

# Cancer Research

## High Expression of Lymphocyte-Associated Genes in Node-Negative HER2+ Breast Cancers Correlates with Lower Recurrence Rates

Gabriela Alexe, Gul S. Dalgin, Daniel Scandfeld, et al.

*Cancer Res* 2007;67:10669-10676. Published online November 15, 2007.

**Updated Version**

Access the most recent version of this article at:  
doi:[10.1158/0008-5472.CAN-07-0539](https://doi.org/10.1158/0008-5472.CAN-07-0539)

**Supplementary Material**

Access the most recent supplemental material at:  
<http://cancerres.aacrjournals.org/content/suppl/2007/11/07/67.22.10669.DC1.html>

**Cited Articles**

This article cites 33 articles, 13 of which you can access for free at:  
<http://cancerres.aacrjournals.org/content/67/22/10669.full.html#ref-list-1>

**Citing Articles**

This article has been cited by 14 HighWire-hosted articles. Access the articles at:  
<http://cancerres.aacrjournals.org/content/67/22/10669.full.html#related-urls>

**E-mail alerts**

[Sign up to receive free email-alerts](#) related to this article or journal.

**Reprints and Subscriptions**

To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at [pubs@aacr.org](mailto:pubs@aacr.org).

**Permissions**

To request permission to re-use all or part of this article, contact the AACR Publications Department at [permissions@aacr.org](mailto:permissions@aacr.org).

# High Expression of Lymphocyte-Associated Genes in Node-Negative HER2+ Breast Cancers Correlates with Lower Recurrence Rates

Gabriela Alexe,<sup>1,2</sup> Gul S. Dalgin,<sup>3</sup> Daniel Scandell,<sup>1</sup> Pablo Tamayo,<sup>1</sup> Jill P. Mesirov,<sup>1</sup> Charles DeLisi,<sup>4</sup> Lyndsay Harris,<sup>5</sup> Nicola Barnard,<sup>7</sup> Maritza Martel,<sup>6</sup> Arnold J. Levine,<sup>2,8</sup> Shridar Ganesan,<sup>8</sup> and Gyan Bhanot<sup>2,4,8,9</sup>

<sup>1</sup>The Broad Institute of the Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts; <sup>2</sup>The Simons Center for Systems Biology, Institute for Advanced Study, Princeton, New Jersey; <sup>3</sup>Molecular Biology, Cellular Biology and Biochemistry Program, Department of Biomedical Engineering, Boston University, Boston, Massachusetts; <sup>4</sup>Yale Cancer Center and <sup>5</sup>Department of Pathology, Yale University School of Medicine, New Haven, Connecticut; <sup>6</sup>Robert Wood Johnson University Hospital, and <sup>7</sup>Cancer Institute of New Jersey, New Brunswick, New Jersey; and <sup>8</sup>BioMaPS Institute and Department of Biomedical Engineering, Rutgers University, Piscataway, New Jersey

## Abstract

**Gene expression analysis has identified biologically relevant subclasses of breast cancer. However, most classification schemes do not robustly cluster all HER2+ breast cancers, in part due to limitations and bias of clustering techniques used. In this article, we propose an alternative approach that first separates the HER2+ tumors using a gene amplification signal for Her2/neu amplicon genes and then applies consensus ensemble clustering separately to the HER2+ and HER2- clusters to look for further substructure. We applied this procedure to a microarray data set of 286 early-stage breast cancers treated only with surgery and radiation and identified two basal and four luminal subtypes in the HER2- tumors, as well as two novel and robust HER2+ subtypes. HER2+ subtypes had median distant metastasis-free survival of 99 months [95% confidence interval (95% CI), 83–118 months] and 33 months (95% CI, 11–54 months), respectively, and recurrence rates of 11% and 58%, respectively. The low recurrence subtype had a strong relative overexpression of lymphocyte-associated genes and was also associated with a prominent lymphocytic infiltration on histologic analysis. These data suggest that early-stage HER2+ cancers associated with lymphocytic infiltration are a biologically distinct subtype with an improved natural history.** [Cancer Res 2007;67(22):10669–76]

## Introduction

Breast cancer (BCA) is a heterogeneous group of diseases. More than 60% of tumors express the estrogen receptor (ER) and are susceptible to treatments targeting the estrogen pathway (1, 2). Twenty percent or more have HER2 amplification and are susceptible to treatments targeting the HER2 pathway (3). Measurement of ER and HER2 levels is a routine part of clinical evaluation and guides treatment. However, there remains significant heterogeneity in treatment response in tumors with similar clinical classification (1).

Gene expression analyses have provided insight into this clinical heterogeneity. One approach is to use supervised analysis of gene

expression data from large clinically annotated data sets to identify prognostic signatures associated with outcome. Van't Veer et al. (4) identified 70 genes whose expression correlated with clinical outcome. Paik et al. (5) found 21 genes that distinguish good/bad prognosis in lymph node-negative ER+ tumors treated with tamoxifen. Wang et al. (6) identified a 76-gene panel correlated with recurrence rate in node-negative patients treated with surgery and radiation.

Another approach has been to use unsupervised clustering of gene expression profiles to stratify samples into natural classes. When Perou et al. (7, 8) used this approach, they found that their samples initially divided into two groups, largely on the basis of the difference in expression of genes in the ER signaling pathway, and which correlated with expression of ER protein. With further clustering, they identified two classes in the ER+ group, which they labeled Luminal A (ER+ with good prognosis) and Luminal B (ER+ with poor prognosis). In the ER- group, they also found two subtypes, labeled Basal-like (ER-, PR-, HER2-) and HER2+ (ER-, HER2+). These subtypes were confirmed by later studies (9–13) and are now widely cited.

However, several basic questions remain unanswered. Is this classification the optimal one for clinical use? Are the classification methods robust to data and clustering method perturbation? Are the subtypes reproducible across multiple data sets? How is the clustering influenced by sampling bias and genes used? How can classification be used in clinical care?

The assignment of samples to subtypes is sensitive to sample and gene set bias. Because most breast tumors are ER+ and many genes are coordinately regulated by ER, data filtering methods (14–18) are strongly biased toward estrogen pathway genes. Clustering applied to such a gene set tends to divide samples along an ER+/ER- split, which might be a misclassification of some samples. For example, using the methods proposed in ref. 7, HER2+ samples split into two groups based on the expression of estrogen pathway genes. Further clustering assorts the HER2+/ER- cases into their own sub-cluster, and the HER2+/ER+ cases classify with the non-Luminal A, ER+ clusters. We will provide evidence from measured recurrence rates that the splitting of HER2+ tumors into ER+/ER- clusters may not reflect a biologically or clinically relevant classification and could simply arise from a combination of sampling bias and choice of clustering technique.

