Abstract

Purpose: Current histopathologic systems for classifying breast tumors require evaluation of multiple variables and are often associated with significant interobserver variability. Recent studies suggest that gene expression profiles may represent a promising alternative for clinical cancer classification. Here, we investigated the use of a customized microarray as a potential tool for clinical practice.

Experimental Design: We fabricated custom 188-gene microarrays containing expression signatures for three breast cancer molecular subtypes [luminal/estrogen receptor (ER) positive, human epidermal growth factor receptor 2 (HER2), and “basaloid”], the Nottingham prognostic index (NPI-ES), and low histologic grade (TuM1). The reliability of these multiple-signature arrays (MSA) was tested in a prospective cohort of 165 patients with primary breast cancer.

Results: The MSA-ER signature exhibited a high concordance of 90% with ER immunohistochemistry reported on diagnosis (P < 0.001). This remained unchanged at 89% (P < 0.001) when the immunohistochemistry was repeated using current laboratory standards. Expression of the HER2 signature showed a good correlation of 76% with HER2 fluorescence in situ hybridization (FISH; ratio ≥2.2; P < 0.001), which further improved to 89% when the ratio cutoff was raised to ≥5. A proportion of low-level FISH-amplified samples (ratio, 2.2–5) behaved comparably to FISH-negative samples by HER2 signature expression, HER2 quantitative reverse transcription-PCR, and HER2 immunohistochemistry. Luminal/ER+ tumors with high NPI-ES expression were associated with high NPI scores (P = 0.001), and luminal/ER+ TuM1-expressing tumors were significantly correlated with low histologic grade (P = 0.002) and improved survival outcome in an interim analysis (hazard ratio, 0.2; P = 0.019).

Conclusion: The consistency of the MSA platform in an independent patient population suggests that custom microarrays could potentially function as an adjunct to standard immunohistochemistry and FISH in clinical practice.

Carcinoma of the breast is a major cause of worldwide morbidity and mortality in females (1). Two important factors in clinical breast cancer classification include determining the estrogen receptor (ER) and the human epidermal growth factor receptor 2 (HER2) status of tumors because both ER and HER2...
platforms (23) and ethnic populations (24), and the reproducibility of gene expression signature–based predictions has been confirmed in replicate experiments (25, 26). Compared with conventional assays, molecular profiling platforms offer the potential advantage of measuring multiple biomarkers and signatures in a single test. Nevertheless, to establish the true clinical utility of molecular profiling, it is essential to validate such signatures in independent, prospectively defined patient cohorts (27). Although some profiling assays have recently undergone clinical validation (28, 29), these studies have been mostly done on U.S. and European cohorts (28, 30), and there is currently no similar project in Asia. Various clinical, epidemiologic, and molecular differences in breast cancer have been reported between different ethnic groups (refs. 24, 31–36; see Discussion), thus raising the need to assess the reliability of these signatures in an Asian cohort as well.

Here, we used electrochemical in situ synthesis to fabricate a custom multi-signature array (MSA) for breast cancer comprising several expression signatures previously defined in our local Asian population (14, 15, 24, 37). We assessed the robustness of the MSA on an independent cohort of 165 predominantly Chinese patients with primary breast cancer. We found the MSA to be highly reliable with respect to several clinical variables, supporting the use of custom arrays to function as a potential adjunct to standard immunohistochemistry and FISH in clinical practice.

**Materials and Methods**

**MSA chip design.** The breast cancer MSA was fabricated using CombiMatrix electrochemical in situ synthesis technology and contains probes for 188 genes representing five previously identified breast cancer molecular signatures. These include signatures for the luminal/ER+, HER2, and basaoid subtypes of breast cancer, which we had previously shown to be present in the Asian breast cancer population (14, 24); Nottingham prognostic index expression signature (NPI-ES; ref. 15), a pathologic stratification staging system for breast cancer prognostication; and TuM1 (37), a signature for low histologic grade that may also serve as a potential predictive biomarker for tamoxifen response. Notably, both the NPI-ES and TuM1 signatures are specific to luminal/ER+ tumors. A complete list of MSA genes is provided in Supplementary Table S1. For each gene, we used vendor-provided software to design five independent (35-mer) probes, and each probe was replicated eight times across the MSA at random locations. We also included 45 control genes from different cellular pathways as internal standards for normalizing RNA expression levels (38). Finally, 25-mer sequences (five replicates per probe) designed against the 20× Eukaryotic Hybridization Control Kit (Affymetrix) were included to facilitate monitoring of the hybridization process.

