[CANCER RESEARCH 64, 2962–2968, May 1, 2004]

Advances in Brief

A Molecular Signature of the Nottingham Prognostic Index in Breast Cancer

Kun Yu,[1] Chee How Lee,[1] Puay Hoon Tan,[2] Ga Sze Hong,[1] Siew Bok Wee,[1] Chow Yin Wong,[1] and Patrick Tan[1,3]

[1]National Cancer Center and [3]Defence Medical and Environmental Research Institute, Republic of Singapore, and [2]Department of Pathology, Republic of Singapore

Abstract

The Nottingham prognostic index (NPI) is a widely used clinicopathological staging system for breast cancer prognostication. Using a step-wise classification approach where breast tumor expression profiles were first divided into general "molecular subtypes" [estrogen receptor (ER)+, ER−, ERBB2+], followed by an independent analysis of each subtype, we identified a 62-gene expression signature (NPI-ES) highly correlated to the NPI in ER+ tumors. The NPI-ES classified the majority of ER+ tumors into two distinct groups with high confidence and was significantly correlated to NPI status in two independent sets of ER+ tumors derived from different centers. The NPI-ES is comparable to the classical NPI in identifying patients likely to exhibit a poor clinical prognosis, as well as to a recently described "prognosis signature" for breast cancer. Our findings demonstrate how expression profiling can complement classical staging systems employing histopathological parameters scored over a continuous range of values.

Introduction

The Nottingham prognostic index (NPI) is a clinicopathological classification system based on tumor size, histological grade, and lymph-node status widely used in Europe and the United Kingdom for breast cancer prognostication (1–4). Despite its utility, it is also acknowledged that currently used histopathological parameters such as tumor grade and cellular morphology are also associated with certain limitations, *e.g.,* many of these variables (such as grade) are subject to significant interobserver variability even after standardization attempts (5), and appropriate cutoff points are often difficult to define when the parameter being measured is scored over a continuous range of values (6). As a potential alternative, several groups have recently used gene expression profiles to classify breast tumors (7–11), and shown that a significant proportion of the intrinsic gene expression variation in breast cancer can be attributed to tumors belonging to different "molecular subtypes" (*e.g.,* estrogen receptor (ER)+ and ER− tumors; Refs. 7–9, 15). In this study, we used a step-wise subtype-specific classification approach to analyze a series of in-house breast cancer expression profiles, and we identified an expression signature of 62 genes (NPI-ES) that was highly correlated to the NPI in ER+ tumors. Confidence in the reliability of this expression signature was obtained by validating it against two independent breast cancer data sets from different centers. We propose that the NPI-ES can potentially function as a molecular surrogate of the NPI in ER+ tumors.

Materials and Methods

Breast Tissues and Clinical Information. Human breast tissues were obtained from the National Cancer Center Tissue Repository, after appropriate approvals from the National Cancer Center Repository and Ethics Committees. Detailed descriptions of sample collection, archiving, and histological assessment of tumors, including techniques and parameters [tumor size, nodal status, histological grade and type, ER/progesterone receptor (PR), and ERBB2 measurements] are presented in the supplementary data. Profiled samples contained at least 50% tumor content. NPI status was calculated as follows: tumor size (cm)*0.2 + grade + lymph-node points (negative nodes = 1 point; positive nodes, 1 to 3 positive = 2 points; positive nodes, ≥4 = 3 points). As tumor size in the Stanford data set was defined using a categorical system, we assigned an approximate value for each categorical grade (*i.e.,* T_1 = 2 cm, T_2 = 3.5, T_3 = 5, T_4 = 3.5), using an "average" value of 3.5 cm (T_2) to describe T_4 tumors (T_4 indicating that the size of the tumor is not known). The end point criteria for disease-free survival (DFS) in the Stanford and Rosetta data sets are described in the original reports (10, 15). The Stanford data set contained seven patients whose deaths were not attributable to cancer or who died within 5 years without a relapse. These patients were excluded from Table 3.

Sample Preparation and Microarray Hybridization. RNA was extracted from tissues using Trizol reagent and processed for Affymetrix Genechip hybridizations using U133A Genechips according to the manufacturer's instructions.

Data Processing and Analysis. Raw Genechip scans were quality controlled using GeneData Refiner and analyzed using Genedata Expressionist or conventional spreadsheet applications. The unsupervised data set (Fig. 1, *A* and *B*) contains genes exhibiting SD > 1.5 across all well-measured samples. Minor variations of the variation filter yielded similar results.[4] Redundant probes for the same gene were removed from analysis, leaving one probe per gene. Average-linkage hierarchical clustering using Pearson correlation distance metrics was performed using CLUSTER and displayed using TREEVIEW software (12). Unsupervised clustering was used to compare gene sets across different array technologies. Gene selections were performed using significance analysis of microarrays (SAM; Ref. 13), with "false discovery rates" being 0.1% for Fig. 1*C* and 15% for Fig. 2. For Fig. 2, similar results were obtained using false discovery rates ranging from 5–30%.[4] Supervised approaches using leave-one-out cross-validation (LOOCV) assays were used to calculate the confidence of a specific class assignment. Weighted voting and prediction strengths were calculated as in Golub *et al.* (Ref. 14; supplementary data). For each leave-one-out "training" data set in the LOOCV assay, the optimal NPI cutoff value was reselected by determining the cutpoint yielding the highest number of differentially expressed genes (supplementary data) from an initial data set comprising all well-measured genes (~14,000 genes). This set of differentially expressed genes was then used to classify the remaining tumor sample. This analysis was repeated for all ER+ tumors, and the overall classification accuracy was determined by summing the individual classifications. Kaplan-Meier survival curves were created using SPSS (SPSS Inc. Chicago, IL). Statistical associations between gene expression and clinical variables were determined by χ^2 analysis. McNemar's test was used to compare the classification accuracies of the NPI-ES and NPI.

