

RESEARCH ARTICLE

Open Access



The effects of lymph node status on predicting outcome in ER+ /HER2- tamoxifen treated breast cancer patients using gene signatures

Jessica G. Cockburn¹, Robin M. Hallett², Amy E. Gillgrass¹, Kay N. Dias¹, T. Whelan¹, M. N. Levine¹, John A. Hassell² and Anita Bane^{1,3*}

Abstract

Background: Lymph node (LN) status is the most important prognostic variable used to guide ER positive (+) breast cancer treatment. While a positive nodal status is traditionally associated with a poor prognosis, a subset of these patients respond well to treatment and achieve long-term survival. Several gene signatures have been established as a means of predicting outcome of breast cancer patients, but the development and indication for use of these assays varies. Here we compare the capacity of two approved gene signatures and a third novel signature to predict outcome in distinct LN negative (-) and LN+ populations. We also examine biological differences between tumours associated with LN- and LN+ disease.

Methods: Gene expression data from publically available data sets was used to compare the ability of Oncotype DX and Prosigna to predict Distant Metastasis Free Survival (DMFS) using an *in silico* platform. A novel gene signature (Ellen) was developed by including patients with both LN- and LN+ disease and using Prediction Analysis of Microarrays (PAM) software. Gene Set Enrichment Analysis (GSEA) was used to determine biological pathways associated with patient outcome in both LN- and LN+ tumors.

Results: The Oncotype DX gene signature, which only used LN- patients during development, significantly predicted outcome in LN- patients, but not LN+ patients. The Prosigna gene signature, which included both LN- and LN+ patients during development, predicted outcome in both LN- and LN+ patient groups. Ellen was also able to predict outcome in both LN- and LN+ patient groups. GSEA suggested that epigenetic modification may be related to poor outcome in LN- disease, whereas immune response may be related to good outcome in LN+ disease.

Conclusions: We demonstrate the importance of incorporating lymph node status during the development of prognostic gene signatures. Ellen may be a useful tool to predict outcome of patients regardless of lymph node status, or for those with unknown lymph node status. Finally we present candidate biological processes, unique to LN- and LN+ disease, that may indicate risk of relapse.

Keywords: Breast cancer, Lymph node status, Gene signature, Estrogen receptor, Prognosis, Oncotype DX, Prosigna

* Correspondence: bane@hhsc.ca

¹Department of Oncology, Juravinski Hospital and Cancer Centre, Hamilton, Canada

³Department of Pathology, Juravinski Hospital and Cancer Centre, Hamilton, Canada

Full list of author information is available at the end of the article



Background

Axillary lymph node (LN) status is the most important prognostic variable in the management of patients with primary estrogen receptor positive (ER+) breast cancer, which accounts for the majority of diagnosed cases. Node positive breast cancer patients have been shown to have a worse prognosis than those with node negative disease. These observations have led, in part, to the development of a Tumour Nodal Metastases (TNM) staging system that incorporates tumour size, nodal involvement, including the absolute number of involved nodes, and the presence or absence of systemic metastases into an incremental staging system [1, 2]. Each stage of disease has specific survival characteristics and is thought to represent the natural progression of a tumour, from its origins in the breast to its metastasis through the lymphatic system to regional lymph nodes and ultimately through the circulatory system to distant sites. Clinicians use the TNM staging system to guide the management of breast cancer patients. Most breast cancer patients with involved axillary lymph nodes, in the absence of significant co-morbidities, are currently offered adjuvant systemic chemotherapy [3, 4].

However, the biological significance of nodal metastases is poorly understood. It is hypothesised that involvement of axillary lymph nodes is an indicator of tumour chronology such that the longer a tumour has been growing in the breast the more likely it is to metastasize to regional axillary nodes. Furthermore, it is thought that breast cancers first metastasize to these nodes and then secondarily to other sites [5, 6]. In support of this hypothesis, there is an established correlation between larger tumour size and lymph node involvement; indeed more timely intervention and resection of smaller primary tumours is associated with a reduced incidence of spread to regional lymph nodes [7]. More importantly, the absence of lymph node involvement is significantly associated with a better prognosis.

An alternative hypothesis suggests that some metastatic tumours avoid the lymphatic system, and instead spread primarily through the circulatory system [8, 9]. The evidence for this theory stems from the knowledge that 30 % of patients who are lymph node negative (LN-) at diagnosis will eventually succumb to metastatic breast disease, even after optimal treatment [10]. Conversely, there is a subset of patients who present with lymph node positive (LN+) disease that never develop distant recurrence, even in the absence of adjuvant treatment [9, 11]. It is likely that the biology of a primary tumour at diagnosis contributes to whether it remains at the primary site, spreads to regional lymph nodes, or metastasizes to distant sites via lymph node spread or through the vascular circulation. It is increasingly recognised that clinical pathological factors alone

are limited in their ability to predict who will develop recurrent cancer or respond to treatment. To this end, a number of genomic signatures have been developed which have shown to be both prognostic (predict risk of distant recurrence) and predictive (predict response to chemotherapy) [12, 13]. It is thought that these signatures detect biological differences in primary tumours indicative of whether a tumour is likely to metastasize.

Here, we explore the relationship between stage and tumour biology to outcome in ER+ breast cancer, in the context of prognostic gene signatures, namely Oncotype DX and Prosigna [14–17]. Specifically, we compared the capacity of Oncotype DX, developed exclusively on and for LN negative (LN-) ER+ patients [17], and Prosigna, developed on all clinical subtypes of breast cancer including those with and without lymph involvement [18], for their capacity to predict outcome in patients with ER+/LN- and ER+/LN+ tumours. Furthermore, we examine the biological pathways represented in patient tumours with and without LN involvement that have good survival versus those that have developed systemic metastases. Finally, using this knowledge, a novel prognostic gene signature, called ‘Ellen’ was developed *in silico* for both LN+ and LN- ER+ breast cancer.

Methods

Patients and samples

All data was publicly available and downloaded from the Gene Expression Omnibus (GEO), NCBI [19] (<http://ncbi.nlm.nih.gov/geo>). Three independent experimental cohorts, GSE17705 [20] and GSE6532 [21] (which comprises 2 separate cohorts), were used for discovery and training and are briefly described in Table 1. Patients in all three cohorts were known to have ER+ tumours, were treated with surgical excision of the primary tumour and axillary dissection followed by 5 years of adjuvant tamoxifen. Limited pathological information is available for each sample, but ER and LN status is provided. The development of distant metastases was recorded over 10-years of clinical follow-up and reported as distant metastases free survival (DMFS). DMFS rates for LN- and LN+ patient subgroups were also reported. Patients with HER2 positive tumours were removed from all cohorts, as HER2 is known to be a poor prognostic variable for both LN+ and LN- tumours. Furthermore, in clinical practice patients with HER2+ ER+ tumours of 1 cm or more commonly receive adjuvant chemotherapy and Herceptin. A tumour was considered HER2 positive if either of the two *HER2* probes on the Affymetrix chip were overexpressed as calculated using previously published methods [22].