**MSA standard operating protocol.** A full description of the MSA standard operating protocol is provided in the Supplementary information S1. Total RNA was extracted from frozen tissue using Trizol (Invitrogen) and RNA quality was assessed using an Agilent Technologies NanoDrop. Samples were processed for MSA profiling if there were clear 185 and 285 peaks with no minor peaks present, whereas samples without the 185/285 peaks were regarded as degraded and excluded.

**MSA data analysis.** Microarray Imager software (CombiMatrix) was used to generate probe-level intensities. Individual MSAs were normalized by median scaling the expression of control genes to the same value. To calibrate the MSAs, we generated MSA profiles for a training set of 16 tumors that had previously been profiled using Affymetrix U133-Plus GeneChips (15). For each gene, we weighted the MSA probes by their strength of correlation to the Affymetrix expression data. Average-linkage hierarchical clustering using Pearson correlation distance metrics was done using Cluster and displayed using TreeView software. Individual MSA profiles in the validation set were classified using Support Vector Machine algorithms (Genedata AG). The MSA gene expression data have been deposited into the Gene Expression Omnibus (accession no. GSE7422).

**Study design and patient selection criteria.** Human breast tissues from the period 2000 to 2004 were obtained from the National Cancer Centre Tissue Repository after appropriate approvals from the National Cancer Centre Research and Ethics Committees and the Singapore General Hospital Ethics Committee. Samples were selected using the following predefined inclusion criteria: histologic diagnosis of invasive ductal carcinoma of the breast, newly diagnosed nonmetastatic breast cancer with no prior treatment, no history of other cancers or the presence of synchronous contralateral breast cancer, availability of frozen tissue, and complete clinical data. Exclusion criteria were tumors with no detectable invasive ductal carcinoma on H&E sections of the frozen sample and degraded RNA. All patients underwent surgical treatment, mastectomy, or breast conservation surgery and were given adjuvant therapy (chemotherapy, hormonal therapy, and chest wall radiotherapy) according to institutional practice guidelines. After surgery, two thirds of patients had adjuvant chemotherapy and two thirds had adjuvant hormonal therapy based on the ER and/or progesterone immunohistochemical status at the time of surgery. Tissue harvesting, preparation for storage, storage, and release of tissue were done by the National Cancer Centre Tissue Repository and under Repository Protocols.

**Clinical and histopathologic variables.** Clinicopathologic variables including tumor size, nodal status, histologic grade (modified Bloom-Richardson grading), lymphovascular invasion, tumor type, ER, and HER2 immunohistochemical status were collected. Repeat immunohistochemical assessments were done by two pathologists. Immunohistochemistry antibodies were SP1 for ER and SP3 for HER2 (Labvision). HER2 immunohistochemical staining was scored according to HercepTest specifications (DAKO). HER2 FISH was done on the frozen tissue samples using the PathVysion kit from Vysis, Inc., in which a HER2/CEP ratio of ≥2.2 was deemed as amplified according to the 2007 American Society of Clinical Oncology/College of American Pathologists guidelines (6). Immunohistochemistry and FISH were done in the Department of Pathology, Singapore General Hospital, a College of American Pathologists accredited laboratory. Concordance between HER2 FISH and immunohistochemistry obtained by our laboratory reported 100% for immunohistochemistry 3+ and 85% for immunohistochemistry 2+ in the recent review of consecutive routine cases of invasive breast cancer done on paraffin sections. This percentage is well within the variables recommended for HER2 testing by Wolff et al. (6). FISH is not routinely done for immunohistochemistry-negative and 1+ cases in our institution, and hence these data are not available.