Our approach for avoiding this problem is to separate HER2+ samples using clinical markers. Because the HER2+ tumors represent a fraction of BCA cases, this does not impact the clustering of the remaining tumors. More importantly, however, such a separation allows the identification of the natural

**Note:** Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

G. Alexe and G.S. Dalgin are joint first authors.

**Requests for reprints:** Gyan Bhanot, BioMaPs Institute, Rutgers University, 271 Hill Center, Piscataway, NJ 08854. Phone: 732-235-9545; E-mail: gyanbhanot@gmail.com and Shridar Ganesan, Cancer Institute of New Jersey, 195 Little Albany Street, New Brunswick, NJ 08903. Phone: 732-235-5211; E-mail: ganesash@umdnj.edu.

©2007 American Association for Cancer Research.

doi:10.1158/0008-5472.CAN-07-0539

substructure of the HER2+ group. Our clustering method uses principal component analysis (PCA) and consensus ensemble clustering (19, 20). We use the data set of Wang et al. (6) with 286 BCA samples on Affymetrix U133a chips from patients who were treated only with surgery and radiation and had a median follow-up of 86 months. This data set was chosen because the tumors were all of similar stage (node-negative); all tumors were treated similarly (no systemic adjuvant treatment was administered), and there was long-term (up to 150 months) clinical follow-up. Thus, any differences in natural history among these tumors is likely to reflect the intrinsic biology and less likely to be confounded by differences in stage and treatment.

Assessment of HER2 status by immunohistochemistry or fluorescence *in situ* hybridization were not available in the data set. However, a positive correlation between the mRNA and protein levels and gene amplification of the genes in the chr17 HER2 amplicon has been shown in several studies (21, 22). We therefore identify HER2+ tumors in this data set using co-amplification of three or more of the *Her2/neu* amplicon genes: *Her2/neu*, *GRB7*, *STARD3*, and *PPARB*. After separating HER2+ and HER2- samples, each set is separately analyzed by PCA and consensus ensemble clustering (19, 20) to identify potential substructure. This analysis identified a total of eight robust BCA subtypes.

Within the HER2- cases, using the labeling schema of Perou et al. (7, 8), we found one Luminal A cluster (ER+, good prognosis), three Luminal B (ER+ with mixed prognosis; labeled LB<sub>1</sub>, LB<sub>2</sub>, LB<sub>3</sub>) and two Basal-like clusters (BA<sub>1</sub>, BA<sub>2</sub>).

A startling observation was that the HER2+ cases clustered into two clear subtypes with significantly different distant metastasis-free survival rates. The major expression signature that separates these two subtypes of HER2+ tumors was not estrogen signaling related. Instead, there was strong up-regulation of a wide variety of lymphocyte-associated genes in the good prognosis subtype. This observation suggests our main result: the presence of a lymphocytic infiltrate in node-negative HER2+ tumors is correlated with an improved natural history. The validation of our observed correlation of natural survival rate with lymphocytic infiltrate in HER2+ breast tumors, the elucidation of the biological basis of such an effect, and its impact on cure and recurrence rates following chemotherapy and adjuvant therapy are important future research directions.

## Materials and Methods

**Consensus clustering of BCA data from Wang et al.** The data set of Wang et al. (6) was obtained from the National Center for Biotechnology Information GenBank GEO database<sup>10</sup> (series entry GSE2034). Clinical information about distant metastasis, along with ER and lymph node status was available. We identify the HER2+ samples using mRNA levels of the chr17 amplicon genes *Her2/neu*, *GRB7*, *STARD3*, and *PPARB* (22). After removing the HER2+ samples, we use PCA and consensus ensemble clustering (19, 20) to split the data into two groups, which are identified as Luminal-like and Basal-like tumors based on mRNA levels of ER pathway genes and ER protein status. Next, consensus ensemble clustering (19, 20) was applied separately to the HER2+, Luminal, and Basal core clusters to identify further substructure using statistical measures to find the optimum number of sub-clusters. The assignment of samples to clusters was rendered insensitive (robust) to perturbation through bootstrapping.

Figure 1 shows the flowchart of the analysis. All details of the procedures used are given as appendices in the Supplementary Material. Appendix 1 describes the data normalization and clustering techniques used. Appendix 2 describes how the HER2+ samples were identified using HER2 amplicon gene levels. Appendix 3 describes how the gene markers that distinguish the subtypes were identified. Appendix 4 describes how to stratify an unknown sample into our subtypes.

**Consensus clustering of data of HER2+ tumors from Harris et al.** Gene expression data obtained from 21 tumors that were part of a neoadjuvant trial of trastuzumab and vinorelbine in HER2+ BCA was obtained from a previously described study (23). Semi-supervised clustering using a set of genes that stratified the two HER2+ groups in the Wang et al. data set was applied to these new data as described in Appendix 5.

**Analysis of histologic sections.** H&E-stained frozen sections of pretreatment core biopsies were available for 21 tumors described in Harris et al. (23) for which gene expression data had been obtained. H&E-stained formalin-fixed paraffin sections from the diagnostic core biopsies were also available for nine of these tumors. Each frozen section was examined independently in a blinded fashion by two BCA pathologists (N.B. and M.M.) and scored for the presence of a lymphocytic infiltrate. A rating of 1 to 3 was used, with 1 signifying minimal to no lymphocytes associated with tumor regions; 2 signifying moderate lymphocytic infiltrate present; and 3+ signifying a marked lymphocytic infiltrate. Fisher's exact test was used to generate *P* values for combined pathologists' scores.

## Results

Forty-two samples (18 ER+ and 24 ER-) had up-regulation of both *Her2/neu* and of two or more of the genes *GRB7*, *STARD3*, and *PPARB* from the *Her2/neu* amplicon. Ten samples (7 ER+ and 3 ER-) had *Her2/neu*, and one other amplicon gene was up-regulated. The first set (42 samples) was identified as a "core HER2+ cluster", and the second (10 samples) was identified as "potentially HER2+". The second set was set aside and classified only after subtype classification was completed and the rules for subtype classification for all subtypes were identified. The core HER2+ cluster was used to identify potential HER2+ substructure.

The 234 remaining samples were analyzed by PCA to identify 1,780 genes on 165 principal components that represented 85% of the variation in the data. Consensus hierarchical clustering was applied to these samples by averaging more than 50 data sets obtained from 80% random resamplings using these 1,780 genes. This identified two strong clusters. The first had 171 samples all with high levels of the estrogen pathway genes (*ESR1*, *GATA3*, *SCUBE2*, *XBPI*, *LIV-1*, *CAI2*; ref. 10) and the second had 63 samples all with down-regulated ER pathway genes.