GSE17705 was used as a training cohort for feature discovery in the generation of the Ellen signature and comprises Affymetrix U133A chip microarray expression

Table 1 Summary of GEO cohort characteristics

	GSE17705	GSE6532-C
<i>n</i>	230	132
LN Positive	91 59 % DMFS at 10 years	89 70 % DMFS at 10 years
LN Negative	139 85 % DMFS at 10 years	43 75 % DMFS at 10 years
Age	NR	61
Age Range	NR	40–88
Distant Metastasis at 10 years (<i>n</i>)	62	33
Overall 10 year DMFS	58.7 %	71.1 %
Location	MD Anderson, US	Guy's Hospital, UK, John Radcliffe Hospital, UK & Uppsala University Hospital, Sweden
Submission Group	Hatzis, Nuvera Biosciences, Woburn Mass	Loi et al., Institut Jules Bordet, Belgium
Affymetrix Chip	U133.A	U133.A & U133.2
NR-Not Reported		

data from 230 ER+/HER2- primary breast cancers, ~40 % of which were LN+. Two additional independent cohorts, GSE6532-A and GSE6532-2, were combined (GSE6532-C) and used to examine the Oncotype DX and Prosigna assays, and to validate the Ellen signature derived from the training cohort. The GSE6532-C cohort contained Affymetrix U133A and U133 Plus 2.0 microarray expression data from 132 ER+/HER2- primary tumours, ~67 % of the patients were lymph node positive. Specific demographic information for GSE17705 and GSE6532 can be found on the GEO website and in previously published reports [19, 20].

Data preparation

To extract the data from these cohorts, the raw intensity files (.CEL) comprising each dataset were downloaded and normalized using the Robust Multichip Algorithm (RMA) [23, 24] to generate a single intensity value for each probeset, using GenePattern (Broad Institute, Cambridge, Massachusetts). This preprocessing method has also been shown to yield concordance with qRT-PCR values and has been used in similar studies [24, 25]. Intensity was standardized using a Z score, where probe intensity was averaged among all samples and subtracted from the probe intensity from a single sample, which was then divided by the standard deviation of the probe intensities. Several other peer reviewed articles refer to a similar method to mimic qRT-PCR based assays using microarray gene expression data [25].

Oncotype DX analysis

To simulate the Oncotype DX assay, only probesets corresponding to the prognostic genes comprising the Oncotype DX gene list were selected. The Oncotype DX recurrence score (RS) is calculated by taking a modified weighted average for each functionally distinct group of genes, which were then combined [17]. The use of *ACTB*, *GAPDH*, and *TFRC* transcripts was excluded as data had been initially normalized using RMA. It is important to note that the range of recurrence scores differs between qRT-PCR (quantitative Real Time-Polymerase Chain Reaction) (RS are greater than 0) and expression microarray platforms (RS normally distributed around zero), as qRT-PCR data distribution is cumulative and microarray data is continuous.

Prosigna analysis

To simulate the Prosigna assay, expression values from only the available (*n* = 45) Affymetrix probe sets corresponding to the 50 Prosigna genes were used. Six genes (*ANLN*, *CDCA1*, *CXXC5*, *FOXCl*, *TMEM45B*, *UBE2T*) from the Prosigna assay, representing both pro- and anti-tumour functions were excluded from the analysis because probesets representing these genes were not represented on the Affymetrix chips. Standardized expression microarray values were used, in place of Nanostring nCounter expression data. The risk of recurrence (ROR) score was calculated using the Spearman correlation of prognostic gene expression to predetermined coefficients relating to the expected expression of each gene based on the intrinsic molecular subtypes as described [18].

Signature performance

Cox Proportional Hazards Regression analysis was used to determine the non-parametric association of continuous signature scores to patient outcome over time. The Cox PH package in R (R Foundation for Statistical Computing, Vienna, Austria) was used to calculate Concordance (C), hazard ratio (HR), p values, and confidence intervals (CI) for each signature. Analysis of signatures was simultaneously performed using all eligible tumours irrespective of patient outcome. Signature performance was compared using statistical variables alone and in the absence of prior knowledge to signature performance in the test cohort. Significant differences between outcome groups were determined by statistical alpha values being less than or equal to 0.05 for each test or the CI range excluding 1, as appropriate. Kaplan-Meier survival curves were generated using the median cut-point for each signature scores to visually represent outcome of patients at high versus low risk of distant metastasis.

Gene set enrichment analysis

Gene set enrichment analysis (GSEA) from Gene Pattern (Broad Institute, Cambridge, Massachusetts), was used to evaluate the biological mechanisms represented by sets of genes associated with distant metastasis free survival (DMFS) in patients with ER+ breast cancer, as previously described [26, 27]. Briefly, LN- and LN+ patient groups were classed by outcome (presence or absence of metastases) and associated Affymetrix data was used to enrich for gene sets. The GSEA algorithm ranks all genes by expression level in either class of samples. It then compares the pattern and frequency of gene expression in each class to previously published gene lists using an iterative approach to find the most related gene sets. An enrichment score (ES) is calculated for each gene set in each cohort, which can then be extrapolated to biological significance. Reported functions of individual genes are from the Gene Ontology Consortium (Release date April 2016, <http://geneontology.org>) [28].

Development and validation of the Ellen signature

Identification of prognostic genes

Prediction Analysis of Microarrays (PAM) [29] was used for feature selection and 10-fold cross-validation was used to estimate the optimal number of features (genes) to comprise the gene signature. DMFS was used as the clinical end-point.

Validation of gene signature

To calculate a final prognostic index, gene Z scores were averaged by outcome association and then subtracted such that the average of poor outcome probesets was subtracted from the average of good outcome probesets, resulting in positive correlation to DMFS. Again, 10 year DMFS was used as the clinical endpoint and Cox PH Regression, C, and HRs were used to evaluate signature performance.

Results

In silico validation

We independently verified the ability of Oncotype DX to predict recurrence in LN- patients in the training cohort using microarray expression data to ensure the validity of our *in silico* strategy ($p < 1.2 \times 10^{-2}$, HR: 3.58) (Table 2). Similar *in silico* approaches have previously been used to replicate gene signatures, including Oncotype DX and Prosigna [30–32].