To calculate the correlation of subtype classification between MSA and Affymetrix platforms and between MSA and clinical variables (FISH, immunohistochemistry, and NPI), k tests were used. To compare associations for other clinical variables between patient subject groups, t test, one-way ANOVA, χ² test, and Fisher’s exact test were used. Logistic regression was used to assess the association of individual NPI components (tumor size, histologic grade, and lymph node status) with the MSA NPI-ES signature. Survival differences were plotted using Kaplan-Meier curves and Cox regression was used for the univariate analysis. Cox regression with stepwise forward hierarchical selection was used for the multivariate analysis. The probability of inclusion into the multivariate model was 0.1.

**Quantitative RT-PCR for HER2 gene expression.** Total RNA was reverse transcribed using Superscript II Reverse Transcriptase (Invitrogen) and quantitative PCR was done using HER2/ERBB2 TaqMan probes (Hs00170433_m1) on a 7500 Real-time system (Applied Biosystems). TaqMan glyceraldehyde-3-phosphate dehydrogenase...
probes (Hs99999905_m1) were used as internal controls. All tumor samples were run in triplicates and compared against negative controls, which were either water controls or tissue samples from reduction mammoplasty for benign breast hypertrophy.

**Results**

**Optimization of the breast cancer MSA.** Using a standard operating protocol, we carried out several assays to assess the technical variability of the MSA, including replicate sample assays by the same operator, replicate sample assays by different operators, and replicate assays by different operators at different times. The assays were done both using RNA from breast cancer cell lines (MCF7) and primary tumors. As shown in Fig. 1A, we attained good reproducibility in the MSA standard operating protocol with correlation coefficients ranging from 0.92 to 0.96 across the different assays (Supplementary Table S2). To assess long-term MSA reproducibility, a set of RNA samples was extracted from primary tumors, and four pairs of replicate hybridizations were done in which each member of a replicate pair was spaced ~12 months apart. The replicate profiles also showed a high degree of concordance [correlation coefficients, 0.94 (0.85-0.99)], indicating that the MSA is likely to have good long-term technical reproducibility. This level of consistency is similar to that achieved by other microarray platforms (25, 39) and suggests that the technical performance of the MSA is likely to be robust toward different operators and sample types.

To calibrate the MSA against primary tumor samples, we profiled 16 breast tumors (the "training set") previously assayed using standard Affymetrix GeneChip technology. Using a hierarchical clustering algorithm, we computed weights for each MSA probe such that the 16 training set tumors segregated into similar sample clusters (luminal/ER+, HER2, and basaloid) as the Affymetrix arrays (Fig. 1B).

**Patient characteristics of the validation set.** Two hundred sixty-seven cases met our predefined inclusion criteria (see Materials and Methods). On histopathologic assessment of tumor content, 38 tumors with no invasive ductal carcinoma were excluded, and a further 61 tumors were excluded for degraded RNA. There was insufficient RNA in three cases and, hence, these were also excluded. In total, 165 (62%) tumors were finally selected as the validation test set. The mean presence of invasive tumor was 72% (SD, 22%) and the RNA Bioanalyzer ratio was 1.5 (SD, 0.3) in the validation test set of 165 cases.

The clinicopathologic variables of the breast cancer patients and their tumors are summarized in Table 1. Importantly, there were no significant differences in these clinical variables between the training set, the “high-quality” validation test set, and the excluded cases (P > 0.05), suggesting that the
Validation cases were classified as luminal/ER positive by the MSA and 57 cases as ER negative. A correlation to ER immunohistochemistry revealed a 90% concordance between the MSA and immunohistochemistry classifications \( (P < 0.001, \kappa \text{ test; Table 2A}) \), indicating a good correlation between gene expression and immunohistochemistry. Similar associations about breast tumor ER status have previously been reported (11). To investigate the discrepant classifications, we retrieved paraffin-embedded tumor blocks for all 165 validation set cases from the hospital pathology archives and subjected these tissues to a repeat ER immunohistochemistry procedure; only 2 cases did not have remaining tissue for reassessment. The correlation of the repeat immunohistochemistry with the ER MSA remained good at 89% \( (P < 0.001, \kappa \text{ test}) \). Of the 14 cases with discrepant immunohistochemistry and ER MSA classifications, 6 reassessments supported the ER MSA and the other 8 did not.

