
Mucinous	0	0	0	1
Invasive Papillary	3	1	3	1
Mixed NST And Lobular	12	0	11	2
Mixed NST And A Special Type	15	0	12	2
Other rare types	2	0	2	0
Total	518	43	472	87

* $p=0.505$, † $p=0.035$

Table 3 The relationship between OATP2 and FABP7 and biomarker expression

		OATP2			FABP7		
		Negative	Positive	p value	Negative	Positive	p value
ER	Negative	145	29	<0.001	137	36	0.002
	Positive	338	13		310	38	
	Total	483	42		447	74	
PgR	Negative	199	33	<0.001	183	45	0.006
	Positive	284	9		259	33	
	Total	483	42		442	78	
HER2	Negative	423	35	0.565	390	69	0.647
	Positive	22	2		17	3	
	Total	445	37		407	72	
EGFR	Negative	315	21	0.020	299	41	0.035
	Positive	74	12		70	18	
	Total	389	33		369	59	
CK18	Negative	55	11	0.009	49	15	0.026
	Positive	384	27		362	54	
	Total	439	38		411	69	
p53	Negative	389	26	0.037	349	58	0.119
	Positive	123	16		116	27	
	Total	512	42		465	85	
P-cadherin	Negative	146	5	0.012	139	10	0.002
	Positive	224	24		208	43	
	Total	370	29		347	53	
CK5/6	Negative	377	26	0.001	351	53	0.025
	Positive	72	15		65	19	
	Total	449	41		416	72	
CK14	Negative	398	25	<0.001	363	53	0.001
	Positive	40	14		36	15	
	Total	438	39		399	68	
CK17	Negative	453	34	0.002	406	70	0.006
	Positive	36	10		32	14	
	Total	489	44		438	84	
FABP7	Negative	394	24	<0.001			
	Positive	68	16				
	Total	462	40				

		OATP2			FABP7		
		Negative	Positive	p value	Negative	Positive	p value
Basal Phenotype (CK5/6 and/or CK14)	Non-Basal	374	20	<0.001	353	40	<0.001
	Basal-like	138	23		112	43	
	Total	512	43		465	83	
Triple Negative (TN)	Non-TN	403	33	0.827	354	71	0.502
	TN	81	6		75	12	
	Total	484	39		429	83	
Basal-like tumours	ER positive	68	19	0.010	65	24	0.657
	ER negative	58	4		42	13	
	Total	126	23		107	37	

Table 4 The relationship between OATP2 and FABP7 and basal/triple negative phenotypes

Table 5 Cox proportional hazards analysis for predictors of breast cancer specific survival (BCSS) and disease-free interval (model including FABP7 expression)

Variable	BCSS			Disease Free Interval		
	Hazard ratio	95% CI	<i>p</i> -value	Hazard ratio	95% CI	<i>p</i> -value
Grade ^a	1.582	1.211-2.066	0.001	1.145	0.920-1.424	0.225
Size >1.5cm ^b	1.377	0.933-2.032	0.107	1.350	0.977-1.864	0.069
LN stage ^c	1.678	1.358-2.072	<0.001	1.765	1.450-2.149	<0.001
Basal phenotype status (+ve)	2.306	1.578-3.369	<0.001	2.268	1.612-3.192	<0.001
ER status (+ve)	1.358	0.922-2.000	0.121	0.888	0.632-1.249	0.496
FABP7 status (+ve)	0.578	0.333-1.004	0.052	0.612	0.390-0.960	0.033

^afitted as linear term, ie. Increase in risk for change in grade of one unit

^bcompared with tumour size ≤ 1.5 cm

^cfitted as linear term, ie. Increase in risk for change in lymph node stage of one unit.