Received 8/6/03; revised 2/11/04; accepted 2/24/04.

Grant support: This work was supported by a grant (to P. Tan) from Agenica Research.

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.

Note: Supplemental data for this article can be found at Cancer Research Online (http://cancerres@aacrjournals.org).

Requests for reprints: Patrick Tan, National Cancer Center/Defence Medical and Environmental Research Institute, 11 Hospital Drive, Singapore 169610, Republic of Singapore. Phone: 65-6-436-8345; Fax: 65-6-226-5694; E-mail: cmrtan@nccs.com.sg.

[4] K. Yu, unpublished observations.

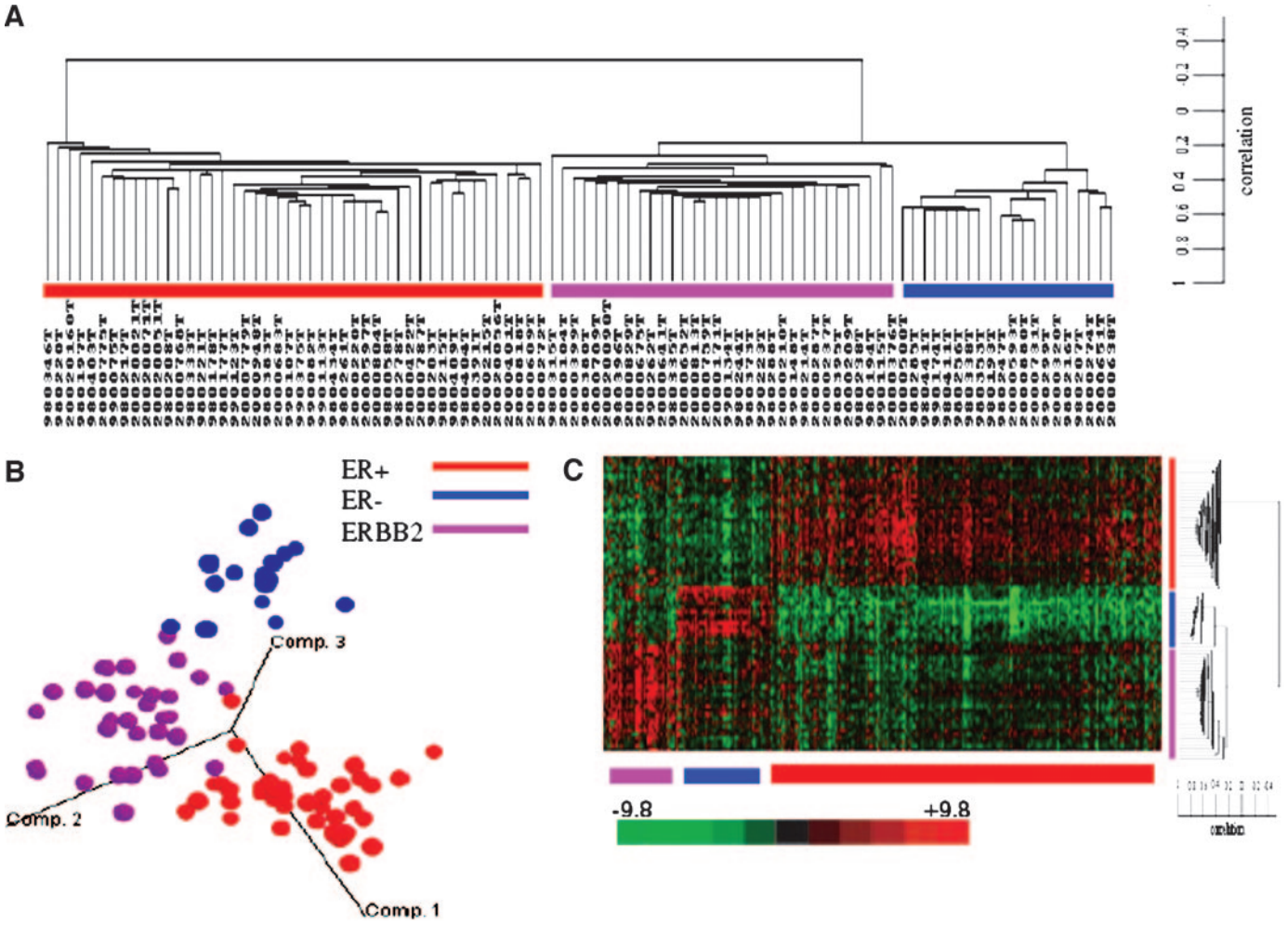

Fig. 1. Clustering of sporadic breast tumors by global expression profiles. *A,* unsupervised hierarchical clustering of 98 breast tumors using the top 376 genes exhibiting the highest variation in gene expression. Correlation distances are depicted as a *bar graph to the right*. *B,* principal component analysis using the 376 gene set. Similar molecular groupings are observed (as in *A*). *C,* hierarchical clustering of samples using the SAM-409 gene set that consists of genes that are significantly regulated between tumor subtypes. Approximately two-thirds of the genes in the SAM-409 gene set exhibit increased expression in estrogen receptor positive tumors.

