

Distinct Patterns of DNA Copy Number Alteration Are Associated with Different Clinicopathological Features and Gene-Expression Subtypes of Breast Cancer

Anna Bergamaschi,^{1,2} Young H. Kim,² Pei Wang,^{3†} Therese Sørli,¹ Tina Hernandez-Boussard,⁴ Per E. Lonning,⁵ Robert Tibshirani,^{3,6} Anne-Lise Børresen-Dale,^{1,7} and Jonathan R. Pollack^{2*}

¹Department of Genetics, Institute for Cancer Research, Rikshospitalet-Radiumhospitalet Medical Center, Oslo, Norway

²Department of Pathology, Stanford University, Stanford, California

³Department of Statistics, Stanford University, Stanford, California

⁴Department of Biochemistry, Stanford University, Stanford, California

⁵Department of Medicine (Oncology), Haukeland University Hospital, Bergen, Norway

⁶Health Research & Policy, Stanford University, Stanford, California

⁷Medical Faculty, University of Oslo, Oslo, Norway

Breast cancer is a leading cause of cancer-death among women, where the clinicopathological features of tumors are used to prognosticate and guide therapy. DNA copy number alterations (CNAs), which occur frequently in breast cancer and define key pathogenetic events, are also potentially useful prognostic or predictive factors. Here, we report a genome-wide array-based comparative genomic hybridization (array CGH) survey of CNAs in 89 breast tumors from a patient cohort with locally advanced disease. Statistical analysis links distinct cytoband loci harboring CNAs to specific clinicopathological parameters, including tumor grade, estrogen receptor status, presence of *TP53* mutation, and overall survival. Notably, distinct spectra of CNAs also underlie the different subtypes of breast cancer recently defined by expression-profiling, implying these subtypes develop along distinct genetic pathways. In addition, higher numbers of gains/losses are associated with the “basal-like” tumor subtype, while high-level DNA amplification is more frequent in “luminal-B” subtype tumors, suggesting also that distinct mechanisms of genomic instability might underlie their pathogenesis. The identified CNAs may provide a basis for improved patient prognostication, as well as a starting point to define important genes to further our understanding of the pathobiology of breast cancer. This article contains Supplementary Material available at <http://www.interscience.wiley.com/jpages/1045-2257/suppmat>. © 2006 Wiley-Liss, Inc.

INTRODUCTION

Breast cancer is a leading cause of cancer-related morbidity and mortality among women. For patients with breast cancer, clinical parameters and histopathological features of the tumor are used to prognosticate clinical outcome. Poor prognostic indicators for localized breast cancer include metastasis to draining axillary lymph nodes, large tumor size, high tumor grade (i.e., poor differentiation, nuclear pleomorphism, and high mitotic index), *ERBB2* (Her2/neu) gene amplification or protein overexpression, and estrogen receptor (ER) negativity (for short term prognosis) (Subramaniam and Isaacs, 2005). The presence of such features is used to determine which patients to treat with adjuvant chemotherapy following surgical resection of their tumor. Additional molecular prognostic markers have been suggested and await further evaluation (Gradishar, 2005; Subramaniam and Isaacs, 2005).

Histopathological and molecular features of tumors are also used to predict tumor response to specific therapies, and thereby select optimal therapeutic regimens. For example, ER or progesterone receptor (PR) positivity predicts response to hormonal therapy with selective estrogen receptor modulators (e.g. tamoxifen) (McGuire, 1978; EBCTCG,

Abbreviations: CGH, comparative genomic hybridization; CLAC, CLuster Along Chromosomes; CNA, copy number alteration; ER, estrogen receptor; FDR, false discovery rate.

Supported by: NIH; Grant number: CA97139, CBCRP; Grant number: 8KB-0135; The Norwegian Cancer Society; Grant number: D-99061 NFR; Grant number: 155218/300; Lillemor Grobstock legacy.

*Correspondence to: Jonathan R. Pollack, M.D., Ph.D., Department of Pathology, Stanford, California 94305-5176, USA. E-mail: pollack1@stanford.edu

†Present address: Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA.

Received 6 April 2006; Accepted 27 June 2006

DOI 10.1002/gcc.20366

Published online 8 August 2006 in

Wiley InterScience (www.interscience.wiley.com).

2005). Amplification/overexpression of *ERBB2* predicts responsiveness to trastuzumab (a humanized monoclonal antibody targeting Her2/neu) (Yeon and Pegram, 2005) and also to dose-dependent use of anthracyclines (Muss et al., 1994), likely due to co-amplification of *TOP2A* (Jarvinen et al., 2000).