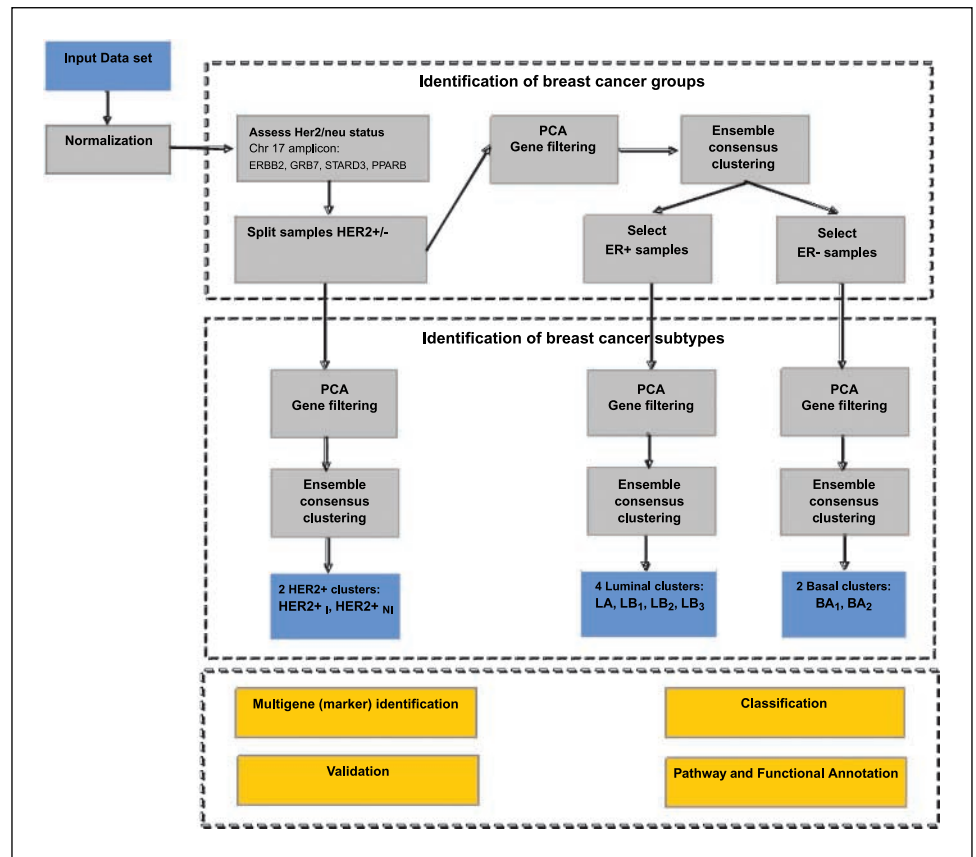
In the first cluster, 163 samples with consistent ER+ signatures from both mRNA levels and protein status were labeled "core Luminal" and 8 samples with ER+ mRNA levels but ER- by clinical testing were labeled "potentially Luminal" and set aside for later classification. In the second cluster, 49 samples with consistent ER- status were labeled "core Basal-like", and the remaining 14 samples ambiguous ER- mRNA and protein signatures were labeled "potentially Basal-like" and set aside for later classification.

We found 252 up-regulated gene probes (121 in the ER- group and 131 in the ER+ group), which distinguished the groups by signal-to-noise ratio [SNR;  $P < 10^{-5}$ , see Appendix 3 and false discovery rate (FDR)  $< 0.005$ ]. The top genes up-regulated in the ER-/ER+ groups were the known markers *CDH3*, *DSC2*, *KRT16*, and *KRT6B* for the Basal subtype and *ESR1*, *GATA3*, *SCUBE2*, *TFF3*, *FOXAI*, *CAI2*, etc. for the Luminal subtype (7, 8).

Figure 2 shows the Kaplan-Meier distant metastasis-free survival curves for the initial Luminal, Basal, and HER2+ subtypes identified

<sup>10</sup> <http://www.ncbi.nlm.nih.gov/geo/>

**Figure 1.** Flowchart of the clustering method. After normalization, mRNA levels of the genes *Her2/neu*, *GRB7*, *STARD3*, and *PPARB* on the *Her2/neu* amplicon are used to isolate the HER2+ samples. PCA and consensus ensemble clustering on the remaining samples stratify them into two clusters identified as Luminal (mostly ER+) and Basal (mostly ER-). Using consistency between gene and protein levels, we define core clusters for these two subtypes and stratify them and the HER2+ cluster further by PCA and consensus ensemble clustering. The optimum number of clusters is estimated a priori through the gap statistic, the Gini index, and the silhouette scores and validated a posteriori through class membership accuracy. The ensemble approach assigns samples to clusters based on an agreement matrix from averaging over clustering techniques (partitioning, agglomerative, probabilistic) applied to bootstrapped data sets.



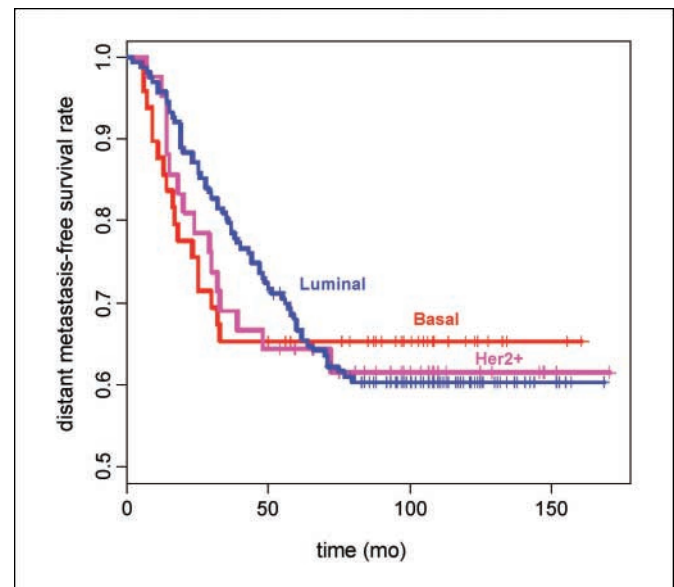
in the data. Although the Luminal subtype has the lowest and the Basal the highest 4-year disease-free survival, these differences disappear at longer follow-up for this group of early-stage tumors. These data suggest that, at least in this cohort of early-stage tumors treated without adjuvant therapy, long-term disease-free survival is not well predicted by just the three global subtypes “Luminal”, “Basal-like”, and “HER2+”.

To identify possible further substructure in the data, we repeated the PCA and robust consensus clustering analysis (19, 20) using all genes separately for each of the Luminal (163 samples), Basal (49 samples), and HER2+ (42 samples) clusters. Gap statistics (24), Gini index (19), and silhouette score (25) identified that the optimum number of sub-clusters in these clusters was 4, 2, and 2, respectively.