Results and Discussion

Defining Molecular Subtypes of Breast Cancer Using Unsupervised Clustering. Different molecular subtypes of breast cancer are associated with highly distinctive expression profiles (7–9, 15). It has been proposed that this may reflect their arising from different cells of origin (7), raising the possibility that different tumor subtypes might use distinct genetic circuits to regulate various aspects of tumorigenesis. From an analytical standpoint, the presence of dramatic gene expression differences between subtypes (inter-subtype differences) might also potentially obscure more subtle patterns of variation within subtypes (intra-subtype differences). Indeed, in an initial analysis where tumors were treated irrespective of subtype, we could not convincingly identify an expression signature correlated to the NPI (supplementary data). To circumvent this problem, we implemented a step-wise methodology in which each molecular subtype was treated as an independent data set. Briefly, a variety of unsupervised clustering techniques were first used to broadly segregate a set of breast tumor expression profiles according to their respective molecular subtype categories. Second, tumors within each subtype were then independently analyzed to define expression signatures that might be correlated to the NPI or its constituent elements.

Using Affymetrix U133A Genechips, we generated expression profiles for 98 sporadic breast tumors derived from our local predominantly Chinese patient population. After data normalization and preprocessing, we applied a SD filter to identify a set of 367 genes (SD-367) exhibiting a high degree of gene expression variation across the tumor series. We used this gene set and unsupervised hierarchical clustering to group the tumor expression profiles on the basis of their overall similarity. The breast tumors self-segregated into three major subgroups, referred to as ER+, ER−, and ERBB2+, respectively (Fig. 1*A*). There was good agreement between these molecular subgroups and conventional immunohistochemisty (supplementary data). A similar segregation pattern was also obtained when the tumor profiles were reclustered using principal components analysis, an independent analytical technique (Fig. 1*B*).

As the SD filter used in the unsupervised analysis does not discriminate between inter-subtype and intra-subtype differences, the SD-367 gene set consequently contains the following two broad classes of genes: those exhibiting significant expression variation between the tumor subtypes (inter-subtype) and those exhibiting expression variations within the subtypes (intra-subtype). Examples of genes exhibiting significant intra-subtype variation included those related to immune function (*e.g., IGLJ3*, *IGLλ*), and stromal tissue remodelling (*e.g., COL11A1*, *MMP13*). To identify genes specifically exhibiting inter-subtype variation, we then used SAM (14) at a false discovery rate of 0.1%, to select 409 genes (SAM-409) that were

