

Genetic Reclassification of Histologic Grade Delineates New Clinical Subtypes of Breast Cancer

Anna V. Ivshina,¹ Joshy George,¹ Oleg Senko,⁵ Benjamin Mow,^{2,3} Thomas C. Putti,³ Johanna Smeds,⁶ Thomas Lindahl,⁶ Yudi Pawitan,⁷ Per Hall,⁷ Hans Nordgren,⁸ John E.L. Wong,^{2,4} Edison T. Liu,¹ Jonas Bergh,⁶ Vladimir A. Kuznetsov,¹ and Lance D. Miller¹

¹Genome Institute of Singapore; ²Department of Haematology-Oncology, National University Hospital; Departments of ³Pathology and ⁴Medicine, National University of Singapore, Singapore; ⁵Russian Computer Center, Russian Academy of Sciences, Moscow, Russia; ⁶Department of Oncology and Pathology, Radiumhemmet, Karolinska Institute and Hospital; ⁷Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden; and ⁸Department of Pathology, Akademiska sjukhuset, University of Uppsala, Uppsala, Sweden

Abstract

Histologic grading of breast cancer defines morphologic subtypes informative of metastatic potential, although not without considerable interobserver disagreement and clinical heterogeneity particularly among the moderately differentiated grade 2 (G2) tumors. We posited that a gene expression signature capable of discerning tumors of grade 1 (G1) and grade 3 (G3) histology might provide a more objective measure of grade with prognostic benefit for patients with G2 disease. To this end, we studied the expression profiles of 347 primary invasive breast tumors analyzed on Affymetrix microarrays. Using class prediction algorithms, we identified 264 robust grade-associated markers, six of which could accurately classify G1 and G3 tumors, and separate G2 tumors into two highly discriminant classes (termed G2a and G2b genetic grades) with patient survival outcomes highly similar to those with G1 and G3 histology, respectively. Statistical analysis of conventional clinical variables further distinguished G2a and G2b subtypes from each other, but also from histologic G1 and G3 tumors. In multivariate analyses, genetic grade was consistently found to be an independent prognostic indicator of disease recurrence comparable with that of lymph node status and tumor size. When incorporated into the Nottingham prognostic index, genetic grade enhanced detection of patients with less harmful tumors, likely to benefit little from adjuvant therapy. Our findings show that a genetic grade signature can improve prognosis and therapeutic planning for breast cancer patients, and support the view that low- and high-grade disease, as defined genetically, reflect independent pathobiological entities rather than a continuum of cancer progression. (Cancer Res 2006; 66(21): 10292-301)

Introduction

In breast cancer care, treatment decisions are guided by efforts to determine the metastatic potential of tumors. Clinical variables

that reflect metastatic potential (e.g., lymph node status, tumor size, histologic grade) or predict for endocrine responsiveness (e.g., estrogen and progesterone receptors) are routinely used to classify tumors into subtypes predictive of outcome. However, these variables are unable to predict with sufficient accuracy which patients will do well without adjuvant treatment, benefit from adjuvant treatment, or respond poorly to current treatment modalities. Some tumor subtypes, despite phenotypic homogeneity, are associated with substantial clinical heterogeneity confounding their clinical utility. Recent studies using DNA microarrays indicate that such clinical heterogeneity may be resolvable at the molecular level (1–6). Indeed, some have shown that gene expression signatures underlying specific biological properties of cancer cells may provide better stratification of metastatic potential than established prognostic variables (1, 2, 7).

Histologic grading in breast cancer seeks to integrate measurements of cellular differentiation and replicative potential into a composite score that quantifies the aggressive behavior of a tumor. The most studied and widely used method of breast tumor grading is the Elston-Ellis modified Scarff, Bloom, Richardson grading system, also known as the Nottingham Grading System (8, 9). The Nottingham Grading System is based on a microscopic evaluation of morphologic and cytologic features of tumor cells, including degree of tubule formation, nuclear pleomorphism, and mitotic count (9). The sum of these scores stratifies breast tumors into grade 1 (G1; well-differentiated, slow-growing), grade 2 (G2; moderately differentiated), and grade 3 (G3; poorly differentiated, highly proliferative) malignancies.

Multivariate analyses in large patient cohorts have consistently shown that the grade of invasive breast cancer is a powerful indicator of disease recurrence and patient death, independent of lymph node status and tumor size (9–12). Untreated patients with G1 disease have a ~95% 5-year survival rate, whereas those with G2 and G3 malignancies have survival rates at 5 years of ~75% and ~50%, respectively. The value of histologic grade in patient prognosis, however, has been questioned by reports of substantial interobserver variability among pathologists (13–16), leading to much debate over the role that grade should play in therapeutic planning (17, 18). Furthermore, where the prognostic significance of G1 and G3 disease is of more obvious clinical relevance, it is less clear what the prognostic value is of the more heterogeneous, moderately differentiated G2 tumors, which comprise ~50% of all breast cancer cases (12, 18, 19).

In this study, we hypothesized that a gene expression signature capable of discriminating low- and high-grade tumors might

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

Requests for reprints: Lance D. Miller, Genome Institute of Singapore, Genome Building 02-01, 60 Biopolis Street, Singapore 13872, Singapore. Phone: 65-6478-8100; Fax: 65-6478-9060; E-mail: MillerL@gis.a-star.edu.sg and Vladimir A. Kuznetsov, Genome Institute of Singapore, Genome Building 02-01, 60 Biopolis Street, Singapore 13872, Singapore. Phone: 65-6478-8198; Fax: 65-6478-9058; E-mail: kuznetsov@gis.a-star.edu.sg

©2006 American Association for Cancer Research.
doi:10.1158/0008-5472.CAN-05-4414

provide a more objective and clinically valuable measure of tumor grade with prognostic significance for patients with moderately differentiated cancer. Through analysis of breast cancer expression data from multiple independent cohorts, we demonstrate the existence of a genetic grade signature that distinguishes new biological and clinical subtypes of breast cancer with important prognostic implications.

Materials and Methods

Patients and tumor specimens. Clinical characteristics of patient and tumor samples of the Uppsala, Stockholm, and Singapore cohorts are summarized in Supplementary File 1A. All cohorts were of unselected populations, and the original tumor material was demonstrated at the time of surgery, frozen in dry ice or liquid nitrogen, and stored under liquid nitrogen or at -70°C .

Uppsala cohort. The Uppsala cohort originally composed of 315 women representing 65% of all breast cancers resected in Uppsala County, Sweden, from January 1, 1987, to December 31, 1989. Information pertaining to breast cancer therapy, clinical follow-up, and sample processing are described elsewhere (20). For histologic grading, new tumor sections were prepared from the original paraffin blocks and stained with eosin (with the exception of a few original van Gieson-stained sections). All sections were graded in a blinded fashion (H.N.) according to the Nottingham Grading System (9) as follows:

Tubule Formation: 3 = poor, if <10% of the tumor showed definite tubule formation, 2 = moderate, if $\geq 10\%$ but $\leq 75\%$, and 1 = well, if $>75\%$.

Mitotic Index: 1 = low, if <10 mitoses, 2 = medium, if 10 to 18 mitoses, and 3 = high, if >18 mitoses (per 10 high-power fields). Field diameter was 0.57 mm.

Nuclear Grade: 1 = low, for little variation in size and shape of nuclei, 2 = medium for moderate variation, and 3 = high for marked variation and large size.

Tumors with summed scores ranging from 3 to 5 were classified as G1; 6 to 7 as G2; and 8 to 9 as G3.