More recently, molecular profiling methods have been used to identify clinically-relevant tumor features not previously appreciated by pathologists. For example, discovery-based (i.e. unsupervised) gene expression profiling studies have defined several distinct subtypes of breast cancer, including the so-called “luminal epithelial-like” subtypes A and B, a “basal epithelial-like” subtype, an *ERBB2*-amplification associated subtype, and a “normal breast-like” subtype (Perou et al., 2000; Sorlie et al., 2001). While luminal A and B subtypes are both ER-positive, luminal B cases are associated with less favorable outcome, as is the *ERBB2* subtype and particularly the ER-negative basal-like subtype (Perou et al., 2000; Sorlie et al., 2001). More directed (i.e. supervised) expression-profiling efforts have also defined gene signatures that predict disease recurrence and survival (van't Veer et al., 2002; Paik et al., 2004); it remains to be seen whether such signatures yield benefit over conventional markers (Eden et al., 2004).

Genomic DNA copy number alterations (CNAs) also provide potentially useful molecular markers for breast cancer prognostication or prediction of treatment response. Frequently observed CNAs include gain of chromosomal regions 1q, 8q, 17q, and 20q, and loss of 1p, 8p, 13q, and 17p (Knuutila et al., 2000). Sites of localized high-level DNA amplification harboring known oncogenes include 7p12 (*EGFR*), 8q24 (*MYC*), 11q13 (*CCND1*), 12q14 (*MDM2*), 17q12 (*ERBB2*), 20q12 (*AIB1*), and 20q13 (*ZNF217*) [(Al-Kuraya et al., 2004), and references therein]. Deletions with known tumor suppressor genes (TSGs) include 13q12 (*BRCA2*), 17p13 (*TP53*), and 17q21 (*BRCA1*). Cytogenetic studies have identified gains on 8q, 17q12, and 20q13 to be associated with poor overall survival (Isola et al., 1995; Tanner et al., 1995; Ross and Fletcher, 1998). DNA amplification of *ERBB2* at 17q12 also predicts response to trastuzumab and high-dose anthracyclines. Since genomic DNA is more stable than mRNA, and since CNAs define key genetic events driving tumorigenesis, such genomic alterations are potentially advantageous as prognostic/predictive factors.

Here, we have used array-based comparative genomic hybridization (array CGH) to profile CNAs genome-wide at high-resolution for 89 locally ad-

vanced primary breast tumors. We report the identification of distinct loci of CNA associated with different clinicopathological features, including tumor grade, ER status, *TP53* mutation, gene-expression subtype, and overall survival. The identified CNAs may provide a basis for improved patient prognostication, as well as a starting point to define important genes contributing to breast cancer development and progression.

MATERIALS AND METHODS

Breast Cancer Specimens

Breast tumor specimens were derived from 89 patients with locally advanced (T3/T4 and/or N2) breast cancer receiving either doxorubicin or fluorouracil-mitomycin based neoadjuvant chemotherapy (Geisler et al., 2001; 2003). Specimens were snap-frozen immediately following surgical excision, then stored at -80°C . Genomic DNA was isolated by chloroform/phenol extraction followed by ethanol precipitation (Nuclear Acid Extractor 340A; Applied Biosystem) exactly as described (Geisler et al., 2001). For 84 of the 89 cases, array CGH analysis was performed using genomic DNA from specimens obtained prior to neoadjuvant therapy. ER status was determined by ligand-binding assay, and *TP53* mutations were previously identified by temporal temperature gradient gel electrophoresis (TTGE) followed by DNA sequencing as described (Geisler et al., 2001). cDNA microarray-based gene expression profiling data for 87 of the tumor specimens were previously published (Sorlie et al., 2001; 2003), as well as their assignments to gene-expression subtypes using the nearest centroid method (Sorlie et al., 2003).

Array CGH

cDNA microarrays were obtained from the Stanford Functional Genomics Facility and included 39,632 human cDNAs, representing 22,488 mapped human genes [18,040 UniGene clusters (Schuler, 1997), together with 4,112 additional mapped ESTs not assigned UniGene IDs]. We performed array CGH according to our published protocols (Pollack et al., 1999; 2002). Briefly, 4 μg of genomic DNA from each tumor specimen was random-primer labeled with Cy5 and cohybridized to the microarray along with 4 μg of Cy3-labeled normal female leukocyte reference DNA from a single donor. Following overnight hybridization and washing, arrays were imaged using a GenePix 4000B scanner (Molecular Devices). Fluorescence ratios were extracted using

SpotReader software (Niles Scientific), and the data uploaded into the Stanford Microarray Database (Gollub et al., 2003) for storage, retrieval, and analysis. Note, array CGH analysis had been previously carried out on 29 of the 89 samples included in the current study, using smaller (~6,700 gene) cDNA microarrays (Pollack et al., 1999; 2002).

