Gene expression profiling of breast tumours from New Zealand patients

Anita Muthukaruppan, Annette Lasham, Cherie Blenkiron, Kathryn J Woad, Michael A Black, Nicholas Knowlton, Nicole McCarthy, Michael P Findlay, Cristin G Print, Andrew N Shelling

ABSTRACT

AIMS: New Zealand has one of the highest rates of breast cancer incidence in the world. We investigated the gene expression profiles of breast tumours from New Zealand patients, compared them to gene expression profiles of international breast cancer cohorts and identified any associations between altered gene expression and the clinicopathological features of the tumours.

METHODS: Affymetrix microarrays were used to measure the gene expression profiles of 106 breast tumours from New Zealand patients. Gene expression data from six international breast cancer cohorts were collated, and all the gene expression data were analysed using standard bioinformatic and statistical tools.

RESULTS: Gene expression profiles associated with tumour ER and ERBB2 status, molecular subtype and selected gene expression signatures within the New Zealand cohort were consistent with those found in international cohorts. Significant differences in clinicopathological features such as tumour grade, tumour size and lymph node status were also observed between the New Zealand and international cohorts.

CONCLUSIONS: Gene expression profiles, which are a sensitive indicator of tumour biology, showed no clear difference between breast tumours from New Zealand patients and those from non-New Zealand patients. This suggests that other factors may contribute to the high and increasing breast cancer incidence in New Zealand compared to international populations.

Breast cancer is the most common cancer among women and is the leading cause of cancer death in women worldwide. It is a multi-factorial disease, with considerable inter-patient heterogeneity, and complex aetiology involving genetic, endocrinological, environmental and lifestyle factors.[1] New Zealand has one of the highest breast cancer incidences in the world.[2] The age-standardised breast cancer incidence rate is 92.1 per 100,000 population,[3] and there has been a steady increase in reported breast cancer incidence from 1978–2004,[4] most likely due to improved diagnostic rates through breast cancer screening. Although breast cancer mortality in New Zealand has declined over the last 20 years,[5] breast cancer remains a significant cause of death in New Zealand women,[2] with an age-standardised rate of 20.3 per 100,000 population.[3] While New Zealand breast cancer incidence is comparable to its neighbouring country Australia, New Zealand women appear to have higher breast cancer mortality,[5,6] possibly due to the high cancer mortality rates for Māori and Pacific women, and the relatively slow funding of new drug treatments in New Zealand.

Gene expression profiling using primary breast tumours has transformed understanding of the molecular aspects of this disease,[7,8] especially when gene expression data has been integrated with data about DNA mutation and copy number, epigenetic factors such as microRNAs and gene methylation, plus protein expression and *in vitro* experimental data.[9–12] These integrated studies now have enabled a re-classification of breast tumours into 10 molecular subtypes from the initial five molecular

subtypes,[13,14] as well as an understanding of pathway signalling in breast cancer,[15] association between DNA copy number variation, gene expression and patient survival,[16] identification of drivers of proliferation in the luminal molecular subtype of breast cancer,[17] confirmation of the heterogeneity of breast tumours[18] and the development of molecular tests for breast cancer prognostication and stratified therapy.[19,20]

In this study, we aimed to generate gene expression profiles of breast tumours from a cohort of 106 New Zealand patients. We compared these profiles and their associated clinical and pathological data with data from non-New Zealand patient cohorts, and have made this new dataset available to the breast cancer research community (see Methods). We found that the statistical associations between gene expression profiles and clinicopathological parameters seen in New Zealand breast tumours are remarkably similar to the associations observed in breast tumours from non-New Zealand breast cancer cohorts.

Materials and methods

Ethics statement

All women provided written consent to their tissue being utilised and their clinical records accessed for this project. This study was approved by the New Zealand Multi-Region Ethics Committee (project MEC-09/06/060) and the Northern X Regional Ethics Committee (NTX-05/08/096). In addition, the study was reviewed by Auckland District Health Board Research Committee (project A+4567), the Christchurch Cancer Society Tissue Bank Board (project 10101PS) as well as the Māori Research Review Committees associated with the Auckland District Health Board and Cancer Society Tissue Bank.

New Zealand primary breast tumour collection

Breast tumour samples were obtained from 47 female patients (Auckland) and 59 female patients (Christchurch) who provided written consent (Supplementary Table 1). Patients included in the study were females aged ≥18 years with breast cancer, had no previous breast malignancies and had not received neoadjuvant treatment prior to surgery. Resected tumours ranged from 6–100mm in size, and were frozen at -80°C within 60 min of surgical resection. The Auckland samples were collected in two stages, 30 between 2005 and 2007, and 17 in 2010. The Christchurch samples were collected between 2003 and 2005. Clinical and pathological data, including up to eight years of follow-up data, were obtained from patient notes collected by the Auckland Breast Cancer Registry for Auckland (http://www.adhb.govt.nz/AucklandBreastCancerRegister/) and the Cancer Society Tissue Bank for Christchurch (http://www.otago.ac.nz/mackenzie-cancer/tissue-bank/). The oestrogen receptor (ER) status of the tumours were determined using standard diagnostic immunohistochemistry (IHC) methods. The ERBB2 status of the tumours was also determined using IHC, and where the IHC results were equivocal, they were determined using fluorescence *in situ* hybridisation (FISH).

RNA extraction and microarray hybridisation

Breast tumour tissues were homogenised in TRIzol (Thermo Fisher Scientific Inc., Waltham, MA, USA), and total RNA was isolated from TRIzol homogenates using chloroform and purified using RNA affinity mini columns (manufacturer's protocols; RNeasy Mini Kit, Qiagen, Germany; PureLink *Pro* 96 Total RNA Purification Kit, Thermo Fisher Scientific Inc.). RNA yields and purity were determined using the NanoDrop spectrophotometer (NanoDrop Technologies Inc., DE, USA). RNA integrity was assessed (Agilent 2100 Bioanalyser, Agilent Technologies Inc., CA, USA; Experion, Bio-Rad Laboratories, CA, USA) and the average 260/280 ratio of total RNA was 2.0 (range 1.8–2.2) and the average RIN/RQI was 8.68 (range 5.6–9.7). Total RNA was labelled, fragmented and hybridised to Affymetrix Human Genome (HG) U133 Plus 2.0 arrays (manufacturer's protocol). For samples GSM900586-GSM900662 of Gene Expression Omnibus (GEO) record GSE36771, 100µg RNA was used (MessageAmp Premier RNA Amplification Kit, Thermo Fisher Scientific Inc.), for samples GSM900663-GSM900692 of GSE36771, 5µg RNA was used (Affymetrix GeneChip Kit, Affymetrix, Santa Clara, CA, USA).

Microarray data analysis

Raw and normalised microarray data together with clinicopathological annotations are available as GEO record GSE36771 (microarrays GSM900586-GSM900692). Quality assessment was performed using Affymetrix Expression Console and dChip[21] to exclude low-quality samples. Data were analysed in R using the affy and limma packages.[22–24] Briefly, the data was quantile normalised using the RMA method (without background correction),[25] followed by differential gene expression analysis in limma using Benjamini and Hochberg false discovery rate control,[26] followed by testing for functional enrichment using GeneSetDB,[27] GATHER,[28] and Ingenuity Pathway Analysis (http://www.ingenuity.com). Analyses of the New Zealand tumour gene expression data using limma showed no significant differences between the Auckland and Christchurch tumours, or between tumour RNA purified using the RNeasy or PureLink methods (data not shown). Therefore no batch correction was applied to tumours collected from the two New Zealand regions in subsequent comparisons between New Zealand and non-New Zealand tumours described below.

