TESI DOCTORAL

PHENOTYPIC AND MOLECULAR CHARACTERIZATION OF THE INTRINSIC MOLECULAR SUBTYPES OF BREAST CANCER

ALEIX PRAT APARICIO
TESI DOCTORAL

PHENOTYPIC AND MOLECULAR CHARACTERIZATION OF THE INTRINSIC MOLECULAR SUBTYPES OF BREAST CANCER

DEPARTAMENT DE MEDICINA
UNIVERSITAT AUTÒNOMA DE BARCELONA (UAB)

2013

DOCTORAND
ALEIX PRAT APARICIO

DIRECTOR
Dr. JUAN ALBANEll MESTRES

TUTOR
Dr. JORDI GIRALT LOPEZ DE SAGREDO
Title: PHENOTYPIC AND MOLECULAR CHARACTERIZATION OF THE INTRINSIC MOLECULAR SUBTYPES OF BREAST CANCER

Introduction:

Implementation of screening/prevention programs and novel treatment strategies is decreasing breast cancer mortality. However, more than 120,000 estimated deaths due to breast cancer are expected annually in the US and Europe combined. A plausible explanation for this scenario is, in part, that we still lack a complete enough picture of the biologic heterogeneity of breast cancers with respect to molecular alterations, treatment sensitivity, and cellular composition. Importantly, this complexity is not entirely reflected by the main clinical parameters (age, node status, tumor size, histological grade) and pathological markers (estrogen receptor [ER], progesterone receptor [PR] and human epidermal growth factor receptor 2 [HER2]), all of which are routinely used in the clinic to stratify patients for prognostic predictions and to select treatments.

Studies based on global gene expression analyses have provided additional insights into this complex scenario. During the last 10 years, four molecular ‘intrinsic’ subtypes of breast cancer (Luminal A, Luminal B, HER2-enriched, and Basal-like) and a Normal Breast-like group have been identified and intensively studied. Known as the ‘intrinsic subtypes of breast cancer’, these groups of tumors have revealed critical differences in incidence, survival, and response to treatment. Importantly, the information provided by the intrinsic subtypes complements and expands the information provided by classical clinical-pathological markers.

Current knowledge of the biology of breast cancer has provided the basis of the various successful adjuvant and neoadjuvant treatment strategies: endocrine therapy for hormone receptor (HR)-positive disease (with or without chemotherapy), anti-HER2 therapies such as trastuzumab in combination or sequentially after chemotherapy for HER2+ disease, and chemotherapy for patients with triple-negative disease. However, the biological diversity displayed by the breast cancer intrinsic subtypes indicate that further sub-classification of patients into different treatment groups should be considered.

The HR+/HER2− group of tumors is mainly composed of two subtypes: Luminal A (good prognosis, chemoresistant and endocrine sensitive) and Luminal B (poor prognosis, mainly chemoresistant and endocrine less sensitive). As discussed above, a main difference between A vs. B is proliferation status, which is low in Luminal A and high in Luminal B tumors. In this context, genomic prognostic assays such as the OncoTypeDX and the MammaPrint signature (or even the pathological marker Ki-67) have the ability to identify tumors with high risk of recurrence, which are mainly Luminal B tumors. An important issue here will be to find which ER+ patients benefit from chemotherapy. As suggested by data from neoadjuvant clinical trials, Luminal B tumors benefit more from chemotherapy than Luminal A tumors, although only less than ~20% of Luminal B patients eventually achieve a pCR. This increased benefit with the administration of chemotherapy in Luminal B is concordant with data coming from NSABP-B20 trial where only node-negative HR+ patients with high OncoTypeDX RS benefited from adjuvant chemotherapy.
In the HR+/HER2+ group of tumors, two subtypes are mainly identified: Luminal B and HER2-enriched. Here a major challenge will be to elucidate differences between the two molecular subtypes in terms of efficacy of chemotherapy, anti-hormonal therapy, and anti-HER2 therapy. For example, are HR+/HER2+/Luminal B tumors less or more sensitive to anti-HER2 therapies than HR+/HER2+/HER2-enriched tumors, and do they respond better to anti-hormonal therapies than HR+/HER2+/HER2-enriched tumors?

Within HR−/HER2+ tumors, ~50–88% of these fall into the HER2-enriched subtype, followed by the other poor prognostic subtypes. Here the challenge will be to determine if HR−/HER2+ that are not of the HER2-enriched subtype, benefit from anti-HER2 therapies, and if HER2+ tumors that are not of the HER2-enriched subtype show similar or different response rates to trastuzumab when compared to HER2+/HER2-enriched tumors. Finally, within triple-negative disease, Basal-like and Claudin-low are the most frequent subtypes identified. Further studies focusing on the efficacy of particular chemotherapies and/or targeted therapies such as the PARP inhibitors and/or anti-CSC therapies in these subgroups of patients are warranted. It will be important to determine if Basal-like and Claudin-low tumors show similar responses to common therapies as they may given their expression similarities, or they may not given their differences including vast differences in proliferation rates.

**Summary of Results:**

In the first study, we evaluated the ability of six clinically relevant genomic signatures to predict relapse in patients with ER+ tumors treated with adjuvant tamoxifen only. To accomplish this, we combined four microarray datasets, and we evaluated research-based versions of PAM50 intrinsic subtyping and risk of relapse (PAM50-ROR) score, 21-gene recurrence score (OncotypeDX), Mammaprint, Rotterdam 76 gene, index of sensitivity to endocrine therapy (SET) and an estrogen-induced gene set. Distant relapse-free survival (DRFS) was estimated by Kaplan–Meier and log-rank tests, and multivariable analyses were done using Cox regression analysis. Harrell's C-index was also used to estimate performance.

Our results showed that all signatures were prognostic in patients with ER+ node-negative tumors, whereas most were prognostic in ER+ node-positive disease. Among the signatures evaluated, PAM50-ROR, OncotypeDX, Mammaprint and SET were consistently found to be independent predictors of relapse. A combination of all signatures significantly increased the performance prediction. Importantly, low-risk tumors (>90% DRFS at 8.5 years) were identified by the majority of signatures only within node-negative disease, and these tumors were mostly luminal A (78%–100%).

Thus, we concluded that most established genomic signatures were successful in outcome predictions in ER+ breast cancer and provided statistically independent information. From a clinical perspective, multiple signatures combined together most accurately predicted outcome, but a common finding was that each signature identified a subset of luminal A patients with node-negative disease who might be considered suitable candidates for adjuvant endocrine therapy alone.
In our second study, we showed that three genes (i.e. biomarkers) do not fully recapitulate the entire biological diversity displayed by breast cancer, and that the PAM50 subtype predictor is better. The reason behind this study is that it has recently been proposed that a three-gene model (SCMGENE) that measures ESR1, ERBB2, and AURKA identifies the major breast cancer intrinsic subtypes and provides robust discrimination for clinical use in a manner very similar to a 50-gene subtype predictor (PAM50). However, the clinical relevance of both predictors was not fully explored, which is needed given that a ~30 % discordance rate between these two predictors was observed.

Using the same datasets and subtype calls provided by Haibe-Kains and colleagues, we compared the SCMGENE assignments and the research-based PAM50 assignments in terms of their ability to (1) predict patient outcome, (2) predict pathological complete response (pCR) after anthracycline/taxane-based chemotherapy, and (3) capture the main biological diversity displayed by all genes from a microarray. In terms of survival predictions, both assays provided independent prognostic information from each other and beyond the data provided by standard clinical-pathological variables; however, the amount of prognostic information was found to be significantly greater with the PAM50 assay than the SCMGENE assay. In terms of chemotherapy response, the PAM50 assay was the only assay to provide independent predictive information of pCR in multivariate models. Finally, compared to the SCMGENE predictor, the PAM50 assay explained a significantly greater amount of gene expression diversity as captured by the two main principal components of the breast cancer microarray data. Our results show that classification of the major and clinically relevant molecular subtypes of breast cancer are best captured using larger gene panels.

Finally, in our third study, we improved the corrent immunohistochemical (IHC)-based definitions of luminal A breast cancer. The reason behind this study is that Luminal A and B IHC-based definitions are imperfect when compared with multigene expression-based assays. To accomplish this, we collected gene expression and pathologic features from primary tumors across five independent cohorts: British Columbia Cancer Agency (BCCA) tamoxifen-treated only, Grupo Español de Investigación en Cáncer de Mama 9906 trial, BCCA no systemic treatment cohort, PAM50 microarray training data set, and a combined publicly available microarray data set. Optimal cutoffs of percentage of progesterone receptor (PR) -positive tumor cells to predict survival were derived and independently tested. Multivariable Cox models were used to test the prognostic significance.