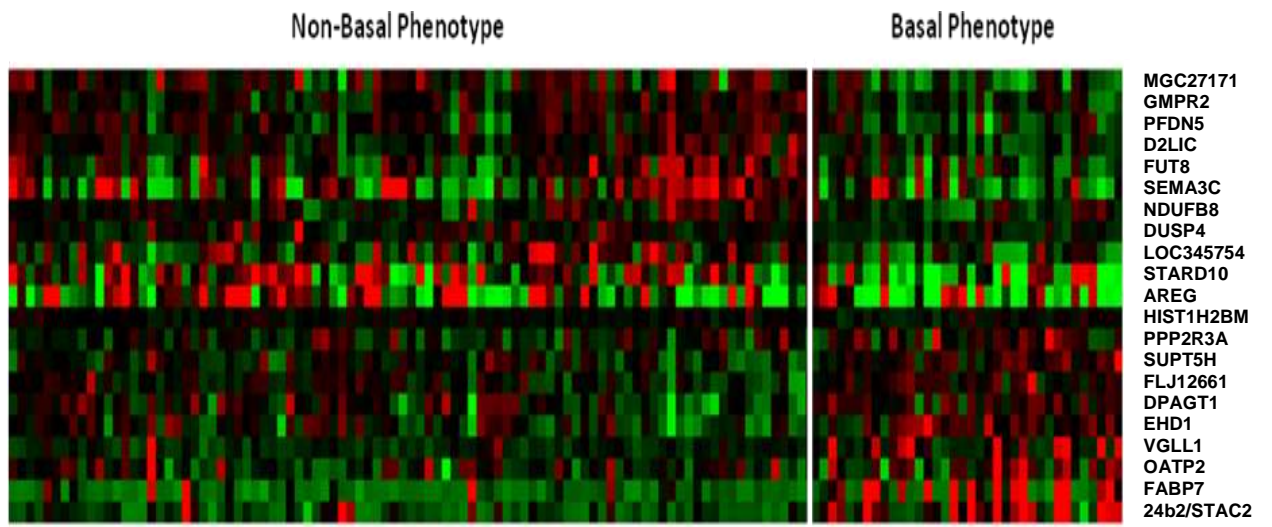


Fig 1

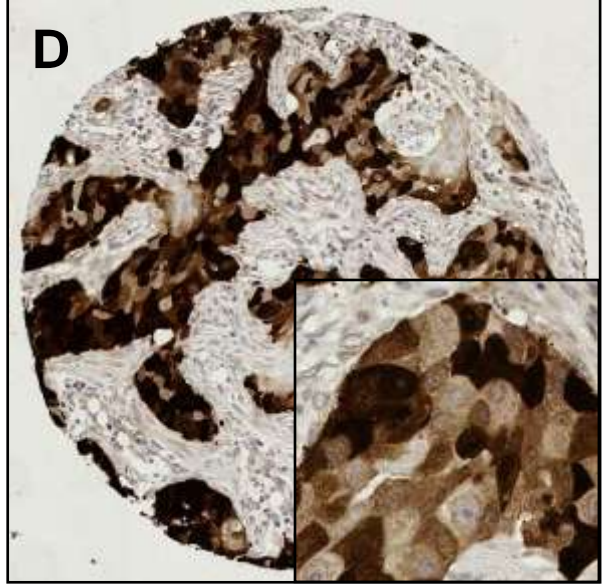
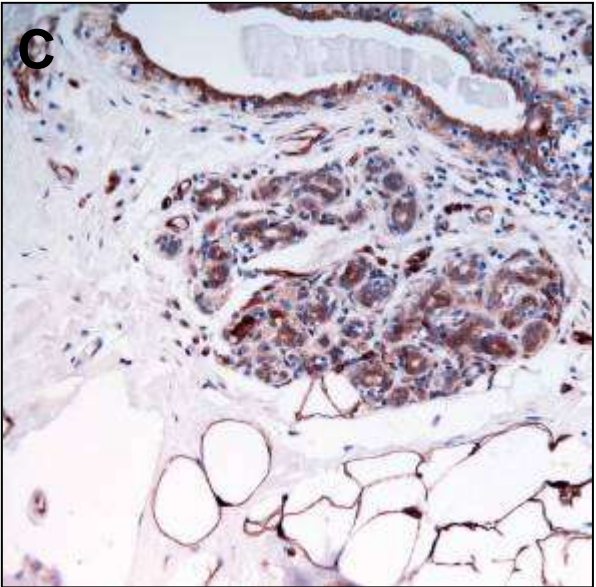