Meta-analysis of microarray data from primary breast tumours from multiple cohorts

Microarray data for 1,034 primary breast tumours were assembled from raw Affymetrix HG U133 ".cel" files from GEO records; the New Zealand cohort consisted of GSE36771 (n=106), and the non-New Zealand cohort consisted of cohorts: GSE1456 (n=159, Stockholm, Sweden),[29] GSE3494 (n=232, Uppsala, Sweden),[30] GSE4922 (n=40, Singapore),[31] GSE6532 (London and Oxford, UK),[32,33] GSE7390 (n=198, France),[34] and GSE36772 (n=57, Singapore, unpublished) (Supplementary Table 1). Duplicate samples were removed from all cohorts and 22,277 probe sets common to the U133 Plus 2.0 and U133A arrays were used for this analysis, as described previously.[35] Array quality assessment was performed using the 'AffyQCReport' package in R, and normalised using RMA (without background correction) and loess splining. Statistical meta-analysis was performed using the metaMA package in R,[36] and differential gene expression was assessed from normalised microarray data using limma. Genes and probe sets were hierarchically clustered using Euclidean distance and the Ward agglomeration method of the hclust function in R. Molecular subtypes were assigned to each tumour using the Single Sample Predictor algorithm applied per cohort.[37] The algorithm uses pre-computed subtype centroids and calculates the correlation between each tumour and each subtype centroid. A tumour was assigned the subtype that it was most highly correlated with, unless none of the correlations were above a certain threshold; a threshold of 0.1 was used. Differences between selected clinicopathological features of the New Zealand and the non-New Zealand cohorts were analysed using Pearson's Chi-squared tests and Student's t-tests. Principal components for the expression data were calculated by singular value decomposition using the svd function in R, after the data had been zero centered and scaled to unit variance.[38] For visualisation purposes, expression values for each probe set were transformed into Z-scores by mean centreing the data then expressing variation above and below the mean on a scale of standard deviation, using the apply function in R. Differences in time to a distant metastatic event (up to eight years in New Zealand cohort and up to 12 years in the non-New Zealand cohort) were assessed visually via Kaplan-Meier curves and statistically via log rank tests and Cox proportional hazards models using the survival package in R.[39] The visualisations for many of the analyses presented in this paper were scripted using the shiny package from R Studio (http://www.rstudio.com/products/shiny/shiny-server/).

Results

Analysis of gene expression profiles of New Zealand breast tumours

Using normalised gene expression data, we compared gene expression differences between two designated groups using limma, within the New Zealand cohort. We first compared the mRNAs differentially expressed between ER+ and ER- breast tumours, as determined using IHC in the clinic. This analysis identified 64 and 13 probe sets significantly differentially expressed at an absolute log fold

change ≥1.5, between ER+ and ER- tumours (adj-$P \leq 1x10^{-12}$). Due to redundancy in the microarray probe sets, these 77 differentially expressed probe sets represented 50 unique annotated genes (38 with increased and 12 with decreased expression between the ER+ and ER- tumours, respectively). The list of mRNAs differentially expressed between ER+ and ER- tumours are provided in Supplementary Table 2, and a heatmap depicting differentially expressed mRNAs between ER+ and ER- tumours are provided in Supplementary Figure 1. The proteins encoded by these differentially expressed mRNAs included known ERα targets such as the progesterone receptor (encoded by the gene *PGR*), trefoil factor 1 (encoded by the gene *TFF1*), anterior gradient 2 homolog (encoded by the gene *AGR2*) and carbonic anhydrase XII (encoded by the gene *CA12*), as well as proteins that function together with ERα such as forkhead box A1 (encoded by the gene *FOXA1*), epidermal growth factor receptor 4 (encoded by the gene *ERBB4*) and GATA binding protein 3 (encoded by the gene *GATA3*).

Next, the mRNAs differentially expressed between 14 ERBB2+ (also known as *HER2/neu*) and 48 ERBB2- tumours were compared; ERBB2 status was available for only 62 tumours. This analysis identified 28 and two probe sets (representing 18 unique annotated genes) that showed significantly increased or decreased expression respectively (adj-P=0.0001). The list of differentially expressed mRNAs are provided in Supplementary Table 3 and a heatmap depicting significantly regulated mRNAs between ERRB2+ and ERRB2- tumours are provided in Supplementary Figure 2. Interestingly, 17 of the 28 significantly upregulated probe sets in ERBB2+ tumours represented 10 mRNAs, including *ERBB2*, located at locus 17q21 (encoded by the genes: *CDK12, CWC25, FBXL20, GRB7, GSDMB, MIEN1, ORMDL3, PCGF2* and *PGAP3*).

When we compared the gene expression between tumours of histological grade 1 (n=11) and grade 3 (n=53), no differentially expressed genes were identified (adj-$P \leq 0.05$; data not shown).

Comparison of the gene expression profiles of New Zealand breast tumours and international breast tumours

We next compared gene expression profiles of New Zealand breast tumours with gene expression profiles of breast tumours from women recruited from other parts of the world: Sweden, Singapore, France and the UK (See Supplementary Table 1). Gene expression data from six published international breast cancer cohorts were combined (n=927; called "non-New Zealand"), and a number of clinicopathological characteristics were analysed to identify if there were any gene expression differences between the breast tumours from the New Zealand and non-New Zealand cohorts, based on the selected clinicopathological characteristics (summarised in Table 1).

The average tumour sizes (P=0.0001) and patients' ages (P=0.03) in the New Zealand cohort were significantly larger than the non-New Zealand cohort (Table 1). There were also significant differences in the proportions of lymph node positive and lymph node negative patients between the New Zealand and non-New Zealand cohort (P=0.00001; Table 1). There were differences between the proportions of histological grade 1, grade 2 and grade 3 tumours between the New Zealand and non-New Zealand cohort (P=0.001; Table 1). Further analyses revealed significant differences in the proportion of grade 3 tumours compared to grade 2 (P=0.009), and in the proportion of grade 3 tumours compared to grade 1 (P=0.002), between the New Zealand and non-New Zealand cohort. Analysis of the composition of the cohorts based on ER status and molecular subtype showed that there were no significant differences in these variables.

We have previously reported that *ESR1* (which encodes the ERα protein) mRNA expression progressively increased from basal-like to ERBB2 to normal-like to luminal tumours.[40] After classifying our breast tumour gene expression data into their molecular subtypes using the PAM50

Table 1: Comparison of selected clinicopathological features of New Zealand and non-New Zealand breast cancer cohorts.

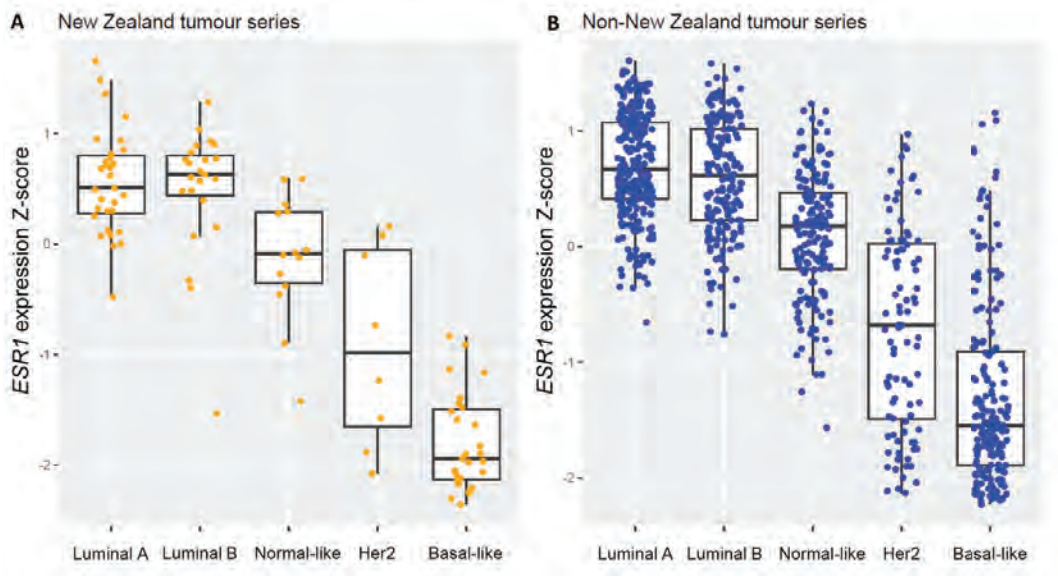
Clinical parameter	**NZ cohort (n=106)**	**Non-NZ cohort (n=927)**	***P*-value (P)**
	Mean (range)	**Mean (range)**	
Patient age (y)*	60 (31–94)	57 (24-93)	0.0271
Patient tumour size (mm)*	29 (6–100)	24 (1–130)	**0.0001**
Tumour ER status[a]	**Numbers (% of total)**	**Numbers (% of total)**	
ER positive	79 (74.5)	743 (74.5)	0.082
ER negative	27 (25.5)	168 (18.1)	
Histological tumour grade[b]			
Grade 1	11 (10.4)	169 (18.2)	**0.00121**
Grade 2	42 (39.6)	403 (43.5)	
Grade 3	53 (50.0)	288 (31.1)	
Patient lymph node status[c]			
Lymph node positive	59 (55.7)	225 (24.3)	**0.00001**
Lymph node negative	45 (42.5)	528 (57.0)	
Tumour molecular subtype			
Luminal A	31 (29.2)	260 (28.0)	0.456
Luminal B	23 (21.7)	187 (20.2)	
Normal-like	14 (13.2)	162 (17.5)	
Her2	8 (7.5)	92 (9.9)	
Basal-like	27 (25.5)	181 (19.5)	
None	3 (2.8)	45 (4.9)	