Our results showed that the clinicopathologic comparisons among luminal A and B subtypes consistently identified higher rates of PR positivity, human epidermal growth factor receptor 2 (HER2) negativity, and histologic grade 1 in luminal A tumors. Quantitative PR gene and protein expression were also found to be significantly higher in luminal A tumors. An empiric cutoff of more than 20% of PR-positive tumor cells was statistically chosen and proved significant for predicting survival differences within IHC-defined luminal A tumors independently of endocrine therapy administration. Finally, no additional prognostic value within hormonal receptor (HR) -positive/HER2-negative disease was observed with the use of the IHC4 score when intrinsic IHC-based subtypes were used that included the more than 20% PR-positive tumor cells and vice versa. We concluded that semiquantitative IHC expression of PR adds prognostic value within the current IHC-based luminal A definition by improving the
identification of good outcome breast cancers. The new proposed IHC-based definition of luminal A tumors is HR positive/HER2 negative/Ki-67 less than 14%, and PR more than 20%.

Overall, these results suggest that the information provided by the intrinsic subtypes, when combined with the current clinical-pathological markers, helps to further explain the biological complexity of breast cancer, increases the efficacy of current and novel therapies, and ultimately improves outcomes for breast cancer patients.
Concordance among gene expression-based predictors for ER-positive breast cancer treated with adjuvant tamoxifen

A. Prat1,2, J. S. Parker1, C. Fan1, M. C. U. Cheang1, L. D. Miller3, J. Bergh4,5, S. K. L. Chia6, P. S. Bernard7, T. O. Nielsen6,8, M. J. Ellis9, L. A. Carey1,10 & C. M. Perou1,11,12*

1Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, USA; 2Department of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain; 3Department of Cancer Biology, Comprehensive Cancer Center, Wake Forest School of Medicine, Winston Salem, USA; 4Department of Oncology-Pathology, Karolinska Institutet & Cancer Center Karolinska, Stockholm, Sweden; 5Department of Medical Oncology, Paterson Institute, Christie Hospital and Manchester University, Manchester, UK; 6British Columbia Cancer Agency, Vancouver, Canada; 7Department of Pathology, University of Utah Health Sciences Center, Salt Lake City, USA; 8Department of Pathology, University of British Columbia, Vancouver, Canada; 9Department of Medicine, Division of Oncology, Siteman Cancer Center at Washington University, St. Louis; 10Department of Medicine, Division of Hematology and Oncology, University of North Carolina, Chapel Hill; 11Departments of Genetics; 12Pathology & Laboratory Medicine, University of North Carolina, Chapel Hill, USA

Received 16 December 2011; revised 9 February 2012; accepted 10 February 2012

Background: ER-positive (ER+) breast cancer includes all of the intrinsic molecular subtypes, although the luminal A and B subtypes predominate. In this study, we evaluated the ability of six clinically relevant genomic signatures to predict relapse in patients with ER+ tumors treated with adjuvant tamoxifen only.

Methods: Four microarray datasets were combined and research-based versions of PAM50 intrinsic subtyping and risk of relapse (PAM50-ROR) score, 21-gene recurrence score (OncotypeDX), Mammaprint, Rotterdam 76 gene, index of sensitivity to endocrine therapy (SET) and an estrogen-induced gene set were evaluated. Distant relapse-free survival (DRFS) was estimated by Kaplan–Meier and log-rank tests, and multivariable analyses were done using Cox regression analysis. Harrell’s C-index was also used to estimate performance.

Results: All signatures were prognostic in patients with ER+ node-negative tumors, whereas most were prognostic in ER+ node-positive disease. Among the signatures evaluated, PAM50-ROR, OncotypeDX, Mammaprint and SET were consistently found to be independent predictors of relapse. A combination of all signatures significantly increased the performance prediction. Importantly, low-risk tumors (>90% DRFS at 8.5 years) were identified by the majority of signatures only within node-negative disease, and these tumors were mostly luminal A (78%–100%).

Conclusions: Most established genomic signatures were successful in outcome predictions in ER+ breast cancer and provided statistically independent information. From a clinical perspective, multiple signatures combined together most accurately predicted outcome, but a common finding was that each signature identified a subset of luminal A patients with node-negative disease who might be considered suitable candidates for adjuvant endocrine therapy alone.

Key words: breast cancer, genomics, luminal, mammaprint, oncotype, PAM50

introduction

Gene expression-based assays have been developed that can successfully predict outcomes in early-stage ER-positive (ER+) breast cancer beyond standard clinicopathological variables [1–5]. OncotypeDX recurrence score (GHI)2 and Mammaprint* (NKI70)3 are clinically available and currently being evaluated in two large prospective clinical trials (TAILORx and MINDACT) [6, 7]. Since then, other prognostic predictors such as the Rotterdam 76-gene signature (ROT76) [8, 9] and the risk of relapse (ROR) score based on the PAM50 subtype assay [10] have been developed using two different node-negative and adjuvant treatment-naïve populations.

Previous studies have also shown that many of these expression signatures are concordant for predicting outcomes [11, 12]. However, it is currently unknown if these findings are still valid in a more contemporary ER+ population treated with adjuvant endocrine therapy only [13]. Moreover, recent signatures specifically designed to track hormonal responsiveness, such as the estrogen-induced gene set (IE-IIE) [14] and the genomic index of sensitivity to endocrine therapy (SET) [15], can also predict survival in early-stage ER+ disease. Thus, estrogen-regulated gene signatures could be tracking ER+ tumors with high endocrine sensitivity.

*Correspondence to: Prof. C. M. Perou, Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, NC 27599, USA. Tel: +1-919-843-5740; Fax +1-919-843-5718; E-mail: cperou@med.unc.edu

© The Author 2012. Published by Oxford University Press on behalf of the European Society for Medical Oncology. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
From a clinical perspective, genomic assays are helping to identify patients with early-stage ER+ breast cancers who do not need chemotherapy and are effectively treated with adjuvant endocrine agents alone. Alternatively, they could identify groups of patients with ER+ tumors who are more likely to be biologically homogenous and/or who might benefit from specific treatment strategies. In this report, we evaluated the relapse prediction abilities of six independent genomic signatures using a cohort of ER-positive breast cancer patients treated with adjuvant tamoxifen only.

**methods**

**patients and samples**

Four different publicly available microarray datasets were combined together to create a single large set of 594 ER+ patients, all of whom received appropriate local therapy and adjuvant tamoxifen only (see supplemental Figure S1, available at *Annals of Oncology* online). Thousand fifty-three Affymetrix U133A CEL files from various publicly available microarray datasets (GSE17705 [MDACC298] [15], GSE6532 [LOI327] [16, 17], GSE12093 [ZHANG136] [18], GSE1456 [PAWITAN159] [19] and MDACC133 [20]) were processed using MAS 5.0 (R/Bioconductor) to generate probe-level intensities with a median array intensity of 600, and each expression value was log; transformed. To batch correct the gene expression data [21, 22], the probeset medians in each individual dataset were adjusted to the MDACC133 reference set accounting for differences in the proportion of clinical ER+/− samples; after batch correction, all ER− tumors were removed, as were all ER+ tumors not treated with tamoxifen only, thus leaving 594 tumors per microarrays.

**genomic predictors**

The four different publicly available microarray datasets were combined to create a single large set of 594 ER+ patients, all of whom received appropriate local therapy and adjuvant tamoxifen only (see supplemental Figure S1, available at *Annals of Oncology* online). A total of 594 patients were included in the current analysis. The clinical characteristics of the patients are shown in Table 1. The median follow-up time was censored at 8.5 years since it was the longest follow-up time in the PAW1159 [19] dataset. Univariate and multivariable analyses (MVA) were calculated using a Cox proportional regression model.

MVA prognostic models including all the signatures as independent continuous variables were built and assessed using a Cox model with the penalized least absolute shrinkage and selection operator (LASSO) method approach [24]. In each case, a training set (2/3 of the dataset) was randomly used to build a model, which was then applied to the testing set (i.e. the remaining 1/3). We repeated this procedure 200 times as previously carried out [5]. In each testing set, the prognostic performance of each model and each individual signature was estimated by calculating the concordance index (C-index) [25]. All statistical computations were carried out in R v.2.8.1 (http://cran.r-project.org).

**results**

**clinicopathological characteristics of the combined microarray and qRT-PCR PAM50 dataset**

We created a large dataset of 1380 ER-positive breast cancer patients treated with adjuvant tamoxifen only using publicly available microarray data (n = 594) and PAM50 qRT-PCR data only (n = 786) from the Nielsen series [4, 15–19]. Six hundred and ten and 699 patients were identified as having node-negative and node-positive disease, respectively (Table 1). As expected, luminal subtypes predominated (n = 1171, 84.9%). Non-luminal subtypes (HER2-enriched and basal-like) represented 9.1% (n = 125) of the patients. The normal breast-like samples were not further considered as these specimens are predominantly composed of normal breast tissue, which precludes the correct assignment to a tumor subtype for meaningful outcome predictions [1, 10].