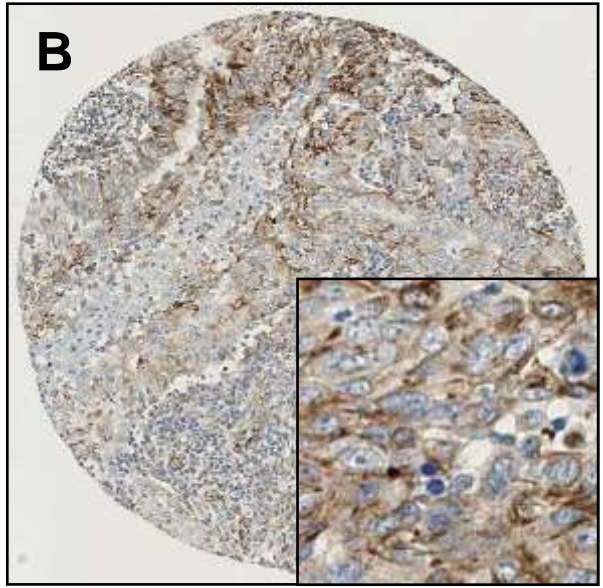
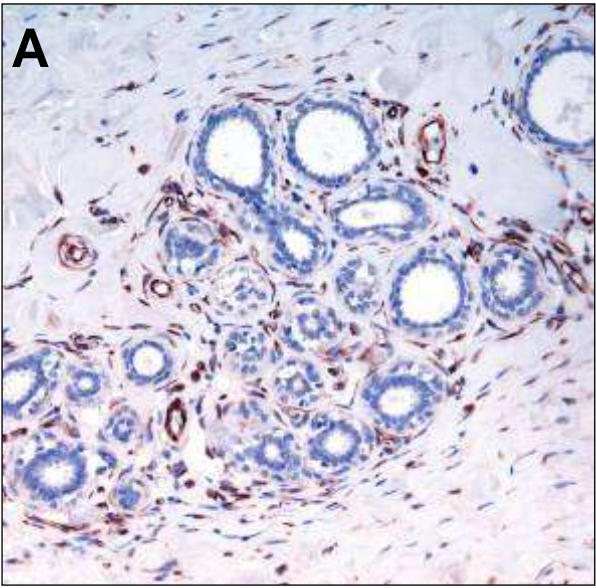
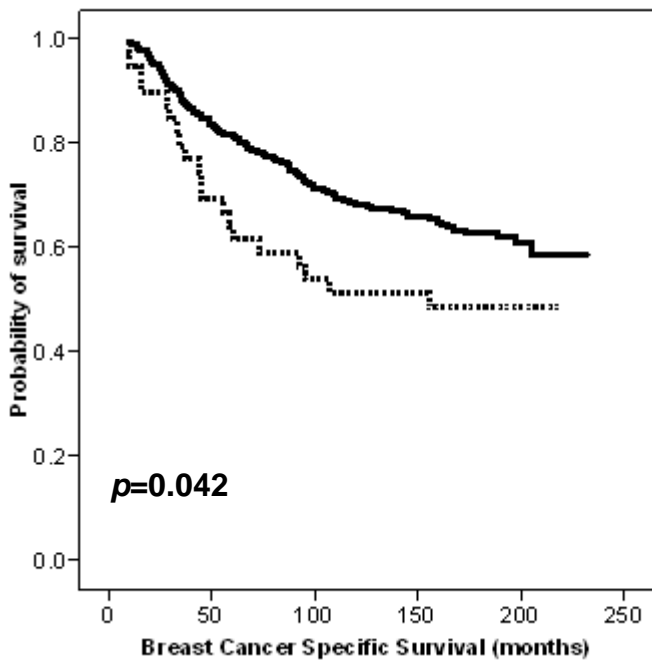