Data Analysis

Background-subtracted fluorescence ratios were normalized for each array by setting the average fluorescence ratio for all array elements equal to 1. Genes were considered reliably measured if the fluorescence intensity for the Cy3 reference channel was at least 1.4-fold above background. Map positions for arrayed cDNA clones were assigned using the NCBI genome assembly, accessed through the UCSC genome browser database (NCBI Build 35). For genes represented by multiple arrayed cDNAs, the average fluorescence ratio was used. DNA gains and losses were identified using the CLuster Along Chromosomes method (CLAC; <http://www-stat.stanford.edu/~wp57/CGH-Miner>) (Wang, 2004). Briefly, the CLAC algorithm builds a hierarchical cluster-style tree along each chromosome, such that neighboring genes with positive and negative ratios are separated into different clusters. DNA gains and losses are then called significant based on the height and width of clusters, and a false discovery rate is estimated by comparison to normal-normal hybridization data. To facilitate comparison with clinicopathological parameters, the ~22,000 mapped human genes were collapsed into 780 cytobands (boundaries defined by NCBI Build 35). For each specimen, cytobands displaying gain or loss were defined as those harboring at least two genes with gain or loss (respectively) called by CLAC, and the magnitude of the CNA was defined as the average fluorescence ratio for all genes residing within the cytoband. Cytobands displaying high-level DNA amplification, here scored only as present or absent, were defined as those called by CLAC and harboring at least two genes with tumor/normal ratios greater than 3. Significant associations between cytobands and clinicopathological parameters were identified using the Significance Analysis of Microarrays (SAM) method (Tusher et al., 2001), which is based on a modified t-statistic (for two-class comparisons) or Cox score (for survival analysis), and uses random permutations of class labels to estimate a false discovery rate (FDR). Kaplan–Meier survival analysis was performed using WinSTAT (R. Finch software).

RESULTS

Array CGH Profiling of Primary Breast Carcinomas

To explore the relationship between CNAs and clinicopathological parameters, we profiled CNAs by array CGH in a series of 89 locally advanced breast tumors [clinicopathological features summarized in Table S1 (Supplementary material for this article can be found at <http://www.interscience.wiley.com/jpages/1045-2257/suppmat>)]. In this cohort, high grade, ER negativity, and *TP53* mutation each showed the expected association with shorter overall survival ($P < 0.05$, Kaplan–Meier analysis).

Array CGH was carried out using cDNA microarrays representing ~22,000 human genes, thereby providing on average a mapping resolution of less than 70 kb (with mapping resolution paralleling gene density for this gene-based array platform). We observed numerous recurrent CNAs (summarized in Fig. S1), the spectrum of which was consistent with prior chromosome-based CGH studies (Knuutila et al., 2000). The most frequent aberrations included gains on 1q (35% of cases), 8q (35%), 11q (26%), and 16p (14%), and losses on 4q (58%), 5q (54%), 6q (43%), 8p (48%), and 14q (48%).

Specific CNAs Are Associated with Clinicopathological Parameters

To discover associations between CNAs and pathological features, for each tumor we first “collapsed” CNA calls (i.e., gain, loss, or no change) for the ~22,000 genes surveyed into CNA calls for the 780 cytoband loci represented by those genes (see Materials and Methods). The analysis of cytobands, which integrates information across neighboring genes, appeared more robust in preliminary analyses. We also considered loci with high-level DNA amplification [ratios >3 , corresponding to at least fivefold amplification (Pollack et al., 1999)] separately from those scored with gain, because they may be mechanistically distinct (Lengauer et al., 1998).

Overall frequencies of gain/loss varied among breast tumors with different clinicopathological features (Table 1). In particular, gains/losses were more frequent (borderline-significant) in ER-negative tumors ($P = 0.06$, Student's *t* test), and high-level DNA amplifications were more common (strong trend) in high-grade ($P = 0.08$) and *TP53*-mutant ($P = 0.13$) tumors.

To identify associations between specific CNAs and pathological parameters, we used the SAM method (Tusher et al., 2001), which corrects for multiple hypothesis (loci) testing in determining statistical significance (see Materials and Methods).

TABLE I. Average Total CNAs for Clinicopathological Parameters

	Grade		ER		TP53		Subtypes			
	Low	High	Pos	Neg	WT	Mut	Lum A	Lum B	ERBB2	Basal-like
Gain	40	48	40	59	41	45	46	45	31 ^a	62 ^b
Loss	54	60	52	80	54	59	59	46	39 ^a	89 ^b
Gain/loss	94	108	92	140	95	103	105	91	70 ^a	152 ^b
Amplification	13	18	14	18	13	17	10	24 ^{c,d}	8	13

^aP < 0.05 (vs. Lum-A or Basal-like).

^bP < 0.05 (vs. ERBB2).