### Table 1. Clinicopathologic features of patients and tumors in the training, validation and excluded sets

<table>
<thead>
<tr>
<th>All invasive ductal carcinoma of the breast</th>
<th>Training</th>
<th>Test</th>
<th>Excluded</th>
<th>( P )</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. cases</td>
<td>16</td>
<td>165</td>
<td>102</td>
<td>0.953</td>
</tr>
<tr>
<td>Age (y)</td>
<td>52</td>
<td>56</td>
<td>59</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>30-79</td>
<td>33-83</td>
<td>35-87</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>11</td>
<td>12</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Age group (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;54 y old</td>
<td>11 (69)</td>
<td>82 (50)</td>
<td>44 (43)</td>
<td>0.143</td>
</tr>
<tr>
<td>&gt;54 y old</td>
<td>5 (31)</td>
<td>83 (50)</td>
<td>58 (57)</td>
<td></td>
</tr>
<tr>
<td>Tumor size (cm)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>3.7</td>
<td>3.6</td>
<td>3.7</td>
<td>0.545*</td>
</tr>
<tr>
<td>Range</td>
<td>1.6-6.0</td>
<td>0.6-12.5</td>
<td>0.3-12.0</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>1.3</td>
<td>1.9</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>T stage (%) ( \dagger )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( T_2 )</td>
<td>4 (25)</td>
<td>36 (22)</td>
<td>13 (13)</td>
<td>0.452</td>
</tr>
<tr>
<td>( T_3 )</td>
<td>10 (62)</td>
<td>102 (61)</td>
<td>72 (70)</td>
<td></td>
</tr>
<tr>
<td>( T_4 )</td>
<td>2 (13)</td>
<td>23 (14)</td>
<td>16 (16)</td>
<td></td>
</tr>
<tr>
<td>Grade</td>
<td>0 (0)</td>
<td>5 (3)</td>
<td>1 (1)</td>
<td></td>
</tr>
<tr>
<td>Lymph node status (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>6 (38)</td>
<td>66 (40)</td>
<td>51 (50)</td>
<td>0.245</td>
</tr>
<tr>
<td>Positive</td>
<td>10 (62)</td>
<td>100 (60)</td>
<td>51 (50)</td>
<td></td>
</tr>
</tbody>
</table>

*Comparison between test and excluded cases only.
\( \dagger \) T stage: tumor-node-metastasis classification based on the American Joint Committee on Cancer staging manual, 6th edition.

MSA classification and validation against Affymetrix GeneChip technology. We generated MSA profiles for all 165 validation set tumors. Eighty-nine (54%) of the cases were luminal/ER+, 28 (17%) were HER2, and 29 (18%) were basaloid; 19 (11%) were indeterminate. To compare the robustness of the MSA profiles to a different technology platform, RNA from a subset of the validation set tumors \( \{ \text{training set, validation and excluded} \} \). Of the 83 cases, 9 (10%) were associated with indeterminate calls on either the MSA or Affymetrix platforms. Of the remaining 74 cases, there was a high concordance of 95% between both technologies, with the MSA classifying 51 tumors as luminal/ER+, 12 as HER2, and 11 as basaloid, and the Affymetrix platform predicting 51 as luminal/ER+, 14 HER2, and 9 basaloid (Supplementary Table S3). The \( \kappa \) test confirmed that the level of observed concordance was highly significant \( (\kappa = 0.89, P < 0.001) \). This result suggests that the signature genes on the MSA are likely to have good cross-platform transportability.