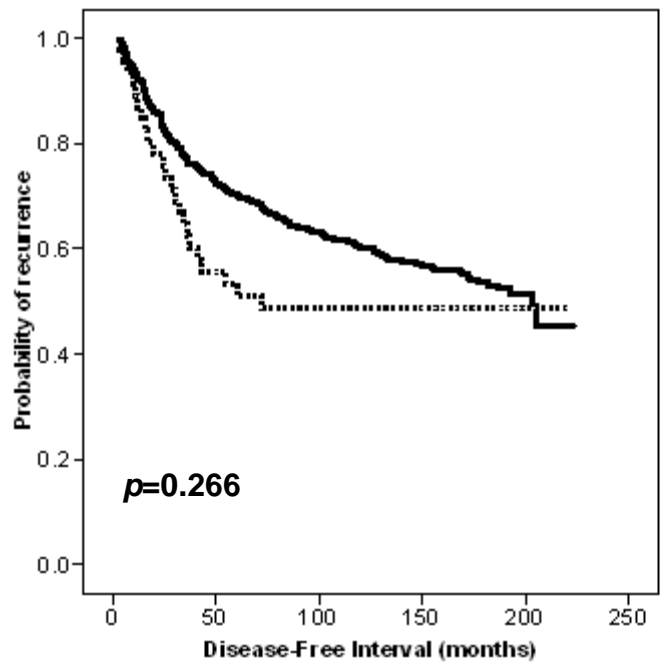
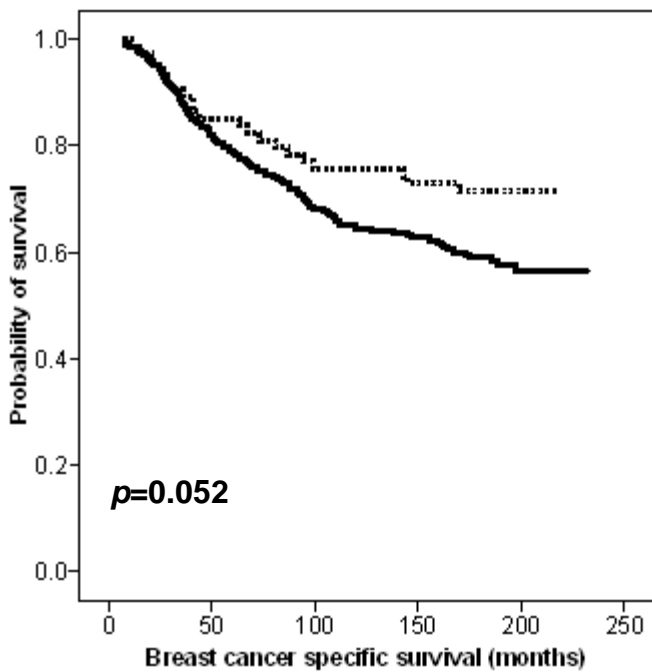
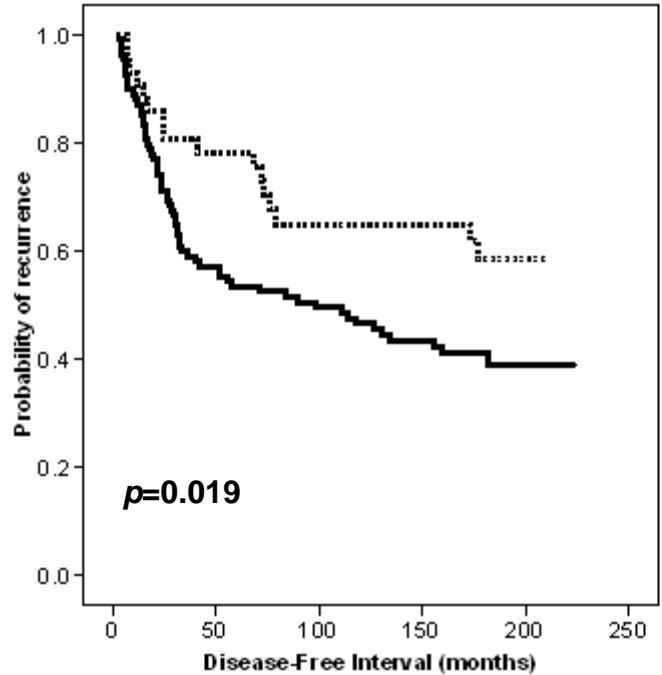