^cP < 0.001 (vs. Lum-A or ERBB2).

^dP < 0.05 (vs. Basal-like).

Since CNAs are known to often span cytobands, the finding of two or more adjacent cytobands associated with a particular clinicopathological parameter (emphasized in the results below) further increased our confidence in the results (as being biologically sensible), which are summarized in Figure 1. We identified several CNAs associated with tumor grade, including loss at 3p14, 4q31-q35, and 5q13-q23 in high-grade tumors. We also found CNAs associated with ER status, where ER-negative tumors exhibited more frequent loss at 5q11-q35 and 12q14-23, and gain at 6p21-p25 and 7p12. Additionally, we identified loci associated with *TP53* mutation status, including gain at 1q21-q32 with wild-type *TP53* and loss at 5q14-q23 with mutant *TP53*.

To define associations between CNAs and clinical outcome, we performed survival analysis using the SAM method (Fig. 1). We identified loss at 6q22-q23 and 13q12-q13 to be associated with favorable outcome, while gain at 7p11-p14, 13q12-q13, and 21q22 were associated with unfavorable overall survival. We also identified high-level DNA amplifications associated with unfavorable outcome, including at 6q22, 15q23, 17q12-q21, and 20q13.

Distinct CNAs Are Associated with Gene-Expression Subtypes

Previous DNA microarray studies had defined clinically-relevant breast cancer subtypes based on distinct patterns of gene expression, including luminal epithelial-like subtypes A and B, a basal epithelial-like subtype, and an *ERBB2*-amplification associated subtype (Perou et al., 2000; Sorlie et al., 2001). Each of the breast tumors in our study had been previously assigned to a gene-expression subtype, by finding the best match of its expression profile to the published average profile of each of the subtypes (Sorlie et al., 2003). To determine whether different gene-expression subtypes were associated with distinct CNAs, we used the two-

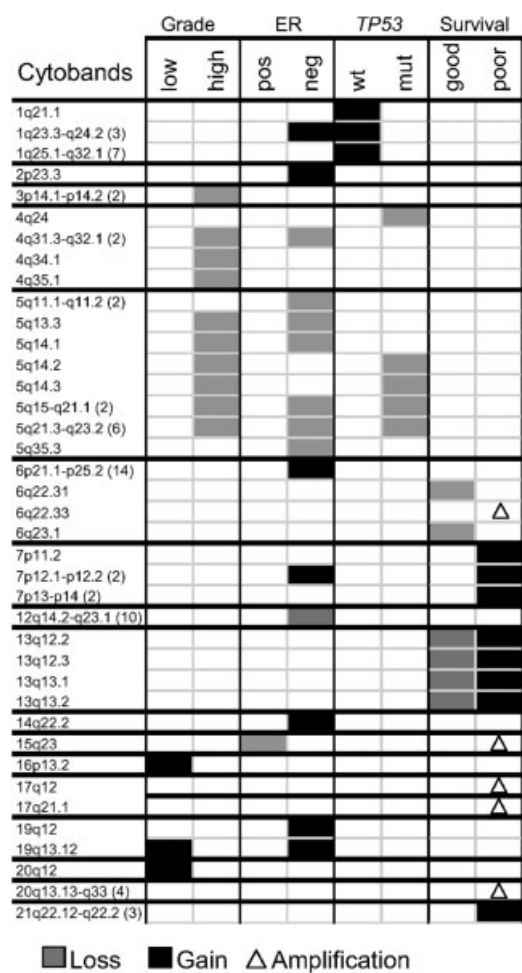


Figure 1. CNAs characterizing clinicopathological parameters. Summary of cytoband loci of DNA loss (gray fill), gain (black fill), and high-level amplification (triangle) significantly associated with tumor grade, ER status, and *TP53* status. For significant cytoband intervals, the number of cytobands is indicated in parentheses. False discovery rates (FDRs) for gain/loss (considered together) are 8% (Grade), 6% (ER status and *TP53* status), and 13% (survival), and for high-level amplification, 13% (survival).

class SAM method (i.e. one subtype versus all others); results are displayed in Figure 2 and summarized in Figure 3.