The PAM50 intrinsic subtypes were prognostic for DRFS within node-negative and node-positive patients (Figure 1A and B). In node-negative disease, luminal A tumors showed a better outcome than luminal B [hazard ratio (HR) = 0.313, P < 0.0001], HER2-enriched (HR = 0.256, P < 0.0001) and basal-like (HR = 0.168, P < 0.001) subtypes. However, no statistical significant differences in DRFS were observed among the poor prognostic luminal B, HER2-enriched and basal-like subtypes. In node-positive disease, the PAM50 subtypes were also prognostic; of note, DRFS of both luminal subtypes was significantly lower compared with their counterparts in node-negative disease (luminal A, HR = 3.29 and luminal B, HR = 2.26, P < 0.0001 for both comparisons). Regardless of nodal status, both luminal subtypes had continued risk of relapse after 5 years; even the lowest risk node-negative luminal A subtype had 5-year DRFS of 96% that dropped to 91% by 8.5 years. A tendency for worse outcomes was also observed in node-positive HER2-enriched tumors compared with node-negative HER2-enriched tumors (HR = 1.91, P = 0.099).

**genomic relationships and biological significance**

For comparisons across different predictors, the combined dataset was confined to the 594 samples/tumor represented by Affymetrix microarray data. We first compared the gene overlap between any two signatures and found that ≤25% of the genes were shared between signatures (supplemental Table S2, available at *Annals of Oncology* online), except for 9 and 11 genes of the GHI signature (n = 21) that were present...
in the IE-IIE and PAM50, respectively, and 15 genes of the IE-IIE signature that were present in PAM50. In spite of relatively little gene overlap, all predictors were significantly correlated (Pearson correlation range 0.36–0.79; \( P < 0.0001 \) for each comparison), with PAM50-RORS, IE-IIE and GHI showing the highest correlation between them (>0.72, \( P < 0.0001 \), Pearson correlation; supplemental Table S2, available at Annals of Oncology online).

The observed correlations suggested that most predictors are tracking tumors with similar biology. To further explore this hypothesis, we evaluated the scores of each signature as a continuous variable and as group categories across the four major intrinsic subtypes (as defined by the PAM50 assay [10]). As expected, each predictor discriminated luminal A tumors from the luminal B subtype and from the rest of the subtypes \( (P < 0.0001, \text{Student’s } t\text{-test} \text{ supplemental Figure S3 and Table S3, available at Annals of Oncology online}) \). The high hormonal sensitivity groups (SET-high and IE-like) and low risk of recurrence groups (PAM50-RORS-low, PAM50-RORP-low, GHI-low, ROT76-good and NKI-good) were largely composed of luminal A tumors (>71%–100%).

**survival analyses within node-negative and node-positive disease**

Univariate DRFS analyses revealed that each signature, evaluated as a continuous variable or as group categories, was highly prognostic in patients with node-negative disease (supplemental Figure S4 and Table S4, available at Annals of Oncology online). As expected, Kaplan–Meier survival analyses showed highly significant differences in DRFS across the groups predicted to have good or intermediate or poor prognosis (PAM50-RORS, PAM50-RORP, GHI, ROT76 and NKI70) or the groups predicted to have high or intermediate versus low expression of ER-regulated genes (SET and IE-IIE). Importantly, all predictors identified groups of node-negative
patients with 93.7%–97.9% and 88.4%–96.2% DRFS at 5.0 and 8.5 years, respectively, although the number of patients in each group differed (Table 2): when limited to the combined microarray dataset and across the predictors with three risk categories (GHI, SET, PAM50-RORS and PAM50-RORP), the PAM50-RORS identified the largest number of low-risk patients (n = 140, 41%), followed by PAM50-RORP (n = 82, 24%), GHI (n = 47, 14%) and SET (n = 27%). Inclusion of the 786 ER+ patient qRT-PCR PAM50 Nielsen series data confirmed that both PAM50-RORP and PAM50-RORS identified 21%–36% of all node-negative patients (n = 551) as low risk [or alternatively they identify 41%–70% of all node-negative luminal A tumors (n = 280) as low risk], and the PAM50-RORP-low and PAM50-RORS-low groups showed a DRFS at 8.5 years of 96.09% and 91.21%, respectively (Table 2 and supplemental Figure S5, available at Annals of Oncology online).

In node-positive disease, univariate DRFS analyses revealed that most signatures were barely significant when evaluated as continuous variables (supplemental Figure S6 and Table S4, available at Annals of Oncology online). When evaluated as group categories, low risk of relapse or high expression of ER-regulated gene groups showed either no statistical significance or borderline significance in terms of DRFS compared with the predicted poor prognostic or low expressers of ER-regulated gene groups. More importantly, no predictor identified a clear node-positive group of patients treated with tamoxifen alone with a DRFS at 8.5 years >90%. Although these results could be related to the sample size, data for PAM50-RORS and PAM50-RORP in node-positive disease confirmed this finding when the qRT-PCR PAM50 Nielsen series was included for a total of 676 patients (supplemental Figure S5, available at Annals of Oncology online). Finally, similar to node-negative disease, the predicted low-risk outcome groups in node-positive disease were predominantly comprised of luminal A tumors (71%–100%; Table 2).

**prognostic prediction performance**

C-index values were calculated to estimate the performance of each genomic signature for predicting outcome (Figure 2). The C-index is a measure of the probability of concordance between the predicted and the observed survival, ranging from 0.5 (random) to 1 (perfect). Although no clear cut-off value has been defined, values >0.70 are indicative of good prediction accuracy [25]. In node-negative disease, the vast majority of signatures showed similar predictive abilities (mean C-index range of 0.70–0.73), except PAM50-PROLIF index (mean C-index of 0.69) and NKI70 (mean C-index of 0.64). Conversely, in node-positive disease, all predictors carried out worse than in node-negative (mean C-index range of 0.56–0.63).

Despite comparable prognostic performance of these signatures and high correlation values among them, we observed that these signatures generally retained their prognostic significance independent of each other when testing two signatures at a time in multivariate analyses (Table 3). Thus, we sought to determine if we could improve prognostic performance by integrating information from all signatures into a single model; we determined that the combined model was better at predicting outcome than individual signatures in node-negative disease (Figure 2A) but failed in node-positive disease (Figure 2B). However, the absolute increase in performance of the combined model within node-negative disease was modest (range 0.02–0.11).

**prognostic predictions within the intrinsic subtypes**

We explored the predictive ability of each signature within the predominant luminal A and B subtypes. In node-negative luminal A disease (n = 185), ROT76 and SET (Figure 3A) were found to be prognostic in univariate analyses, and patients with the low-risk group showed a DRFS at 8.5 years of 94%–96% (supplemental Table S5, available at Annals of Oncology online). When limited to the microarray dataset, the PAM50-

### Table 2. Low-risk group comparison among signatures

<table>
<thead>
<tr>
<th></th>
<th>Node-negative</th>
<th>Node-positive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low-risk group</td>
<td>Low-risk group</td>
</tr>
<tr>
<td>N</td>
<td>% of luminal A</td>
<td>DRFS at 8.5 years</td>
</tr>
<tr>
<td>RORP (PAM50)</td>
<td>82 (24%)</td>
<td>100%</td>
</tr>
<tr>
<td>RORS (PAM50)</td>
<td>140 (41%)</td>
<td>100%</td>
</tr>
<tr>
<td>PROLIF (PAM50)</td>
<td>72 (22%)</td>
<td>100%</td>
</tr>
<tr>
<td>GHI</td>
<td>47 (14%)</td>
<td>94%</td>
</tr>
<tr>
<td>ROT76&lt;sup&gt;b&lt;/sup&gt;</td>
<td>164 (48%)</td>
<td>85%</td>
</tr>
<tr>
<td>IE-IIE&lt;sup&gt;b&lt;/sup&gt;</td>
<td>235 (69%)</td>
<td>72%</td>
</tr>
<tr>
<td>NKI70&lt;sup&gt;b&lt;/sup&gt;</td>
<td>136 (40%)</td>
<td>78%</td>
</tr>
<tr>
<td>SET</td>
<td>26 (8%)</td>
<td>81%</td>
</tr>
<tr>
<td>RORP (PAM50)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>116 (21%)</td>
<td>100%</td>
</tr>
<tr>
<td>RORS (PAM50)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>197 (36%)</td>
<td>100%</td>
</tr>
<tr>
<td>PROLIF (PAM50)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>142 (26%)</td>
<td>99%</td>
</tr>
</tbody>
</table>

<sup>a</sup>Since proliferation (PROLIF) index does not have previously defined cut-offs, patients in the low-risk group are the ones with the lowest quartile expression.

<sup>b</sup>qRT-PCR PAM50 data from the Nielsen series have been included. N, number of patients in the low-risk group and the percentage from the total number of patients based on node status.

DRFS, distant relapse-free survival.
RORS and PAM50-RORP were trending toward significance (supplemental Table S5, available at Annals of Oncology online) and both were significant when the Nielsen series was included for a total of 280 luminal A patients (supplemental Table S5, available at Annals of Oncology online and PAM50-RORP in Figure 3B).