**Stratification of luminal BCAs.** In the Luminals, 1,674 gene probes on 114 principal components represented 85% of the variation. Consensus ensemble clustering on these probes found sub-clusters of size 48, 26, 47, and 42. The core of each sub-cluster was defined as the subset with a pair-wise agreement score across bootstrap experiments exceeding 75%. These core subgroups had 44, 22, 38, and 28 samples, respectively. The remaining (non-core) samples from each subgroup were put aside, and their classification was attempted at the end based on the patterns identified for the core subsets.

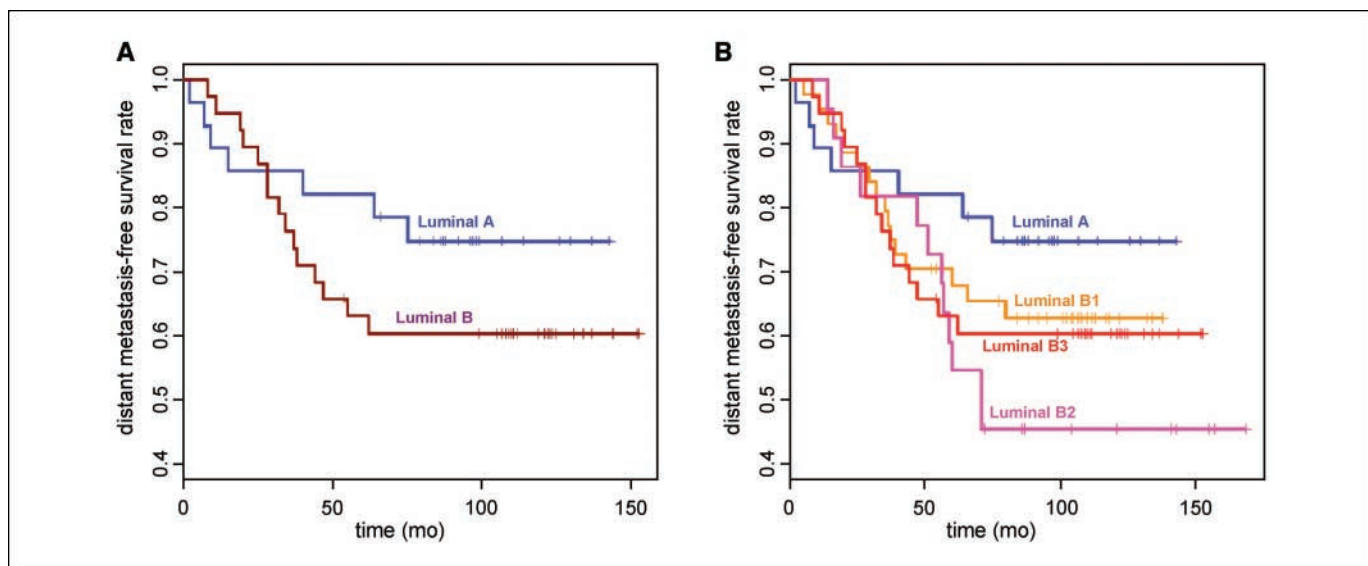
One core sub-cluster had an average distant metastasis-free survival rate above 79%. All others had average distant metastasis-free rates below 60%. Following Sorlie et al. (8), we labeled the “good prognosis” core sub-cluster Luminal A (LA) and the three “poor prognosis” clusters collectively as Luminal B (LB) or separately as LB<sub>1</sub>, LB<sub>2</sub>, and LB<sub>3</sub>. The Kaplan-Meier curves for the Luminal A and

Luminal B subtypes are shown in Fig. 3A and B. For classification, we identified probes that distinguish each core Luminal subgroup from the rest individually (based on SNR,  $P = 0.01$ , FDR = 0.1) or in combinations (also called patterns, see ref. 18).



**Figure 2.** Kaplan-Meier curves for the core Luminal, Basal, and HER2+ samples. The Luminal subtype has the best short-term distant metastasis-free survival rate, but the long-term rate is poor because of the presence of non-Luminal A subtypes (see Fig. 3).





**Figure 3.** A, Kaplan-Meier curves for Luminal A versus Luminal B.  $P = 0.14$ . B, stratification of the Luminal B into LB<sub>1</sub>, LB<sub>3</sub>, and LB<sub>2</sub> shows distinct distant metastasis-free survival.

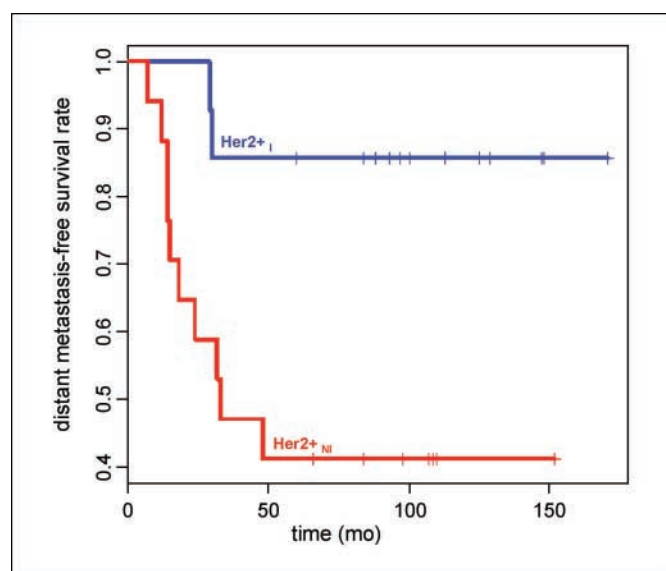
**Stratification of Basal-like BCAs.** Ensemble clustering on the 1,116 genes identified by PCA on 36 principal components stratified the core Basal cluster into two sub-clusters with 24 and 25 samples. Within these, pair-wise agreement score across bootstrap experiments identified two core subgroups BA<sub>1</sub> and BA<sub>2</sub>, with 15 and 22 samples and average recurrence rates of 32% and 41%, respectively. The BA<sub>1</sub> subset had up-regulation of genes in the Wnt signaling pathway, Immunity and defense, nucleoside, nucleotide, and nucleic acid metabolism, whereas the BA<sub>2</sub> set had up-regulated genes in the integrin signaling pathway, cell adhesion, developmental processes, cell structure, and motility (26, 27). However, the recurrence-free survival curves for BA<sub>1</sub> and BA<sub>2</sub> were not significantly different (log-rank  $P$  value = 0.6; Supplementary Figure SA6-1 in the Supplementary Material). Thus, these basal subtypes may be biologically distinct entities, but display no difference in natural history after local treatment alone.