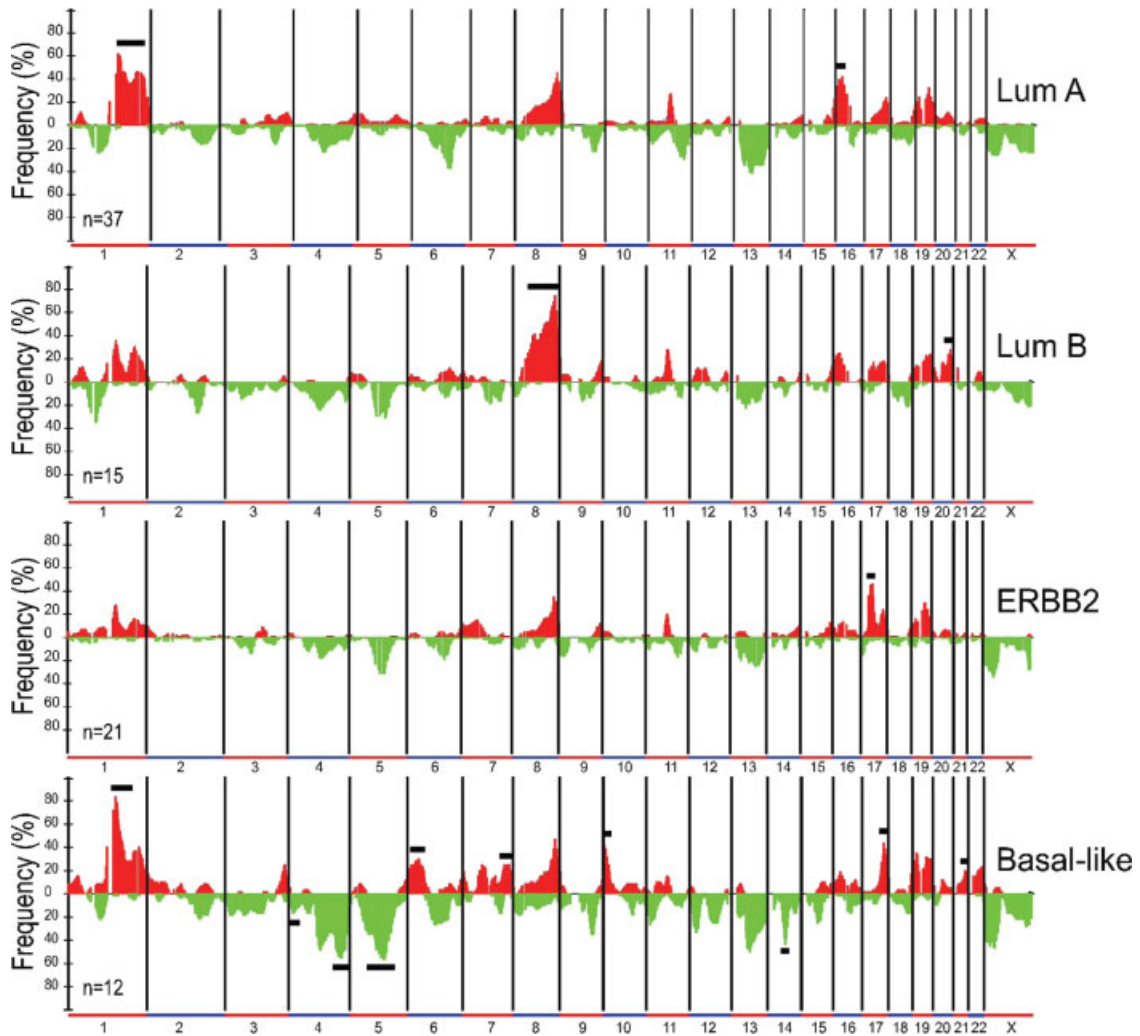


Figure 2. Spectra of CNAs among different gene-expression subtypes. Frequency plots summarizing distribution of CNAs (cytoband loci) in luminal-A, luminal-B, ERBB2-associated, and basal-like breast cancer subtypes. Red and green indicate gain and loss, respectively. Significant subtype-specific CNAs are indicated by black bar. Number (n) of specimens assigned to each subtype is indicated. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

As expected, ERBB2 subtype tumors, characterized by the amplification and resultant overexpression of *ERBB2* and its neighbors (Perou et al., 2000), exhibited more frequent amplification at 17q12-q21 (harboring *ERBB2*). Luminal-A group tumors were associated with gain at 1q12-q41 and 16p12-p13. Luminal B tumors exhibited more frequent loss at 3q12, gain at 8q11-q24 and 20q13, and high-level amplification at 7p22, 8q11-24, 19q13, and 20q13. CNAs associated with basal-like subtype tumors included loss at 3q12, 4p15-p32, 4q31-q35, 5q11-q31, and 14q22-q23, and gain at 1q12-q41, 6p12-p25, 7q22-q36, 10p12-p15, 17q25, and 21q22. Only 2 of the 87 cases (for which gene-expression data were available) were assigned to the normal breast-like subtype, precluding a meaningful analysis of this subgroup.

Interestingly, overall frequencies of gain/loss and high-level amplification also varied among breast tumor subtypes (Table 1). In particular, gains/losses were more frequent in basal-like tumors ($P = 0.02$; compared to ERBB2-associated tumors) and less common in ERBB2-associated tumors ($P = 0.02$; compared to luminal-A or basal-like tumors). High-level DNA amplifications were more frequent in luminal-B tumors, compared to luminal-A or ERBB2-associated tumors ($P < 0.001$), or to basal-like tumors ($P < 0.05$).