In node-positive luminal A disease \((n = 81)\) on the microarray dataset, GHI, NKI70 and IE-IIE were prognostic when evaluated as a continuous variable, and the combined low and intermediate risk GHI groups identified a group of significantly low-risk node-positive luminal A tumors \((n = 30, 37\%)\) with an outstanding DRFS at 8.5 years \((96\%, P < 0.01; \text{Figure 3C})\). When we included the qRT-PCR PAM50 Nielsen series dataset \((n = 326)\), PAM50-RORS and PAM50-RORP were found prognostic as a continuous variable and as group categories, with the low-risk PAM50-RORP group achieving a DRFS at 8.5 years of 84.02\% \((P < 0.01; \text{Figure 3D})\).

Within the luminal B subtype \((n = 120)\), the vast majority of signatures were found to be prognostic when evaluated as a continuous variable in node-negative disease (supplemental Table S6, available at Annals of Oncology online); however, no statistically significant group of patients with >90\% DRFS at 8.5 years was identified by any of the predictors (supplemental Table S6, available at Annals of Oncology online); similar findings were obtained when we included the qRT-PCR PAM50 Nielsen series dataset. Finally, no significant prognostic ability was found within node-positive luminal B tumors, with \((n = 285)\) or without \((n = 70)\) the Nielsen series, respectively (supplemental Table S6, available at Annals of Oncology online).

**discussion**

Our data indicates that (i) clinically used signatures and ER-regulated gene signatures are tracking tumors with similar underlying biology (luminal A versus not) and show significant agreement in outcome predictions; (ii) the performance of these signatures is most relevant in node-negative disease; and (iii) some single genomic signatures can perform nearly as well as a combination of two or more signatures, although a combination of multiple signatures was statistically the best. Importantly, this is the first report to show that groups of patients with >95\% DRFS at 8.5 years might only be consistently identified within node-negative and luminal A disease. Alternatively, for patients with luminal B cancer treated only with tamoxifen, additional therapies should be offered, which, as of today, would suggest chemotherapy.

These results also demonstrate that most of the signatures evaluated in this study can provide similar outcome predictions, although significant differences across predictors...
are present. This result is harmonious with our previous observation of concordance between intrinsic subtypes, NKI70 and GHI in a cohort of heterogeneously treated ER+ and ER− breast cancer patients [12]. Importantly, here, we show that these and other signatures are tracking ER+ tumors with a similar biology. Indeed, the vast majority of ER+ tumors identified here as having either basal-like, HER2-enriched or luminal B subtypes were correctly classified by the other signatures as having a poor prognosis. On the other hand, luminal A tumors were mostly identified as having good outcome and showing high expression of ER-regulated signatures. Interestingly, a recent neoadjuvant aromatase inhibitor clinical trial reported that the luminal A subtype benefits the most from endocrine therapy [26].

The performance of each predictor in node-positive disease was significantly worse when compared with node-negative disease, and almost no group of patients with node positivity had a DRFS >90% at 8.5 years by any predictor; the only exceptions being GHI within luminal A disease. In two previously published node-positive ER+ cohorts receiving adjuvant endocrine treatment only (TransATAC and SWOG-8814), the 9-year DRFS and 10-year disease-free survival estimates were 83% and 60% for the low-risk groups of the GHI, respectively [27, 28]. A plausible biological explanation is that in advanced luminal A primaries, a small percentage of cells within the bulk of the tumor have already metastasized and/or acquired endocrine resistance. Indeed, the presence of these subclones is supported by data from a neoadjuvant endocrine trial [29]. However, within node-positive luminal A tumors, some patients with the low and intermediate risk score of GHI had a DRFS at 8.5 years >90%. Hence, future studies are warranted to determine if these or other predictors can identify, within the luminal A subtype, a group of node-positive patients whose survival with endocrine therapy could preclude the administration of adjuvant chemotherapy. The MINDACT [6] trial, which has completed accrual, and the RxPONDER trial (NCT01272037) will address this issue, particularly for patients with one to three positive lymph nodes.

Multivariate analyses including two predictors at a time revealed that, in most cases, many of these correlated predictors, in particular the PAM50-RORP, GHI, NKI70 and SET, remained statistically independent of each other (Table 3). This finding suggests that these predictors are not the same. In fact, at the individual level, the risk group assignment concordance among these predictors was found to be 36% for PAM50-RORP versus GHI, 54% for PAM50-RORP (low/med versus high) versus NKI70 and 74% for GHI (low/intermediate versus high) versus NKI70. Cohen’s kappa coefficients between risk group assignments were also indicative of slight to fair agreement (range 0.11–0.42) [30, 31]. The clinical relevance of this finding is currently unknown. However, a plausible explanation is that these signatures might be tracking different poor outcome luminal/ER+ subtypes; support for this

Figure 3. Kaplan–Meier DRFS analysis of selected gene signatures within luminal A disease treated with adjuvant tamoxifen only. (A) SET index within node-negative luminal A tumors; (B) PAM50-RORP within node-negative luminal A tumors (Nielsen series included); (C) GHI within node-positive luminal A tumors; (D) PAM50-RORP within node-positive luminal A tumors (Nielsen series included). The complete survival analyses can be found in supplemental Tables S5 and S6, available at Annals of Oncology online. DRFS, distant relapse-free survival.
heterogeneity comes from Parker et al. [10], where five statistically significant groups of luminal tumors were identified. Nonetheless, when we built a model here using all predictors, we only observed modest improvements in performance. This finding suggests that gene expression profiling may be reaching its maximum prognostic power.

There are several important caveats to our analyses that must be recognized and always kept in mind when interpreting ‘across platform’ genomic studies. First, although we strove to implement each predictor as published, signatures developed on platforms other than the Affymetrix U133A were suboptimally implemented. This is because when taking a predictor from one technology and applying it to another platform, different oligonucleotide probes/sequences are used to represent each gene (and thus may not behave identically), and each technology has unique normalization methods. Second, changing platforms/technologies almost always causes a loss of genes (see supplemental Table S1, available at Annals of Oncology online), and this loss was significantly present for PAM50 (6/50) and NKI70 (12/60), which likely explains the observed lower performance of this predictor with respect to others. Nonetheless, many of the across platform evaluated predictors carried out well including the PAM50-ROR and GHI; the survival outcomes of the GHI low-risk group within node-negative disease was highly concordant to previous publications [32] despite that the absolute survival rates are highly dataset dependent. Finally, we could not compare the prognostic ability of these signatures versus standard clinicopathological variables since these variables were not available from most microarray datasets. This highlights the need for centralized collection of clinical and pathology data in all genomic studies.

To conclude independently derived genomic predictors of breast cancer recurrence perform similarly and are tracking tumors with similar biology. However, most predictors were statistically independent from the others and thus, these should not be considered to be interchangeable assays. From a clinical perspective, adding genomic signatures together provided modest improvements in outcome prediction, but may not be practical given cost.

acknowledgements

This work was presented as an oral presentation (Abstract #502) at the American Society for Clinical Oncology Annual Meeting in Chicago, 2011.

funding

NCI Breast SPORE program (P50-CA58223-09A1); (ROI-420 CA138255) Breast Cancer Research Foundation, the Sociedad Española de Oncología Médica (SEOM) and the V Foundation for Cancer Research. AP is affiliated to the Medicine PhD program of the Autonomous University of Barcelona, Spain.

disclosure

CMP, PSB, TON and MJE are equity stock holders of University Genomics and BioClassifier LLC. CMP, PSB, MCUC, TON, MJE and JSP have filed a patent on the PAM50 assay. AP, CF, LDM, JB, SKLC and LAC have declared no conflicts of interest.

References

Explanatory factors of sexual function in sexual minority women breast cancer survivors

U. Boehmer¹, A. Timm², A. Ozonoff² & J. Potter³

¹Departments of Community Health Sciences, Boston University School of Public Health, Boston; ²Departments of Biostatistics, Boston University School of Public Health, Boston; ³Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, USA

Received 7 October 2011; revised 21 January 2012; accepted 16 February 2012

Background: The sexual function of sexual minority women (women with female partners) who are breast cancer survivors is mostly unknown. Our objective is to identify explanatory factors of sexual function among sexual minority women with breast cancer and compare them with a control sample of sexual minority women without cancer.

Patients and methods: Using a conceptual framework that has previously been applied to heterosexual breast cancer survivors, we assessed the relationship of each explanatory factor to sexual function in sexual minority women. Using generalized estimating equations, we identified explanatory factors of sexual function and identified differences by case and control status.

Results: Self-perception of greater sexual attractiveness and worse urogenital menopausal symptoms explain 44% of sexual function, after controlling for case and control status. Focusing only on partnered women, 45% of sexual function was explained by greater sexual attractiveness, postmenopausal status, and greater dyadic cohesion.

Conclusions: All of the relevant explanatory factors for sexual function among sexual minority survivors are modifiable as has been suggested for heterosexual survivors. Sexual minority survivors differ from heterosexual survivors in that health-related quality of life is less important as an explanatory factor. These findings can guide adaptation of interventions for sexual minority survivors.