Table 2 Oncotype DX validation on GSE17705

	HR	p value	CI upper	CI lower
Combined Patients	1.74	5.5E-02	0.99	3.07
LN- Patients	3.58	1.2E-02	1.38	9.27
LN+ Patients	1.16	6.8E-01	0.57	2.34

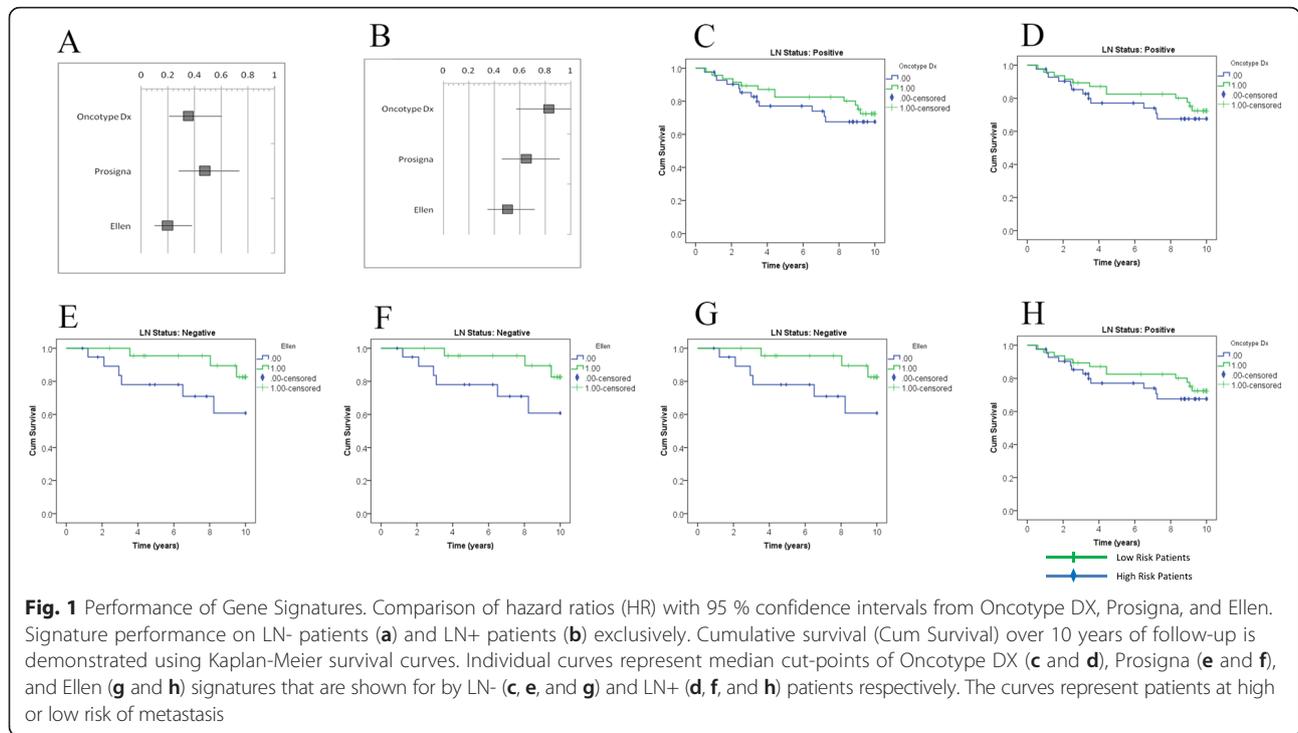
Signature comparison

We examined the performance of the Oncotype DX and Prosigna gene signatures on transcript profiles of breast cancer patients with either LN- or LN+ disease. To do so, the Oncotype DX algorithm was replicated *in silico* using Affymetrix gene expression data as described above. We subsequently tested the prognostic ability of the simulated algorithm on ER+ tumours from LN+ and LN- patients. As expected, the simulated Oncotype DX algorithm was able to significantly predict outcome for ER+ LN- patients ($p < 1.26 \times 10^{-4}$, HR: 0.36, C:0.78) (Fig. 1 and Table 3) which confirms its prognostic capacity in this group of patients. We also used the modified Oncotype DX algorithm, to predict outcome of ER+ LN+ patients. Oncotype DX was unable to predict risk of recurrence for ER+ LN+ patients from GSE6532-C ($p > 0.30$) (Fig. 1 & Table 3).

We subsequently simulated the Prosigna gene assay *in silico* using Affymetrix gene expression data, as described in the methods. As expected, the simulated Prosigna signature was able to significantly predict outcome for ER+ LN- patients ($p < 8.07 \times 10^{-4}$, HR: 0.48, C:0.79) (Fig. 1 and Table 3), as well as in ER+ LN+ patients ($p < 1.34 \times 10^{-2}$, HR: 0.65, C: 0.62) (Fig. 1 and Table 3).

We then developed an independent signature, known as “Ellen”, using both LN- and LN+ patients from the training cohort, and demonstrated that it was able to more significantly predict outcome of LN- and LN+ cohorts than either the Oncotype DX or Prosigna gene signatures. For LN- patients, Ellen scores were associated with the ability to predict risk of relapse with a concordance of 0.85 and hazard ratio of 0.20 ($p < 1.27 \times 10^{-6}$) (Fig. 1 and Table 3). Similarly, for LN+ patients Ellen score was able to predict risk of distant metastasis with a concordance of 0.71 and hazard ratio of 0.50 ($p < 1.74 \times 10^{-4}$).

The Ellen gene signature comprises 57 genes; expression of 33 of these genes is associated with a low risk of distant metastasis whereas expression of 24 is associated with high risk (Table 4). The biological processes of the genes present in all three signatures (Ellen, Oncotype DX and Prosigna) were functionally annotated using the Gene Ontology Consortium (Fig. 2 and Table 4). All three signatures included genes with functions related to gene expression, proliferation, immune response, cell migration, cell cycle, and post translational modification (PTM) and trafficking. Ellen and Prosigna each contained genes that represented unique biological processes; namely epigenetic and angiogenic processes for Ellen and DNA repair and replication processes for Prosigna (Table 5). Direct comparison of gene lists showed that there are 11 overlapping genes between Oncotype DX and Prosigna (BAG1, BCL2, BIRC5, CCNB1, ERBB2, ESR1, GRB7, MKI67, MMP11, MYBL2, PGR) and no



additional overlapping genes between Ellen and either of the other two signatures.

Biological differences between LN status and outcome

Gene Set Enrichment Analysis (GSEA) was used to identify biological processes potentially related to outcome in ER+ tumours with and without lymph node involvement. The GSEA algorithm was performed independently on LN+ and LN- samples, using systemic recurrence as the phenotypic class variable. Based on these findings, biological pathways that are related to outcome in LN- (Table 6) and LN+ (Table 7) patients groups were identified. Additional information pertaining to specific overlapping genes and statistical parameters is available in the Additional file 1. A number of cancer-related pathways were enriched in each subgroup of patient samples, including proliferation, epithelial-mesenchymal transition (EMT), epigenetic modification, and immunity [33]. Poor outcome LN- patient tumours were

enriched for proliferation, growth factor signalling and epigenetic modification gene sets (Table 6). Whereas, poor outcome LN+ patient tumours were enriched for gene sets associated with EMT, migration, differentiation, and apoptosis. The tumours from patients with good survival, both LN- and LN+, were enriched for immune response gene sets. This was particularly evident for patients with LN+ disease where 6 of the top 10 gene sets, associated with good outcome were comprised of 649 immune response related genes (Table 7).