DISCUSSION

The main objective of our study was to explore associations between array CGH-detected CNAs and clinicopathological parameters in breast can-

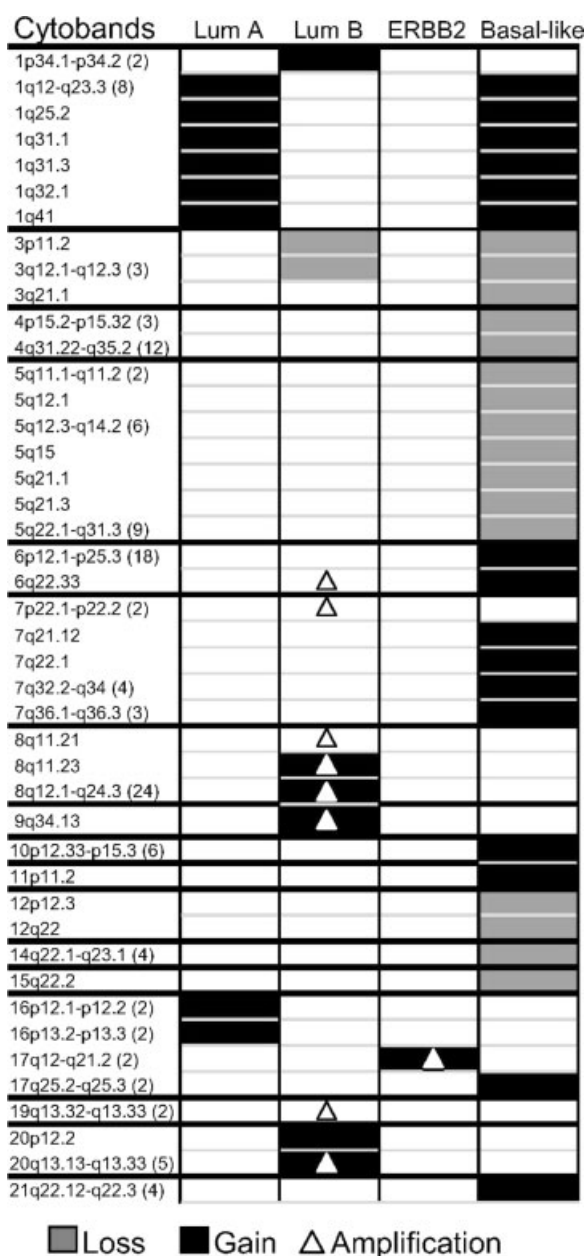


Figure 3. CNAs characterizing gene-expression subtypes. Summary of cytoband loci of DNA loss (gray fill), gain (black fill), or high-level amplification (triangle) significantly associated with gene-expression tumor subtypes. For significant cytoband intervals, the number of cytobands is indicated in parentheses. False discovery rates (FDRs) for gain/loss (considered together) are 3.6% (luminal-A), 2.5% (luminal-B), 7.8% (ERBB2-associated), and 1.6% (basal-like), and for high-level amplification, 3.1% (luminal-B) and 13.2% (ERBB2-associated).

cer. By SAM analysis of array CGH data, we discovered associations between CNAs at various cytoband loci and clinicopathological parameters, including tumor grade, ER status, *TP53* mutation, gene-expression tumor subtype, and overall survival. Although most of the observed associations are novel, the validity of our findings is supported

by the previous identification of some of the same associations by cytogenetic and molecular methods. In a chromosome-based CGH study of invasive breast tumors (Richard et al., 2000), loss on 5q was also found among the changes more frequent in high-grade tumors, as were gain on 2p and 6p and loss on 5q and 12q in ER-negative tumors. More recently, a BAC array-based CGH study of breast cancer (Loo et al., 2004) also identified gain on 1q and loss on 5q among the changes more frequent in ER-negative tumors. While in both these studies there were discordant loci as well, this may in part reflect differences in methodology or patient cohorts. Also in concordance with our study, a quantitative analysis of chromosome CGH data previously linked 5q15-q21 deletion to *TP53* mutation (Jain et al., 2001).