**Substructure of HER2+ BCAs.** In the HER2+ core cluster, PCA identified 651 genes on 31 principal components representing 85% variation. Consensus ensemble clustering on these genes split the HER2+ cluster into two subgroups with 21 samples each. The core subgroups identified within these (denoted HER2<sub>+I</sub> and HER2<sub>+NI</sub>) had 14 and 17 samples with distinct distant metastasis-free survival rates of 89% and 42%, respectively. Figure 4 shows the metastasis-free survival curves for the two HER2+ subgroups. The difference in metastasis rates between HER2<sub>+I</sub> and HER2<sub>+NI</sub> is highly significant ( $P = 0.01$ ) compared with the difference between the Luminal A and the Luminal B subtypes ( $P = 0.14$ ). In comparison, we found that recurrence in HER2+ group was not well separated by ER status ( $P = 0.34$ , see Supplementary Fig. SA6-2 in Supplementary Material). These results suggest that ER status is not a good discriminator of recurrence in this set of early-stage HER2+ cases treated with local therapy alone.

Using SNR to find the genes that distinguish HER2<sub>+I</sub> from HER2<sub>+NI</sub>, we found that immunoglobulin genes were highly up-regulated in HER2<sub>+I</sub> with  $P < 0.001$  and FDR  $< 0.05$ . A gene ontology analysis showed an enrichment of the immune support pathway involving T-cell activation, inflammation-mediated chemokine and cytokine signaling, and B cell activation ( $P < 0.01$ ). Intriguingly, the

cluster of chemokine genes adjacent to the HER2 amplicon, which include *CCL2*, *CCL5*, *CCL8*, *CCL13*, *CCL18*, and *CCL23*, are among the set of genes up-regulated in HER2<sub>+I</sub>. However, most of the up-regulated lymphocyte-associated genes are not located near the HER2 amplicon (Supplementary Table S3). Our main conclusion is that node-negative HER2+ tumors with the up-regulation of immunoglobulins and other lymphocyte-associated genes (the HER2<sub>+I</sub> cluster) have an improved natural survival rate compared with HER2+ tumors that lack this signature (the HER2<sub>+NI</sub> cluster).

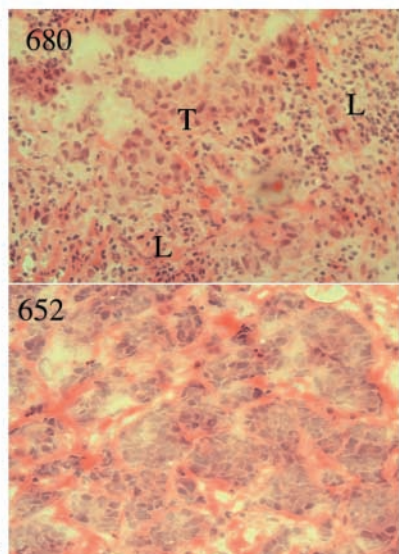
**The HER2<sub>+I</sub> subset of HER2+ cancers is associated with a prominent lymphocytic infiltrate.** The tumor RNA used to generate the gene expression profiles in the Wang et al. data set, and indeed most BCA gene expression data sets, comes from bulk tumor. Microarray analysis of these tumor specimens tests not only



**Figure 4.** Kaplan-Meier curves for the HER2+ subtypes (89% and 42% long-term distant metastasis-free survival for HER2<sub>+I</sub> and HER2<sub>+NI</sub>;  $P = 0.01$ ).

**Figure 5.** HER2<sub>+</sub> BCAs have a prominent lymphocytic infiltrate. Gene expression data were used to stratify a subset of HER2<sub>+</sub> BCA specimens from a neoadjuvant phase II trial of trastuzumab and vinorelbine (18) into HER2<sub>+</sub> and HER2<sub>+NI</sub> subclasses. H&E-stained sections of each case were then independently scored by two BCA pathologists for the presence of lymphocytic infiltration as described in Materials and Methods. Scores: 1, minimal to no tumor-associated lymphocytes; 2, presence of a moderate lymphocytic infiltrate; 3, presence of a very robust lymphocytic infiltrate. *NE*, specimen was not of adequate quality for proper evaluation. The table contains the gene expression classification, histologic scoring, and clinical response of each specimen.  $P < 0.0001$ , Fisher exact test, differences in scores between the two tumor classes is statistically significant. *Right*, typical histologic image of a tumor with a prominent lymphocytic infiltrate (*top right*) and a tumor with no lymphocytic infiltrate (*bottom right*). *Top right*, T, region with tumor cells; L, region with prominent lymphocytes.

Tumor	Path CR	Subtype by gene expression	Lymphocytic infiltration	
			Pathologist 1	Pathologist 2
680	No	HER2 <sub>+</sub>	2	2
568	No	HER2 <sub>+</sub>	2	3
334	Yes	HER2 <sub>+</sub>	3	3
438	No	HER2 <sub>+</sub>	NE	NE
422	Yes	HER2 <sub>+</sub>	3	3
611	No	HER2 <sub>+</sub>	NE	NE
<hr/>				
514	No	HER2 <sub>+NI</sub>	1	1
652	No	HER2 <sub>+NI</sub>	1	1
405	No	HER2 <sub>+NI</sub>	1	2
512	No	HER2 <sub>+NI</sub>	NE	NE
278,253	No	HER2 <sub>+NI</sub>	1	1
641	No	HER2 <sub>+NI</sub>	1	1
637	No	HER2 <sub>+NI</sub>	1	1



the tumor cells but also nontumor components of the tumor microenvironment. One explanation for the enrichment of lymphocyte-associated genes in HER2<sub>+</sub> is that these tumors have a strong lymphocytic infiltration. To directly test this hypothesis, we analyzed an independent set of HER2<sub>+</sub> tumors for which both histology and gene expression data were available. These data were obtained from a previously reported (23) set of HER2<sub>+</sub> tumors, which were part of a phase II neo-adjuvant trial of treatment with trastuzumab and vinorelbine. Before treatment, core biopsies were obtained, and separate cores were processed for histology and for RNA extraction, amplification, and hybridization to Affymetrix UI33a arrays. Gene expression data were available for 21 tumors.

We analyzed these array data by semi-supervised ensemble hierarchical clustering using a set of lymphocyte-associated genes identified by our previous analysis of the HER2<sub>+</sub> samples in the Wang et al. data. This analysis split the new tumor samples into three clusters; cluster 1 had six a strong lymphocyte-associated signature and corresponded to HER2<sub>+</sub>; cluster 3 had seven samples and little to no expression of these lymphocyte-associated genes and corresponded to HER2<sub>+NI</sub>. Cluster 2, with eight samples, had a mixed signature and was not easily classifiable as either HER2<sub>+</sub> or HER2<sub>+NI</sub>. Core biopsies, being by nature small specimens, may have more variable sampling of the tumor microenvironment and biases introduced by cDNA amplification. We believe that these factors are the cause of the difficulty in clearly stratifying the subset of HER2<sub>+</sub> tumors in cluster 2.