Fig 2

A**B**

OATP 2
 — Negative
 ··· Positive

C**D**

FABP 7
 — Negative
 ··· Positive

Fig 3

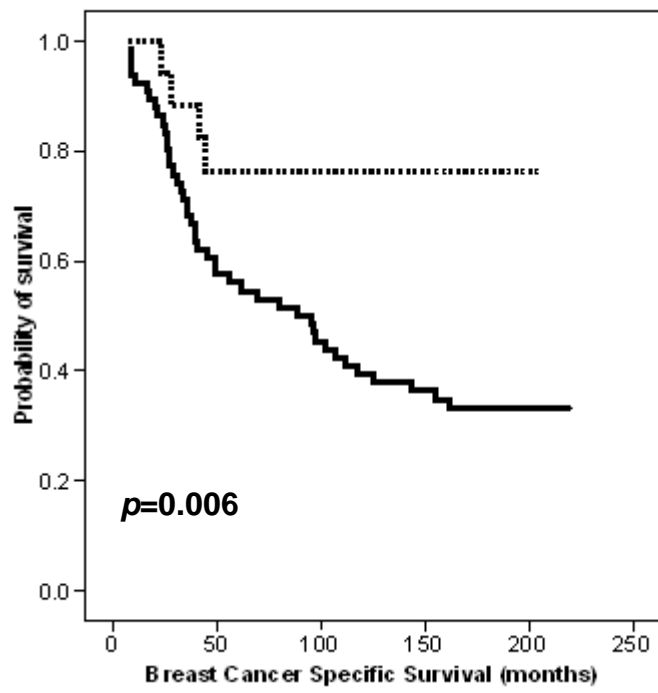
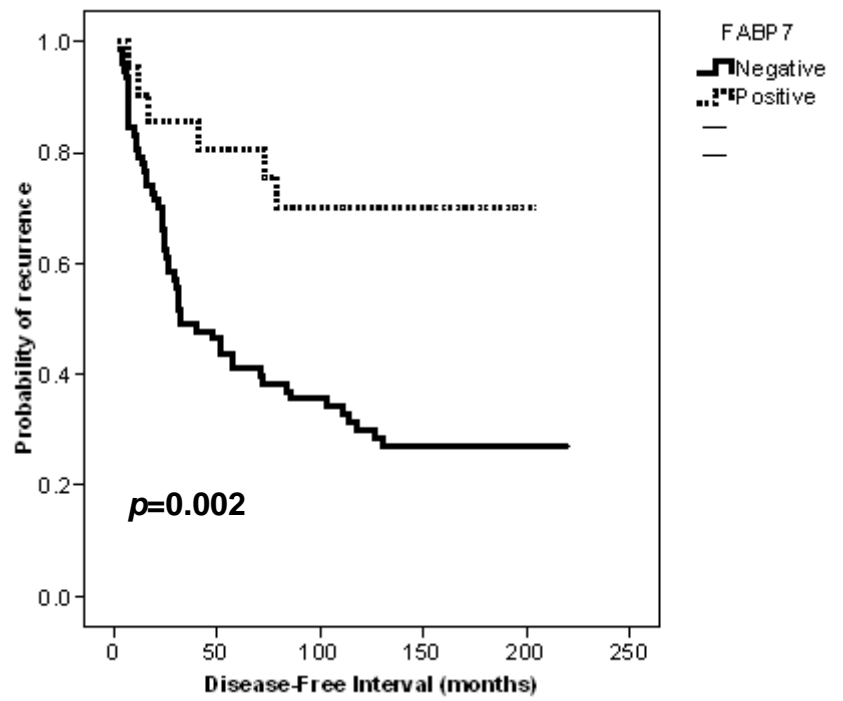
A**B**

Fig 4

Figure legends

Fig 1. Heatmap colours represent relative RNA expression levels (green=low expression, red=high expression).

Fig 2. Representative photographs of OATP2 and FABP7 immunoreactivity in breast determined using immunohistochemistry. OATP2 expression in A) normal breast, B) breast cancer; FABP7 expression in C) normal breast, B) breast cancer. All 20x magnification; inset 40x magnification

Fig 3. OATP2 and FABP7 expression in relation to A and C) breast cancer specific survival and B and D) disease free interval (DFI) in breast tumours.

Fig 4. FABP7 expression in relation to A) breast cancer specific survival and B) disease free interval (DFI) in breast tumours showing basal phenotype