Discussion

Lymph node status is the most prognostic variable for determining outcome in patients with ER+ breast cancer. However, it is unknown whether lymph node involvement is simply an indication of tumour progression over time or whether a primary tumour’s ability to metastasize is pre-determined by tumour biology. Gene signatures are an attractive option to predict outcome and several have been validated for use on ER+ breast cancer patients. Oncotype DX is a prognostic (and predictive) gene signature developed and validated using ER+ LN- tumours exclusively, whereas the development of the Prosigna gene signature included LN+ tumour samples. We wanted to examine the performance of Oncotype DX and Prosigna on LN+ patients and hypothesized that if lymph node involvement is merely a

Table 3 Oncotype DX, Prosigna, and Ellen performance

	LN- Patients		LN+ Patients	
	p value	Concordance	p value	Concordance
Oncotype DX	1.26E-04	0.78	3.06E-01	0.58
Prosigna	8.07E-04	0.79	1.34E-02	0.62
Ellen	1.27E-06	0.85	1.74E-04	0.71

Table 4 Number of Ellen genes associated with different biological pathways

	Low Risk of Metastasis Genes	High Risk of Metastasis Genes	Total
Gene Expression	13	2	15
	EGR1, FOS, JUN, NAT10, RPL11, ZFP36, EEF2, LITAF, POLR2E, POLR3E, RPLP2, RPS15, RPS23	RPL38, RPS11	
Proliferation	5	6	11
	KIDINS220, PIK3R1, ZFP36L2, CDIPT, CXCL12	JTB, SERPINB3, NUCKS1, SNRPE, SPDEF, TXN	
Immune Response	5	3	8
	FOS, CXCL12, HLA-DPA1, JAK1, PCBP2	FKBP4, MTDH, NUCKS1	
Cell Migration	3	4	7
	SPTBN1, CYFIP1, CXCL12	S100P, ARF6, CSTA, NUCKS1	
Apoptosis	6	1	7
	JUN, SGK1, LITAF, TNFRSF10B, GLTSCR2, ITM2B	S100G(-)	
Stress Response	5	0	5
	DUSP1, ABAT, CIRBP, GLTSCR2, GPX4		
Metabolism	1	4	5
	GPX4	FLOT1, SQLE, COX5B, GPR172A	
Epigenetics	2	2	4
	NCOR1, SMARCA2	NAT10, H3F3A	
Angiogenesis	0	2	2
		ACTC1, MB	
Differentiation	2	0	2
	DPYSL2, RAI2		
Cell Cycle	0	1	1
		SFN	
PTM and Trafficking	1	0	1
	DUSP1		
Cytoskeleton	0	1	1
		KRT10	

function of tumour progression, then the signatures developed using LN- patient samples (Oncotype DX) should similarly be able to predict outcome for LN+ patients.

The Oncotype DX signature was developed using weighted averages of 16 genes (excluding housekeeping genes) known to be associated with outcome in ER+ LN- breast cancer using a qRT-PCR platform [17]. This

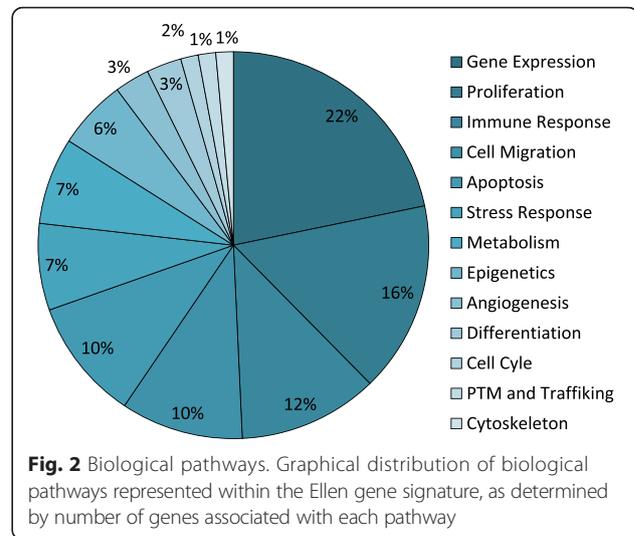


Fig. 2 Biological pathways. Graphical distribution of biological pathways represented within the Ellen gene signature, as determined by number of genes associated with each pathway

21 gene signature has been validated and FDA approved for its ability to predict outcome in an independent cohort of ER+ LN- breast cancer patients [34, 35]. We simulated the Oncotype DX algorithm *in silico* using Affymetrix gene expression data and tested the prognostic ability of the simulated algorithm on ER+ tumours from LN+ and LN- patients. As expected, the simulated Oncotype DX algorithm was able to significantly predict outcome for ER+ LN- patients, confirming its prognostic capacity in this group of patients and supporting the validity of our *in silico* approach to assess Oncotype DX performance. Furthermore, the *in silico* approach we utilized has been used by others to compare gene expression data from different platforms including qRT-PCR and expression microarrays and to simulate gene signatures such as Oncotype DX and Prosigna [24, 29–31, 33–35].

In our *in silico* study, Oncotype DX was unable to significantly predict risk of recurrence for ER+ LN+ patients (Fig. 1 and Table 3), suggesting that a signature such as Oncotype DX, developed and validated on ER+ LN- patients, is not optimal for predicting outcome in ER+ LN+ patients. We cannot exclude the possibility that there is a subset of LN+ patients for whom Oncotype DX might be an appropriate prognostic assay, but further exploration in this area is needed. As such, there are several ongoing clinical trials, including SWOG S1007 and RxPONDER aimed at validating the prognostic utility of Oncotype DX for ER+ breast cancer patients with limited LN+ disease, the results from these studies are eagerly awaited [36, 37].