All calculations were using Pearson's Chi square tests except *patients' age and *tumour size, where a Student's t-test was used. [a]Unavailable ER status information for four patients in GSE3494; exclusion of 12 ER-/PGR+ patients (GSE1456=6, GSE4922=5, GSE36772=1). [b]Unavailable tumour grade information for 67 patients (GSE1456=12, GSE6532=51, GSE7390=2, GSE3494=2); [c]Unavailable lymph node status information for 176 patients (GSE1456=159, GSE36771=2, GSE6532=6, GSE3494=9). See Supplementary Table 1 for patient clinical information.

predictor (see Methods), we compared the distribution of *ESR1* mRNA expression between the five molecular subtypes in the New Zealand and non-New Zealand tumour samples. We found that New Zealand and non-New Zealand tumours showed similar patterns of *ESR1* mRNA expression within and between each molecular subtype, whereby *ESR1* mRNA expression decreases from the luminal subtypes, to normal-like, to ERBB2 to basal-like tumour subtypes (Figure 1).

These box plots show the levels of *ESR1* mRNA in each of the molecular subtypes in the New Zealand (A) and non-New Zealand (B) breast tumour cohorts. The molecular subtypes are shown on the x-axis, and the Z-transformed level of *ESR1* mRNA (based on *ESR1* microarray probeset 205225_at) is shown on the y-axis. The box plots show the medians, upper and lower quartiles, with whiskers extending to the 5th and 95th percentiles.

Figure 1: The distribution of *ESR1* mRNA expression levels in breast tumours stratified by molecular subtype are similar in the New Zealand and non-New Zealand cohorts.



We next determined whether a number of commonly-used gene expression signatures used to stratify tumours showed similar stratification of breast tumours from New Zealand women as from non-New Zealand cohorts.[41–43] To do this we used the first principal component (PC) calculated from the expression data represented by the genes in the selected signature (see Methods) to act as an **indicator of pathway activity** of that gene signature. We assessed similarities and differences between cohorts in terms of the clinical features that correlated with the 1st PC of three well-known gene signatures (see Figure 2 for a short description of these gene signatures): (1) the PAM50 signature,[41] (2) the Genomic Grade Index (GGI)[42,44] and (3) the ER attractor.[43] The results of the analyses were displayed as heatmaps (see Figure 3 for a description of heatmaps) where the tumours were sorted based on the magnitude of the 1st PC of each signature that was analysed.

Figure 2: Description of gene signatures.

PAM50

The PAM50 signature is a 50-gene subtype predictor that is used to identify the molecular subtype of breast tumours.[41] It has recently been approved for use in the clinic (as the Prosigna diagnostic assay) to indicate long-term risk of recurrence in patients with ER+ breast tumours treated with endocrine therapy, in conjunction with other clinicopathological factors.

Genomic Grade Index

The Genomic Grade Index (GGI) is a 97 gene signature that is strongly associated with histological grade 1 and grade 3 breast tumours.[42] It also stratifies histological grade 2 breast tumours into high and low risk of recurrence within five years and is currently being evaluated in clinical trials. The GGI is also able to identify better endocrine therapy-treated patients with very low risk of distant recurrence at 10 years, better than histological grade.[44]

ER Attractor

Gene expression signatures can serve as surrogates of cancer phenotypes or signalling pathways. The ER attractor gene signature is a surrogate signature for the ER signalling pathway, and the top 50 ranked genes in this signature consist of numerous genes that are strongly co-expressed with *ESR1*, such as *CA12*, *TFF1*, *XBP1*, *NAT1*, *GATA3* and *FOXA1*.[43]

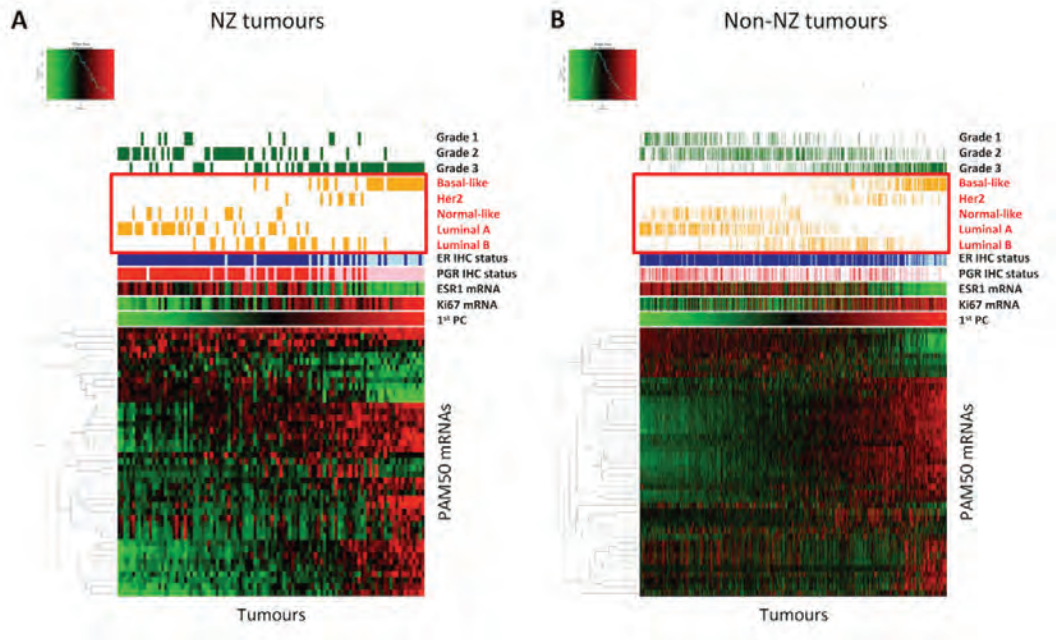
Figure 3: Guide to heatmaps.

Heatmaps are useful for visualising the expression (ie, the use) of a set of genes in a set of tumours. They show individual genes as rows and individual tumours as columns, with clinical features of each tumour such as ER status, tumour grade and molecular subtype indicated above the heatmap. For each gene in each tumour, heatmaps will show the level of expression (ie, the level of use) represented visually by a gradient of colours from red (high gene expression) through black (median gene expression) to green (low gene expression). In heatmaps the genes and the tumours are sometimes clustered based on their similarity, with the relationships between them summarised in the form of a tree diagram (dendrogram) at the top and/or side of the heatmap. Clinical and pathological information can be added to heatmaps, to visualise links between gene expression and clinicopathological information.

Figures 4A and B showed that the PAM50 signature did indeed stratify the New Zealand breast tumours into molecular subtypes (boxed in red), similar to that seen in the non-New Zealand cohort. For example, in both cohorts the basal-like breast tumours (boxed in red) were associated with a higher magnitude of the 1st PC (green) or higher inferred activity of the PAM50 signature and these tumours also tended to be histological grade 3 (Figure 4). Similarly, for both cohorts the luminal A tumours were mostly ER+, associated with a lower magnitude of the 1st PC (red) or lower inferred activity of the PAM50 signature, and tended to be of lower histological grade (Figure 4).