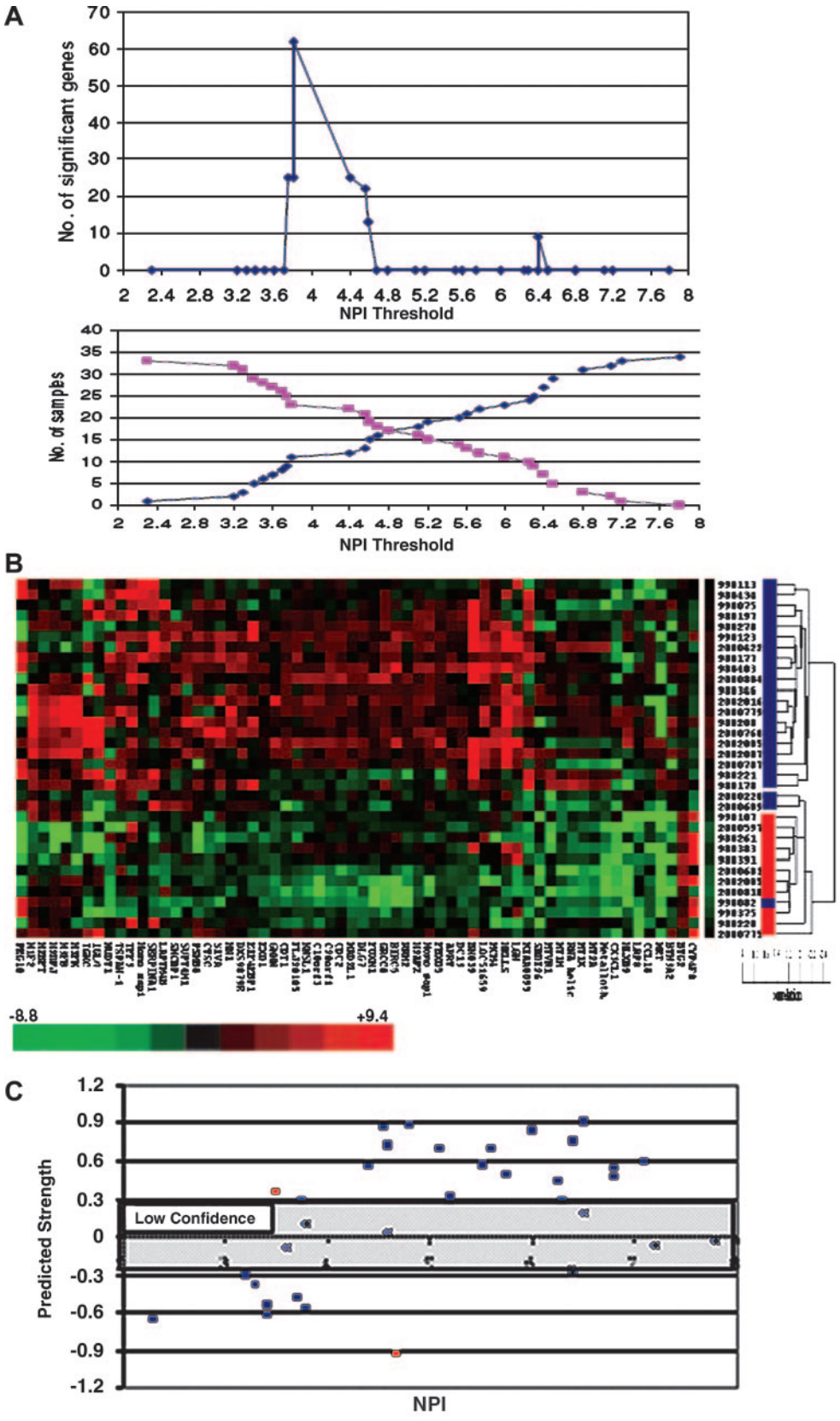

Fig. 2. Identification of an expression signature correlated to the NPI [(Nottingham prognostic index; NPI-expression signature (NPI-ES)]. *A,* identification of differentially expressed genes using a moving NPI threshold. The *upper graph* depicts the number of genes (*Y*-axis) exhibiting significant differential expression identified at each threshold value (*X*-axis). The *lower graph* depicts the number of tumors in each group. *Pink*, high-NPI; *blue*, low-NPI. A threshold of four delivers the highest number of differentially regulated genes. *B,* hierarchical clustering of estrogen receptor (ER)+ samples using the NPI-ES. The *red bar* indicates samples of low NPI (<4) whereas the *blue bar* indicates samples of high NPI. *C,* classification and prediction confidence of ER+ tumor samples using the NPI-ES. Samples are sorted by their NPI value (*X*-axis). Weighted voting was used to classify the samples and the prediction strengths of each sample (*Y*-axis). Sample classifications with a prediction strength of <0.3 are considered "uncertain" or "low-confidence" (*gray area*).

significantly regulated in a subtype-specific manner (100 genes were found in common between the SD-367 and SAM-409 gene sets, see supplementary data; Fig. 1c). In this analysis, SAM was used in a descriptive fashion to identify genes that appear to be largely responsible for defining the clusters. As such, the SAM false discovery rate control should not be interpreted in this situation as providing a true statistical inference procedure because the clusters to which it was applied were derived from the observed data. Approximately 69% of the SAM-409 gene set consisted of genes exhibiting high expression in the ER+ subgroup compared with the other two subtypes, including the estrogen receptor gene *ESR1* and estrogen-regulated genes such as *LIV1*, *TFF1*, and *MYB* (supplementary data). In agreement with other studies, high expression levels of *GATA3*, *HNF3α*, *Annexin A9*, and *XBP1*, were also observed in this subtype (7–9, 11). In contrast, the ER− subgroup was associated with high expression of basal mammary epithelia markers (*keratin 5* and *17*), the basement membrane protein *ladinin 1*, and the serine protease inhibitor *maspin*, a tamoxifen-inducible gene that is expressed in an inverse fashion to ER (16). Finally, the ERBB2+ subtype was associated with high expression levels of the *ERBB2* receptor and other genes physically linked to the 17q21 locus, such as *GRB7* and *PNMT* (15), suggesting the presence of DNA amplification. Taken collectively, our results validate and confirm previous reports that the majority of breast tumors can indeed be subdivided into distinct molecular subtypes on the basis of their global gene expression profiles.

Identification of a Molecular Signature Correlated to the NPI in ER+ Tumors. We focused on the 34 tumors belonging to the ER+ subtype and identified genes within this subtype whose expression was correlated to NPI status. Classically, breast cancer patients are typically stratified by the NPI into three major groups, *e.g.,* "good" (NPI $<$ 3.4), "moderate" (NPI 3.4–5.4), and "poor" prognosis (NPI $>$ 5.4; Ref. 2). Possibly reflecting the effects of variability across different scoring pathologists, other studies have proposed slightly different values for the cutoff values defining these groups (17). To avoid any potential bias in determining the appropriate NPI cutoff value, we conducted a moving threshold analysis where the ER+ tumors were divided into a series of binary groups by a NPI threshold that was steadily increased from 2.3 to 7.8. At each threshold value, genes exhibiting significant variation in expression between the two groups were identified. We found that using an NPI cutoff value of 3.8 to 4.6 yielded a gene set of 62 differentially expressed genes (Fig. 2*A*), the majority of which exhibited increased expression in the ER+ samples with a high NPI (Fig. 2*B*). We refer to this 62-member gene set as an "NPI Expression Signature" or NPI-ES (supplementary data). The genes belonging to the NPI-ES are associated with a wide variety of cellular functions implicated in oncogenesis, including DNA replication and cell division (*APRT*, *MCM4*, *KNSL 1*, *CDC2*), cellular signaling (*chemokine ligand 1*, *Met*, *ShC*), apoptosis (*survivin, CD27-binding protein*), and cellular adhesion (*discs-large homologue 7*, *tetraspan 1*). Of the individual NPI components (tumor size, tumor grade, lymph-node status), tumor grade appears to represent the predominant contributor to the molecular makeup of the NPI-ES (supplementary data).