Prosigna was approved as a prognostic assay for distant metastasis-free survival for patients with ER+ disease with 0–3 positive lymph nodes. The 50 disease associated-genes comprising the Prosigna assay were derived from the intrinsic molecular subtype signatures

Table 5 Comparison of biological processes associated with each gene signature

	Biological Pathway	Signature		
		Ellen	Oncotype DX	Prosigna
Common	Gene expression	15	2	5
	Proliferation	11	4	15
	Immune response	8	3	1
	Cell migration	7	4	9
	Cell cycle	1	4	12
	PTM and trafficking	1	2	4
Unique	Epigenetics	4	0	0
	Angiogenesis	2	0	0
	DNA repair	0	0	4
	DNA replication	0	0	3
Other	Apoptosis	7	0	1
	Metabolism	5	0	2
	Stress response	5	0	2
	Differentiation	2	0	2
	Cytoskeleton	1	0	5
	Intracellular signalling	0	3	6
	Survival	0	2	2
	Drug metabolism	0	1	1

discovered in 2000 [18, 38]; both LN- and LN+ breast cancer samples were used to develop and validate the Prosigna assay ([39], TransATAC and ABCSG8 clinical trials). The simulated Prosigna signature, described here was able to significantly predict outcome for ER+ LN- and LN+ patients separately. This suggests that including LN+ patient samples in signature development will improve signature performance when applied to LN+ patient tumour samples.

The Ellen signature, which was developed using both LN- and LN+ patients, was able to more significantly predict outcome of LN- and LN+ cohorts than either the Oncotype DX or Prosigna gene signatures. It is possible that the increased significance, concordance, and hazard ratios derived from the Ellen signature are related to it being both trained and validated using Affymetrix data and we recognize that our results need to be validated using an independent cohort of patients. Alternatively, the increased significance of Ellen could be reflective of the importance of the biological processes, represented by the signature genes, to outcome in ER+ breast cancer. As detailed in Table 5, Ellen, Oncotype DX, and Prosigna signatures each represent common biological processes including: gene expression, proliferation, immune response, cell migration, cell cycle, and PTM and Trafficking. However, genes

Table 6 Gene sets enriched in lymph node negative patients

LN-	Gene Set Name	Inferred Biological Activity	Genes Represented	ES
Long Term Remission	SHEN_SMARCA2_TARGETS_UP	Epigenetic	347	0.65
	SCHUETZ_BREAST_CANCER_DUCTAL_INVASIVE_UP		321	0.60
	BOQUEST_STEM_CELL_UP	Stemness	239	0.57
	ANASTASSIOU_CANCER_MESENCHYMAL_TRANSITION_SIGNATURE	EMT	60	0.71
	SMID_BREAST_CANCER_LUMINAL_A_UP		78	0.67
	PICCALUGA_ANGIOIMMUNOBLASTIC_LYMPHOMA_UP	Immune	171	0.58
	TURASHVILI_BREAST_DUCTAL_CARCINOMA_VS_LOBULAR_NORMAL_DN		52	0.71
	DACOSTA_UV_RESPONSE_VIA_ERCC3_COMMON_DN	Radiation	409	0.52
	FLECHNER_BIOPSY_KIDNEY_TRANSPLANT_OK_VS_DONOR_UP	Immune	480	0.51
	REN_ALVEOLAR_RHABDOMYOSARCOMA_DN	PAX3-FOXO1 down	386	0.52
Distant Metastasis	SHEN_SMARCA2_TARGETS_DN	Epigenetic	312	-0.54
	RICKMAN_HEAD_AND_NECK_CANCER_F		47	-0.72
	SOTIRIOU_BREAST_CANCER_GRADE_1_VS_3_UP		130	-0.54
	ROSTY_CERVICAL_CANCER_PROLIFERATION_CLUSTER	Proliferation	124	-0.51
	KEGG_OLFACTORY_TRANSDUCTION	Signalling	66	-0.58
	REACTOME_STRIATED_MUSCLE_CONTRACTION		27	-0.70
	KUNINGER_IGF1_VS_PDGF_TARGETS_UP	IGF	58	-0.58
	XU_HGF_TARGETS_REPRESSED_BY_AKT1_DN	AKT/HGF	81	-0.51
	MIKKELSEN_MEF_ICP_WITH_H3K27ME3	Epigenetic	118	-0.48
	MIKKELSEN_MCV6_LCP_WITH_H3K27ME3	Epigenetic	15	-0.76

Table 7 Gene Sets enriched in lymph node positive patients

LN+	Gene Set Name	Inferred Biological Activity	Genes Represented	ES	
Long Term Remission	SMID_BREAST_CANCER_NORMAL_LIKE_UP		420	0.65	
	WIELAND_UP_BY_HBV_INFECTION	Immune	91	0.77	
	WALLACE_PROSTATE_CANCER_RACE_UP	Increased Risk	259	0.66	
	MCLACHLAN_DENTAL_CARIES_UP	Immune	225	0.59	
	KIM_LRRC3B_TARGETS	Immune	28	0.85	
	FLECHNER_BIOPSY_KIDNEY_TRANSPLANT_REJECTED_VS_OK_UP	Immune	78	0.67	
	VANTVEER_BREAST_CANCER_ESR1_DN	ER down	197	0.59	
	TURASHVILI_BREAST_DUCTAL_CARCINOMA_VS_DUCTAL_NORMAL_DN		144	0.60	
	ICHIBA_GRAFT_VERSUS_HOST_DISEASE_D7_UP	Immune	89	0.66	
	LEE_DIFFERENTIATING_T_LYMPHOCYTE	Immune	138	0.61	
	Distant Metastasis	NIKOLSKY_BREAST_CANCER_8Q12_Q22_AMPLICON		86	-0.68
		NIKOLSKY_BREAST_CANCER_8Q23_Q24_AMPLICON		83	-0.66
		ANASTASSIOU_CANCER_MESENCHYMAL_TRANSITION_SIGNATURE	EMT	60	-0.70
TURASHVILI_BREAST_DUCTAL_CARCINOMA_VS_DUCTAL_NORMAL_UP			28	-0.65	
FARMER_BREAST_CANCER_CLUSTER_5			18	-0.74	
TURASHVILI_BREAST_LOBULAR_CARCINOMA_VS_DUCTAL_NORMAL_UP			52	-0.55	
MILI_PSEUDOPODIA_HAPTOTAXIS_UP		Migration	348	-0.41	
DING_LUNG_CANCER_EXPRESSION_BY_COPY_NUMBER		CNVs	81	-0.49	
IIZUKA_LIVER_CANCER_PROGRESSION_G2_G3_UP		Differentiation	24	-0.64	
HAMAI_APOPTOSIS_VIA_TRAIL_UP		Apoptosis	463	-0.38	

related to angiogenesis and epigenetics are unique to Ellen. Both of these processes have been demonstrated to be important for outcome in ER+ breast cancer [6, 14, 33, 40–43]. Additional multivariable studies are being conducted, using an independent cohort of patients, to assess the relationship between these biological features and other clinical variables, including tumour size, grade, and histological subtype to validate the prognostic potential of Ellen.