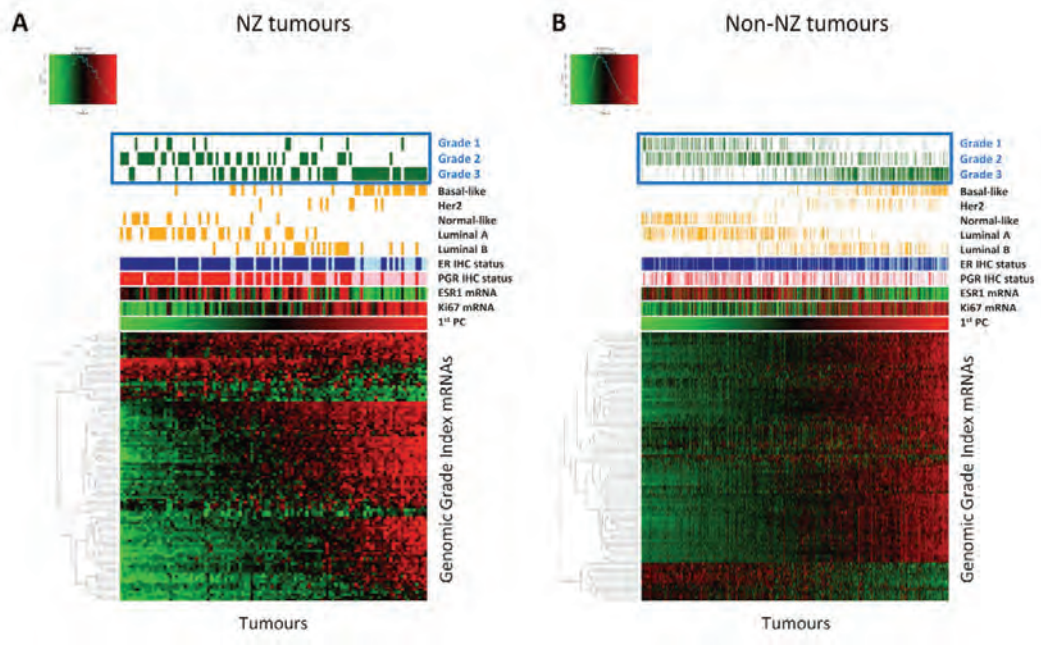
The expression profiles of 50 genes comprising the PAM50 signature are shown as heatmaps for (A) New Zealand and (B) non-New Zealand cohorts.[41] The data are sorted by the magnitude of the 1st PC (principal component) of the PAM50 signature from low (green) to high (red). The PAM50 genes were hierarchically clustered using Euclidean distance and the Ward agglomeration method. Gene expression data were Z-transformed and expression levels mapped to colours on a red-black-green scale as indicated by the colour key at the top left of the plot. Shown above the heatmaps, indicated by multi-coloured solid bars are clinical and pathological information, with histologic grade, molecular subtype, ER status (IHC), PGR status (IHC), *ESR1* mRNA expression (microarray), Ki67 mRNA expression (microarray; Ki67 protein is a marker of proliferation) and the 1st PC magnitude for the PAM50 signature genes for each tumour.

Figure 4: The PAM50 signature stratifies both the New Zealand and non-New Zealand cohort of tumours into their molecular subtypes.



NZMJ 27 October 2017, Vol 130 No 1464
ISSN 1175-8716
www.nzma.org.nz/journal

Figure 5: The Genomic Grade Index signature stratifies the New Zealand and non-New Zealand cohort of tumours into their molecular tumour grades.



Analysis of the Genomic Grade Index (GGI) signature showed that this gene signature similarly separated both the New Zealand breast tumours and the non-New Zealand tumours into three molecular grades (boxed in blue) in Figures 5A and B. Similar to Figure 4, in both cohorts the basal-like breast tumours were associated with a higher magnitude of the 1st PC (green) or higher inferred activity of the GGI signature and these tumours also tended to be histological grade 3 (Figure 5).

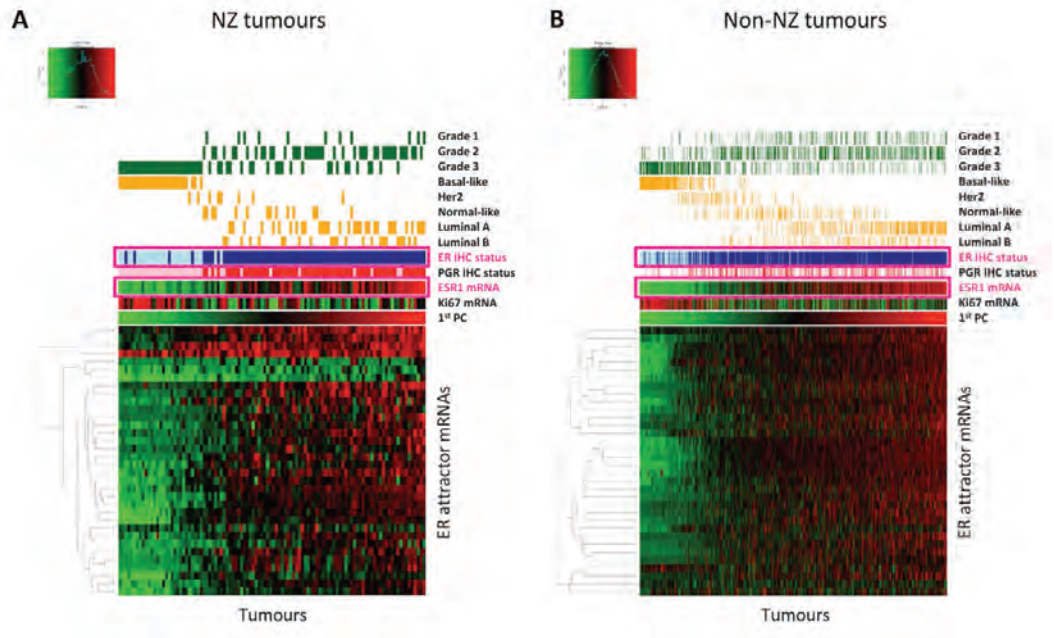
The expression profiles of 97 genes comprising the Genomic Grade Index (GGI) are shown as heatmaps for (A) New Zealand and (B) non-New Zealand cohorts.[42] The data are sorted by the magnitude of the 1st PC (principal component) of the GGI signature from low (green) to high (red). The GGI genes were hierarchically clustered using Euclidean distance and the Ward agglomeration method. Gene expression data were Z-transformed and expression levels mapped to colours on a red-black-green scale as indicated by the colour key at the top left of the plot. Shown above the heatmaps, indicated by multi-coloured solid bars are clinical and pathological information, with histologic grade, molecular subtype, ER status (IHC), PGR status (IHC), *ESR1* mRNA expression (microarray), Ki67 mRNA expression (microarray; Ki67 protein is a marker of proliferation) and the 1st PC for the GGI signature for each tumour.

Analysis of the expression of genes associated with the ER attractor[43] signature predominantly stratified both New Zealand and non-New Zealand cohorts of breast tumours by ER status, as indicated by both ER positivity by IHC and by expression of *ESR1* mRNA (boxed in pink) in Figures 6A and B. As seen in Figure 6, almost all IHC ER-negative tumours were of the basal-like subtype.