Classification of Tumors by the NPI-ES Defines Two Discrete Molecular Groups. One advantage in the use of molecular profiles for tumor classification is the ability to mathematically quantify the confidence level of the classification (12), which is particularly important if the classification affects the subsequent course of treatment. Notably, although the ER+ samples in our data set were associated with a continuous spectrum of classical NPI values (2–8), the clustering analysis using the NPI-ES appeared to separate the ER+ tumors into two apparently discrete groups (Fig. 2*B*), raising the possibility that samples exhibiting continuous values based on histopathological parameters may be nevertheless separable into discrete categories at the molecular level. To test this hypothesis, we used a supervised learning algorithm, weighted voting, with LOOCV assays to classify the tumors into "high NPI" and "low NPI" categories based on their global gene expression profiles. In addition to classification accuracy, quantitative metrics (prediction strengths) were also calculated to provide an assessment of prediction confidence (Fig. 2*C*). Briefly, the ER+ samples were divided into a series of thirty-four "leave-one-out" data sets, each consisting of all ER+ tumors except one, and the optimal NPI cutoff value was reselected for each data set. Confirming the original analysis, an NPI cutoff value of 4.0 was identified as the optimal cutpoint for every leave-one-out data set (supplementary data), and all 34 individual 60 member predictor gene sets generated by the LOOCV assay exhibited substantial overlaps (mean overlap 93%, range 72–98%) with the original 62 gene NPI-ES signature identified in the previous section. Twenty-eight of the original 34 tumors (82%) were correctly classified by the LOOCV assay, and 24 of these 28 tumors (or 70% of all tumors) were classified with high confidence (Fig. 2*C*). Taken collectively, these results suggest that the NPI-ES can be used to classify the majority of the ER+ tumors in our data set into discrete groups with high confidence.

Application of the NPI-ES Across Multiple Independent Breast Cancer Expression Data Sets. To test the ability of the NPI-ES to predict both NPI status and disease prognosis in a series of blind "test sets," we used two independent publicly available breast cancer data sets. The first data set (referred to as the Rosetta data set) consists of 78 lymph-node-negative breast tumors profiled using oligonucleotide-based microarrays, and the duration of DFS (the time from initial tumor diagnosis to the appearance of a new distant metastasis) for each patient (10). Importantly, there are already several published studies confirming the prognostic value of the NPI in node-negative tumors (18, 19). The second data set consists of 78 breast carcinomas profiled using cDNA microarrays with overall patient survival information (referred to as the Stanford data set; Ref. 15). The availability of these data sets allowed us to independently test the predictive power of the NPI-ES, as the Rosetta and Stanford data sets are different from our data set in multiple ways, including (*a*) patient population, (*b*) sample handling protocols, (*c*) scoring pathologist, (*d*) adjuvant treatment selection (more than half of the Stanford patients were subsequently treated with tamoxifen, whereas in contrast only 5 of 78 patients in the Rosetta data set were treated with systemic adjuvant therapy), and (*e*) choice of array technology and probe sets (two-color in the Rosetta and Stanford data sets and single color in ours).

Rosetta Breast Cancer Data Set. Of the 409 genes identified by SAM analysis defining the ER+, ER−, and ERBB2+ subtypes, 276 genes (67%) were found on the Rosetta microarray. We applied this gene set to the 78 Rosetta tumor profiles and identified 49 tumors belonging to the ER+ molecular subtype (supplementary data). We

Table 1 *Association of NPI-ES[a] expression and NPI status in Rosetta and Stanford ER+ tumors*

The 1st column represents the number of tumors expressing high or low levels of the NPI-ES. The significant correlation between NPI-ES expression and NPI status in the Stanford data set remains valid even when the T_4 tumors are removed from analysis ($P = 0.04$; Kun Yu, unpublished observations).

	Student's *t* test (continuous)		χ^2 (binary)		
Rosetta	mean (variance)	$P = 0.0004$	Low (<3.4)	High	$P = 0.0087$
High (24[b])	3.1 ± 0.4		13	11	
Low (25)	2.3 ± 0.6		22	3	
Stanford		$P = 0.001$			
High (13)	5.3 ± 0.5				
Low (33)	4.5 ± 0.6				

[a] NPI-ES, Nottingham prognostic index-expression signature; ER, estrogen receptor.
[b] Figure in parenthesis represents the no. of samples.