Given that the three signatures examined performed with various levels of accuracy in LN+ and LN- patient populations, we were interested in exploring the biological processes that might be related to outcome in ER+ LN+ and LN- tumours separately, using GSEA. Patients with good outcome (irrespective of their original LN status) had tumours with expression profiles enriched for immune related genes (Tables 6 and 7). This was particularly striking for LN+ tumours where 6 of the 10 gene sets associated with good outcome were immune related. This enrichment of immune related gene sets may be indicative of immune cell infiltration in some tumours and suggests that a subset of ER+ breast cancer patients have a robust anti-tumour immune response and that this in turn may be associated with improved survival [39, 44, 45].

We examined the ontology of genes comprising the Ellen signature to determine whether their functions overlap with those identified using the GSEA and found that 11 % of the Ellen genes are related to immune response. This further supports an important role for immune response in ER+ tumours and the utility of the signature. For example, we found that CXCL12 and JAK1 are both more highly expressed in low risk tumours. It has been reported that increased expression of CXCL12 is a strong positive prognostic factor that correlates with disease free and overall survival in both ER+ and ER- tumours [46, 47]. JAK1 is a protein tyrosine kinase involved in the response to interferons; recently the closely related JAK2 family member was found to be associated with improved outcome in breast cancer [48]. In addition, the expression of HLA-DPA1, which is normally expressed on antigen presenting cells, may indicate the presence of immune infiltrate [49]. Overall, the presence of these immune related genes in low risk tumours indicates that immune response is an important factor in the progression of breast cancer.

Patients with poor outcome showed enrichment for different gene sets depending on whether their tumour was LN+ or LN- at diagnosis. For example, poor outcome LN- patient tumours were enriched for proliferation, growth

factor signalling, and epigenetic modification gene sets, also represented by individual genes comprising the Ellen signature (Table 6). Proliferation in ER+ breast cancer is a poor prognostic factor and correlates with the Luminal B subtype [39]. Epigenetic modification is thought to have some role in tumour progression, as global hypermethylation of the tumour genome has been associated with poor outcome [50–52]. In addition there are several studies reporting that HDAC inhibitor usage may be useful as adjuvant chemotherapeutics in this high risk group [53, 54]. Whereas, patients with LN+ disease and poor outcome had tumours enriched for EMT and migration suggesting a migratory phenotype [9, 55].

Taken together, the different biological processes highlighted for LN- and LN+ groups may explain why gene signatures developed for one group would not necessarily be predictive of outcome in the other.

Conclusion

In summary, we have shown that by comparing Onco-type DX and Prosigna with a novel gene signature, it is important to include patients with both LN+ and LN- status when developing prognostic gene signatures. Furthermore, we have identified candidate biological processes that imply how tumour biology can be related to outcome. This is particularly evident for LN+ tumours with good outcome, where there is enrichment in immune response gene expression, and for LN- tumours with poor outcome, where there is an enrichment for genes involved in epigenetic modification. We developed and characterized Ellen, a gene signature that is designed to be predictive of outcome for all patients with ER+ breast cancer without distant spread, using an unbiased gene selection process. The genes represented in this signature are similar to those whose pathways were found to be enriched using GSEA, further suggesting that Ellen would be suitable for use in a variety of biologically unique ER+ breast tumours. Work is currently underway to validate the performance of Ellen using an alternate platform and with additional independent cohorts. Further, the clinical information available for the training and validation cohorts was limited, so it is difficult to know whether there are other confounding variables. Ultimately, this study shows that gene expression of primary tumours can be informative about metastatic potential and can be distinguished between LN- and LN+ patients. In addition Ellen, once validated, would be able to provide prognostic information for patients with tumours accompanied by small lymph node metastasis, such as isolated tumour cells or micrometastases, those with incomplete lymph node dissections (ie sentinel node only), or those who have no lymph node information.

Additional file

Additional file 1: (RAR 837 kb)

Abbreviations

BC, breast cancer; C, concordance; CI, confidence interval; CoxPH, Cox proportional hazards; DMFS, distant metastasis free survival; EMT, epithelial mesenchymal transition; ER, estrogen receptor; ES, enrichment score; GEO, gene expression omnibus; GO, gene ontology; GSEA, gene set enrichment analysis; HR, hazard ratio; LN, lymph node; PAM, prediction analysis of microarrays; PTM, post translational modification; qRT-PCR, quantitative real time polymerase chain reaction; RMA, robust multichip algorithm; ROR, risk of recurrence; RS, recurrence score; TNM, tumour node metastasis

Funding

Funding for this project was provided by an operating grant (AB, JH) and fellowship (JGC, AEG) from the Canadian Breast Cancer Foundation.

Availability of data and materials

All datasets used for this study are available from the Gene Expression Omnibus repository hosted by NCBI. GSE17705: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse17705> GSE6532: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=gse6532> All specific statistical software required for analysis and its availability is denoted in the main text of the manuscript.

Authors' contributions

JGC conception and design, acquisition of data, interpretation, and preparation of manuscript. RMH developed methods, interpretation, and preparation of manuscript. AEG interpretation, preparation, and review of manuscript. KND conception and design, interpretation, and preparation of manuscript. TW conception and design of project. MNL conception and design of project. JAH conception and design, interpretation, and preparation of manuscript. AB conception and design, interpretation, and preparation of manuscript. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Author details

¹Department of Oncology, Juravinski Hospital and Cancer Centre, Hamilton, Canada. ²Department of Biochemistry and Biomedical Sciences, Centre for Functional Genomics, McMaster University, Hamilton, Canada. ³Department of Pathology, Juravinski Hospital and Cancer Centre, Hamilton, Canada.

Received: 8 November 2014 Accepted: 4 July 2016

Published online: 28 July 2016

References

1. CCO. Surgical Management of Early-Stage Invasive Breast Cancer Overview Guideline Report History. 2011.
2. NCCN. Practice Guidelines in Oncology. 2012.
3. Fisher B, Dignam J, Wolmark N, DeCillis A, Emir B, Wickerham DL, Bryant J, Dimitrov NV, Abramson N, Atkins JN, Shibata H, Deschene L, Margolese RG. Tamoxifen and chemotherapy for lymph node-negative, estrogen receptor-positive breast cancer. *J Natl Cancer Inst.* 1997;89:1673–82.
4. Muss HB, Woolf S, Berry D, Cirincione C, Weiss RB, Budman D, Wood WC, Henderson IC, Hudis C, Winer E, Cohen H, Wheeler J, Norton L. Adjuvant chemotherapy in older and younger women with lymph node-positive breast cancer. *JAMA.* 2005;293:1073–81.
5. Capulli M, Angelucci A, Driouch K, Garcia T, Clement-Lacroix P, Martella F, Ventura L, Bologna M, Flamini S, Moreschini O, Lidereau R, Ricevuto E, Muraca M, Teti A, Rucci N. Increased expression of a set of genes enriched in oxygen binding function discloses a predisposition of breast cancer