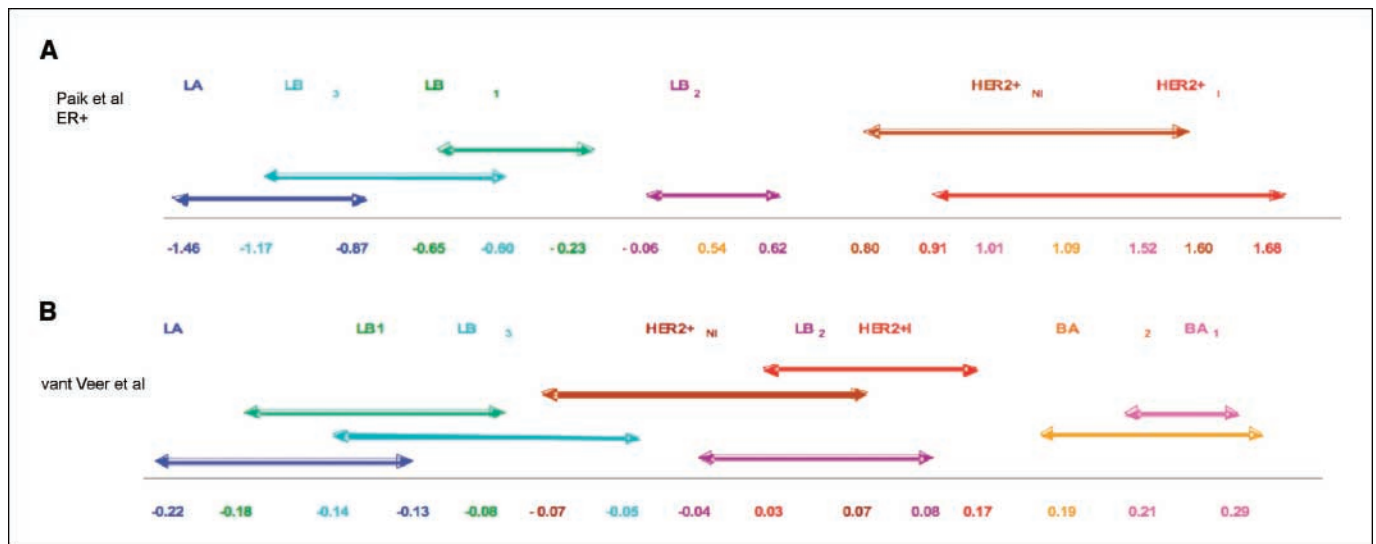
Histologic specimens of each tumor in cluster 1 (HER2<sub>+</sub>) and cluster 3 (HER2<sub>+NI</sub>) were independently evaluated for the presence of a lymphocytic infiltrate by two BCA pathologists who had no knowledge of the gene expression classification. The tumors classified as HER2<sub>+</sub> all scored as having a moderate to marked lymphocytic infiltrate, and almost all the HER2<sub>+NI</sub> scored as having minimal lymphocytic infiltrate (see Fig. 5). The differences in histologic scoring of lymphocytic infiltration between these groups was highly statistically significant based on the Fisher's exact test (overall  $P < 0.0001$ ). These data are consistent with the HER2<sub>+</sub> tumors being associated with a prominent lymphocytic infiltrate.

We note that of the six tumors in cluster 1 that corresponds to HER2<sub>+</sub>, two had a documented complete pathologic response (pCR) to neoadjuvant treatment with trastuzumab and vinorelbine. Of the seven tumors in cluster 3, which corresponds to HER2<sub>+NI</sub>, there were no pCRs. In cluster 2, there was one documented pCR; this tumor had histologic evidence of lymphocyte infiltration in the original diagnostic biopsy, but a mixed gene expression profile.

**Comparison with existing gene panels.** Several panels of genes (4–6) have been proposed to distinguish good prognosis from bad prognosis BCA patients under different conditions and treatments. Strictly speaking, these apply only under their specific conditions; e.g., the Oncotype DX panel of Paik et al. (5) is specific to recurrence in node-negative, ER<sub>+</sub> patients treated with tamoxifen. Nevertheless, it is instructive to apply these panels to our subtypes.

Figure 6A shows the ranges of raw scores for the ER<sub>+</sub> samples in our core clusters for the Oncotype DX panel (5). To make a precise comparison, it is necessary to have the measured reverse transcription-PCR values for the 21 genes in this for all the samples correct for possible variance bias for each gene and compute a normalized score using the scaling in ref. 5. In the absence of these detailed data, we used the following procedure. Using only ER<sub>+</sub> samples, we averaged all correlated probe values for each gene per array followed by a standard normalization of each gene across samples before using the weights in ref. 5 to compute a raw score for each sample. This is shown in Fig. 6A and should give the correct relative recurrence risk score, which is an increasing function of the unnormalized (raw) score. The order of increasing risk predicted by the panel is Luminal A, LB<sub>1</sub>, LB<sub>3</sub>, LB<sub>2</sub>, HER2<sub>+NI</sub>, HER2<sub>+</sub> subtype. Figures 6B shows the ranges of subtype scores using the van't Veer (4) panel. Although the general trend from Luminal A to Basal is correct, the details are not. In particular, the relative recurrence risk of the HER2<sub>+</sub> and HER2<sub>+NI</sub> subtypes is inverted compared with measured values.

A striking feature of Fig. 6A and B is that the range of scores within subtypes is small; most of the variability being between subtypes. This suggests that the Oncotype DX and van't Veer stratification is a surrogate for the subtypes identified in the



**Figure 6.** Paik and van't Veer panel scores for the BCA subtypes. Arrows, 95% confidence intervals for each core clusters. The relative risk increases from left to right.

present paper and may not be adding much prognostic information within subtypes.

In Appendix 4, Supplementary Table S1, we provide a simple (but not necessarily complete or optimal) classifier based on 120 genes that can stratify a sample into our subtypes of Supplementary Material. This classifier, if adequately validated, could be of clinical use in stratifying patients by recurrence risk. We have used this scheme to classify all the samples set aside as potentially Luminal or HER2+ or Basal. These assignments are listed at the end of Supplementary Table S2, which gives the assignment of all samples into one of the eight subtypes.

In Supplementary Fig. SA4-1, we show a heat map for 10 samples chosen randomly from each subtype over the 120 genes in Supplementary Table S1 to visually illustrate its accuracy.