**MSA validation against ER status.** Eighty-nine of 146 (61%) validation cases were classified as luminal/ER positive by the MSA and 57 cases as ER negative. A correlation to ER immunohistochemistry revealed a 90% concordance between the MSA and immunohistochemistry classifications \( (P < 0.001, \kappa \text{ test; Table 2A}) \), indicating a good correlation between gene expression and immunohistochemistry. Similar associations about breast tumor ER status have previously been reported (11). To investigate the discrepant classifications, we retrieved paraffin-embedded tumor blocks for all 165 validation set cases from the hospital pathology archives and subjected these tissues to a repeat ER immunohistochemistry procedure; only 2 cases did not have remaining tissue for reassessment. The correlation of the repeat immunohistochemistry with the ER MSA remained good at 89% \( (P < 0.001, \kappa \text{ test}) \). Of the 14 cases with discrepant immunohistochemistry and ER MSA classifications, 6 reassessments supported the ER MSA and the other 8 did not.

### Table 2. Validation of MSA expression signatures

#### (A) ER status by immunohistochemistry

<table>
<thead>
<tr>
<th>MSA signature</th>
<th>ER+</th>
<th>ER-</th>
</tr>
</thead>
<tbody>
<tr>
<td>ER+</td>
<td>85</td>
<td>4</td>
</tr>
<tr>
<td>ER-</td>
<td>10</td>
<td>47</td>
</tr>
</tbody>
</table>

\( \kappa = 0.795, P < 0.001 \)

#### (B) HER2 status by FISH

<table>
<thead>
<tr>
<th>MSA signature</th>
<th>FISH ratio &lt;2.2</th>
<th>FISH ratio ≥2.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>HER2+</td>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>HER2-</td>
<td>86</td>
<td>31</td>
</tr>
</tbody>
</table>

\( \kappa = 0.433, P < 0.001 \)

#### (C) HER2 status by FISH

<table>
<thead>
<tr>
<th>MSA signature</th>
<th>FISH ratio &lt;5</th>
<th>FISH ratio ≥5</th>
</tr>
</thead>
<tbody>
<tr>
<td>HER2+</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>HER2-</td>
<td>109</td>
<td>8</td>
</tr>
</tbody>
</table>

\( \kappa = 0.646, P < 0.001 \)

#### (D) NPI

<table>
<thead>
<tr>
<th>NPI-ES</th>
<th>NPI &lt;4</th>
<th>NPI ≥4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>20</td>
<td>17</td>
</tr>
<tr>
<td>High</td>
<td>12</td>
<td>40</td>
</tr>
</tbody>
</table>

\( \kappa = 0.316, P = 0.001 \)

#### (E) Tumor grade

<table>
<thead>
<tr>
<th>TuM1</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>15</td>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>Low</td>
<td>2</td>
<td>22</td>
<td>20</td>
</tr>
</tbody>
</table>

\( P = 0.002, \text{ Fisher's exact test} \)

Note: D and E are evaluated in a subset of 89 luminal/ER+ breast tumors.

*ER expression signature correlation with ER immunohistochemical status.

\( \dagger \) HER2 expression signature correlation with HER2 copy number status as measured by HER2/CEP FISH ratio. HER2/CEP FISH ratio ≥2.2 was deemed amplified.

\( \ddagger \) Highly amplified FISH samples (ratio ≥5) considered as one group was compared with low-amplified FISH positive (ratio 2.2-5) and FISH-negative samples as the second group.

\( 
\text{FISH ratio of} \geq 5 \text{ for medium and high positive HER2}.

\( \ddagger \) Correlation of NPI-ES expression signature against NPI using cutoff of 4.

\( \dagger \) Correlation of TuM1 gene signature against tumor grade.
classification, with some tumors originally classified as ER+ by immunohistochemistry, now supporting an ER-negative immunohistochemistry classification and vice versa (Fig. 2; Supplementary Table S4).

MSA validation against HER2 receptor status. Twenty-eight of 146 (19%) of cases were classified as HER2 positive by the MSA and the rest as HER2 negative. There was a good agreement of 76% between HER2 status by MSA and FISH using the established HER2/CEP FISH ratio cutoff of ≥2.2 recommended by the American Society of Clinical Oncology/College of American Pathologists guidelines (ref. 6; \( P < 0.001 \), \( \kappa = 0.433 \); Table 2B). Only one case did not have sufficient tissue for reassessment by FISH. Thus, there seems to be a strong positive correlation between MSA HER2 signature expression and high-level HER2 FISH positivity.