The expression profiles of 50 genes comprising the ER attractor signature are shown as heatmaps for (A) New Zealand and (B) non-New Zealand cohorts.[43] The data are sorted by the magnitude of the 1st PC (principal component) of the ER attractor signature from low (green) to high (red). The ER attractor genes were hierarchically clustered using Euclidean distance and the Ward agglomeration method. Gene expression data were Z-transformed and expression levels mapped to colours on a red-black-green scale as indicated by the colour key at the top left of the plot. Shown above the heatmaps, indicated by multi-coloured solid bars are clinical and pathological information, with histologic grade, molecular subtype, ER status (IHC), PGR status (IHC), ESR1 mRNA expression

NZMJ 27 October 2017, Vol 130 No 1464
ISSN 1175-8716

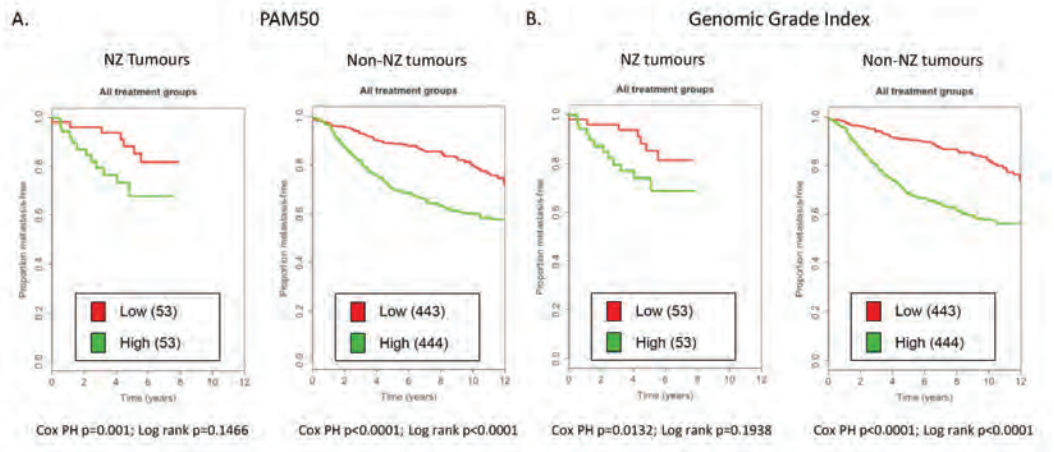
Figure 6: The ER attractor gene signature stratifies both the New Zealand and non-New Zealand cohort of tumours by ER status and ESR1 mRNA expression.



(microarray), Ki-67 mRNA expression (microarray; Ki67 protein is a marker of proliferation) and the 1st PC for the ER attractor signature for each tumour.

We next evaluated how two of these prognostic signatures are associated with breast cancer patient prognosis in New Zealand women and in the non-New Zealand cohort. We analysed the 1st PC of genes comprising the PAM50 and GGI signatures. In both cases the PC magnitude was significantly associated with distant-metastasis events, as illustrated by Kaplan-Meier curves (Figure 7). These results showed that even with small numbers of New Zealand samples, both the PAM50 (Figure 7A) and the GGI (Figure 7B) signatures, were significantly associated with patient prognosis in the New Zealand and non-New Zealand cohorts. As described in the Methods, we assessed time to a distant metastastic event for up to eight years in New Zealand cohort (median follow-up time=4.43 years, n=106), and up to 12 years in the non-New Zealand cohort (median follow-up time=6.17 years, n=756). For this analyses, patients in each cohort (New Zealand and non-New Zealand) were divided into two groups based on the level of expression of each gene signature for both the PAM50 and GGI.

Figure 7: Kaplan-Meier curves showing prognosis of New Zealand and non-New Zealand breast cancer patients classified using the PAM50 and Genomic Grade Index signatures.



Kaplan-Meier curves were plotted using the 1st principal component (PC) of (A) the PAM50 or (B) the GGI signatures for patients from all treatment groups. Patients with tumours that have lower than the median 1st PC of either of the signatures ("low") are shown as red curves (New Zealand n=53, non-New Zealand n=443), and patients with tumours that have greater than the median 1st PC of either the signatures ("high") are shown as green curves (New Zealand n=53, non-New Zealand n=444).

Discussion

In this study, we have generated a gene expression dataset of 106 prospectively collected fresh frozen breast tumours from New Zealand women using standard RNA extraction methods and microarray analysis techniques. This dataset, with associated clinical data, is available for other investigators to use. Certain clinico-pathological parameters such as patient age, patient tumour size, lymph node status and proportion of histological grades differed significantly between this cohort and non-New Zealand cohorts (Table 1). However, when analysed at the level of mRNA gene expression, we observed that the New Zealand and non-New Zealand cohorts share multiple clinical associations with common gene expression signatures involving the ERα signalling pathway, the ERBB2 signalling pathway, proliferation-based pathways and distribution patterns of *ESR1* mRNA expression between breast tumour subtypes.

Our study is the first to analyse the gene expression of breast tumours from New Zealand women. The gene expression data within our New Zealand cohort showed significant gene expression differences between histopathology-determined ER+ and ER- tumours, and ERBB2+ and ERBB2- tumours. The genes differentially expressed between ER+ and ER- tumours in our New Zealand cohort (Supplementary Table 2) consisted of genes that have already been described in the literature as differentially expressed in breast tumours dependent on ER status, such as *PGR, TFF1, AGR2, CA12, ERBB4, FOXA1* and *GATA3*.[45–48] Some of these genes are overexpressed in breast carcinomas (*CA12, FOXA1, GATA3, TFF1, TFF3*),[49] some are co-expressed with the ER and PGR (*EVL, SLC39A6, TBC1D9*)[50] and other genes have been shown to be GATA3 targets (*DACH1, THSD4*).[51]

The genes differentially regulated between ERRB2+ and ERRB2- tumours included 17 upregulated probe sets, and represented 10 genes including *ERBB2* itself, that were all located at locus 17q21 (*CDK12, CWC25, FBXL20, GRB7, GSDMB, MIEN1, ORMDL3, PCGF2* and *PGAP3*; Supplementary Table 2). These findings are consistent with other published studies that have shown that the genes on this particular locus, together with *ERBB2* are both overexpressed and amplified in breast tumours,[52–54] and breast cancer cell lines,[55,56] when analysed using various methods: IHC, FISH, array comparative genomic hybridisation and gene expression. Many of these genes are also required for the proliferation and survival of ERBB2+ breast cancer cells.[55]

Strikingly similar associations between each tumour's pathology and underlying gene expression makeup were observed in a parallel analysis between our New Zealand patient cohort and a compiled international non-New Zealand cohort.

These similarities include multi-way relationships between the expression of breast cancer-associated gene sets (PAM50Genomic Grade Index and ER attractor), with ER status, histological tumour grade, *ESR1* mRNA expression, molecular subtype and Ki67 mRNA (*MKI67* gene; Figures 4–6). The expression of the Ki67 protein in breast tumours is a useful proliferation marker, and is used clinically to assess for prognosis and response to endocrine therapy in patients.[59–61] Our analyses show that individually, the expression of *MKI67* mRNA is associated with both histological grade and with the Genomic Grade Index, and inversely associated with *ESR1* mRNA expression. We also noted similarities in expression patterns of *ESR1* mRNA between the molecular subtypes in both cohorts (Figure 1) and similarities in patient outcome in both cohorts when stratified using the PAM50 and GGI signatures (Figure 7). However, despite this similarity between New Zealand and non-New Zealand tumours at the gene pathway level, a comparison between tumours of histological Grade 1 (n=11) and Grade 3 (n=53) in the New Zealand cohort failed to identify genes significantly enriched for proliferation-based functions, possibly due to the small cohort size.

We note that like many non-New Zealand breast cancer gene expression datasets, the New Zealand breast tumours were

biased towards a larger size than the tumours commonly identified today by mammographic screening. Although tumour size appears not to strongly influence breast tumour gene expression patterns,[57] it is possible that small breast tumours may have a different biology, which reduces their clinical progression.[58] Therefore, additional prospective gene expression profiling studies using smaller tumours than were collected historically in New Zealand may be needed to fully understand the smaller and biologically different breast cancers detected by screening programs. As there might be difficulties in obtaining diagnostically spare tumour tissue from small breast tumours for research purposes, prospective studies that utilise tissue from biopsies such as fine needle aspiration (FNA) may need to be considered.