While CNAs associated with tumor grade, ER status, and *TP53* mutation have the potential to inform pathobiology, the identification of CNAs associated with clinical outcome might be more directly beneficial in improved patient prognostication. We identified several loci of CNA correlated with unfavorable overall survival, including the previously reported amplified loci 17q12 (*ERBB2*) (Ross and Fletcher, 1998) and 20q13 (*ZNF217*) (Tanner et al., 1995). Of the novel loci identified, the 7p11.2 gain harbors *EGFR*, encoding an oncogenic receptor tyrosine kinase and target for molecularly-directed therapies (Agrawal et al., 2005); although expression of *EGFR* has been evaluated there has been no consensus on its prognostic utility (Rampaul et al., 2005). Clearly additional studies on independent patient cohorts (and including earlier stages of breast cancer) are warranted to validate the prognostic utility of identified CNAs, and to assess whether they might provide an improvement over currently used prognostic factors.

An important finding of our study was the association between selected CNAs and the different gene-expression subtypes of breast cancer. While most associations are novel, a recent SNP array-based loss of heterozygosity (LOH) study of breast cancer also reported LOH on 4p and 5q to be associated with breast tumors with "basal-like" expression profiles (Wang et al., 2004). Another study also identified X-chromosomal abnormalities specific to basal-like breast tumors, but most reflected X-chromosome isodisomy with no net DNA gain or loss that would be detectable by CGH (Richardson et al., 2006). Notably, that the gene-expression subtypes in our study exhibited distinct spectra of CNAs suggests they develop along different genetic pathways. The constellation of genetic

alterations might directly specify the gene-expression phenotype, as 17q12 (*ERBB2*) amplification appears to do so for the *ERBB2*-amplification associated subtype. Alternatively, specific breast epithelial cell types or progenitors (e.g., with basal or luminal characteristics) might be more susceptible to transformation via a particular constellation of genetic changes.

Several of the subtype-specific CNAs are noteworthy. For example, gain on 8q was more frequent in luminal-B subtype tumors, which are a subset of ER-positive tumors characterized by higher proliferation rates and associated unfavorable outcomes. Among the genes on 8q, the *MYC* oncogene plays a key role in promoting cell proliferation (Adhikary and Eilers, 2005), and this association may in part explain the higher proliferation rates of these tumors. Within the 6p21-p25 gain associated with ER-negative and basal-like tumors reside several candidate oncogenes, including *DEK*, *E2F3*, *NOTCH4*, *PIM1*, and *CCND3*.

Another prominent finding was the loss on 5q associated with high grade, ER-negative, *TP53*-mutant, and basal-like breast tumors. Since this constellation of clinicopathological features often occurs together (Sorlie et al., 2001), it is not possible to infer the direct causal associations, e.g., whether a 5q TSG(s) directly cooperates with *TP53* mutation to effect tumorigenesis, or whether loss on 5q is more generally associated with aggressive tumor characteristics. Intriguingly, basal-like breast tumors share features with breast tumors arising in *BRCA1* mutation carriers, which are also typically high-grade, ER-negative, and HER2-negative (Johannsson et al., 1997), and exhibit in common gene-expression patterns (Sorlie et al., 2003) and expression of basal cytokeratins (Foulkes et al., 2003). In this regard, it is of interest that *BRCA1*-associated tumors also exhibit frequent loss of 5q (Tirkkonen et al., 1997), and that a *BRCA1*-modifier locus for hereditary breast cancer penetrance has been mapped to 5q (Nathanson et al., 2002). Several interesting tumor suppressor gene candidates map to 5q, including *RAD17*, *XRCC4*, *APC*, and *RAD50*.

It is also notable that the different breast cancer gene-expression subtypes exhibited different overall frequencies of CNA, with gain/loss more common in basal-like tumors and less so in *ERBB2*-associated tumors, and high-level amplification more frequent in luminal-B tumors. This finding suggests the possibility that different mechanisms of genomic instability underlie the pathogenesis of the different breast tumor subtypes, and may contribute to their distinct biological and clinical behaviors.

In summary, we have here defined array-CGH detected CNAs associated with distinct clinicopathological features, including clinical outcome and gene-expression subtypes. Our findings support the potential utility of CNA-based prognostication in breast cancer. The discovery of CNAs associated with clinicopathological features, including tumor grade, ER status, *TP53* mutation, gene-expression subtype, and overall survival, also provides a starting point to identify the underlying genes and further our understanding of the pathobiology of breast cancer.

ACKNOWLEDGMENTS

We thank the Stanford Functional Genomics Facility for producing high-quality DNA microarrays, and the Stanford Microarray Database group for providing outstanding database support. We also thank the members of the Pollack and Børresen-Dale labs for helpful comments and discussions.