Interestingly, we observed a further improvement in concordance between the HER2 FISH and MSA results (from 76% to 89%) when the tumors were divided into either tumors with negative or low-level HER2 amplification (FISH ratio <5) or tumors with medium to high-level HER2 amplification (FISH ratio ≥5; \( \kappa = 0.646, P < 0.001 \); Table 2C). To ask if this could be due to poor MSA sensitivity, we independently measured HER2 gene and protein expression in a subset of FISH-negative, low-level amplified, and high-level amplified samples by quantitative RT-PCR, a more sensitive method for assaying gene expression. We thus, there seems to be a strong positive correlation between MSA HER2 signature expression and high-level HER2 FISH positivity.

To ask if this could be due to poor MSA sensitivity, we independently measured HER2 gene and protein expression in a subset of FISH-negative, low-level amplified, and high-level amplified samples by quantitative RT-PCR, a more sensitive method for assaying gene expression. We found that HER2 gene expression in the FISH-negative and low-level amplified samples were highly similar even by quantitative RT-PCR (\( P = 0.53 \)), whereas high-level amplified samples exhibited significantly elevated HER2 gene expression (\( P = 0.04 \) and \( P = 0.007 \), compared with FISH-negative and low-level amplified samples, respectively; Fig. 3A). At the protein level, > 90% of low-level amplified cases showed absent or marginal (immunohistochemistry 1+) HER2 protein expression similar to FISH-negative samples (Fig. 3B; Supplementary Table S5B). In contrast, two thirds of high-level amplified cases showed moderate (immunohistochemistry 2+) to strong (immunohistochemistry 3+) HER2 protein overexpression (Fig. 3B; Supplementary Table S5B; \( \chi^2 \) test, \( P < 0.001 \)). To confirm this finding, we subjected all available cases in the validation set to a repeat HER2 immunohistochemical assessment. The slides were scored and reviewed by two pathologists under current criteria and compared with FISH. Even with the increase in the sample size, our observation remained unchanged. It is noted that HER2 amplification in immunohistochemistry-negative and immunohistochemistry 1+ cases has previously been reported in other series (42).

These results suggest that the apparent bias of MSA HER2 predictions for high-level FISH amplifications is unlikely to be simply due to a lack of MSA sensitivity but may be due to certain low-level amplification tumors behaving more like FISH-negative samples (see Discussion).

MSA validation against the NPI and low tumor grade. The NPI is a clinicopathologic staging system that incorporates tumor size, tumor grade, and lymph node status for breast cancer prognostication. Fifty-two of 89 (58%) luminal/ER+ cases were classified by the MSA as expressing high levels of the NPI-ES expression signature, whereas 37 cases were classified as NPI-ES negative. Using a previously defined NPI cutoff value of 4 (15), with lower NPI values indicating good prognosis and high NPI values indicating poor prognosis, a concordance of 67% between NPI-ES expression and NPI status (\( \kappa = 0.316, P = 0.001 \); Table 2D) was observed. Of the individual NPI components, only histologic grade (\( P = 0.004 \)) was significantly associated with NPI-ES expression in this set of tumors, whereas lymph node status (\( P = 0.103 \)) and tumor size (\( P = 0.096 \)) were
not (Supplementary Table S6). This result validates the association between the NPI and NPI-ES in an independent patient population.