Estrogen and progesterone receptors were assessed by Abbott's quantitative enzyme immunoassay (Abbott Laboratories, Chicago, IL) and deemed positive if >0.05 fmol/ μg DNA. Vascular endothelial growth factor (VEGF) was measured in tumor cytosol by a quantitative immunoassay kit (Quantikine-human VEGF; R&D Systems, Minneapolis, MN) as described (21). Protein levels of Ki67 were analyzed using anti-Ki67 antibody (MIB-1) by the grid-graticula method with cutoffs: low = 2, medium >2 and <6 , high = 6. Cyclin E was measured using the antibody HE12 (Santa Cruz Biotechnology, Inc., Santa Cruz, CA) with cutoffs: low = 0% to 4%, medium = 5% to 49%, and high = 50% to 100% stained tumor cells (22). Vascular growth was determined by routine staining of tumor sections. P53 mutational status was determined by cDNA sequencing as previously described (20). The Uppsala tumor samples were approved for microarray profiling by the ethical committee at the Karolinska Institute, Stockholm, Sweden.

Stockholm cohort. The Stockholm samples were derived from breast cancer patients operated on at the Karolinska Hospital from January 1, 1994, through December 31, 1996, and identified in the Stockholm-Gotland breast cancer registry (23). Information on patient age, tumor size, number of metastatic axillary lymph nodes, hormonal receptor status, distant metastases, site and date of relapse, initial therapy, and date and cause of death were obtained from patient records and the Stockholm-Gotland Breast Cancer Registry. Tumor sections were graded by H.N. in the same fashion as the Uppsala tumors. Only histologic G2 samples were evaluated in this study. The Stockholm samples were approved for microarray profiling by the ethical committee at the Karolinska Hospital, Stockholm, Sweden.

Singapore cohort. The Singapore samples were derived from patients operated on at the National University Hospital (Singapore) from February 1, 2000, through January 31, 2002. Routine clinical data were obtained from

pathology reports, but no information on recurrence or cause of death was available. Tumor sections were graded by T.C.P. according to the Nottingham grading system as applied to the Uppsala and Stockholm cohorts, with the following exception: Mitotic Index: 1 = low, if <8 mitoses; 2 = medium, if 9 to 16 mitoses; and 3 = high, if >16 mitoses (per 10 high-power fields); field diameter was 0.55 mm. Only histologic G2 samples were evaluated in this study. The Singapore samples were approved for microarray profiling by the Singapore National University Hospital ethics board.

After exclusions based on tissue availability, RNA amount, RNA integrity, clinical annotation, and microarray quality control, expression profiles of 249, 58, and 40 tumors from the Uppsala, Stockholm, and Singapore cohorts, respectively, were deemed suitable for further analysis.

Microarray profiling. All tumor samples were profiled on the Affymetrix U133A&B genechips. Microarray analysis of the Uppsala and Singapore samples was carried out at the Genome Institute of Singapore. The Stockholm samples were analyzed at Bristol-Myers Squibb (Princeton, NJ; ref. 23). All microarray data are accessible at National Center for Biotechnology Information (NCBI) Gene Expression Omnibus.⁹ Uppsala and Singapore data can be accessed via series accession number GSE4922; Stockholm data is accessible via series accession GSE1456. RNA preparation, microarray hybridization, and data processing were carried out essentially as described (23, 24). All data were normalized using the global mean method (MAS5), and probe set signal intensities were natural log transformed and scaled by adjusting the mean signal to a target value of log 500.

Class prediction. Prediction analysis of microarrays (PAM) and statistically weighted syndromes (SWS) were used side by side in this study to allow a performance comparison between two robust but mathematically distinct class prediction algorithms in terms of classification accuracies and total number of genes required for maximum accuracy. PAM is a modification of the nearest-centroid method and was applied as previously described (25). SWS is a supervised, combinatorial pattern recognition method based on a statistical voting procedure that uses selected subsets of predictors that act alone or in combination (26, 27). Briefly, the methodology is based on several steps: (a) optimal recoding of the given covariates to obtain discrete-valued variables; (b) selection of the most informative and statistically robust of these discrete-valued variables and their combinations (termed syndromes) that best characterize the classes of interest; (c) tallying the statistically weighted "votes" of these syndromes to compute the value of the outcome prediction function (by leave-one-out cross-validation). SWS was applied as previously described (26).

Other data sets. The Sotiriou et al. (5) data was kindly provided by C. Sotiriou (Jules Bordet Institut, Brussels, Belgium). The van't Veer et al. (3) and van de Vijver et al. (4) microarray data and clinical annotations were downloaded from the Rosetta Inpharmatics publications archive. All microarray probe sequences were mapped to UniGene build 186. For hierarchical clustering, log expression values were mean centered, and genes and tumors were clustered using Pearson correlation (uncentered) and average linkage (CLUSTER and TREEVIEW software).¹⁰

The PAM 264-gene classifier. Initially, we ran the PAM algorithm with all 44,928 probe sets as input in the Uppsala G1 to G3 comparisons and acquired a minimal set of 18 probe sets, which gave the lowest misclassification (error) rate: 3 of 68 for G1 and 3 of 55 for G3 predictions. Alternatively, we identified a secondary minimum on the error curve at 264 probe sets, which misclassified only 4 of 68 G1 and 4 of 55 G3 tumors. All 264 probe sets are differentially expressed between G1 and G3 tumors by T-test and permutation χ^2 -test (SWS) at $P < 0.01$ after adjusting for false discovery (Bonferroni).

Statistical analysis of gene ontology terms. Gene ontology (GO) analysis was facilitated by PANTHER software (28).¹¹ Selected gene lists were statistically compared (Mann-Whitney) with a reference list (i.e., NCBI

⁹ <http://www.ncbi.nlm.nih.gov/geo/>.

¹⁰ <http://rana.lbl.gov/EisenSoftware.htm>.

¹¹ <http://panther.appliedbiosystems.com/>.

Build 35) composed of all genes represented on the microarray to identify significantly overrepresented and underrepresented GO terms.

Survival analysis. The Kaplan-Meier estimate was used to compute survival curves, and the *P* value of the likelihood-ratio test was used to assess the statistical significance of the resultant hazard ratios. Disease-free survival in the Uppsala, Stockholm, and van de Vijver (4) cohorts was defined as the time interval from surgery until the first recurrence (local, regional, or distant) or last date of follow-up. For the Sotiriou and van't Veer data sets, the disease-free survival event was distant metastasis, as previously published (3, 5). Cases with contralateral disease or events occurring beyond 10 years were censored. Multivariate analysis by Cox proportional hazard regression and all survival statistics were done in the R survival package.

Scoring by the Nottingham prognostic index. Nottingham prognostic index (NPI) scores were calculated as follows: $[0.2 \times \text{tumor size (cm)}] + \text{grade (1, 2, or 3)} + \text{lymph node stage (I, II, or III)}$. Tumor size was defined as the longest diameter of the tumor. Lymph node stage was I if lymph node negative, II if ≤ 3 nodes involved, and III if > 3 nodes involved (29). As the number of cancerous lymph nodes were not available for the Uppsala cohort, a lymph node stage score of II was assigned if one or more nodes were involved and a score of III if nodal involvement showed evidence of periglandular growth. For genetic grade NPI (ggNPI) calculations, grade scores (1, 2, or 3) were replaced by genetic grade predictions (1 or 3). NPI scores ≤ 2.4 = EPG (excellent prognostic group); < 3.4 = GPG (good prognostic group); scores of 3.4 to 5.4 = MPG (moderate prognostic group); scores > 5.4 = PPG (poor prognostic group).

Descriptive statistics. For intergroup comparisons using the clinicopathologic measurements, Mann-Whitney *U* test statistics were used for continuous variables and one-sided Fisher's exact test used for categorical variables (Statistica-6 and StatXact-6 software).