Key words: breast neoplasm, case–control study, female, homosexuality, sexual dysfunctions

Introduction

Sexual dysfunction or difficulties remain a persistent concern of breast cancer survivors (BCS) [1–3]. Sexual dysfunction is common and distressing, affecting ~50% of BCS [3–5]. Depending on the dimension of sexual function (desire, arousal, orgasm, frequency of sexual activity) measured, the incidence of sexual dysfunction varies from 15% to a high of 64% [4, 5]. Broeckel et al. [6] demonstrated worse sexual functioning in long-term BCS compared with controls, including greater lack of sexual interest, inability to relax and enjoy sex, difficulty becoming aroused, and difficulty achieving orgasm. Study findings are inconsistent when sexual frequency is used as the measure of sexual functioning: Ganz et al. [4, 7] found no
PAM50 assay and the three-gene model for identifying the major and clinically relevant molecular subtypes of breast cancer

A. Prat · J. S. Parker · C. Fan · C. M. Perou

Received: 11 June 2012 / Accepted: 15 June 2012 / Published online: 3 July 2012 © The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract It has recently been proposed that a three-gene model (SCMGENE) that measures ESR1, ERBB2, and AURKA identifies the major breast cancer intrinsic subtypes and provides robust discrimination for clinical use in a manner very similar to a 50-gene subtype predictor (PAM50). However, the clinical relevance of both predictors was not fully explored, which is needed given that a ~30% discordance rate between these two predictors was observed. Using the same datasets and subtype calls provided by Haibe-Kains and colleagues, we compared the SCMGENE assignments and the research-based PAM50 assignments in terms of their ability to (1) predict patient outcome, (2) predict pathological complete response (pCR) after anthracycline/taxane-based chemotherapy, and (3) capture the main biological diversity displayed by all genes from a microarray. In terms of survival predictions, both assays provided independent prognostic information from each other and beyond the data provided by standard clinical-pathological variables; however, the amount of prognostic information was found to be significantly greater with the PAM50 assay than the SCMGENE assay. In terms of chemotherapy response, the PAM50 assay was the only assay to provide independent predictive information of pCR in multivariate models. Finally, compared to the SCMGENE predictor, the PAM50 assay explained a significantly greater amount of gene expression diversity as captured by the two main principal components of the breast cancer microarray data. Our results show that classification of the major and clinically relevant molecular subtypes of breast cancer are best captured using larger gene panels.

Keywords Breast cancer · Microarrays · PAM50 · Prognosis · Gene expression

Introduction

Over the years, global gene expression analyses have identified at least four intrinsic subtypes of breast cancer (Luminal A, Luminal B, HER2-enriched, and Basal-like) and a normal-like group with significant differences in terms of their risk factors, incidence, baseline prognoses and responses to systemic therapies [1–4]. In 2009, we reported a clinically applicable gene expression-based predictor that robustly identifies these main intrinsic subtypes by quantitative measurement of 50 genes (i.e., PAM50) [1]. Identification of these molecular subtypes using pathology-based surrogate definitions based upon hormone receptors (HRs), HER2 and Ki-67 expressions has been adopted by the 2011 St. Gallen Consensus Conference for treatment decision-making in early breast cancer [5], however, controversy exists as to whether these complex...
molecular subtypes can be effectively captured using four or less biomarkers.

Recently, Haibe-Kains et al. [6] reported a mRNA expression predictor that classifies tumors into four molecular entities (ER+/HER2−/Low Proliferative, ER+/HER2−/High Proliferative, HER2+/High Proliferative, ER−/HER2−) by quantitative measurement of three genes (ESR1, ERBB2 and AURKA). Similar to the PAM50 subtype predictions, the molecular entities identified by the SCMGENE predictor were found significantly associated with survival outcome [6]. However, a direct head-to-head comparison between both predictors was not performed despite that fact that the concordance (i.e., κ score) between these two predictors was 0.59 (0.58–0.61), which is considered moderate agreement and similar to the κ scores obtained when histological grade is evaluated by two independent observers [7].

In this study, we compared the SCMGENE assignments and the research-based PAM50 assignments in terms of their ability to (1) predict patient outcome, (2) predict pathological complete response (pCR) after anthracycline/taxane-based chemotherapy, and (3) capture the main biological diversity displayed by all genes from a microarray.

Materials and methods

Clinical and gene expression data

We used the clinical (Supplemental file: jnci-JNCI-11-0924-s02.csv) and gene expression data (http://www.compbio.dfci.harvard.edu/pubs/sbtpaper/data.zip) as provided by Haibe-Kains et al. [6]. For survival predictions, we used distant metastasis-free survival as the endpoint since it provides the largest number of patients that can be evaluated across 13 datasets (CAL [8], EMC2 [9], DFHCC [10], MAINZ [11], MDA5 [12], MSK [13], NKI [14], TAM [15], TRANSBIG [16], UCSF [17], UNT [18], VDX [19] and VDX3 [20]). None of the datasets (or samples) used for survival (or response prediction) were used to derive the SCMGENE or the PAM50 subtype predictor.

To compare chemotherapy response data, we used the clinical data of one of the datasets (MAQC2 [GSE20194] [21]) evaluated by Haibe-Kains et al. [6], which is composed of 230 pre-treatment samples with annotated response data (pCR vs. residual disease [RD]) after neo-adjuvant anthracycline/taxane-based chemotherapy. Samples that received trastuzumab were excluded.

Combined microarray dataset

Eighteen Affymetrix and Agilent-based datasets (CAL [8], DFHCC [10], DUKE [22], EORTC10994 [23], EXPLO [24], KOO [25], MAINZ [11], MAQC2 [21], MDA4 [26], MSK [13], NKI [14], PNC [27], STK [28], TRANSBIG [16], UNC337 [29], UNT [18], UPP [30] and VDX [19]) as provided in Haibe-Kains et al. [6] and with an appropriate distribution of ER+ (50–90 %, as defined by IHC) versus ER− tumors were combined into a single gene expression matrix. Probes mapping to the same gene (Entrez ID as defined by the manufacturer) were averaged to generate independent expression estimates. In each cohort, genes were median centered and standardized to zero mean and unit variance.

Statistical analyses

Distant metastasis-free survival univariate and multivariate analysis were calculated using a Cox proportional regression model. Likelihood ratio statistics of subtypes defined by the PAM50 or the SCMGENE predictors were also evaluated after accounting for clinical–pathological variables (age at diagnosis, nodal status, and tumor size) and type of systemic adjuvant treatment (chemotherapy, endocrine, and none). Models were first conditioned on one predictor and the clinical–pathological variables, and then the significance of the other was tested. Chemotherapy response (pCR vs. RD) predictions of each variable were evaluated using univariate and multivariate logistic regression analyses. Finally, $R^2$ values of each predictor (SCMGENE or PAM50) for each principal component (PC) were calculated using a simple linear regression model. All statistical computations were performed in R v.2.8.1 (http://www.cran.r-project.org).

Results

Outcome prediction

To compare the ability of the SCMGENE and PAM50 assays to predict patient outcome, we performed Cox proportional hazard regression analyses using the entire combined dataset as provided by Haibe-Kains et al. [6]. In the multivariate model (MVA), both predictors were found significantly associated with distant metastasis-free survival (Table 1) and the Luminals A and B segregation of the PAM50 assay was found significantly associated with outcome, whereas the ER+/HER2−/Low Proliferative and ER+/HER2−/High Proliferative segregation of the SCMGENE predictor was not. Conversely, distant metastasis-free survival differences of the ER−/HER2− versus the ER+/HER2−/Low Proliferative groups were found significant, whereas the Basal-like versus Luminal A segregation was not.

To compare the amount of independent prognostic information provided by each predictor, we estimated the likelihood ratio statistic of each predictor in a model that already included clinical–pathological variables (age,
tumor size, treatment and nodal status) and the other predictor. The results revealed that the PAM50 subtypes provide a larger amount of independent prognostic information than the SCMGENE subtypes when using the entire cohort of heterogeneously treated patients (Fig. 1A, B). Similar results were observed when using the subset of patients that did not receive adjuvant systemic therapy (Fig. 1C, D), and in the subset of patients with HR? tumors that received adjuvant tamoxifen-only (Fig. 1E, F).

Chemotherapy response prediction

To compare the ability of the PAM50 and SCMGENE assays to predict response to chemotherapy, we evaluated the MAQC2 (GSE20194) [21] dataset included in Haibe-Kains et al. [6] analyses. This cohort is composed of 226 pretreatment samples with annotated response data (pCR vs. RD) after neoadjuvant anthracycline/taxane-based chemotherapy (without trastuzumab for HER2? disease). As shown in Table 2, although both assays predicted response in univariate analysis, the PAM50 assay was the only one to provide independent predictive information in the MVA model.

Of note, the association of the PAM50 subtype with response was strengthened when PAM50 subtyping of the MAQC2 dataset was performed after median centering the PAM50 genes/rows (Supplemental Table 1). In fact, we and others have previously proposed median gene centering to minimize technical bias and allow the correct identification of ER-, ER+, and HER2+ samples [31, 32]. Median gene centering of the UNC337 dataset before PAM50 or SCMGENE predictions also improved the survival classifications (Supplemental Fig. 1).