As previously mentioned, most of the studies on breast cancer in New Zealand have mainly focused on health system inequalities, various life style factors, epidemiology and ethnic and socio-economic associations for breast cancer risk. The small number of studies that have attempted to investigate any differences in breast cancer biology between the various ethnic populations such as between New Zealand Māori and New Zealand European women have been contradictory and inconclusive; albeit with small patient numbers for robust investigation.[62–67] In the future, carefully designed and adequately powered genomic studies, using census-compatible ethnicity data may be able to exclude or confirm whether there are indeed biological factors associated with low breast cancer survival of New Zealand Māori women.[68] However, given the past negative impact on Māori of some New Zealand genomic studies analysing ethnicity,[69–71] future studies will require careful design to capture ancestry and ethnicity accurately, careful interpretation, and should be led by Māori researchers, to avoid inaccurate conclusions or misinterpretation.

In conclusion, in this study we describe the results of the first microarray analysis conducted on a large number of breast tumours from New Zealand women and make the data available for others to use. The clinical-gene expression relationships evident in New Zealand patients were consistent with published breast cancer gene expression data from outside New Zealand, highlighting the similarities between breast cancer in New Zealand and other regions. This consistency is reassuring, and it provides confidence that the gene expression data generated from our New Zealand cohort can be collated with other published, international gene expression cohorts and used in validation studies. With the introduction of internationally-developed genomic tests into the clinic in New Zealand, it is reassuring that breast tumours from New Zealand women exhibit similar molecular features to international cohorts, suggesting these tests should have the same relevance to clinical practices in New Zealand as they do overseas. Since the technically robust diagnostic test Prosigna (the PAM50 classifier) is clearly associated with ER+ patient prognosis in the clinic, our findings invite further investigation of the potential clinical utility of the Prosigna assay and related tests in New Zealand. We hope that future analysis of this New Zealand breast tumour data, potentially alongside non-New Zealand datasets, will provide valuable insights into breast cancer that ultimately improve patient outcomes in New Zealand.

Clinical summary

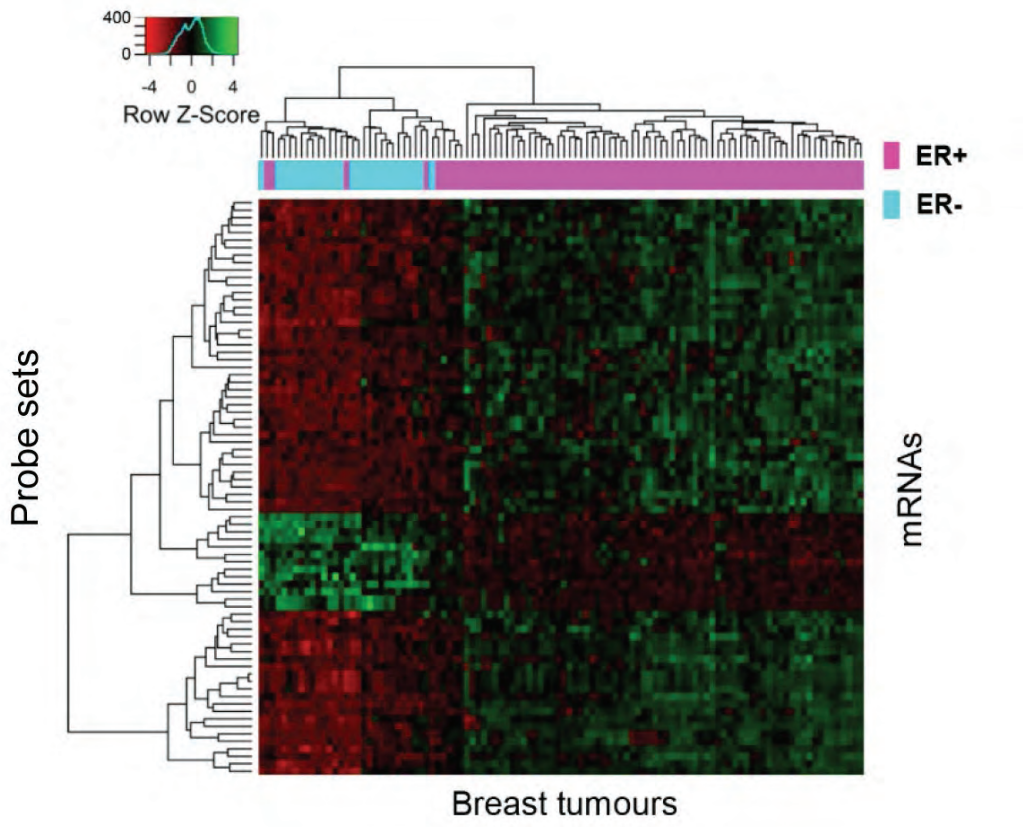
When analysed at the gene expression level, breast tumours from New Zealand and non-New Zealand cohorts share multiple clinical associations with common gene expression signatures involving ER status and ERBB2 status.

These similarities include multi-way relationships between the expression of genes that constitute the PAM50 signature, Genomic Grade Index (GGI), and the ER attractor (an oestrogen pathway-associated gene signature), as well as similar distribution patterns of *ESR1* mRNA expression between breast tumour subtypes and patient outcome using the PAM50 and GGI signatures.

Our findings suggest that breast tumours from New Zealand women exhibit similar molecular features to international cohorts, suggesting that genomic tests and gene expression signatures should have the same relevance to clinical practice in New Zealand as they do overseas.

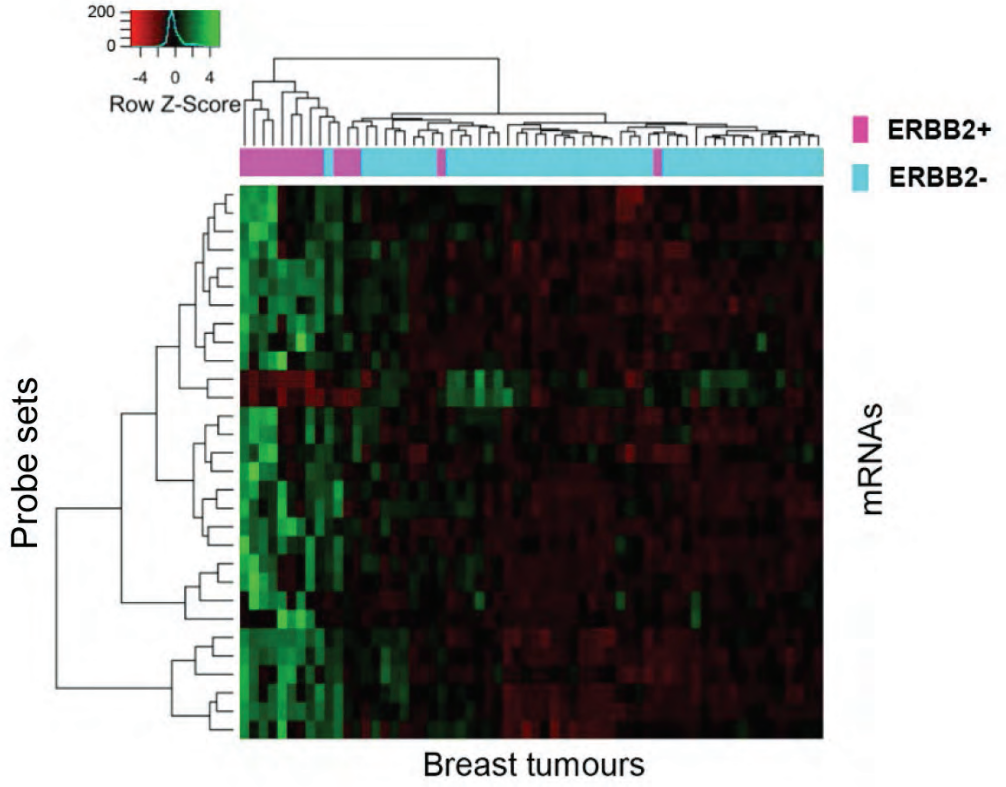
Appendix

Supplementary Figure 1: Differentially expressed mRNAs between ER+ and ER- tumours in the New Zealand cohort.



Heatmap depicting significantly regulated mRNAs between IHC-determined ER+ and ER- tumours in the New Zealand breast cancer cohort (statistical cutoffs used for differential expression were adj-$P \leq 1 \times 10^{-12}$ and absolute log fold change ≥1.5). ER+ tumours are represented by magenta bars above heatmap; ER- tumours are represented by light blue bars above heatmap. Probe sets were hierarchically clustered using Euclidean distance and the Ward agglomeration method (probe sets listed in Supplementary Table 2). Expression data for each gene was Z-transformed across tumours and expression levels mapped to colours on a red-black-green scale as indicated by the colour key at the top left of the plot.