- bone metastases to generate metastasis spread in multiple organs. *J Bone Miner Res.* 2012;27:2387–98.
6. Van den Eynden GG, Van Laere SJ, Van der Auwera I, Gilles L, Burn JL, Colpaert C, van Dam P, Van Marck EA, Dirix LY, Vermeulen PB. Differential expression of hypoxia and (lymph)angiogenesis-related genes at different metastatic sites in breast cancer. *Clin Exp Metastasis.* 2007;24:13–23.
 7. Clarke R, Skaar TC, Bouker KB, Davis N, Lee YR, Welch JN, Leonessa F. Molecular and pharmacological aspects of antiestrogen resistance. *J Steroid Biochem Mol Biol.* 2001;76:71–84.
 8. Jatoi I, Hilsenbeck SG, Clark GM, Osborne CK. Significance of Axillary Lymph Node Metastasis in Primary Breast Cancer. *J Clin Oncol.* 1999;17:2334.
 9. Tseng SPLWW. Micrometastatic Cancer Cells in Lymph Nodes, Bone Marrow, and Blood. *CA Cancer J Clin.* 2014;64:195–206.
 10. Ursaru M, Jari I, Naum A, Scripcariu V, Negru D. Causes of death in patients with stage 0-II breast cancer. *Rev Med Chir Soc Med Nat Iasi.* 2015;119:374–8.
 11. Singh SK, Clarke ID, Terasaki M, Bonn VE, Hawkins C, Squire J, Dirks PB. Identification of a cancer stem cell in human brain tumors. *Cancer Res.* 2003;63:5821–8.
 12. Loi S, Haibe-Kains B, Desmedt C, Wirapati P, Lallemand F, Tutt AM, Gillet C, Ellis P, Ryder K, Reid JF, Daidone MG, Pierotti MA, Berns EM, Jansen MP, Foekens JA, Delorenzi M, Bontempi G, Piccart MJ, Sotiriou C. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC Genomics.* 2008;9:239.
 13. Prat A, Parker JS, Fan C, Cheang MCU, Miller LD, Bergh J, Chia SKL, Bernard PS, Nielsen TO, Ellis MJ, Carey LA, Perou CM. Concordance among gene expression-based predictors for ER-positive breast cancer treated with adjuvant tamoxifen. *Ann Oncol.* 2012;23:2866–73.
 14. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* 2002;415:530–6.
 15. Blohmer JU, Rezaei M, Kümmel S, Kühn T, Warm M, Friedrichs K, Benkow A, Valentine WJ, Eiermann W. Using the 21-gene assay to guide adjuvant chemotherapy decision-making in early-stage breast cancer: a cost-effectiveness evaluation in the German setting. *J Med Econ.* 2013;16:30–40.
 16. Habel LA, Shak S, Jacobs MK, Capra A, Alexander C, Pho M, Baker J, Walker M, Watson D, Hackett J, Blick NT, Greenberg D, Fehrenbacher L, Langholz B, Quesenberry CP. A population-based study of tumor gene expression and risk of breast cancer death among lymph node-negative patients. *Breast Cancer Res.* 2006;8:R25.
 17. Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Baehner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med.* 2004;351:2817–26.
 18. Filipits M, Rudas M, Jakesz R, Dubsy P, Fitzal F, Singer CF, Dietze O, Greil R, Jelen A, Sevelde P, Freibauer C, Müller V, Jänicke F, Schmidt M, Kölbl H, Rody A, Kaufmann M, Schroth W, Brauch H, Schwab M, Fritz P, Weber KE, Feder IS, Hennig G, Kronenwett R, Gehrman M, Gnant M. A new molecular predictor of distant recurrence in ER-positive, HER2-negative breast cancer adds independent information to conventional clinical risk factors. *Clin Cancer Res.* 2011;17:6012–20.
 19. Barrett T, Wilhite S, Ledoux P, Evangelista C, Kim I, Tomashevsky M, Marshall K, Phillip P, Holko M, Yefanov A, Lee H, Zhang N, Roberston C, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* 2013;41:D991–5.
 20. Symmans WF, Hatzis C, Sotiriou C, Andre F, Peintinger F, Regitnig P, Daxenbichler G, Desmedt C, Domont J, Marth C, Delalogue S, Bauernhofer T, Valero V, Booser DJ, Hortobagyi GN, Pusztai L. Genomic index of sensitivity to endocrine therapy for breast cancer. *J Clin Oncol.* 2010;28:4111–4119.
 21. Loi S, Haibe-Kains B, Desmedt C, Lallemand F, Tutt AM, Gillet C, Ellis P, Harris A, Bergh J, Foekens JA, Klijn JG, Larsimont D, Buyse M, Bontempi G, Delorenzi M, Piccart MJ, Sotiriou C. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. *J Clin Oncol.* 2007;25:1239–46.
 22. Hallett RM, Dvorkin A, Gabardo CM, Hassell JA. An algorithm to discover gene signatures with predictive potential. *J Exp Clin Cancer Res.* 2010;29:120.
 23. McCall MN, Bolstad BM, Irizarry RA. Frozen robust multiarray analysis (fRMA). *Biostatistics.* 2012;11:242–253.
 24. Györfy B, Molnar B, Lage H, Szallasi Z, Eklund AC. Evaluation of microarray preprocessing algorithms based on concordance with RT-PCR in clinical samples. *PLoS One.* 2009;4:e5645.
 25. Dvorkin-Gheva A, Hassell JA. Identification of a novel luminal molecular subtype of breast cancer. *PLoS One.* 2014;9:e103514.
 26. Aravind Subramanian PT. Gene Set Enrichment Analysis. 2014.
 27. Hallett RM, Dvorkin-Gheva A, Bane A, Hassell JA. A Gene Signature for Predicting Outcome in Patients with Basal-like Breast Cancer. *Sci Rep.* 2012;2:227.
 28. The Gene Ontology Consortium. Gene Ontology Consortium: going forward. *Nucleic Acids Res.* 2014;43:D1049–56.
 29. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A.* 2002;99:6567–72.
 30. Györfy B, Benke Z, Lániczky A, Balázs B, Szállási Z, Timár J, Schäfer R. RecurrenceOnline: an online analysis tool to determine breast cancer recurrence and hormone receptor status using microarray data. *Breast Cancer Res Treat.* 2012;132:1025–34.
 31. Elloumi F, Hu Z, Li Y, Parker JS, Gulley ML, Amos KD, Troester MA. Systematic bias in genomic classification due to contaminating non-neoplastic tissue in breast tumor samples. *BMC Med Genomics.* 2011;4:54.
 32. Naoi Y, Kishi K, Tsunashima R, Shimazu K, Shimomura A, Maruyama N, Shimoda M, Kagara N, Baba Y, Kim SJ, Noguchi S. Comparison of efficacy of 95-gene and 21-gene classifier (Oncotype DX) for prediction of recurrence in ER-positive and node-negative breast cancer patients. *Breast Cancer Res Treat.* 2013.
 33. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2010;144:646–74.
 34. Holt S, Bertelli G, Humphreys I, Valentine W, Durrani S, Pudney D, Rolles M, Moe M, Khawaja S, Sharaiha Y, Brinkworth E, Whelan S, Jones S, Bennett H, Phillips CJ. A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-negative or pN1mi, ER-positive breast cancer in the UK. *Br J Cancer.* 2013;108:2250–8.
 35. Joh JE, Esposito NN, Kiluk JV, Laronga C, Lee MC, Loftus L, Soliman H, Boughey JC, Reynolds C, Lawton TJ, Acs PI, Gordan L, Acs G. The effect of Oncotype DX recurrence score on treatment recommendations for patients with estrogen receptor-positive early stage breast cancer and correlation with estimation of recurrence risk by breast cancer specialists. *Oncologist.* 2011;16:1520–6.
 36. Albain KS, Barlow WE, Shak S, Hortobagyi GN, Livingston RB, Yeh I-T, Ravdin P, Bugarini R, Baehner FL, Davidson NE, Sledge GW, Winer EP, Hudis C, Ingle JN, Perez EA, Pritchard KI, Shepherd L, Gralow JR, Yoshizawa C, Allred DC, Osborne CK, Hayes DF. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncol.* 2010;11:55–65.
 37. Saghatelyan M, Mook S, Pruneri G, Viale G, Glas AM, Guerin S, et al. Additional prognostic value of the 70-gene signature (MammaPrint®) among breast cancer patients with 4-9 positive lymph nodes. *Breast.* 2013.
 38. 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. Molecular portraits of human breast tumours. *Nature.* 2000;406:747–52.
 39. Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T, Davies S, Fauron C, He X, Hu Z, Quackenbush JF, Stijleman IJ, Palazzo J, Marron JS, Nobel AB, Mardis E, Nielsen TO, Ellis MJ, Perou CM, Bernard PS. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol.* 2009;27:1160–7.
 40. Jovanovic J, Rønneberg JA, Tost J, Kristensen V. The epigenetics of breast cancer. *Mol Oncol.* 2010;4:242–54.
 41. Huynh KT, Hoon DSB. Epigenetics of regional lymph node metastasis in solid tumors. *Clin Exp Metastasis.* 2012;29:747–56.
 42. Lo P-K, Sukumar S. Epigenetics and breast cancer. *Pharmacogenomics.* 2009; 9:1879–902.
 43. Kamalakaran S, Varadan V, Giercksky Russnes HE, Levy D, Kendall J, Janevski A, Riggs M, Banerjee N, Synnsetvedt M, Schlichting E, Kåresen R, Shama Prasad K, Rotti H, Rao R, Rao L, Eric Tang M-H, Satyamoorthy K, Lucito R, Wigler M, Dimitrova N, Naume B, Borresen-Dale A-L, Hicks JB. DNA methylation patterns in luminal breast cancers differ from non-luminal subtypes and can identify relapse risk independent of other clinical variables. *Mol Oncol.* 2011;5:77–92.