Capturing the main biological diversity

Finally, to compare both predictors in terms of their ability to capture the main biological diversity displayed by all genes in a breast cancer microarray, we first combined 18 datasets evaluated by Haibe-Kains et al. [6] and identified the two

### Table 1 Distant metastasis-free survival Cox proportional hazards models of primary breast cancer patients

<table>
<thead>
<tr>
<th>Variables</th>
<th>Univariate analysis</th>
<th>Multivariate analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HR</td>
<td>Lower 95 %</td>
</tr>
<tr>
<td>Age (cont. variable)</td>
<td>0.989</td>
<td>0.983</td>
</tr>
<tr>
<td>Node status</td>
<td>1.176</td>
<td>0.851</td>
</tr>
<tr>
<td>Tumor size T2–T4 versus T0–T1</td>
<td>1.305</td>
<td>1.104</td>
</tr>
<tr>
<td>Treatment (yes vs. no)</td>
<td>0.973</td>
<td>0.845</td>
</tr>
<tr>
<td>PAM50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luminal A</td>
<td>1.0</td>
<td>–</td>
</tr>
<tr>
<td>Luminal B</td>
<td>1.797</td>
<td>1.503</td>
</tr>
<tr>
<td>HER2-E</td>
<td>2.677</td>
<td>2.120</td>
</tr>
<tr>
<td>Basal-like</td>
<td>2.144</td>
<td>1.737</td>
</tr>
<tr>
<td>Normal-like</td>
<td>1.073</td>
<td>0.670</td>
</tr>
<tr>
<td>Three-gene signature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ER+/HER2−/Low Prolif</td>
<td>1.0</td>
<td>–</td>
</tr>
<tr>
<td>ER+/HER2−/High Prolif</td>
<td>1.852</td>
<td>1.531</td>
</tr>
<tr>
<td>HER2+</td>
<td>2.785</td>
<td>2.196</td>
</tr>
<tr>
<td>ER−/HER2−</td>
<td>2.536</td>
<td>2.041</td>
</tr>
</tbody>
</table>

HER2-E HER2-enriched, Prolif proliferation, HR hazard ratio
Compared to the SCMGENE subtypes, the PAM50 subtypes explained substantially more variation in gene expression for both PC1 and PC2 (Fig. 2a, b), with these components being especially prominent for the separation of the Luminal A (or ER+/HER2- /Low Proliferative) and Luminal B (or ER+/HER2- /High Proliferative) subtypes. To confirm these findings, we also evaluated all PCs in each normalized dataset provided by Haibe-Kains et al. [6] and observed that among 483 PCs significantly explained by either one of the predictors, the PAM50 explained 2.27 times more independent variation in expression than the SCMGENE assay.

### Discussion

Our results presented here, using the same data provided by Haibe-Kains et al. [6], suggest that (1) the SCMGENE and the PAM50 predictors should not be considered the same in terms of outcome prediction; (2) both provide independent...
prognostic information; (3) the amount of prognostic information provided by the PAM50 predictor is greater than the information provided by the SCMGENE predictor; and (4) the PAM50 assay is the only independent predictor of neoadjuvant chemotherapy response.

A potential explanation of our findings is that the biological diversity of breast cancer is better captured using the quantitative measurement of the 50 PAM50 gene set compared to the 3 genes of the SCMGENE assay. This finding is further supported by our previous data during the PAM50 assay development, where the minimum number of genes required to identify the intrinsic molecular subtypes, as defined by subtype classifications based upon the ~1,900 intrinsic gene list with a 93 % accuracy, was the final selected 50 genes [1]. In fact, gene sets with less than 50 genes showed significantly worse accuracies, particularly for tumors of the Luminal B and HER2-enriched subtypes (Supplemental Fig. 2). Importantly, only 33.3 % (12/36) of all microarray datasets evaluated in Haibe-Kains et al. [6] had all the PAM50 genes available, whereas 100 % of the datasets had all three genes of the SCMGENE assay, thus highlighting another caveat of this study.

In total, these analyses show that a combination of ER, HER2, and a single proliferation biomarker (i.e., AURKA) is prognostic, but is suboptimal to capture the biological diversity of breast cancers, which has similar implications for the capture of this biological diversity using IHC-based methods. Although a head-to-head comparison of both assays in terms of their clinical utility might be warranted in the future, our results suggest that classification of the major and clinically relevant molecular subtypes is better achieved using larger gene sets that capture a greater proportion of the biological diversity of breast cancers.

Acknowledgments This study was supported by funds from the NCI Breast SPORE Program (P50-CA58223-09A1), by RO1-CA138255, by the Breast Cancer Research Foundation, and the Sociedad Española de Oncología Médica (SEOM). A. Prat is affiliated to the Medicine PhD Program of the Autonomous University of Barcelona (UAB), Spain.

Conflict of interest C. M. P. is a stock holder of BioClassifier LLC. C. M. P. and J. S. P. have filed a patent on the PAM50 assay. A. P. and C. F. have declared no conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References
Prognostic Significance of Progesterone Receptor–Positive Tumor Cells Within Immunohistochemically Defined Luminal A Breast Cancer

Aleix Prat, Maggie Chon U. Cheang, Miguel Martin, Joel S. Parker, Eva Carrasco, Rosalía Caballero, Scott Tyldesley, Karen Gelmon, Philip S. Bernard, Torsten O. Nielsen, and Charles M. Perou

ABSTRACT

Purpose
Current immunohistochemical (IHC)-based definitions of luminal A and B breast cancers are imperfect when compared with multigene expression-based assays. In this study, we sought to improve the IHC subtyping by examining the pathologic and gene expression characteristics of genomically defined luminal A and B subtypes.

Patients and Methods
Gene expression and pathologic features were collected from primary tumors across five independent cohorts: British Columbia Cancer Agency (BCCA) tamoxifen-treated only, Grupo Español de Investigación en Cáncer de Mama 8906 trial, BCCA no systemic treatment cohort, PAM50 microarray training data set, and a combined publicly available microarray data set. Optimal cutoffs of percentage of progesterone receptor (PR)–positive tumor cells to predict survival were derived and independently tested. Multivariable Cox models were used to test the prognostic significance.

Results
Clinicopathologic comparisons among luminal A and B subtypes consistently identified higher rates of PR positivity, human epidermal growth factor receptor 2 (HER2) negativity, and histologic grade 1 in luminal A tumors. Quantitative PR gene and protein expression were also found to be significantly higher in luminal A tumors. An empiric cutoff of more than 20% of PR-positive tumor cells was statistically chosen and proved significant for predicting survival differences within IHC-defined luminal A tumors independently of endocrine therapy administration. Finally, no additional prognostic value within hormonal receptor (HR)–positive/HER2-negative disease was observed with the use of the IHC4 score when intrinsic IHC-based subtypes were used that included the more than 20% PR–positive tumor cells and vice versa.

Conclusion
Semiquantitative IHC expression of PR adds prognostic value within the current IHC-based luminal A definition by improving the identification of good outcome breast cancers. The new proposed IHC-based definition of luminal A tumors is HR positive/HER2 negative/Ki-67 less than 14%, and PR more than 20%.

J Clin Oncol 31:203-209. © 2012 by American Society of Clinical Oncology

INTRODUCTION

Hormonal receptor (HR)–positive breast cancer is a clinically and biologically heterogeneous entity.1-3 Studies based on gene expression profiling have identified at least two major groups of HR–positive tumors, known as the luminal A and B intrinsic subtypes of breast cancer. These two molecular entities have shown significant differences in baseline prognosis and sensitivity to cytotoxic therapies.4-6 Currently, a gene expression–based assay known as the PAM50 subtype predictor identifies the intrinsic molecular subtypes of breast cancer and provides a risk of relapse (ROR) score in a fashion similar to the Oncotype DX (Genomic Health, Redwood City, CA) recurrence score (RS).4,6 These two assays provide valuable and independent prognostic information beyond standard clinicopathologic variables. However, standardized gene expression–based tests are not readily available in most of the world as a result of cost, assay turnaround times, and other logistic issues. Thus surrogate definitions of the intrinsic subtypes and/or risk of relapse groups developed using routine pathology and clinical parameters could be of great practical value.7,8
We have previously reported an immunohistochemical (IHC)-based surrogate definition of the luminal A (IHC-luminal A) and luminal B/human epidermal growth factor receptor 2 (HER2) -negative (IHC-luminal A/HER2-negative) subtypes based on the quantitative expression of the proliferation-related marker Ki-67 within HR-positive/HER2-negative disease. This definition has now been adopted by the 2011 St Gallen Expert Consensus Panel Recommendation Guidelines for the systemic treatment of early breast cancer, which recommend adjuvant endocrine therapy alone for patients with IHC-luminal A tumors and the addition of chemotherapy for patients with IHC-luminal A/HER2-negative tumors. Here we further refine the IHC-based definition of luminal A and B through the use of quantitative progesterone receptor (PR) expression.