The TuM1 expression signature was previously identified by our group as a potential biomarker for low histologic grade and predictor of tamoxifen response (37). Forty-five of 89 (51%) luminal/ER+ cases were classified as expressing high levels of the TuM1 expression signature, whereas 44 cases were classified as TuM1 negative. A comparison with histologic grade (modified Bloom-Richardson grading) revealed that 15 of 17 (88%) grade 1 tumors exhibited high levels of TuM1 expression compared with 30 of 72 grade 2 and 3 tumors (P = 0.002, Fisher’s exact test). These figures are consistent with our previous findings (37) and validate the association between TuM1 expression and tumor grade. Finally, the patient follow-up data in our validation series are still relatively short, with a mean follow-up time of only 4.4 years (SD, 1.5 years) and disease-free survival period of 4.1 years (SD, 1.7 years). Nevertheless, in an interim analysis, we found that patients with luminal/ER+ tumors expressing high levels of the TuM1 signature were associated with significantly improved disease-free survival compared with patients with low-TuM1 expressing tumors [hazard ratio (HR), 0.24; P = 0.01; Fig. 4]. In a univariate analysis involving patient age (≤54 or >54 years, representing premenopausal/perimenopausal and postmenopausal categories, respectively), tumor stage, lymph node status, histologic grade, lymphovascular invasion, NPI, TuM1, and NPI-ES, age (P = 0.046), tumor stage (P = 0.015), grade (P = 0.083), and TuM1 (P = 0.01) were significant predictors of disease-free survival. However, a subsequent multivariate analysis of these variables with a probability of inclusion of 0.1, which also included the use of adjuvant chemotherapy (P = 0.088), showed that only TuM1 (HR, 0.20; P = 0.019) remained as the only independent prognostic factor (Supplementary Table S7). Notably, the use of adjuvant chemotherapy (HR, 0.7; P = 0.225) in this cohort did not significantly affect disease-free survival. Because patients in this validation set who had ER+ tumors were the recipients of antihormonal therapy, this result is consistent with our initial hypothesis that TuM1 may predict response to antihormonal therapy.

The effect of archival protocol affects sample tissue quality. Of the 267 cases that were initially evaluated for this study, 165 (62%) met our inclusion criteria and were profiled on the MSA, whereas the remaining cases were excluded. We examined the excluded samples and found a striking association between the likelihood of a sample being excluded and the year of archival. Specifically, whereas the percentage of excluded cases for either degraded RNA or lack of tumor content was 56% for samples archived in 2000, this number was only 5% for samples archived from 2004, where only one sample (5%) was excluded for degraded RNA (Supplementary Table S8). This finding is consistent with an improvement.
of archival protocol over the years. Moreover, a review of the 19 cases wherein the MSA produced indeterminate calls showed that the tissue had good tumor content (72%; SD, 25%) and good quality RNA (Bioanalyzer ratio, 1.5; SD, 0.3), indicating that the inability to classify these samples is likely not to be a result of poor tissue quality. Thus, while not ignoring the strong necessity for improving and standardizing protocols for tissue handling and preservation, these results suggest that our standard operating protocol for the MSA platform may have general applicability to the current clinical setting with a low frequency of sample loss. One obvious way to increase number of cases suitable for MSA profiling would be to inspect the tissue block at the time of accrual to ensure adequacy of sampling.

Discussion

The primary objective of our study was to validate a series of gene expression signatures that we had previously described in our local Asian breast cancer population. These include signatures for molecular subtype classification (luminal/ER+, HER2, and basaloid), the NPI (NPI-ES), and low histologic grade (TuM1). Although signatures for breast cancer molecular subtype have already been repeatedly observed in a number of U.S. and European studies (28, 29), our study is novel because validation studies for the NPI-ES and TuM1 signatures have not been previously reported. We found the MSA to be highly reliable with respect to several clinical variables, supporting the use of custom arrays to function as a potential adjunct to standard immunohistochemistry and FISH in clinical practice.

One notable element in our study was the decision to validate these signatures using a customized microarray platform as opposed to a generic genome-wide array. To date, custom designed arrays have only been used in a few validation studies (39, 43, 44), and most of these have examined a single expression signature. In concept, however, the use of a focused array may be advantageous for the actual implementation of such devices in the clinical setting. First, limiting the array to informative genes may increase their cost-effectiveness (e.g., by establishing multiplex arrays in which multiple patient samples are profiled on a single array). Second, having limited gene content may also prove useful in transporting these informative gene sets across other technology platforms that are superior to current solid-substrate arrays in terms of sensitivity, assay speed, and sample quantity (e.g., quantitative RT-PCR). Third, compared with standard assays, the MSA is able to assess multiple biomarkers in a single test. Currently, at least two or three independent assays [immunohistochemistry (for ER and HER2) and FISH (for HER2)] are required to determine the ER and HER2 status of breast tumors, with each assay requiring different sets of technical protocols and analyze reagents. The ability to achieve comparable accuracies of ER and HER2 classification using a single test and technical protocol, coupled with the ability to also discern additional signatures (i.e., NPI-ES and TuM1), highlights the potential for such array platforms in the molecular diagnostic arena.