Results

PAM and SWS class prediction algorithms accurately classify low- and high-grade tumors. To study the relationship

between gene expression and histologic grade, we analyzed the expression patterns of $> 39,000$ transcripts (i.e., 44,928 probe sets on Affymetrix U133A&B arrays) in 347 primary breast tumors. The tumors were derived from three independent population-based cohorts: (a) Uppsala (249 samples), (b) Stockholm (58 samples), and (c) Singapore (40 samples; Supplementary File 14). Beginning with the Uppsala data set, composed of 68 G1, 126 G2, and 55 G3 tumors, we examined the performance of two different class prediction algorithms in predicting histologic grade as defined by the Nottingham Grading System guidelines: prediction analysis of microarrays (PAM) and statistically weighted syndromes (SWS). Both algorithms rank order the genes according to specific algorithmic criteria for assessing differential expression between classes. Then, a posterior probability is iteratively estimated for each sample by classic leave-one-out cross-validation. In two-group comparisons, high misclassification error rates were observed in the G1-G2 and G2-G3 predictions by both methods (data not shown), whereas G1-G3 classification accuracy was very high, suggesting that G2 tumors are not molecularly distinct from those of low or high grade. For the G1-G3 comparisons, maximal prediction accuracies were obtained with 18 probe sets (representing 18 genes) by PAM and only six probe sets (representing five genes) by SWS (Table 1). Both methods correctly classified 96% (65 of 68) of the G1s and 95% (52 of 55) of the G3s. Notably, all genes of the classifiers (Table 1) are expressed at higher levels in G3 tumors, with the exception of three genes of the PAM classifier, which are expressed at higher levels in the G1 tumors. The posterior probability (Pr) calculated by each method is an estimate of the likelihood that a sample belongs to one class (termed "G1-like") or the other (i.e., "G3-like"). Both SWS and PAM scored the vast majority of G1 and G3 tumors with high probabilities of class membership. The Pr scores for the SWS class assignments

Table 1. Affymetrix probe sets comprising the SWS and PAM genetic grade signatures

Unigene name (build 186)	Gene symbol	Genbank accession no.	Affy ID
Barren homologue (<i>Drosophila</i>)*	BRRN1	D38553	A.212949_at
Cell division cycle associated 8	CDC48	BC001651	A.221520_s_at
V-myb myeloblastosis viral oncogene homologue (avian)-like 2	MYBL2	NM_002466	A.201710_at
Hypothetical protein FLJ11029	FLJ11029	BG165011	B.228273_at
FBJ murine osteosarcoma viral oncogene homologue B [†]	FOSB	NM_006732	A.202768_at
Chromosome 6 open reading frame 173	C6orf173	BG492359	B.226936_at
Serine/threonine-protein kinase 6	STK6	BC027464	A.208079_s_at
Anillin, actin binding protein (scraps homologue, <i>Drosophila</i>)	ANLN	AK023208	B.222608_s_at
Centromere protein E, 312 kDa	CENPE	NM_001813	A.205046_at
TTK protein kinase	TTK	NM_003318	A.204822_at
Signal peptide, CUB domain, EGF-like 2 [†]	SCUBE2	AI424243	A.219197_s_at
V-fos FBJ murine osteosarcoma viral oncogene homologue [†]	FOS	BC004490	A.209189_at
TPX2, microtubule-associated protein homologue (<i>Xenopus laevis</i>)	TPX2	AF098158	A.210052_s_at
Kinetochore protein Spc24	Spc24	AI469788	B.235572_at
Forkhead box M1	FOXM1	NM_021953	A.202580_x_at
Maternal embryonic leucine zipper kinase	MELK	NM_014791	A.204825_at
Cell division cycle associated 5	CDC45	BE614410	B.224753_at
Cell division cycle associated 3	CDCA3	NM_031299	A.221436_s_at
Serine/threonine-protein kinase 6[‡]	STK6	NM_003600	A.204092_s_at

*Probe sets in boldface comprise the SWS signature.

[†] Higher expression in G1 tumors.

[‡] All probe sets comprise the PAM signature, except STK6 (A.204092_s_at).

are shown in Supplementary File 1B. Notably, 95% of the tumors showed >75% probability of belonging to either the G1-like or G3-like class, indicating a highly discriminant statistical basis for class prediction (see also Fig. 1A, top).

Separation of G2 tumors by genetic grade. We next applied the SWS and PAM grade classifiers to the 126 G2 tumors of the Uppsala cohort to ask if these genetic determinants of low and high grade might resolve moderately differentiated G2 tumors into separable classes. Interestingly, we observed that the G2 tumors separated well into G1-like ($n = 83$) and G3-like ($n = 43$) classes with few tumors ($n = 5$) exhibiting intermediate Pr scores (Supplementary File 1C). As observed for the G1 and G3 tumors, we found that a high percentage of the G2 tumors (96%) were assigned by the SWS classifier (and 94% by the PAM classifier, data not shown) to either the G1-like or G3-like classes with >75% probability. To test whether this observation might reflect a cohort-specific selection bias, we applied the classifiers directly to G2 tumors of the Stockholm ($n = 58$) and Singapore ($n = 40$) cohorts and observed similar results (Supplementary File 1D-E), indicating that almost all G2 tumors can be molecularly well separated into distinct low- and high-grade-like classes (henceforth called G2a and G2b genetic grades).

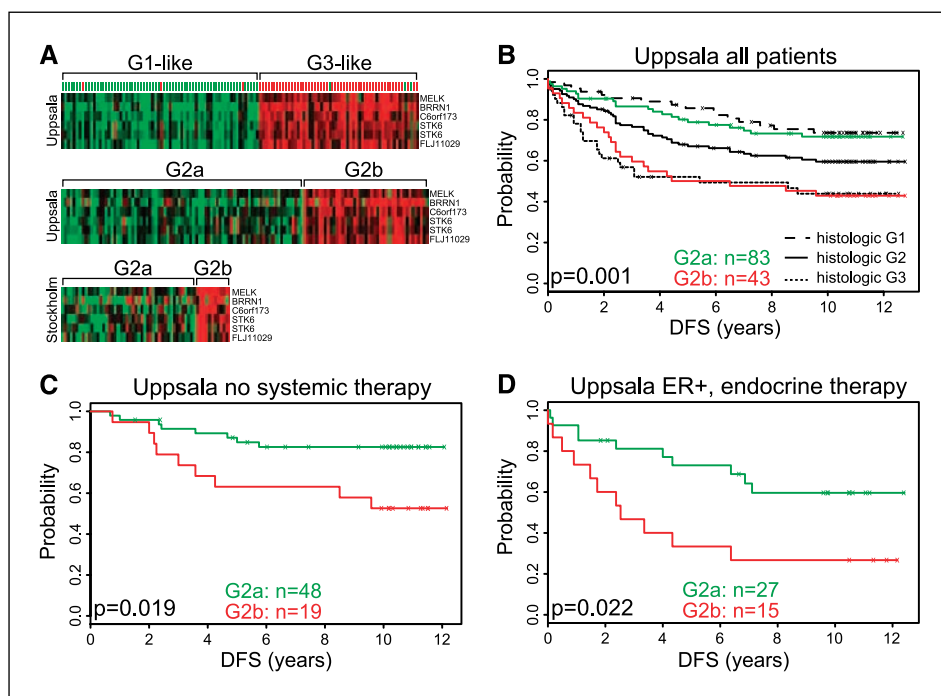
Genetic grade is prognostic of disease recurrence in moderately differentiated tumors. To determine if the genetic grade classification correlates with patient outcome, we examined the disease-free survival of patients with histologic G2 tumors classified as G2a or G2b by the SWS algorithm (Fig. 1A, middle and bottom). (Due to space limitations and high concordance between the SWS and PAM classifiers, only data for the SWS classifier are presented henceforth.) Overall, Uppsala patients with G2a tumors showed significantly less disease recurrence than those with G2b disease ($P = 0.001$; Fig. 1B). Notably, no significant difference was observed between the G2a and G1 curves, or the G2b and G3 curves. The G2a-G2b survival difference was further observed in specific therapeutic contexts, including patients who received no

adjuvant systemic therapy ($P = 0.019$; Fig. 1C) and those with estrogen receptor-positive tumors who received endocrine therapy only ($P = 0.022$; Fig. 1D). In a similar fashion, the genetic grade classifier was also predictive of recurrence in the Stockholm (G2) patients who received systemic therapy (i.e., chemotherapy, endocrine therapy or both; $P = 0.027$; Supplementary File 1F) and those with estrogen receptor-positive disease who received only endocrine treatment ($P = 0.032$; Supplementary File 1G).