PATIENTS AND METHODS

Patients, Samples, and Clinical Data

Multiple different and independent data sets were used to assess the significance of PR IHC results. Gene expression and/or clinicopathologic features were evaluated across five different data sets: (1) a combined genomic data set of nine publicly available microarray cohorts (GSE18229, GSE18864, GSE22219, GSE25066, GSE2990, GSE4922, GSE7390, GSE7849, and NK1295), (2) the PAM50 microarray-based subtype predictor training data set (PAM50-training, GSE10886), (3) a British Columbia Cancer Agency (BCCA) tamoxifen-treated cohort (BCCA-tamoxifen), (4) the Grupo Español de Investigación en Cáncer de Mama (GEICAM) 9906 trial, and (5) the BCCA no adjuvant systemic therapy (AST) cohort (BCCA-no AST). A detailed CONSORT diagram can be found in Appendix Table A1 (online only).

All patients from the BCCA-tamoxifen cohort had early-stage HR-positive disease and received adjuvant treatment with tamoxifen only. In the GEICAM 9906 phase III trial cohort, patients with node-positive disease were randomly assigned to adjuvant fluorouracil, epirubicin, and cyclophosphamide versus fluorouracil, epirubicin, and cyclophosphamide followed by weekly paclitaxel, and patients with HR-positive disease subsequently received adjuvant endocrine therapy. The BCCA-no AST cohort includes “clinically low risk” patients with primary breast cancer diagnosed between 1986 and 1992 who did not receive adjuvant systemic therapy. Characteristics of both BCCA cohorts and the GEICAM 9906 cohort have been previously described. From the PAM50-training cohort, we performed global and proliferation (ROR-P) as previously described for the BCCA-tamoxifen cohort (training data set), and the optimal cutoff to predict distant relapse–free survival (DRFS) was independently tested in the GEICAM 9906 and BCCA-no AST cohorts.

RESULTS

Gene and Protein Expression Differences Between Luminal A and B Tumors

To identify global and single gene expression differences, we performed a two-class significance analysis of microarrays between prototypical luminal A and B tumors from the PAM50-training cohort. A total of 1,539 genes (348 upregulated and 1,191 downregulated) were found differentially expressed (false discovery rate < 1%) between both subtypes (Appendix Fig A1, online only; Data Supplement). The upregulated gene list in luminal A tumors was found enriched for genes involved in cell differentiation (eg, Kruppel-like factor 4 and Jun proto-oncogene) and cell adhesion (eg, vinculin and collagen, type XV1, α1) biologic processes. Conversely, the downregulated gene list in luminal A tumors (ie, genes highly expressed in luminal B tumors) was found enriched for genes involved in immune response (eg, interleukin 2 receptor α and CD86) and cell-cycle (eg, cyclin B1 and RAD51) biologic processes, which is indicative of the faster proliferation rates known to be part of luminal B tumors.

Among the relatively upregulated genes in luminal A tumors was the progesterone receptor gene (PGR), but not the estrogen receptor gene (ESR1). To further explore these findings, we evaluated the mRNA expression of PGR and ESR1 in two independent studies in which PAM50 was performed using the quantitative reverse-transcriptase polymerase chain reaction platform (GEICAM 9906 and BCCA-tamoxifen) and confirmed that PGR, but not ESR1, was found significantly upregulated in luminal A tumors compared with luminal B tumors (Figs 1A and 1B; P < .001, t test). Interestingly, PGR was found only weakly correlated (Pearson correlation coefficient =
−0.19) with the expression of the Ki-67 gene MKI67, indicating that these two genes may provide different biologic information.

The mRNA expression-based data suggested that semiquantitative scoring of the PR protein, but not ER protein, might help discriminate the genomically defined luminal A from B tumors. To further explore this hypothesis, we compared the percentage of PR-positive and ER-positive tumor cells as assessed by IHC, among luminal A and B tumors in the GEICAM 9906 cohort,11 and observed that only the percentage of PR-positive cells can discriminate luminal A from B tumors (Figs 1C and 1D). However, it is important to note that considerable overlap was observed. Finally, PR protein expression was also weakly anticorrelated with Ki-67 protein expression ($r = −0.20$).

Clinicopathologic Features of Luminal A and B Tumors

To identify clinicopathologic differences among the genomically defined luminal A and B tumors, we evaluated the clinicopathologic features of 2,257 patients with luminal A or B primary breast cancer. Across three independent cohorts (Table 1), luminal A tumors showed significantly higher rates of PR positivity, HER2 negativity, histologic grade 1, and tumor stage T0-T1 compared with luminal B tumors. No significant differences in ER status were observed, with the vast majority (92% to 96%) of luminal A and B tumors being ER positive.

IHC-Based Versus PAM50 Subtype Definitions

Current IHC-based definitions of luminal A and B subtypes are imperfect when compared with multigene expression-based assays.5 To further illustrate this, we evaluated the distribution of the IHC-based definitions within luminal A and B tumors in the BCCA-tamoxifen6 and the GEICAM 9906 cohorts.11 As expected, whereas a large majority (81% to 85%) of luminal A tumors were identified as IHC-luminal A, 35% to 52% of luminal B tumors were also identified as IHC-luminal A (Table 2).

![Figure 1](http://www.jco.org)
Finally, we explored the survival of the luminal A and B subtypes within the IHC-based luminal A and IHC-luminal B/HER2-negative tumors in the BCCA-tamoxifen cohort (Appendix Table A3, online only). In both cases, luminal A tumors showed a significantly better DRFS outcome than non–luminal A tumors. In multivariable Cox model survival analyses adjusted for histologic grade, age at diagnosis, nodal positivity, and tumor size, the hazard ratio for DRFS in PAM50 luminal A tumors compared with PAM50 non–luminal A was 0.642 within IHC-luminal A tumors (95% CI, 0.422 to 0.975, P = .038) and 0.582 within IHC-luminal B/HER2-negative tumors (95% CI, 0.323 to 1.047, P = .071).

Survival Outcomes Based on the Percentage of PR-Positive Cells

These data suggested that (1) further improvements in the IHC-luminal A definition is needed because many PAM50-defined luminal B tumors are erroneously identified as IHC-luminal A and (2) quantitative scoring of PR-positive tumor cells, but not ER-positive tumor cells, might help identify good-outcome breast cancers. To test this hypothesis, we evaluated the association of the visually determined percentage of PR-positive and ER-positive invasive breast carcinoma cells with survival outcomes within IHC-luminal A tumors of the BCCA-tamoxifen cohort. As expected, the percentage of PR-positive cancer cells, but not the percentage of ER-positive cancer cells (data not shown), was associated with DRFS after adjusting for standard clinicopathologic variables, with the optimal PR percentage cutoff to predict outcome being found to be 20% (Appendix Fig A2-A3, online only). In contrast, within IHC-luminal B/HER2-negative tumors (ie, HR-positive/Ki-67 ≤ 14%), semiquantitative expression of either PR or ER was not found to be associated with outcome differences (data not shown).

We then tested the prognostic value of the PR cutoff of more than 20% within IHC-luminal A tumors in two independent cohorts of patients with primary breast cancer (GEICAM 9906 and the BCCA-no AST cohorts). In both data sets, patients with IHC-luminal A tumors having low positive PR-positive tumor cells (< 20%) showed significantly poorer survival compared with tumors with more than 20% of PR-positive tumor cells (Figs 2A and 2B). Multivariable analyses confirmed the independent association between PR expression and survival (Appendix Table A4-A5, online only).

Table 1. Clinicopathologic Characteristics of Luminal A and B Tumors

<table>
<thead>
<tr>
<th>Variable</th>
<th>BCCA-Tamoxifen ER-Positive Only</th>
<th>GEICAM 9906 Node Positive</th>
<th>Combined Microarray Dataset All</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Luminal A</td>
<td>Luminal B</td>
<td>No.</td>
</tr>
<tr>
<td>No.</td>
<td>372</td>
<td>329</td>
<td>—</td>
</tr>
<tr>
<td>Mean age, years</td>
<td>66.6</td>
<td>67.4</td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>25</td>
<td>7</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>186</td>
<td>54</td>
<td>129</td>
</tr>
<tr>
<td>3</td>
<td>135</td>
<td>29</td>
<td>179</td>
</tr>
<tr>
<td>Nodal positivity</td>
<td></td>
<td></td>
<td>&gt; .05</td>
</tr>
<tr>
<td>Tumor size &gt; 2.0 cm</td>
<td>245</td>
<td>72.1</td>
<td>215</td>
</tr>
<tr>
<td>IHC ER-positive status</td>
<td></td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>IHC PR-positive status</td>
<td>248</td>
<td>72</td>
<td>174</td>
</tr>
<tr>
<td>Clinical HER2-positive status</td>
<td>15</td>
<td>4</td>
<td>30</td>
</tr>
</tbody>
</table>

Abbreviations: BCCA, British Columbia Cancer Agency; ER, estrogen receptor; GEICAM, Grupo Español de Investigación en Cáncer de Mama; HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry; PR, progesterone receptor.