The cases in our validation series were derived from a population of predominantly Chinese patients. Interestingly, several reports have described intriguing epidemiologic, clinical, and molecular differences in breast cancer between different ethnic groups (31). For example, whereas the peak of breast cancer incidence in Western populations occurs in the postmenopausal age group and is relatively rare among those <40 years of age, many Asian populations display a striking premenopausal incidence peak (33, 35, 45). This may be due to age-related differences in breast cancer risk factors like parity and body size for women of different ethnicities. At the metabolic and histopathologic level, ethnic differences in the levels of circulating sex steroid hormones (46) and differing frequencies of breast cancer histologic subtypes have also been reported (47). Furthermore, at least one report has observed distinct gene expression patterns associated with African American breast cancer (48), and recent discoveries have shown that lung cancers in Asian women are frequently mutated in the EGFR gene (49). Taken collectively, these observations support the need for considering how ethnicity might influence the performance of such molecular diagnostic assays.

One interesting subpopulation in this validation series involved tumors possessing low levels of HER2 FISH amplification. These cases tended to be immunohistochemistry negative for HER2 and the majority were luminal/ER+ by the MSA. At present, trastuzumab (Herceptin) is clinically recommended in the adjuvant setting for HER2 FISH–positive samples, and the latest 2007 American Society of Clinical Oncology/College of American Pathologist guidelines recommend using a positive FISH ratio of 2.2 as a threshold (6, 42). More recently, RT-PCR has also been shown to be a valuable method for determining HER2 status (50). In this study, we observed a good concordance of 76% between HER2 MSA and FISH at the 2.2 threshold. However, the concordance further improved to 89% when the tumors were divided into negative/low positive (ratio <5) and medium-high level amplified categories (ratio ≥5). One possible explanation is that certain tumors with low-level FISH positivity (ratio, 2.2-5) may behave more similarly to FISH-negative samples than to tumors with high-level amplification (ratio >5), with no protein significant expression. Supporting this, it was reported that in a large series of 6,556 tissues analyzed with
HER2 immunohistochemistry and FISH, FISH amplification could be observed in HER2 immunohistochemistry-negative and immunohistochemistry 1+ cases (42). We confirmed this observation using three different independent assays: MSA HER2 expression signature, HER2 quantitative RT-PCR, and HER2 protein immunohistochemistry. These results suggest that it may prove valuable for future studies to investigate the validity of raising the FISH ratio to >2.2, the threshold that is currently recommended.

In conclusion, our results show the reproducibility and robustness of our gene signatures on the MSA, with a simple operating protocol that can measure multiple biomarkers simultaneously (i.e., the standard ER and HER2 markers), and in addition, the TuM1, a low-grade marker that has prognostic implications. One potential weakness of this study is that the clinical follow-up is still in the early years and immature, and hence the long-term prognostic power of some signatures cannot be conclusively assessed. Nevertheless, at least one signature (TuM1) was associated with improved survival, even in this interim analysis. Future work will involve assessing the long-term outcome of disease in our breast cancer patients and comparing them to similar outcome metrics in the United States and Europe.

Acknowledgments

We thank the Tissue Repository, National Cancer Centre, Singapore for their assistance and Sew Wan Hee for advice in mathematical analysis.

References

42. Owens MA, Herten BC, Da Silva MM. HER2 amplification ratios by fluorescence in situ hybridization...


Clinical Cancer Research

Clinical Validation of a Customized Multiple Signature Microarray for Breast Cancer


*Clin Cancer Res* 2008;14:461-469.

Updated version
Access the most recent version of this article at:
http://clincancerres.aacrjournals.org/content/14/2/461

Supplementary Material
Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2008/01/31/14.2.461.DC1

Cited articles
This article cites 47 articles, 22 of which you can access for free at:
http://clincancerres.aacrjournals.org/content/14/2/461.full.html#ref-list-1

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.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.