To assess the possibility of microarray platform-specific bias, we tested the veracity of our findings using independent data sets generated on different microarray platforms. Three publicly available data sets were used: Sotiriou et al. (a custom spotted cDNA array manufactured at the National Cancer Institute), and van't Veer et al. and van de Vijver et al. (both of which used a custom oligonucleotide array by Agilent Technologies). As shown in Fig. 2, all three data sets confirmed the significant associations of our signature genes with disease recurrence in G2 breast cancer. G2a and G2b tumors, as estimated by hierarchical cluster analysis with the signature genes, displayed significant differences in disease-free survival in estrogen receptor-positive, tamoxifen-treated patients ($P = 0.015$; Fig. 2A); early stage, lymph node-negative patients receiving no systemic therapy ($P = 0.014$; Fig. 2B); and lymph node-positive patients receiving chemotherapy only ($P = 0.0065$; Fig. 2C). Taken together, these findings show that the genetic grade signature is a robust prognostic indicator of disease-free survival in moderately differentiated cancer independent of different treatment modalities and reproducible in multiple unrelated breast cancer cohorts.

Genetic grade is a powerful and independent risk factor. To assess the prognostic importance of the genetic grade signature, we used multivariate Cox regression models to compare its performance to that of conventional prognostic indicators assessed in each of the cohorts analyzed in Figs. 1 and 2, including lymph node status, tumor size, patient age, and estrogen receptor and progesterone receptor status (when available) entered as

Figure 1. Survival differences between G2a and G2b genetic grade subtypes. A, expression profiles of the Uppsala and Stockholm tumors segregated by the SWS genetic grade signature. Green and red vertical bars (top), histologic G1 and G3 tumors, respectively. B to D, Kaplan-Meier survival curves for G2a (green) and G2b (red) subtypes, alone or superimposed on survival curves of histologic grades 1, 2, and 3. Uppsala cohort survival curves for all patients (B), patients who received no systemic therapy (C), and patients positive for estrogen receptor who received endocrine therapy only (D). The likelihood ratio test P value reflects the significance of the hazard ratios.



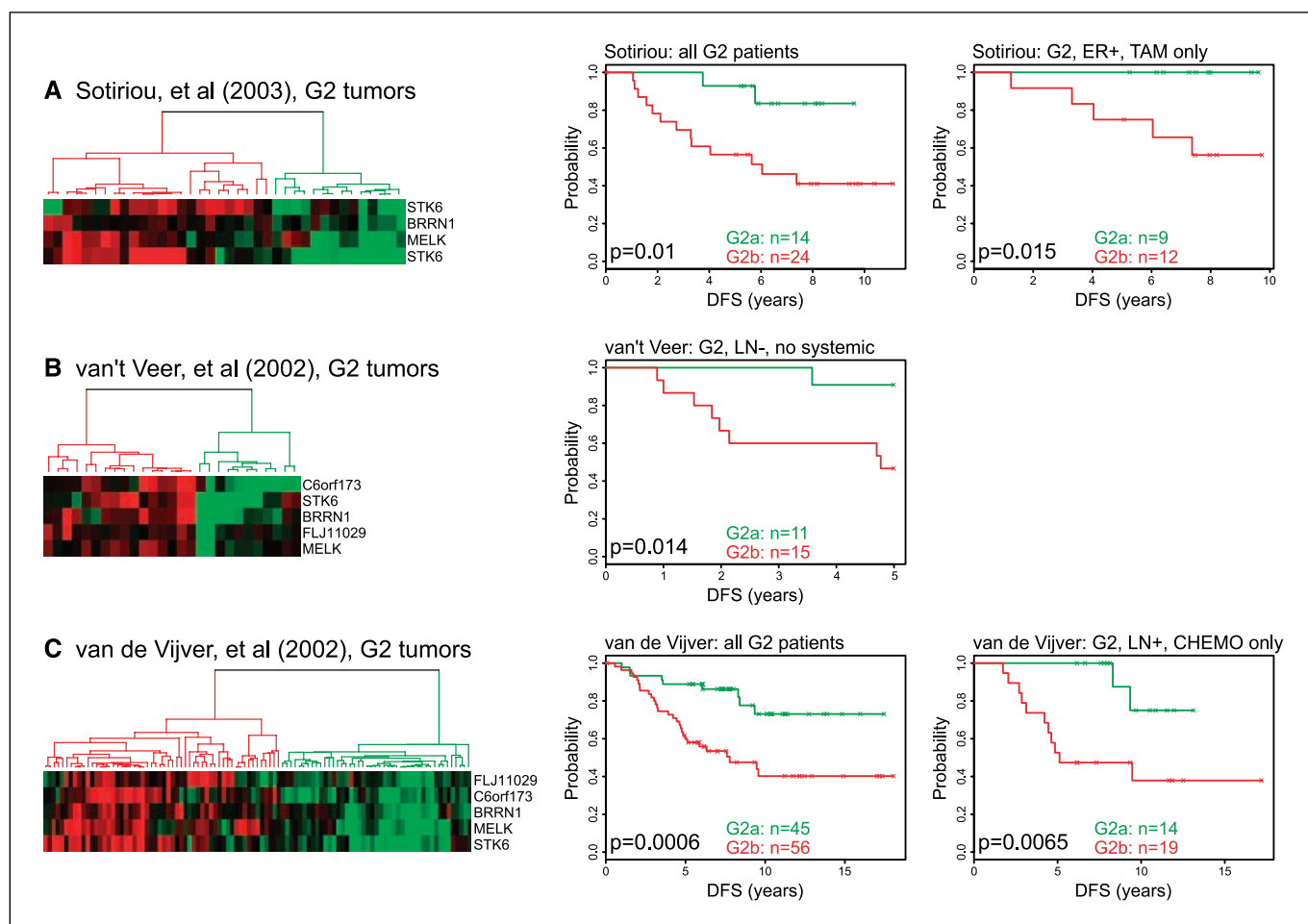


Figure 2. Meta-analysis of genetic grade genes in other patient populations. The SWS grade signature genes were mapped to three previously published breast cancer microarray data sets: (A) Sotiriou et al. (ref. 5; cDNA microarray); (B) van't Veer et al. (ref. 3; long-oligonucleotide array); and (C) van de Vijver et al. (ref. 4; same array as in B). G2 tumors from each study were hierarchically clustered. Primary tumor branches (separating low and high expressors) were used to approximate G2a and G2b tumors for survival analysis. Kaplan-Meier estimates of disease-free survival are shown for all patients (A-C) and treatment-specific subgroups (A, C).

categorical variables. We found that the genetic grade signature not only remained significantly associated with disease recurrence in most cases, but performed equal to or better than lymph node status and tumor size in many patient subgroups (see Supplementary File 2A-E).