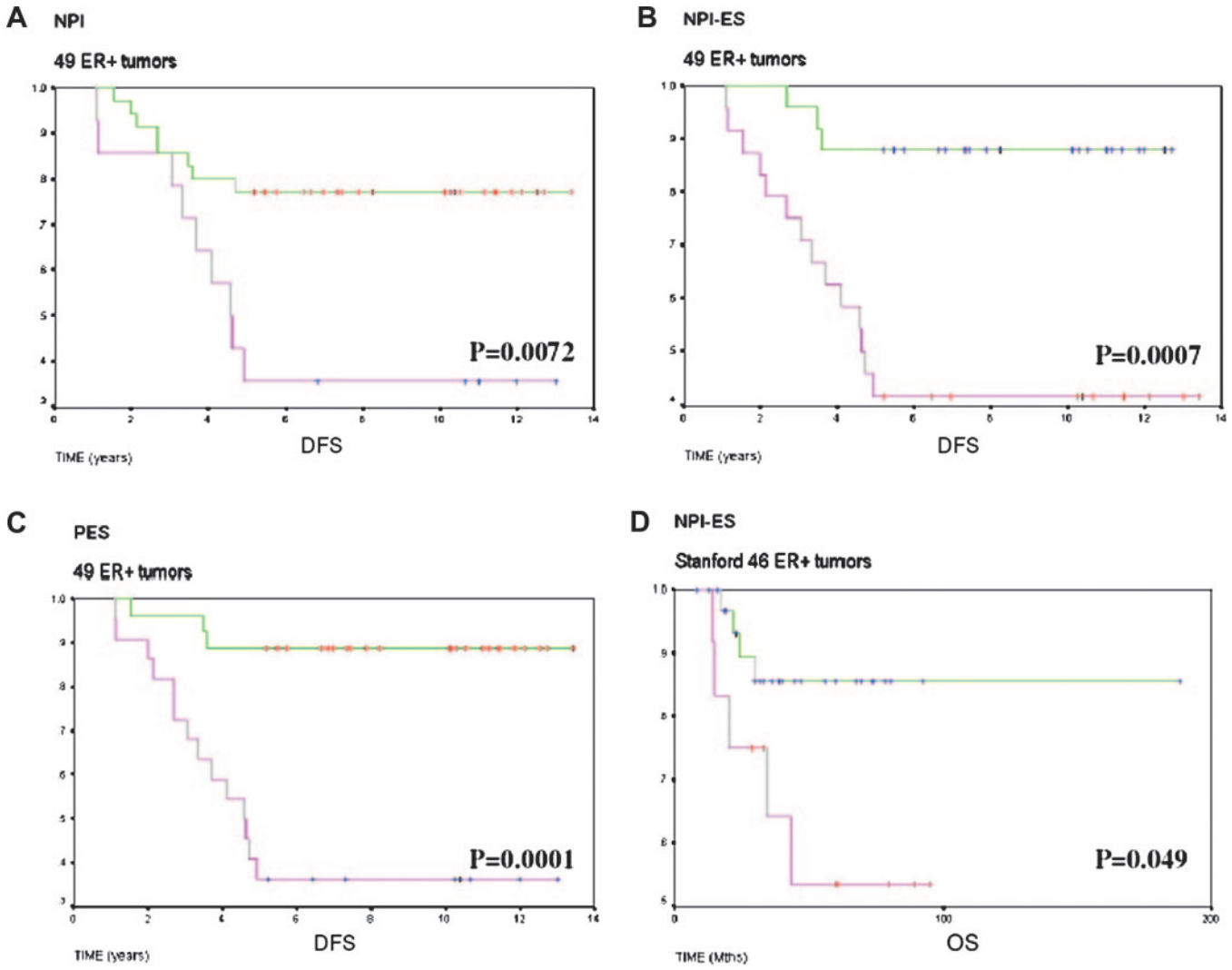

Fig. 3. Kaplan-Meier survival analyses. *Green lines* represent low Nottingham prognostic index (NPI; *A*), low NPI-expression signature (ES) expression levels (*B*), or low "prognosis" signature expression levels (PES; *C*), whereas *pink lines* represent high levels. *A,* 49 Rosetta estrogen receptor (ER)+ tumors stratified by classical NPI into "good" prognosis (NPI < 3.4; 35 tumors) and "moderate" prognosis (NPI > 3.4; 14 tumors) groups. *B,* the same 49 Rosetta ER+ tumors stratified by NPI-ES into groups expressing high (24 tumors) *versus* low levels of the NPI-ES (25 tumors). *C,* the same 49 Rosetta ER+ tumors stratified by the 70-gene prognosis signature into "good prognosis" group (27 tumors) *versus* "poor prognosis" group (22 tumors), respectively. *D,* the 46 Stanford ER+ tumors stratified by NPI-ES into groups expressing high (13 tumors) *versus* low (33 tumors) levels of the NPI-ES. *DFS*, disease free survival; *OS*, overall survival.

then determined that 46 of the 62 NPI-ES genes were also present on the Rosetta microarray. Because the Rosetta data set is based on a different array technology than ours, it is not possible to directly apply the trained weighted voting model developed on our data set to classify the Rosetta tumors. However, following the strategy described in Ramaswamy *et al.* (20) for the comparison of gene sets across different array technologies, we used hierarchical clustering to group the 49 ER+ Rosetta tumors using the overlapping NPI-ES set of 46 genes. The clustering analysis divided the 49 ER+ Rosetta tumors into two groups consisting of 24 and 25 tumors exhibiting "high" and "low" expression levels of the NPI-ES, respectively (supplementary data).