Table 2. Distribution of the IHC-Based Subtypes Across the Luminal A and B Intrinsic Subtypes

<table>
<thead>
<tr>
<th>IHC-Based Subtypes</th>
<th>BCCA-tamoxifen</th>
<th>GEICAM 9906</th>
<th>Combined Microarray Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Luminal A</td>
<td>286</td>
<td>109</td>
<td>231</td>
</tr>
<tr>
<td>Luminal B</td>
<td>109</td>
<td>35.4</td>
<td>65.4</td>
</tr>
<tr>
<td>Luminal A</td>
<td>231</td>
<td>35.4</td>
<td>65.4</td>
</tr>
<tr>
<td>Luminal B</td>
<td>134</td>
<td>45.1</td>
<td>45.1</td>
</tr>
</tbody>
</table>

NOTE. Within hormone receptor–positive/HER2-negative disease, the concordance x value score between the PAM50 luminal A and B definition with the IHC-luminal A and B/HER2-negative definition was 0.196 and 0.407 (slight to fair agreement) in the GEICAM 9906 cohorts and BCCA-tamoxifen, respectively.

Abbreviations: BCCA, British Columbia Cancer Agency; GEICAM, Grupo Español de Investigación en Cáncer de Mama; HER2, human epidermal growth factor receptor 2; IHC, immunohistochemistry.
only). In the BCCA-no AST cohort, the breast cancer-specific survival at 15 years of patients with IHC-luminal A tumors with more than 20% PR-positive tumor cells was 94.0% (95% CI, 91.6% to 98.2%).

We next evaluated the distribution of the gene-expression based intrinsic subtypes (gold standard) within IHC-luminal A tumors in the GEICAM 9906 cohort based on this more than 20% PR cutoff (Table 3). Consistent with the preceding findings, 63% of IHC-luminal A tumors with more than 20% of PR-positive cells were identified as luminal A, whereas 24% of IHC-luminal A tumors with ≤20% of PR-positive cells were identified as luminal A, thus confirming that this definition helps to better discriminate true luminal A tumors from the rest. Finally, although the PR cutoff of 20% increased the percentage of luminal A tumors identified within what would otherwise have been considered IHC-luminal B/HER2-negative tumors from 5.9% to 30.9%, the majority of this group remained composed of luminal B (55.6%) tumors.

**Comparison of Prognostic Values of IHC-Based Subtypes, IHC4 Score, and PAM50-ROR-P Score**

We compared the contribution of the newly proposed IHC-based subtype definitions (IHC-luminal A [HR positive/HER2 negative/Ki-67 < 14%], IHC-luminal B [HR positive/HER2 negative/Ki-67 < 14%/PR ≤ 20% or HR positive/HER2 negative/Ki-67 > 14%]) with a version of the IHC4 score and with PAM50 ROR-P score in the subset of patients with HR-positive/HER2-negative tumors from the GEICAM 9906 cohort (n = 580). All three classifications added significant prognostic information beyond clinical variables (Figs 3A, 3B, and 3C), with IHC-based subtypes and IHC4 score providing similar amounts of prognostic information and PAM50 ROR-P providing the largest amount.

Finally, we evaluated the independent prognostic information that each classification provided when considered in the presence of one of the others. When the IHC4 score was included in the model, adding intrinsic IHC-based subtype did not provide significant independent information (Fig 3D). However, when the IHC-based subtype was included in the model, the IHC4 score did not provide additional information (Fig 3E). On the other hand, inclusion of PAM50 ROR-P provided significant independent prognostic information beyond the information provided by either the IHC4 score or the IHC-based subtypes (Figs 3D and 3E).

**DISCUSSION**

Patients with early breast cancer with tumors that are ER positive and/or PR positive (ie, luminal) have lower risks of recurrence and...
In Dunnwald et al., women with luminal A and B tumors experienced higher risks of mortality compared with women with ER-positive/PR-positive tumors, independent of the various demographic and clinical tumor characteristics. These data are concordant with our centrally reviewed pathology data presented here, which show that PR positivity, and especially high expression of PR protein, is more frequently observed in tumors with a better baseline prognosis (ie, luminal A) than tumors with a poor baseline prognosis (ie, luminal B). It is important to note that a substantial number of luminal B tumors (~50% to 75%) are still PR positive, although the expression of PR may be less than in luminal A tumors.

The ability of ER and/or PR expression to predict benefit to endocrine and/or cytotoxic therapy has also been evaluated. In terms of endocrine sensitivity, a recent patient-level meta-analysis of randomized trials from the Early Breast Cancer Trialists’ Collaborative Group that evaluated adjuvant tamoxifen versus no adjuvant tamoxifen suggested that recurrence and death rate ratio is independent of PR status (or level) in ER-positive disease. Similar data have been observed in another smaller randomized adjuvant study. In addition, PR expression levels have not shown to predict aromatase inhibitors’ efficacy over tamoxifen in ER-expressing tumors in two large adjuvant clinical trials. This is consistent with a recent neoadjuvant trial in which luminal A and B tumors, as defined by the PAM50 assay, did not show significant differences in terms of response to aromatase inhibitors, although luminal A tumors achieved higher rates of Preoperative Endocrine Prognostic Index score of 0, which is a validated biomarker of outstanding outcome after adjuvant endocrine therapy alone. Overall, these data suggest that luminal A and B tumors benefit similarly from endocrine therapies, but that patients with luminal A tumors have a better baseline prognosis than those with luminal B tumors.

In terms of chemotherapy benefit, the majority of adjuvant and neoadjuvant data suggest that HR status is a strong predictor of general chemosensitivity, with HR-positive tumors showing less benefit to cytotoxic drugs than HR-negative tumors. Moreover, in the neoadjuvant setting, luminal A tumors achieve lower rates of pathologic complete response with anthracycline/taxane-based chemotherapy compared with luminal B tumors. In addition, Oncotype DX has shown that within HR-negative disease, those tumors with high RS (ie, non-luminal A tumors) benefit the most from adjuvant chemotherapy. Interestingly, in a retrospective analysis from three adjuvant clinical trials, low expression of both ER and PR, and potentially low expression of PR within ER-positive patients, was found predictive of adding chemotherapy to endocrine therapy. Overall, these data suggest that luminal A tumors are less chemosensitive than luminal Bs.

Fig 3. Disease-free survival log likelihood ratio (LR) statistics of six different predictive models (A–E) in patients of the Grupo Español de Investigación en Cáncer de Mama 9906 cohort with hormone receptor (HR)–positive/human epidermal growth factor receptor 2 (HER2)–negative breast cancer. The variables evaluated were the following: immunohistochemical (IHC)-based scoring of estrogen receptor, progesterone receptor, HER2, and Ki-67 (IHC4 score; continuous variable), IHC-based subtypes (HR positive/HER2 negative/Ki-67 < 14% > 20% [luminal A], HR positive/HER2 negative/Ki-67 < 14% < 20% and HR positive/HER2 negative/Ki-67 > 14% [luminal B]), and PAM50 risk of recurrence score based on subtype and proliferation (ROR-P; continuous variable). (*) P < .05.
luminal A tumors is HR-positive/HER2-negative/Ki-67 less than 14% and PR more than 20%.

REFERENCES


AUTHORS’ DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Although all authors completed the disclosure declaration, the following author(s) and/or an author’s immediate family member(s) indicated a financial or other interest that is relevant to the subject matter under consideration in this article. Certain relationships marked with a “U” are those for which no compensation was received; those relationships marked with a “C” were compensated. For a detailed description of the disclosure categories, or for more information about ASCO’s conflict of interest policy, please refer to the Author Disclosure Declaration and the Disclosures of Potential Conflicts of Interest section in Information for Contributors.

Employment or Leadership Position: None Consultant or Advisory Role: Torsten O. Nielsen, BioClassifier (C) Stock Ownership: Philip S. Bernard, University Genomics, BioClassifier; Charles M. Perou, University Genomics, BioClassifier Honoraria: None Research Funding: None Expert Testimony: None Other Remuneration: None

AUTHOR CONTRIBUTIONS

Conception and design: Aleix Prat, Charles M. Perou Administrative support: Rosalia Caballero Provision of study materials or patients: Maggie Chon U. Cheang, Miguel Martin, Eva Carrasco, Rosalia Caballero, Philip S. Bernard, Torsten O. Nielsen, Charles M. Perou Collection and assembly of data: Aleix Prat, Maggie Chon U. Cheang, Miguel Martin, Eva Carrasco, Rosalia Caballero, Philip S. Bernard, Charles M. Perou Data analysis and interpretation: Aleix Prat, Maggie Chon U. Cheang, Miguel Martin, Joel S. Parker, Scott Tyldesley, Karen Gelmon, Philip S. Bernard, Torsten O. Nielsen, Charles M. Perou Manuscript writing: All authors Final approval of manuscript: All authors