G2a and G2b subtypes are molecularly and pathologically distinct. The clinical data suggest that G2a and G2b genetic grades may represent separate pathologic entities. We investigated this possibility in the Uppsala cohort by several approaches. First, we analyzed the expression levels of the 264 probe sets (i.e., representing ~232 genes) that we identified by PAM as the maximum number of probe sets capable of recapitulating a high G1-G3 classification accuracy (see Materials and Methods). These genes represent the topmost significant differentially expressed genes between G1 and G3 tumors after correcting for false discovery (Supplementary File 3A). As shown in Fig. 3, hierarchical cluster analysis using this set of genes shows a striking separation of the G2 population into two primary tumor profiles highly resembling the G1 and G3 profiles and that separate well into the G2a and G2b classes. Indeed, all but 11 of these 264 probe sets were also differentially expressed (at $P < 0.05$, Wilcoxon rank-sum test) between the G2a and G2b tumors. This finding shows that extensive molecular heterogeneity exists within the G2 tumor

population, and this heterogeneity is robustly defined by the major determinants of G1 and G3 cancer. It also shows that a much larger and pervasive transcriptional program underlies the genetic grade predictions of the SWS signature—despite its composition of a mere five genes. Furthermore, statistical analysis of the GO terms associated with the G2a-G2b differentially expressed genes revealed the significant enrichment of numerous biological processes and molecular functions. Table 2 displays a selected set of significantly enriched GO categories that includes cell cycle, inhibition of apoptosis, cell motility, and stress response, suggesting an imbalance of these cellular processes between the G2a- and G2b-type tumor cells (see Supplementary File 3B for the complete list of GO categories and associated P values).

To extend our analysis beyond the transcriptional level, we investigated the differences between G2a and G2b tumors using conventional clinicopathologic variables measured in the Uppsala cohort (Supplementary File 2F). Of the three histologic grading criteria, both mitotic count and nuclear pleomorphism were found to significantly vary between the G2a and G2b tumors ($P = 0.007$ and $P = 0.05$, respectively). Protein levels of the proliferation marker Ki67 were also found to be significantly different between the G2a and G2b tumors ($P < 0.0001$). These findings, together with those of the GO analysis, suggest that the genetic grade signature

may largely mirror cell proliferation, and thus reflect the replicative potential of breast tumor cells. However, other oncogenic factors were also found to be associated with genetic grade. In the G2b tumors, protein levels of VEGF, a major inducer of angiogenesis, and the degree of vascular growth were both found to be significantly higher compared with the G2a samples ($P = 0.015$ and $P = 0.002$, respectively), suggesting that a difference in angiogenic potential also distinguishes the two genetic grade classes. TP53 mutations were found in only 6% ($n = 5$) of the G2a tumors, whereas 44% ($n = 19$) of the G2b tumors were p53 mutants ($P < 0.0001$) consistent with their higher replicative potential, and likely conferring a further survival advantage to these tumors via decreased apoptotic potential. We also observed higher levels of cyclin E1 protein ($P = 0.04$) in the G2b tumors, which, in addition to contributing to enhanced proliferation (30), may also confer greater genomic instability (31, 32). Finally, we observed a significant difference in hormonal status between the G2a and G2b tumors, with a higher fraction of estrogen receptor-negative ($P = 0.06$) and progesterone receptor-negative ($P = 0.02$) tumors in the G2b class, indicating differences in hormone sensitivity and dependence. Taken together, these results show that multiple tumorigenic properties measured at the RNA, DNA, protein, and cellular levels can subdivide the G2a and G2b tumor subtypes.

G2a and G2b tumors are not identical to histologic G1 and G3 cancers. Both the survival and gene expression data suggest that the G2a and G2b classes may be clinically and molecularly indistinguishable from histologic G1 and G3 tumors, respectively. To address this, we further analyzed the expression patterns of the 264 grade-associated probe sets in the Uppsala cohort. We

discovered 14 genes and 57 genes significantly differentially expressed ($P < 0.01$, Mann-Whitney U test) between the G1 and G2a tumors and the G3 and G2b tumors, respectively. By GO analysis, the differentially expressed genes of the G1-G2a comparison pointed to significant differences in cell cycle-related processes and oncogenesis, whereas differences between the G2a and G3 tumors included cell cycle-related processes, inhibition of apoptosis, oncogenesis, and cell motility (Table 2; Supplementary File 3B).

Statistical analysis of the clinical markers revealed further distinctions in the G1-G2a and the G2b-G3 tumor comparisons. As shown in Supplementary File 2F, G2a tumors displayed significant increases in tumor size, lymph node positivity, cellular mitoses, tubule formation, and Ki67 levels compared with histologic G1 tumors, and the G3 population showed significant increases in tumor size, vascular growth, mitoses, tubule formation, cyclin E1, and estrogen receptor-negative status when compared with the G2b tumors. Taken together, these data indicate that the G2a and G2b populations, although highly similar to G1 and G3 tumors in terms of survival and transcriptional configuration, remain separable by conventional clinical characteristics and GO analysis of differentially expressed genes.

Genetic grade improves prognosis by the NPI. The NPI is a widely accepted method of stratifying patients into prognostic groups (GPG, MPG, and PPG) based on lymph node stage, tumor size, and histologic grade (33). We investigated whether incorporating genetic grade into the NPI could improve patient stratification. A simplified substitution method was explored. For all tumors of the Uppsala and Stockholm cohorts for which NPI