We compared the tumors in these two subgroups to determine whether they were associated with differences in their NPI values. Using two distinct statistical approaches where the tumor NPI values were treated either as a continuous gradient (Student's *t* test), or as two discrete groups (χ^2 analysis, using a classical NPI cutoff value of 3.4), tumors exhibiting high expression of the NPI-ES consistently exhibited a significantly higher NPI value compared with tumors expressing low levels of the NPI-ES ($P = 0.0004$ for continuous analysis, $P = 0.0087$ for binary analysis; Table 1). We also investigated the prognostic performance of the NPI-ES and NPI using Kaplan-Meier survival analysis (Fig. 3). In agreement with other studies, patients with tumors of low NPI (<3.4) exhibited better DFS

Table 2 *Odds ratio for distant metastasis within 5 years as a first event in Rosetta ER+[a] tumors based upon classical NPI staging and NPI-ES expression*

Odds ratios were calculated using a standard two-by-two table.

	ER+ tumors		Odds ratio (95% CI)
	Free >5 yr	<5 yr	
NPI ($P = 0.006$)			6.08 (1.58–23.39)
Low (<3.4)	27	8	
High (≥3.4)	5	9	
	84.4%[b]	52.9%[c]	
NPI-ES ($P < 0.001$)			10.27 (2.40–43.94)
Low	22	3	
High	10	14	
	68.8%	82.4%	

[a] ER, estrogen receptor; NPI, Nottingham prognostic index; ES, expression signature; CI, confidence interval.
[b] Percentage of good prognosis patients correctly predicted.
[c] Percentage of poor prognosis patients correctly predicted.

Table 3 *Odds ratio for relapse within 5 years as a first event in Stanford ER+[a] tumors based upon PES expression and NPI-ES expression*

Seven samples did not possess proper relapse information and were removed from analysis (leaving 39 ER+ tumors).

	ER+ tumors		Odds ratio (95% CI)
	Free	Relapse	
PES ($P = 0.078$)			3.6 (0.83–15.55)
Low	21	7	
High	5	6	
	80.8%	46.2%	
NPI-ES ($P = 0.027$)			4.9 (1.13–21.16)
Low	21	6	
High	5	7	
	80.8%	53.8%	

[a] ER, estrogen receptor; PES, prognosis expression signature; NPI-ES, Nottingham prognostic index-expression signature; CI, confidence interval.

compared with patients of higher NPI (>3.4; $P = 0.007$; Fig. 3*A*). When this same population was restratified by the NPI-ES, patients with tumors of low NPI-ES expression also exhibited improved relapse-free survival compared with patients with tumors expressing high levels of the NPI-ES ($P = 0.0007$). This analysis indicates that expression of the NPI-ES is significantly correlated with classical NPI status and clinical outcome in ER+ tumors even in an independent data set generated by a different array technology.

In a clinical setting, it would be medically important to accurately identify, at the initial point of diagnosis, those patients that are likely to eventually experience a poor clinical outcome. Other investigators have also voiced a similar belief (10), as the early prediction of such "poor prognosis" patients might then allow these patients to be treated with more aggressive or alternative clinical regimens. We observed that in the Rosetta data set, the NPI-ES successfully predicted a greater number of poor prognosis patients than the classical NPI (14 of 17 for NPI-ES *versus* 9 of 17 for NPI; Table 2). Although this difference was not statistically significant because of its small sample size ($P = 0.06$ using McNemar's test), it does suggest that for ER+ tumors, the NPI-ES is at least equivalent to the NPI system of staging for the prediction of such poor prognosis patients. The predictive accuracies of the NPI-ES and the NPI remained comparable even after optimizing the NPI cutoff value for this series of ER+ tumors (supplementary data).

Stanford Data Set. A similar approach was used to test the NPI-ES on the Stanford data set. Of the SAM-409 gene set used to define the ER+, ER−, and ERBB2+ subtypes, 136 genes were found on the Stanford microarray, and these genes were used to cluster the Stanford tumors to identify 46 tumors belonging to the ER+ molecular subtype (supplementary data). These 46 tumors were then clustered using the NPI-ES (31 matches on the Stanford microarray) into high-NPI-ES (13 tumors) and low-NPI-ES groups (33 tumors). Once again, Student's *t* test revealed a significant association ($P = 0.001$) between the high and low expressing NPI-ES subgroups and classical NPI status (Table 1). A Kaplan-Meier survival analysis also demonstrated a significant ($P = 0.0493$) overall survival (OS) advantage in patients with low-NPI-ES expressing tumors compared with patients with high-NPI-ES expressing tumors (Fig. 3*D*). There was no statistically significant difference in survival observed when these patients were restratified by the classical NPI system of staging.[4] Interestingly, there appears to be a strong correlation between ER+ tumors expressing high levels of the NPI-ES and the "luminal C" molecular subtype identified in Sorlie *et al.* (15), although none of the 62 genes belonging to the NPI-ES have been reported to be expressed in the latter (supplementary data).