44. Hsu DS, Kim MK, Balakumaran BS, Acharya CR, Anders CK, Clay T, Lyerly HK, Drake CG, Morse MA, Febbo PG. Immune signatures predict prognosis in localized cancer. *Cancer Invest.* 2010;28:765–73.
45. Oh E, Choi Y-L, Park T, Lee S, Nam SJ, Shin YK. A prognostic model for lymph node-negative breast cancer patients based on the integration of proliferation and immunity. *Breast Cancer Res Treat.* 2012;132:499–509.
46. Mirisola V, Zuccarino A, Bachmeier BE, Sormani MP, Falter J, Nerlich A, Pfeffer U. CXCL12/SDF1 expression by breast cancers is an independent prognostic marker of disease-free and overall survival. *Eur J Cancer.* 2009;45:2579–87.
47. Wendt MK, Cooper AN, Dwinell MB. Epigenetic silencing of CXCL12 increases the metastatic potential of mammary carcinoma cells. *Oncogene.* 2008;27:1461–71.
48. Miller CP, Thorpe JD, Kortum AN, Coy CM, Cheng W-Y, Ou Yang T-H, Anastassiou D, Beatty JD, Urban ND, Blau CA. JAK2 Expression Is Associated with Tumor-Infiltrating Lymphocytes and Improved Breast Cancer Outcomes: Implications for Evaluating JAK2 Inhibitors. *Cancer Immunol Res.* 2014;2:301–6.
49. Englert NA, Spink BC, Spink DC. Persistent and non-persistent changes in gene expression result from long-term estrogen exposure of MCF-7 breast cancer cells. *J Steroid Biochem Mol Biol.* 2011;123:140–50.
50. Koboldt DC, Fulton RS, McLellan MD, Schmidt H, Kalicki-Veizer J, McMichael JF, et al. Comprehensive molecular portraits of human breast tumours. *Nature.* 2012;1–10.
51. Feng Q, Zhang Z, Shea MJ, Creighton CJ, Coarfa C, Hilsenbeck SG, et al. An epigenomic approach to therapy for tamoxifen-resistant breast cancer. *Cell Res.* 2014.
52. Klarmann GJ, Decker A, Farrar WL. Epigenetic gene silencing in the Wnt pathway in breast cancer. *Epigenetics.* 2008;3:59–63.
53. Fan J, Yin W-J, Lu J-S, Wang L, Wu J, Wu F-Y, Di G-H, Shen Z-Z, Shao Z-M. ER alpha negative breast cancer cells restore response to endocrine therapy by combination treatment with both HDAC inhibitor and DNMT inhibitor. *J Cancer Res Clin Oncol.* 2008;134:883–90.
54. Wagner JM, Hackanson B, Lubbert M, Jung M. Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. *Clin Epigenetics.* 2010;1:117–136.
55. Hennessy BT, Stemke-hale K, Gilcrease MZ, Krishnamurthy S, Lee J, Fridlyand J, Agarwal R, Joy C, Liu W, Stivers D, Baggerly K, Lluch A, Monteagudo C, He X, Weigman V, Palazzo J, Hortobagyi GN, Nolden LK, Wang NJ, Valero V, Gray JW, Perou CM, Mills GB. Characterization of a Naturally Occurring Breast Cancer Subset Enriched in Epithelial-to-Mesenchymal Transition and Stem Cell Characteristics. *Cancer Res.* 2009;69:4116–24.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