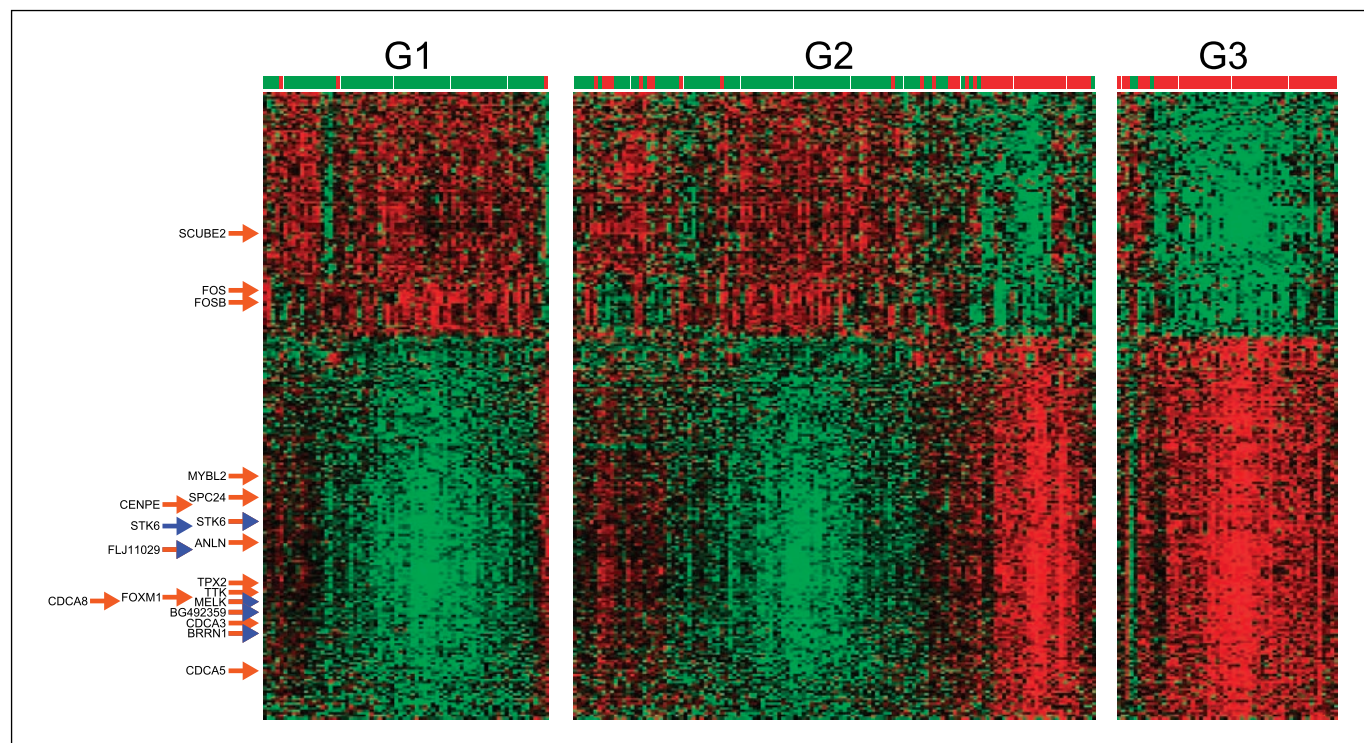


Figure 3. Expression profiles of the top 264 grade (G1-G3) associated gene probe sets. Gene probe sets (rows) and tumors (columns) were hierarchically clustered by Pearson correlation and average linkage, then tumors were grouped according to grade, while maintaining original cluster order within groups. Red, above mean expression; green, below mean expression. The degree of color saturation reflects the magnitude of expression relative to the mean. Vertical bars, G1-like/G2a tumors (green) and G3-like/G2b tumors (red).

Table 2. GO enrichment analysis of differentially expressed genes

	G1 vs G2a				G2a vs G2b				G2b vs G3			
	<i>P</i> *	T [†]	E [‡]	O [§]	<i>P</i>	T	E	O	<i>P</i>	T	E	O
Biological process												
Cell cycle	6.2E-06	853	0.76	7	5.7E-28	853	7.08	50	2.5E-06	853	2.29	12
Chromatin packaging and remodeling	1.3E-02	196	0.18	2	2.5E-02	196	1.63	5				
Mitosis	2.7E-02	287	0.26	2	6.8E-15	287	2.38	22	1.1E-03	287	0.77	5
Inhibition of apoptosis					4.4E-03	127	1.05	5	4.9E-03	127	0.34	3
Oncogenesis	1.6E-02	600	0.54	3	5.5E-04	600	4.98	14	5.5E-03	600	1.61	6
Cell motility					3.6E-02	291	2.42	6	4.4E-02	291	0.78	3
Stress response					5.0E-03	187	1.55	6				
Molecular function												
Kinase activator	1.1E-03	54	0.05	2	7.2E-06	54	0.45	6				
Histone	3.5E-03	99	0.09	2	5.0E-02	99	0.82	3				
Nucleic acid binding	1.3E-02	3014	2.7	7								
Microtubule family cytoskeletal protein					7.6E-07	233	1.93	12	4.2E-04	233	0.63	5
Chemokine									7.5E-03	48	0.13	2
Nonreceptor serine/threonine protein kinase					7.8E-04	289	2.4	9				
Extracellular matrix linker protein					1.9E-02	25	0.21	2				
Pathway												
Insulin/IGF pathway-MAPKK/MAPK cascade	4.9E-02	56	0.05	1								
Apoptosis signaling pathway									4.9E-02	131	0.35	2
Ubiquitin proteasome pathway					3.0E-02	80	0.66	3				

Abbreviations: IGF, insulin-like growth factor; MAPKK, mitogen-activated protein kinase kinase.

**P* value for significance of GO term enrichment.

† Total number of genes with given GO annotation.

‡ Expected number of genes with given GO annotation.

§ Observed number of genes with given GO annotation.

scores and survival information could be obtained ($n = 296$), histologic grade (1, 2, or 3) was replaced by the genetic grade prediction (1 or 3) and new NPI (i.e., ggNPI) scores were computed (see Materials and Methods). The survival of patients stratified into risk groups was then compared between classic NPI and ggNPI. As shown in Fig. 4A, the survival curves of the NPI and ggNPI prognostic groups were highly comparable; however, we observed that the ggNPI tended to shift patients from worse to better prognostic groups. Practical guidelines that use the NPI in therapeutic decision-making often recognize an EPG composed of patients with NPI scores ≤ 2.4 (34, 35). Untreated patients in this group with lymph node-negative disease have a 95% 10-year

survival probability—equivalent to that of an age-matched female population without breast cancer (35). Thus, patients in this group are routinely not recommended for postoperative adjuvant therapy (34–36). We compared the EPGs, as defined by the NPI and ggNPI stratifications, in a subset of 142 lymph node-negative patients who received no adjuvant systemic therapy. Forty and 82 patients were classified into the EPG by the classic NPI and ggNPI, respectively. Of the 40 patients classified into the EPG by the classic NPI, only one was considered different by the ggNPI; whereas of those classified as needing adjuvant therapy by the classic NPI (i.e., scores >2.4), 43 were reclassified by the ggNPI into the EPG. When examined for outcome, the survival curves of the 40 and 82

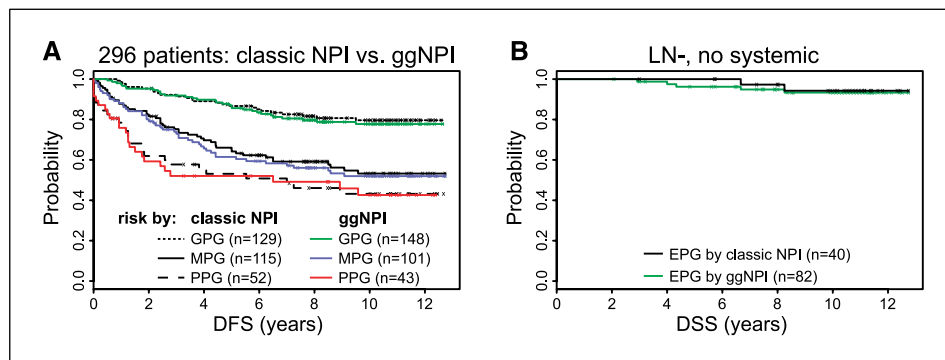


Figure 4. Stratification of patient risk by classic NPI and ggNPI. A, Kaplan-Meier survival curves are shown for risk groups determined by the classic NPI (black curves) and the NPI calculated with genetic grade assignments (ggNPI; colored curves). B, the disease-specific survival curves of node-negative, untreated patients classified into the EPG by classic NPI (black curve) or ggNPI (green curve) are compared.

EPG patients by NPI and ggNPI, respectively, were statistically indistinguishable, both showing ~94% survival at 10 years (Fig. 4B). Thus, twice as many patients could be accurately classified into the EPG by the ggNPI, suggesting that the use of genetic grade can improve prognostication of which patients should be spared adjuvant systemic therapy.

Discussion

Although the clinical subtyping of cancer has historically been based largely on phenotypic properties, comprehensive genomic and transcriptomic analyses are beginning to reveal robust genotypic determinants of tumor subtype. In this report, we show that the genetic essence of histologic grade can be distilled down to the expression patterns of a mere five genes with powerful prognostic implications. We found that G2a and G2b tumors, as distinguished by the five-gene genetic grade signature, displayed markedly different metastatic potentials—a finding that was robustly reproducible in independent breast cancer cohorts and in the context of different microarray platforms. Further analysis revealed that the G2a and G2b classes could be separated by extensive molecular, biological, and tumorigenic differences known to separate low- and high-grade cancer (37), including proliferation rate (mitotic index, Ki67), angiogenic potential (VEGF, vascular growth), p53 mutational status, and estrogen and progesterone dependence.