The Prognostic Capacity of the NPI-ES Is Comparable with a Previously Described "Prognosis Signature" for Breast Cancer. In the same study by Van't Veer *et al.* (10), the authors also identified a 70-gene prognosis expression signature (PES) that predicted the DFS status of breast tumors. Interestingly, there is minimal overlap between the genes belonging to the NPI-ES and the PES, because they share only one gene in common, and only eight genes if compared with the extended PES list of 231 genes (supplementary data). To investigate the prognostic performance of the NPI-ES and the PES on the Rosetta ER+ tumors, we first used Kaplan-Meier survival analysis to compare the DFS of patients stratified either by the NPI-ES (Fig. 3*B*) or the PES (Fig. 3*C*). Both classifiers were associated with statistically significant differences in survival (PES, $P = 0.0001$; NPI-ES, $P = 0.0007$). It is worth noting, however, that the identification of the PES was directly based on the expression profiles and clinical information of these same tumors, and thus the Rosetta tumors are not strictly blinded to the PES. When the PES and NPI-ES were applied to the Stanford ER+ tumors, both molecular signatures delivered highly similar predictive accuracies for poor prognosis tumors (Table 3), suggesting that the prognostic performances of the NPI-ES and PES are relatively comparable.

In summary, we have in this report identified a 62-gene expression signature that can potentially function as a molecular surrogate for the NPI in ER+ tumors. Confidence in the reliability of the NPI-ES was obtained by showing that it was significantly correlated with NPI status for two independent sets of tumors generated by different centers. We note that the NPI-ES was derived to correlate with NPI status and, as a consequence, indirectly with overall survival. We did not use our array data to directly derive a predictor for clinical outcome. Nevertheless, NPI-ES expression was also significantly associated with other clinical parameters, such as disease-free survival (Table 1), and overall survival (Fig. 3*D*), the latter being a clinical outcome historically associated with the NPI. One interesting concept emerging from this study is that samples exhibiting apparently continuous variables at the histopathological level may nevertheless be separable into discrete categories at the molecular level. This may address a major challenge in cancer histopathology, *i.e.,* the difficulty of defining clinically appropriate cutoff values when the parameter being scored is of a continuous nature. We conclude by acknowledging that more work needs to be performed before the clinical utility of the NPI-ES can be fully assessed. First, the predictive power of the NPI-ES obviously needs to be tested against a much larger group of early (node-negative) and late-stage (node-positive) tumors. Second, although we have demonstrated the applicability of the NPI-ES in the ER+ molecular subtype, expression of the NPI-ES does not appear to be correlated as well to NPI values associated with the other molecular subtypes (ER−, ERBB2+; supplementary data). As mentioned earlier, one explanation might be that the genetic circuitry regulating tumor grade, lymph-node status, and tumor size may be distinct between the different subtypes, which will require the identification of subtype-specific molecular signatures for these other subtypes. Given that ER− and ERBB2+ tumors are typically associated with highly aggressive clinical courses, addressing this issue will undoubtedly be a crucial issue for additional research efforts.

Acknowledgments

P. Tan thanks Hui Kam Man for encouragement and support.

References

1. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer: I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;19:403–10.
2. Galea MH, Blamey RW, Elston CW, Ellis IO. The Nottingham prognostic index in primary breast cancer. Breast Cancer Res Treat 1992;22:207–19.
3. Balslev I, Axelsson CK, Zedeler K, Ramussen BB, Carstensen B, Mouridsen HT. The Nottingham prognostic index applied to 9,419 patients from the studies of the Danish Breast Cancer Cooperative Group (DBCG). Breast Cancer Res Treat 1994;32:281–90.

4. Sauerbrei W, Hubner K, Schmoor C, Schumacher M. Validation of existing and development of new prognostic classification schemes in node negative breast cancer. Breast Cancer Res Treat 1997;42:149–63.
5. Gilchrist K, Kalish WL, Gould VE, et al. Interobserver reproducibility of histopathological features in stage II breast cancer. An ECOG study Breast Cancer Res Treat 1985;5:3–10.
6. Buettner P, Garbe C, Guggenmoos-Holzmann I. Problems in defining cutoff points of continuous prognostic factors: Example of tumor thickness in primary cutaneous melanoma J Clin Epidemiol 1997;50:1201–10.
7. Perou CM, Sorlie T, Eisen MB, et al. Molecular portraits of human breast tumors. Nature (Lond) 2000;406:747–52.
8. 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.
9. 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 USA 2003;100:10393–8.
10. van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature (Lond) 2002;415:530–6.
11. West M, Blanchette C, Dressman H, et al. Predicting the clinical status of human breast cancer by using gene expression profiles. Proc Natl Acad Sci USA 2001;98: 11462–7.
12. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 1998;95:14863–8.
13. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci USA 2001;98:5116–21.
14. Golub TR, Slonim DK, Tamayo P, et al. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science (Wash D C) 1999;286:531–7.
15. Sorlie TC, Perou M, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA 2001;98:10869–74.
16. Martin KJ, Kritzman BM, Price LM, et al. Linking gene expression patterns to therapeutic groups in breast cancer. Cancer Res 2000;60:2232–8.
17. Sundquist M, Thorstenson S, Brudin L, Nordenskjold B. Applying the Nottingham prognostic index to a Swedish breast cancer population. Breast Cancer Res Treat 1999;53:1–8.
18. Barbareschi M, Caffo O, Veronese S, et al. Bcl-2 and p53 expression in node-negative breast carcinoma: a study with long-term follow-up. Hum Pathol 1996; 27:1149–55.
19. Frkovic-Grazio S, Bracko M. Long term prognostic value of Nottingham histological grade and its components in early (pT1NOM0) breast carcinoma. J Clin Pathol 2002;55:88–92.
20. Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. Nat Genet 2003;33:49–54.