## Discussion

There is now considerable data (22) showing that HER2+ amplification defines a distinct biological class of BCA. We show how these can be pre-identified in microarray data using a criterion based on the up-regulation of the Her2/neu amplicon. We then used PCA and consensus ensemble clustering technique to find the optimum number of subtypes in HER2+ and HER2- cases. Our methods readily identified the Luminal A and Basal-like subtypes initially described by Perou et al. (7), but found further substructure into four Luminal, two HER2+, and two Basal-like subtypes with distinct genetic signatures and recurrence rates. The HER2+ cases were analyzed separately and found to assort into two groups, one of which is marked by the strong up-regulation of lymphocyte-associated genes and was shown to correlate with lymphocyte infiltration of the primary tumor. By keeping the HER2+ samples separate, we prevented perturbation by the clustering of the Luminal and Basal-like cancers, which strongly separate by ER status.

These data emphasize how the composition of a heterogeneous group of samples greatly affects the outcome of clustering algorithms. Thus, if all tumors are grouped together, in the HER2+ cases, only the ER-, HER2+ subset will separate into a unique group. If, however, as was done in the present analysis, the

HER2+ group is identified separately, the clustering analysis will separate them into two groups that are distinguished not by an ER signature, but by a lymphocytic infiltrate signature. The challenge is to determine how best to apply these rules to large tumor collections. In this paper, we suggest one possible approach, which is to independently identify certain clinical classes of tumors and then apply robust clustering algorithms. Ultimately, the methods that can identify subgroups of BCA patients with unique natural histories, or clear responses to certain therapies, will be the ones useful to determine which stratification approaches are best.

The data presented here also show the value of expression signatures coming from nonmalignant elements of the tumor microenvironment. Gene expression profiling of bulk tumors captures expression signatures of both the tumor cells and nearby nontumor cells. Gene expression signatures coming from elements of the immune system, the inflammatory pathways, and other stromal elements may reflect relevant features of tumor biology that may impact both natural history and response to specific therapies.

In correlating existing prognostic gene expression panels to the BCA subtypes identified in our analysis, we found that individual subtypes have relatively narrow ranges of scores, with Luminal A always scoring as good prognosis and Basal-like and HER2+ as poor prognosis cases. Similar data have been previously reported (7, 28). However, all the panels (4, 5) reverse the HER2+ subtype recurrence risk relative to measured data. These data suggest that existing panels, which were identified as predicting prognosis in mixed collections of BCAs, do not add independent prognostic information within a given BCA subtype. As Figs. 3A and B and 4 show, better stratification by prognosis is obtained by restricting the analysis to within individual BCA subtypes, based on clinical markers or using clustering techniques.

We find two Basal-like tumor subtypes with 32% and 41% natural recurrence rates. Sotiriou et al. (11) also identified two Basal subtypes distinguished by expression of certain proliferation markers and growth factors. Our Basal subtypes also exhibit some of these markers: e.g., we find the proliferation markers, *topoisomerase II $\alpha$*  ( $P = 0.106$ ), *CDC2* ( $P < 0.005$ ), and *proliferating cell nuclear antigen (PCNA)* ( $P = 0.014$ ) up-regulated in BA<sub>1</sub>. However,



unlike the basal subtypes previously reported, there was no difference in many of the other markers, such as *fosB*, *c-fos*, and *caveolin-2*, between BA<sub>1</sub> and BA<sub>2</sub>.

Our analysis found four Luminal BCA subtypes. The Luminal A samples present as a very stable cluster with a strong ER+ signature and the best natural history. The poorer prognosis Luminal subtypes (LB<sub>1</sub>, LB<sub>3</sub>, and LB<sub>2</sub> in order of relative risk) all have worse prognosis than Luminal A. All these tumors are ER+ and most likely HER2-. The biological basis that makes them cluster separately and the observed differences in their natural history are difficult to understand and need further analysis and validation from the investigation of larger tumor collections and correlation with other clinical parameters such as PR status, tumor grade, and proliferation index.

The most significant result of this study is the discovery of two HER2+ subtypes, HER2+<sub>T</sub> and HER2+<sub>NI</sub> in node-negative patients, with recurrence rates of 11% and 58%, respectively. We observe a single pathway correlated with the low distant metastatic rate of the HER2+<sub>T</sub> subtype: the immunoglobulin pathway. As the RNA was extracted from bulk tumor tissue, the high relative expression of immunoglobulin genes and other lymphocyte-associated genes (e.g., *IGLC2*, *IGKC*, *TNFRSF17*, *NKG7*, *MAP4K1*, *XCCL1*, *XCCL2*, *TRATI*, *CTSW*, *IGHA1*, etc.) suggests that these tumors have a robust lymphocytic infiltration. Because immunoglobulin genes are plentiful and likely co-regulated, the prominence of an immunoglobulin signature in HER2+<sub>T</sub> does not necessarily imply that this lymphocytic infiltrate is mostly composed of B cells.

These data support prior clinical observations that the histologic presence of a lymphocyte infiltrate identifies a subgroup of early-stage HER2+ BCA that have a remarkably good natural history when treated with local therapy alone (29, 30). Intriguingly, in both these prior studies, this difference in natural history was limited to node-negative HER2+ tumors and was not present in node-positive HER2+ tumors, suggesting that the presence of lymphocytic infiltrate may affect the likelihood of metastasis in early lesions. Of note, in these studies, the correlation of lymphocytic infiltrate with better outcome was not seen in HER2- tumors. Thus, the biological basis, cellular makeup, and clinical impact of lymphocyte infiltration may be quite different in different subtypes of BCA. A collection of 105 lymphocyte-associated genes that can separate HER2+<sub>T</sub> from HER2+<sub>NI</sub> with high accuracy (>99%) in leave-one-out experiments is shown in Supplementary Table S3. We also show in this table all the genes with  $P \leq 0.005$  and  $FDR \leq 0.05$ .

The presence of specific immune infiltrates have been associated with improved prognosis in other cancer types, including ovarian cancer and colon cancer (31–33). However, the presence of certain T-cell subsets have also shown to negatively impact prognosis in

other studies, suggesting that the exact composition of the immune infiltrate and/or its specific interactions with tumor cell biology may vary significantly in different tumors (34, 35). It may be that certain classes of cancers may induce distinct types of lymphocytic infiltrate that impacts their natural history. In keeping with this, it has been reported that the lymphocyte infiltration in HER2+ tumors has a preponderance of macrophages, whereas lymphocyte infiltrates in HER2- tumors were composed mostly of T cells (29). The biological basis for the lymphocytic infiltrate in these diverse tumor types, their exact nature and composition, and their impact on natural history needs further investigation.

The collection of tumors we examined were all early stage (lymph node negative), and none received adjuvant therapy. Thus, we do not know if the subclasses of tumors we identified will have different responses to adjuvant treatment. For example, the question of whether the two HER2+ subgroups will respond differently to chemotherapy and/or trastuzumab remains open. One recent study suggests that HER2+ tumors with lymphocytic infiltrates may have high response rates when treated with trastuzumab alone (36). In the subset of tumors we analyzed from a neoadjuvant trial of trastuzumab and vinorelbine, there seems to be an association with lymphocytic infiltrate and achieving a complete pathologic response. Although provocative, this hypothesis requires rigorous testing in larger patient populations.