Ma et al. (38) were the first to report a histologic grade signature capable of distinguishing low- and high-grade breast cancer. Using 12K cDNA microarrays to analyze material from 10 G1, 11 G2, and 10 G3 microdissected tumors, they identified 200 genes differentially expressed between G1 and G3 tumors. Using these genes for tumor clustering, they observed that the majority of G2 tumors possessed a hybrid signature intermediate to G1 and G3, with few exceptions (see Fig. 3 of their original report; ref. 38). Notably, this finding is in contrast with our discovery that the majority of G2 tumors do not display hybrid signatures, but rather possess clear G1-like or G3-like genetic features. According to our classifier, only a small percentage (~6%) of the tumors in our study had intermediate genetic grade measurements (i.e., Pr scores <0.75 for G1-like and G3-like). To address this discrepancy, we cross-compared the 200 grade-associated genes in their list to our expanded set of 232 genes, and observed a statistically significant overlap of 35 genes ($P < 1.0 \times 10^{-7}$; Monte Carlo simulation). This overlap, however, represents only a small percentage of either gene list, indicating that the discrepant observations are most likely explained by fundamentally different signature compositions. It is also possible that differences in sample selection and preparation, sample size, RNA purification, microarray analysis, and data normalization could have contributed to the variable results.

It should also be noted that although a small percentage (6%) of the tumors in this study had intermediate genetic grade measurements, too few were discovered to determine the clinical relevance of this intermediate genotype. Furthermore, the origins of these intermediate tumors are unclear and could be biological or technical in nature. They may arise as homogeneous tumors that truly borderline low and high grade, or rather reflect heterogeneous compositions of both low- and high-grade cell types, such as that observed in tubular mixed carcinoma (39). Alternatively, that we observed the same percentage of intermediacy in tumors of all grades and across cohorts suggests that this class may represent a baseline level of uncertainty owing to technical noise.

Whether grade is a continuum through which breast cancer progresses or merely the end point of distinct genetic pathways has been debated (39–44). Studies comparing primary tumors to their subsequent metastases have supported the grade progression model, particularly when multiple recurrences were analyzed (39, 45). However, comparative genomic studies have identified reproducible chromosomal alterations that distinguish low- and high-grade disease, including a 16q deletion unique to G1 carcinomas (43, 44, 46). These studies argue against the progression model and point to genetic origins of histologic grade. In our study, 94% of 347 primary tumors could be molecularly classified with high probability of being G1-like or G3-like. This finding supports the genetic pathways model of grade origin, suggesting that the large majority of breast cancers fundamentally exist in one of two predominant forms marked by the molecular and clinical essence of low or high grade. Whether these forms correlate with grade-specific genomic alterations (43, 44, 46) remains unknown.

In multivariate analyses (Supplementary File 2), not only did the grade signature remain an independent predictor of disease recurrence in most of the patient subgroups analyzed, but in most cases, it was equivalent to or more powerful than lymph node status and tumor size, underscoring its role as a valuable new prognostic indicator. Other studies have reported gene expression signatures prognostic of breast cancer recurrence (2–5, 24, 47, 48); however, the contribution of grade-predictive genes within these signatures is largely unclear. Of our five genetic grade genes, three are components of other prognostic signatures in breast cancer. BRRN1 overlaps with our previously published 32-gene TP53 signature (24). MELK is a component of both the van't Veer et al. (3) 70-gene metastasis predictor and the Sotiriou et al. (5) 485-gene list associated with recurrence. STK6 is a component of the Paik et al. (47) 21-gene metastasis predictor. To what extent these signatures capture the prognostic potential of grade remains unknown. For example, histologic grade was not significant in a multivariate model with the van't Veer (4) 70-gene predictor, whereas grade remained highly significant ($P < 0.001$) in a multivariate model including the Paik 21-gene signature (47). How the genetic grade signature can add value to conventional and signature-based prognostic models should be further considered.

Following submission of this article, Sotiriou et al. (49) published their findings of a 97-gene expression grade index associated with histologic grade and correlated with relapse-free survival in estrogen receptor-positive breast cancer. Their grade index, like our grade signature, could dichotomize the vast majority of G2 tumors into two groups with expression profiles and survival characteristics resembling those of G1 and G3 tumors. Likewise, the prognostic performance of their grade index was found to be independent of lymph node status and tumor size. Comparison of the two gene classifiers revealed that three of our five grade signature genes, and 68 of our larger 232-gene set, overlapped with their 97-gene index. This high degree of overlap suggests that the two signatures may use the same fundamental transcriptional programming for predicting patient outcomes. Whether the two predictors are collinear with respect to patient survival will be an important question moving forward. Nevertheless, that our two studies converge on similar findings reinforces the view that gene expression-based measurements of histologic grade can substantially contribute to patient prognosis.

Our results indicate that G2 invasive breast cancer, at least in genetic terms, does not exist as a single clinical entity. The genetic grade signature dichotomized G2 tumors into biologically and

clinically distinct subtypes that significantly improved patient prognosis. By adding genetic grade to the NPI (ggNPI), we showed that twice as many lymph node-negative patients could be accurately classified with an excellent prognosis, suggesting that the use of genetic grade can improve prediction of which patients should be spared systemic adjuvant therapy thereby minimizing harm due to late adverse health effects. The value of the ggNPI on the prognosis of lymph node-positive patients, however, was less clear as the stratification of these patients by the classic NPI was encumbered by insufficient clinical data (in the Uppsala cohort) regarding number of positive nodes. Further studies in large and sufficiently characterized cohorts will be needed to better determine the value of integrating tumor genetic grade with the NPI.

In conclusion, the genetic grading of moderately differentiated tumors allows the refinement of the G2 subtype into subgrades 2a and 2b with significant clinical ramifications. Although several G1 and G3 tumors were misclassified by the genetic grade signature, too few were observed to study the prognostic significance of the signature in high and low grade disease. Indeed, it is possible that the predictive capacity of the grade signature could extend to

all tumors, regardless of histologic grade, as a continuous scalable variable (7, 47). How to best integrate the genetic grade measurement with other risk factors for improved prognosis warrants further study.

Acknowledgments

Received 12/22/2005; revised 8/7/2006; accepted 8/25/2006.

Grant support: Singapore Agency for Science Technology and Research and grants from the Swedish Cancer Society, the Stockholm Cancer Society, and the King Gustav Fifth Jubilee Fund.

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

We thank Dr. L. Lipovich for gene annotation work, Drs. S. Klaar and J. Bjohle for clinical annotation of specimens and patients, and K.R. Govindarajan for database support.

Author contributions: J. Bergh and J.E.L. Wong provided access to patient material. H. Nordgren, T.C. Putti, and T. Lindahl characterized the tumor specimens. J. Smeds and B. Mow performed microarray experiments. V.A. Kuznetsov provided bioinformatics and statistical design of the study. V.A. Kuznetsov, A.V. Ivshina, J. George, and O. Senko analyzed the data, E.T. Liu, P.H. Hall, and Y. Pawitan helped shape the paper. A.V. Ivshina, V.A. Kuznetsov, and L.D. Miller conceived of the study, provided interpretation of the results, and wrote the paper.