Supplementary Figure 2: Differentially expressed mRNAs between ERBB2+ and ERBB2- tumours in the New Zealand cohort.



Gene expression in New Zealand breast tumours known by pathological analysis to be ERBB2+ (n=14) and ERBB2- tumours (n=48) was compared (adj-*P*=0.0001). ERBB2+ and ERBB2- tumours are represented by magenta bars above heatmap and light blue bars above heatmap respectively. Tumours and genes were hierarchically clustered using Euclidean distance and the Ward agglomeration method (gene probe sets listed in Supplementary Table 3). Gene expression data was Z-transformed and expression levels mapped to colours on a red-black-green scale as indicated by the colour key at the top left of the plot.

NZMJ 27 October 2017, Vol 130 No 1464
ISSN 1175-8716

Competing interests:
Nil.

Acknowledgements:
We wish to thank Mr Liam Williams from the Centre of Genomics and Proteomics, Auckland, New Zealand for his assistance with microarray hybridisation. For New Zealand breast tumour analysis we are grateful to: (i) the New Zealand Breast Cancer Registers supported by the New Zealand Breast Cancer Foundation for their provision of patient follow-up data and Associate Prof Lance Miller (Wake Forest University, North Carolina, USA) for assistance with assembly of breast cancer cohorts, (ii) surgical teams in Auckland led by Mr Wayne Jones and Dr Vanessa Blair for collection of patient samples at the time of surgery following informed patient consent, (iii) Dr Reena Ramsaroop and Dr Robyn Oldfield for pathological assistance, (iv) Mrs Helen Morrin, Dr Bridget Robinson and staff of the Christchurch Tissue Bank for provision of tissue bank breast cancer material and accompanying clinical data. Thank you to Ms Sandra Fitzgerald for assistance with microarray hybridisation and Dr Deborah Wright for assistance with the collation of clinical data. We would also like to thank Breast Cancer Cure, New Zealand, for providing support for this work to Drs Lasham, Blenkiron and Print. We would also like to thank the University of Auckland and The Cancer Society of New Zealand for supporting Dr Muthukaruppan.

Author information:
Anita Muthukaruppan, Research Fellow, Department of Obstetrics and Gynaecology, Faculty of Medical and Health Sciences, The University of Auckland, Auckland;
Annette Lasham, Senior Research Fellow, Department of Molecular Medicine and Pathology, Faculty of Medical and Health Sciences, The University of Auckland, Auckland;
Cherie Blenkiron, Senior Research Fellow, Department of Molecular Medicine and Pathology, Faculty of Medical and Health Sciences, The University of Auckland, Auckland;
Kathryn J Woad, Lecturer, School of Veterinary Medicine and Science, Sutton Bonington Campus, The University of Nottingham, Leicestershire, United Kingdom;
Michael A Black, Associate Professor, Department of Biochemistry, University of Otago, Dunedin; Nicholas Knowlton, PhD Student, Department of Molecular Medicine and Pathology, Faculty of Medical and Health Sciences, The University of Auckland, Auckland;
Nicole McCarthy, Associate Professor, ICON Cancer Care Wesley, Wesley Medical Centre, Queensland, Australia; Michael P Findlay, Professor, Discipline of Oncology, Faculty of Medical and Health Sciences, The University of Auckland, Auckland;
Cristin G Print, Professor, Department of Molecular Medicine and Pathology, Maurice Wilkins Centre, The University of Auckland Bioinformatics Institute, The University of Auckland, Auckland; Andrew N Shelling, Professor, Department of Obstetrics and Gynaecology, The University of Auckland, Auckland.

Corresponding author:
Dr Anita Muthukaruppan, Department of Obstetrics and Gynaecology, The University of Auckland, Room 502–201K, Level 2, Building 502, Faculty of Medical and Health Sciences, The University of Auckland, 85 Park Road, Grafton, Auckland 1023.
a.muthukaruppan@auckland.ac.nz

URL:
http://www.nzma.org.nz/journal/read-the-journal/all-issues/2010-2019/2017/vol-130-no-1464-27-october-2017/7396

REFERENCES:

1. Benson JR, Jatoi I, Keisch M, Esteva FJ, Makris A, Jordan VC. Early breast cancer. Lancet 2009; 373:1463–79.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International journal of cancer Journal international du cancer 2010; 127:2893–917.
3. New Zealand Ministry of Health. Cancer: New registrations and deaths 2012. In: Health WMo, ed. 6/10/2015 ed. Wellington 2015.
4. Cunningham R, Shaw C, Blakely T, Atkinson J, Sarfati D. Ethnic and socioeconomic trends in breast cancer incidence

in New Zealand. BMC cancer 2010; 10:674.

5. Waldon J, Lamb DS, Delahunt B, et al. A comparison of cancer statistics in New Zealand and Australia: 1996–2007. The New Zealand Medical Journal 2014; 127:20–9.
6. Alafeishat L, Elwood M, Ioannides S. Cancer mortality and incidence trends comparing New Zealand and Australia for the period 2000–2007. The New Zealand Medical Journal 2014; 127:9–19.
7. Ellis MJ, Perou CM. The genomic landscape of breast cancer as a therapeutic roadmap. Cancer discovery 2013; 3:27–34.
8. Reis-Filho JS, Pusztai L. Gene expression profiling in breast cancer: classification, prognostication, and prediction. Lancet 2011; 378:1812–23.
9. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016; 534:47–54.
10. Gao Y, Jones A, Fasching PA, et al. The integrative epigenomic-transcriptomic landscape of ER positive breast cancer. Clinical epigenetics 2015; 7:126.
11. Morganella S, Alexandrov LB, Glodzik D, et al. The topography of mutational processes in breast cancer genomes. Nature communications 2016; 7:11383.
12. Tishchenko I, Milioli HH, Riveros C, Moscato P. Extensive Transcriptomic and Genomic Analysis Provides New Insights about Luminal Breast Cancers. PloS one 2016; 11:e0158259.
13. Ali HR, Rueda OM, Chin SF, et al. Genome-driven integrated classification of breast cancer validated in over 7,500 samples. Genome biology 2014; 15:431.
14. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proceedings of the National Academy of Sciences of the United States of America 2003; 100:8418–23.
15. Mertins P, Mani DR, Ruggles KV, et al. Proteogenomics connects somatic mutations to signalling in breast cancer. Nature 2016; 534:55–62.
16. Pereira B, Chin SF, Rueda OM, et al. The somatic mutation profiles of 2,433 breast cancers refines their genomic and transcriptomic landscapes. Nature communications 2016; 7:11479.
17. Gatza ML, Silva GO, Parker JS, Fan C, Perou CM. An integrated genomics approach identifies drivers of proliferation in luminal-subtype human breast cancer. Nature genetics 2014; 46:1051–9.
18. The Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. Nature 2012; 490:61–70.
19. Harbeck N, Sotlar K, Wuerstlein R, Doisneau-Sixou S. Molecular and protein markers for clinical decision making in breast cancer: today and tomorrow. Cancer treatment reviews 2014; 40:434–44.
20. Gyorffy B, Hatzis C, Sanft T, Hofstatter E, Aktas B, Pusztai L. Multigene prognostic tests in breast cancer: past, present, future. Breast cancer research : BCR 2015; 17:11.
21. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection. Proceedings of the National Academy of Sciences of the United States of America 2001; 98:31–6.
22. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL http://www.R-project.org/, Vienna, Austria 2013.
23. Gautier L, Cope L, Bolstad BM, Irizarry RA. affy--analysis of Affymetrix GeneChip data at the probe level. Bioinformatics (Oxford, England) 2004; 20:307–15.
24. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research 2015; 43:e47.
25. Irizarry RA, Hobbs B, Collin F, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics (Oxford, England) 2003; 4:249–64.
26. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing.
27. Araki H, Knapp C, Tsai P, Print C. GeneSetDB: A comprehensive meta-database, statistical and visualisation framework for gene set analysis. FEBS open bio 2012; 2:76–82.
28. Chang JT, Nevins JR. GATHER: a systems approach to interpreting genomic signatures. Bioinformatics (Oxford, England) 2006; 22:2926–33.
29. Pawitan Y, Bjohle J, Amler L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts.