Our results raise multiple questions about the biology of HER2+ BCAs. Why do certain tumors have a lymphocytic infiltrate? Does the size of the HER2 amplicon, i.e., whether it encompasses the nearby cytokine-rich region on 17q, correlate with the lymphocytic infiltration? What is the exact composition of this infiltrate in different tumor subtypes? Is the tumor or tumor-associated stroma secreting key cytokines responsible for lymphocyte chemotaxis? Does the infiltrate reflect an autoimmune recognition of the tumor? Are the infiltrating lymphocytes providing a crucial growth factor for the tumor or tumor-associated stroma? Comparison of gene expression data derived from bulk tumor and then from micro-dissection of separate tumor and nontumor compartments may help answer some of these questions.

## Acknowledgments

Received 2/12/2007; revised 7/30/2007; accepted 9/19/2007.

**Grant support:** S. Ganesan is supported by grants from the National Cancer Institute, University of Medicine and Dentistry of New Jersey Foundation and the Sidney Kimmel Foundation. L. Harris is supported by the Department of Defense Clinical Translational Research Award (W81XWH-04-1-0549).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We are grateful to Arnold Rabson and Edmund Lattime for helpful discussions regarding the data.

## References

- Anim JT, John B, Abdulsathar SS, et al. Relationship between the expression of various markers and prognostic factors in breast cancer. *Acta Histochem* 2005;107:87–93.
- Gruvberger S, Ringner M, Chen Y, et al. Estrogen receptor status in breast cancer is associated with remarkably distinct gene expression patterns. *Cancer Res* 2001;61:5979–84.
- Diermeier S, Horvath G, Knuechel-Clarke R, Hofstaedter F, Szollosi J, Brockhoff G. Epidermal growth factor receptor coexpression modulates susceptibility to Herceptin in HER2/neu overexpressing breast cancer cells via specific erbB-receptor interaction and activation. *Exp Cell Res* 2005;304:604–19.
- van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999–2009.
- Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351:2817–26.
- Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;365:671–9.
- Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumours. *Nature* 2000;406:747–52.
- Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869–74.
- Alexe G, Dalgin GS, Ramaswamy R, DeLisi C, Bhanot G. Data perturbation independent diagnosis and validation of breast cancer subtypes using clustering and patterns. *Cancer Informatics* 2006;2:243–74.
- Sorlie T, Wang Y, Xiao C, et al. Distinct molecular mechanisms underlying clinically relevant subtypes of breast cancer: gene expression analyses across three different platforms. *BMC Genomics* 2006;7:127.



11. Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 2003;100:10393–8.
12. Sotiriou C, Powles TJ, Dowsett M, et al. Gene expression profiles derived from fine needle aspiration correlate with response to systemic chemotherapy in breast cancer. *Breast Cancer Res* 2002;4:R3.
13. Bertucci F, Eisinger F, Houlgatte R, Viens P, Birnbaum D. Gene-expression profiling and identification of patients at high risk of breast cancer. *Lancet* 2002;360:173–4; author reply 4.
14. Gosset WS. The probable error of a mean. *Biometrika* 1908;6:1–25.
15. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98:5116–21.
16. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* (New York, NY) 1999;286:531–7.
17. Lyons-Weiler J, Patel S, Becich MJ, Godfrey TE. Tests for finding complex patterns of differential expression in cancers: towards individualized medicine. *BMC Bioinformatics* 2004;5:110.
18. Alexe G, Hammer PL. Spanned patterns for the logical analysis of data discrete applied mathematics 2006;154:1039–49.
19. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning J* 2003;52:91–118.
20. Strehl A, Ghosh J. Cluster ensembles: a knowledge reuse framework for combining partitionings. Eighteenth National Conference on Artificial Intelligence; 2002 July 28–August 01, 2002; Edmonton, Alberta, Canada; 2002. p. 93–8.
21. Bertucci F, Borie N, Ginestier C, et al. Identification and validation of an ERBB2 gene expression signature in breast cancers. *Oncogene* 2004;23:2564–75.
22. Kauraniemi P, Kallioniemi A. Activation of multiple cancer-associated genes at the ERBB2 amplicon in breast cancer. *Endocr Relat Cancer* 2006;13:39–49.
23. Harris LN, You F, Schnitt SJ, et al. Predictors of resistance to preoperative trastuzumab and vinorelbine for HER2-positive early breast cancer. *Clin Cancer Res* 2007;13:1198–207.
24. Tibshirani R, Walther G, Hastie T. Estimating the number of clusters in a dataset via the Gap statistic. *Journal of the Royal Statistics Society (Series B)* 2001;63:411–23.
25. Kaufmann L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. 1st ed: John Wiley & Sons; 1990.
26. Mi H, Lazareva-Ulitsky B, Loo R, et al. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res* 2005;33:D284–8.
27. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
28. Fan C, Oh DS, Wessels L, et al. Concordance among gene-expression-based predictors for breast cancer. *N Engl J Med* 2006;355:560–9.
29. Pupa SM, Bufalino R, Invernizzi AM, et al. Macrophage infiltrate and prognosis in c-erbB-2-overexpressing breast carcinomas. *J Clin Oncol* 1996;14:85–94.
30. Rilke F, Colnaghi MI, Cascinelli N, et al. Prognostic significance of HER-2/neu expression in breast cancer and its relationship to other prognostic factors. *Int J Cancer* 1991;49:44–9.
31. Zhang L, Conejo-Garcia JR, Katsaros D, et al. Intratumoral T cells, recurrence, and survival in epithelial ovarian cancer. *N Engl J Med* 2003;348:203–13.
32. Pages F, Berger A, Camus M, et al. Effector memory T cells, early metastasis, and survival in colorectal cancer. *N Engl J Med* 2005;353:2654–66.
33. Galon J, Costes A, Sanchez-Cabo F, et al. Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* (New York, NY) 2006;313:1960–4.
34. Wolf D, Wolf AM, Rumpold H, et al. The expression of the regulatory T cell-specific forkhead box transcription factor FoxP3 is associated with poor prognosis in ovarian cancer. *Clin Cancer Res* 2005;11:8326–31.
35. Dranoff G. The therapeutic implications of intratumoral regulatory T cells. *Clin Cancer Res* 2005;11:8226–9.
36. Gennari R, Menard S, Fagnoni F, et al. Pilot study of the mechanism of action of preoperative trastuzumab in patients with primary operable breast tumors overexpressing HER2. *Clin Cancer Res* 2004;10:5650–5.