References

- Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000;403:503-11.
- Sorlie T, Perou CM, Tibshirani R, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 2001;98:10869-74.
- van't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530-6.
- van de Vijver MJ, He YD, van't Veer LJ, et al. A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* 2002;347:1999-2009.
- Sotiriou C, Neo SY, McShane LM, et al. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 2003;100:10393-8.
- Bullinger L, Dohner K, Bair E, et al. Use of gene-expression profiling to identify prognostic subclasses in adult acute myeloid leukemia. *N Engl J Med* 2004;350:1605-16.
- Chang HY, Nuyten DS, Sneddon JB, et al. Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. *Proc Natl Acad Sci U S A* 2005;102:3738-43.
- Bloom HJ, Richardson WW. Histological grading and prognosis in breast cancer; a study of 1409 cases of which 359 have been followed for 15 years. *Br J Cancer* 1957;11:359-77.
- Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 1991;19:403-10.
- Schumacher M, Schmoor C, Sauerbrei W, et al. The prognostic effect of histological tumor grade in node-negative breast cancer patients. *Breast Cancer Res Treat* 1993;25:235-45.
- Roberti NE. The role of histologic grading in the prognosis of patients with carcinoma of the breast: is this a neglected opportunity? *Cancer* 1997;80:1708-16.
- Lundin J, Lundin M, Holli K, et al. Omission of histologic grading from clinical decision making may result in overuse of adjuvant therapies in breast cancer: results from a nationwide study. *J Clin Oncol* 2001;19:28-36.
- Harvey JM, de Klerk NH, Sterrett GF. Histological grading in breast cancer: interobserver agreement, and relation to other prognostic factors including ploidy. *Pathology* 1992;24:63-8.
- Frierson HF, Jr., Wolber RA, Berean KW, et al. Interobserver reproducibility of the Nottingham modification of the Bloom and Richardson histologic grading scheme for infiltrating ductal carcinoma. *Am J Clin Pathol* 1995;103:195-8.
- Robbins P, Pinder S, de Klerk N, et al. Histological grading of breast carcinomas: a study of interobserver agreement. *Hum Pathol* 1995;26:873-9.
- Dalton LW, Pinder SE, Elston CE, et al. Histologic grading of breast cancer: linkage of patient outcome with level of pathologist agreement. *Mod Pathol* 2000;13:730-5.
- Younes M, Laucirica R. Lack of prognostic significance of histological grade in node-negative invasive breast carcinoma. *Clin Cancer Res* 1997;3:601-4.
- Hayes DF, Isaacs C, Stearns V. Prognostic factors in breast cancer: current and new predictors of metastasis. *J Mammary Gland Biol Neoplasia* 2001;6:375-92.
- Trudeau ME, Pritchard KI, Chapman JA, et al. Prognostic factors affecting the natural history of node-negative breast cancer. *Breast Cancer Res Treat* 2005;89:35-45.
- Bergh J, Norberg T, Sjogren S, et al. Complete sequencing of the p53 gene provides prognostic information in breast cancer patients, particularly in relation to adjuvant systemic therapy and radiotherapy. *Nat Med* 1995;1:1029-34.
- Linderholm BK, Lindahl T, Holmberg L, et al. The expression of vascular endothelial growth factor correlates with mutant p53 and poor prognosis in human breast cancer. *Cancer Res* 2001;61:2256-60.
- Lindahl T, Landberg G, Ahlgren J, et al. Overexpression of cyclin E protein is associated with specific mutation types in the p53 gene and poor survival in human breast cancer. *Carcinogenesis* 2004;25:375-80.
- Pawitan Y, Bjohle J, Amler L, et al. Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast Cancer Res* 2005;7:R953-64.
- Miller LD, Smeds J, George J, et al. An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A* 2005;102:13550-5.
- Tibshirani R, Hastie T, Narasimhan B, et al. Diagnosis of multiple cancer types by shrunken centroids of gene expression. *Proc Natl Acad Sci U S A* 2002;99:6567-72.
- Kuznetsov VA, Ivshina AV, Sen'ko OV, Kuznetsova AV. Syndrome approach for computer recognition of fuzzy systems and its application to immunological diagnostics and prognosis of human cancer. *Math Comput Model* 1996;23:92-112.
- Jackson AM, Ivshina AV, Senko O, et al. Prognosis of intravesical bacillus Calmette-Guerin therapy for superficial bladder cancer by immunological urinary measurements: statistically weighted syndrome analysis. *J Urol* 1998;159:1054-63.
- Mi H, Lazareva-Ulitsky B, Loo R, et al. The PANTHER database of protein families, subfamilies, functions and pathways. *Nucleic Acids Res* 2005;33:D284-8.
- Swedish Breast Cancer Cooperative Group. Randomized trial of two versus five years of adjuvant tamoxifen for postmenopausal early stage breast cancer. *J Natl Cancer Inst* 1996;88:1543-9.
- Lukas J, Herzinger T, Hansen K, et al. Cyclin E-induced S phase without activation of the pRb/E2F pathway. *Genes Dev* 1997;11:1479-92.
- Spruck CH, Won KA, Reed SI. Deregulated cyclin E induces chromosome instability. *Nature* 1999;401:297-300.
- Minella AC, Swanger J, Bryant E, et al. p53 and p21 form an inducible barrier that protects cells against cyclin E-cdk2 deregulation. *Curr Biol* 2002;12:1817-27.
- Haybittle JL, Blamey RW, Elston CW, et al. A prognostic index in primary breast cancer. *Br J Cancer* 1982;45:361-6.
- Blamey RW. The design and clinical use of the Nottingham Prognostic Index in breast cancer. *Breast* 1996;5:156-7.
- Stotter A. A prognostic table to guide practitioners advising patients on adjuvant systemic therapy in early breast cancer. *Eur J Surg Oncol* 1999;25:341-3.
- Feldman M, Stanford R, Catcheside A, et al. The use of a prognostic table to aid decision making on adjuvant therapy for women with early breast cancer. *Eur J Surg Oncol* 2002;28:615-9.
- Lacroix M, Toillon RA, Leclercq G. Stable "portrait" of breast tumors during progression: data from biology, pathology and genetics. *Endocr Relat Cancer* 2004;11:497-522.
- Ma XJ, Salunga R, Tuggle JT, et al. Gene expression

- profiles of human breast cancer progression. *Proc Natl Acad Sci U S A* 2003;100:5974-9.
39. Cserni G. Tumour histological grade may progress between primary and recurrent invasive mammary carcinoma. *J Clin Pathol* 2002;55:293-7.
40. Tubiana M, Koscielny S. Natural history of human breast cancer: recent data and clinical implications. *Breast Cancer Res Treat* 1991;18:125-40.
41. Tabar L, Fagerberg G, Chen HH, et al. Tumour development, histology and grade of breast cancers: prognosis and progression. *Int J Cancer* 1996;66:413-9.
42. Millis RR, Barnes DM, Lampejo OT, et al. Tumour grade does not change between primary and recurrent mammary carcinoma. *Eur J Cancer* 1998;34:548-53.
43. Roylance R, Gorman P, Harris W, et al. Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer. *Cancer Res* 1999;59:1433-6.
44. Buerger H, Otterbach F, Simon R, et al. Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. *J Pathol* 1999;189:521-6.
45. Hitchcock A, Ellis IO, Robertson JF, et al. An observation of DNA ploidy, histological grade, and immunoreactivity for tumour-related antigens in primary and metastatic breast carcinoma. *J Pathol* 1989;159:129-34.
46. Buerger H, Mommers EC, Littmann R, et al. Ductal invasive G2 and G3 carcinomas of the breast are the end stages of at least two different lines of genetic evolution. *J Pathol* 2001;194:165-70.
47. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 2004;351:2817-26.
48. Wang Y, Klijn JG, Zhang Y, et al. Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *Lancet* 2005;365:671-9.
49. Sotiriou C, Wirapati P, Loi S, et al. Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. *J Natl Cancer Inst* 2006;98:262-72.