Breast cancer research : BCR 2005; 7:R953–64.

30. Miller LD, Smeds J, George J, et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. Proceedings of the National Academy of Sciences of the United States of America 2005; 102:13550–5.

31. Ivshina AV, George J, Senko O, et al. Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. Cancer research 2006; 66:10292–301.

32. Loi S, Haibe-Kains B, Desmedt C, et al. Definition of clinically distinct molecular subtypes in estrogen receptor-positive breast carcinomas through genomic grade. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2007; 25:1239–46.

33. Loi S, Haibe-Kains B, Desmedt C, et al. Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. BMC genomics 2008; 9:239.

34. Desmedt C, Piette F, Loi S, et al. Strong time dependence of the 76-gene prognostic signature for node-negative breast cancer patients in the TRANSBIG multicenter independent validation series. Clinical cancer research : an official journal of the American Association for Cancer Research 2007; 13:3207–14.

35. Lasham A, Samuel W, Cao H, et al. YB-1, the E2F pathway, and regulation of tumor cell growth. Journal of the National Cancer Institute 2012; 104:133–46.

36. Marot G, Foulley JL, Mayer CD, Jaffrezic F. Moderated effect size and P-value combinations for microarray meta-analyses. Bioinformatics (Oxford, England) 2009; 25:2692–9.

37. Hu Z, Fan C, Oh DS, et al. The molecular portraits of breast tumors are conserved across microarray platforms. BMC genomics 2006; 7:96.

38. West M, Blanchette C, Dressman H, et al. Predicting the clinical status of human breast cancer by using gene expression profiles. Proceedings of the National Academy of Sciences of the United States of America 2001; 98:11462–7.

39. Therneau TM, Grambsch PM. Modeling survival data: extending the Cox model. 13 ed. New York: Springer; 2000.

40. Muthukaruppan A, Lasham A, Woad KJ, et al. Multimodal Assessment of Estrogen Receptor mRNA Profiles to Quantify Estrogen Pathway Activity in Breast Tumors. Clinical breast cancer 2016.

41. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 2009; 27:1160–7.

42. Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. Journal of the National Cancer Institute 2006; 98:262–72.

43. Cheng WY, Ou Yang TH, Anastassiou D. Biomolecular events in cancer revealed by attractor metagenes. PLoS computational biology 2013; 9:e1002920.

44. Ignatiadis M, Azim HA Jr., Desmedt C, et al. The Genomic Grade Assay Compared With Ki67 to Determine Risk of Distant Breast Cancer Recurrence. JAMA oncology 2016; 2:217–24.

45. Li R, Campos J, Iida J. A Gene Regulatory Program in Human Breast Cancer. Genetics 2015; 201:1341–8.

46. Han W, Jones FE. HER4 selectively coregulates estrogen stimulated genes associated with breast tumor cell proliferation. Biochemical and biophysical research communications 2014; 443:458–63.

47. Barnett DH, Sheng S, Charn TH, et al. Estrogen receptor regulation of carbonic anhydrase XII through a distal enhancer in breast cancer. Cancer research 2008; 68:3505–15.

48. Wright TM, Wardell SE, Jasper JS, et al. Delineation of a FOXA1/ERalpha/AGR2 regulatory loop that is dysregulated in endocrine therapy-resistant breast cancer. Molecular cancer research : MCR 2014; 12:1829–39.

49. Davidson B, Stavnes HT, Holth A, et al. Gene expression signatures differentiate ovarian/peritoneal serous carcinoma from breast carcinoma in effusions. Journal of cellular and molecular medicine 2011; 15:535–44.

50. Andres SA, Wittliff JL. Co-expression of genes with estrogen receptor-alpha and progesterone receptor in human breast carcinoma tissue. Hormone molecular biology and clinical investigation 2012; 12:377–90.

51. Cohen H, Ben-Hamo R, Gidoni M, et al. Shift in GATA3 functions, and GATA3 mutations, control progression and clinical presentation in breast cancer. Breast cancer research : BCR 2014; 16:464.

52. Sircoulomb F, Bekhouche I, Finetti P, et al. Genome profiling of ERBB2-amplified breast cancers. BMC cancer 2010; 10:539.

53. Staaf J, Jonsson G, Ringner M, et al. High-resolution genomic and expression analyses of copy number alterations in HER2-amplified breast cancer. Breast cancer research: BCR 2010; 12:R25.

54. Lamy PJ, Fina F, Bascoul-Mollevi C, et al. Quantification and clinical relevance of gene amplification at chromosome 17q12–q21 in human epidermal growth factor receptor 2-amplified breast cancers. Breast cancer research: BCR 2011; 13:R15.

55. Sahlberg KK, Hongisto V, Edgren H, et al. The HER2 amplicon includes several genes required for the growth and survival of HER2 positive breast cancer cells. Molecular oncology 2013; 7:392–401.

56. Shiu KK, Wetterskog D, Mackay A, et al. Integrative molecular and functional profiling of ERBB2-amplified breast cancers identifies new genetic dependencies. Oncogene 2014; 33:619–31.

57. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proceedings of the National Academy of Sciences of the United States of America 2003; 100:10393–8.

58. Welch HG, Prorok PC, O'Malley AJ, Kramer BS. Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness. The New England journal of medicine 2016; 375:1438–47.

59. Duffy MJ, Harbeck N, Nap M, et al. Clinical use of biomarkers in breast cancer: Updated guidelines from the European Group on Tumor Markers (EGTM). European journal of cancer (Oxford, England: 1990) 2017; 75:284–98.

60. Denkert C, Budczies J, von Minckwitz G, Wienert S, Loibl S, Klauschen F. Strategies for developing Ki67 as a useful biomarker in breast cancer. Breast (Edinburgh, Scotland) 2015; 24:S67–72.

61. Sinn HP, Schneeweiss A, Keller M, et al. Comparison of immunohistochemistry with PCR for assessment of ER, PR, and Ki-67 and prediction of pathological complete response in breast cancer. BMC cancer 2017; 17:124.

62. Seneviratne S, Lawrenson R, Scott N, Kim B, Shirley R, Campbell I. Breast cancer biology and ethnic disparities in breast cancer mortality in New Zealand: a cohort study. PloS one 2015; 10:e0123523.

63. Robson B, Purdie G, Cormack D. Unequal Impact II: Maori and Non-Maori Cancer Statistics by Deprivation and Rural-Urban Status 2002–2006. In: Health WMo, ed. Wellington 2010.

64. Teng AM, Atkinson J, Disney G, et al. Ethnic inequalities in cancer incidence and mortality: census-linked cohort studies with 87 million years of person-time follow-up. BMC cancer 2016; 16:755.

65. Seneviratne S, Lawrenson R, Harvey V, et al. Stage of breast cancer at diagnosis in New Zealand: impacts of socio-demographic factors, breast cancer screening and biology. BMC cancer 2016; 16:129.

66. Weston MK, Moss DP, Stewart J, Hill AG. Differences in breast cancer biological characteristics between ethnic groups in New Zealand. Breast cancer research and treatment 2008; 111:555–8.

67. Dachs GU, Kano M, Volkova E, et al. A profile of prognostic and molecular factors in European and Maori breast cancer patients. BMC cancer 2010; 10:543.

68. Sarfati D, Robson B. Equitable cancer control: better data needed for indigenous people. The lancet oncology 2015; 16:1442–4.

69. Whittle PM. Health, inequality and the politics of genes. The New Zealand medical journal 2010; 123:67–75.

70. Merriman T, Cameron V. Risk-taking: behind the warrior gene story. The New Zealand medical journal 2007; 120:U2440.

71. Lea R, Chambers G. Monoamine oxidase, addiction, and the "warrior" gene hypothesis. The New Zealand Medical Journal 2007; 120:U2441.