REFERENCES

- Adhikary S, Eilers M. 2005. Transcriptional regulation and transformation by Myc proteins. *Nat Rev Mol Cell Biol* 6:635–645.
- Agrawal A, Gutteridge E, Gee JM, Nicholson RI, Robertson JF. 2005. Overview of tyrosine kinase inhibitors in clinical breast cancer. *Endocr Relat Cancer* 12(Suppl 1):S135–S144.
- Al-Kuraya K, Schraml P, Torhorst J, Tapia C, Zaharieva B, Novotny H, Spichtin H, Maurer R, Mirlacher M, Kochli O, Zuber M, Dietrich H, Mross F, Wilber K, Simon R, Sauter G. 2004. Prognostic relevance of gene amplifications and coamplifications in breast cancer. *Cancer Res* 64:8534–8540.
- Early Breast Cancer Trialists' Collaborative Group (EBCTCG). 2005. Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: An overview of the randomised trials. *Lancet* 365:1687–1717.
- Eden P, Ritz C, Rose C, Ferno M, Peterson C. 2004. "Good Old" clinical markers have similar power in breast cancer prognosis as microarray gene expression profilers. *Eur J Cancer* 40:1837–1841.
- Foulkes WD, Stefansson IM, Chappuis PO, Begin LR, Goffin JR, Wong N, Trudel M, Akslen LA. 2003. Germline *BRCA1* mutations and a basal epithelial phenotype in breast cancer. *J Natl Cancer Inst* 95:1482–1485.
- Geisler S, Borresen-Dale AL, Johnsen H, Aas T, Geisler J, Akslen LA, Anker G, Lonning PE. 2003. *TP53* gene mutations predict the response to neoadjuvant treatment with 5-fluorouracil and mitomycin in locally advanced breast cancer. *Clin Cancer Res* 9:5582–5588.
- Geisler S, Lonning PE, Aas T, Johnsen H, Fluge O, Haugen DF, Lillehaug JR, Akslen LA, Borresen-Dale AL. 2001. Influence of *TP53* gene alterations and c-erbB-2 expression on the response to treatment with doxorubicin in locally advanced breast cancer. *Cancer Res* 61:2505–2512.
- Gollub J, Ball CA, Binkley G, Demeter J, Finkelstein DB, Hebert JM, Hernandez-Boussard T, Jin H, Kaloper M, Matese JC, Schroeder M, Brown PO, Botstein D, Sherlock G. 2003. The Stanford Microarray Database: Data access and quality assessment tools. *Nucleic Acids Res* 31:94–96.
- Gradishar WJ. 2005. The future of breast cancer: The role of prognostic factors. *Breast Cancer Res Treat* 89 (Suppl 1):S17–S26.
- Isola JJ, Kallioniemi OP, Chu LW, Fuqua SA, Hilsenbeck SG, Osborne CK, Waldman FM. 1995. Genetic aberrations detected by comparative genomic hybridization predict outcome in node-negative breast cancer. *Am J Pathol* 147:905–911.
- Jain AN, Chin K, Borresen-Dale AL, Erikstein BK, Lonning PE, Kaarensen R, Gray JW. 2001. Quantitative analysis of chromosomal CGH in human breast tumors associates copy number abnormalities with p53 status and patient survival. *Proc Natl Acad Sci USA* 98:7952–7957.

- Jarvinen TA, Tanner M, Rantanen V, Barlund M, Borg A, Grenman S, Isola J. 2000. Amplification and deletion of topoisomerase II α associate with ErbB-2 amplification and affect sensitivity to topoisomerase II inhibitor doxorubicin in breast cancer. *Am J Pathol* 156:839–847.
- Johannsson OT, Idvall I, Anderson C, Borg A, Barkardottir RB, Egilsson V, Olsson H. 1997. Tumour biological features of BRCA1-induced breast and ovarian cancer. *Eur J Cancer* 33:362–371.
- Knuutila S, Autio K, Aalto Y. 2000. Online access to CGH data of DNA sequence copy number changes. *Am J Pathol* 157:689.
- Lengauer C, Kinzler KW, Vogelstein B. 1998. Genetic instabilities in human cancers. *Nature* 396:643–649.
- Loo LW, Grove DI, Williams EM, Neal CL, Cousens LA, Schubert EL, Holcomb IN, Massa HF, Glogovac J, Li CI, Malone KE, Daling JR, Delrow JJ, Trask BJ, Hsu L, Porter PL. 2004. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. *Cancer Res* 64:8541–8549.
- McGuire WL. 1978. Steroid receptors in human breast cancer. *Cancer Res* 38:4289–4291.
- Muss HB, Thor AD, Berry DA, Kute T, Liu ET, Koerner F, Cirrincione CT, Budman DR, Wood WC, Barcos M, Henderson IC. 1994. c-erbB-2 expression and response to adjuvant therapy in women with node-positive early breast cancer. *N Engl J Med* 330:1260–1266.
- Nathanson KL, Shugart YY, Omaruddin R, Szabo C, Goldgar D, Rebbeck TR, Weber BL. 2002. CGH-targeted linkage analysis reveals a possible BRCA1 modifier locus on chromosome 5q. *Hum Mol Genet* 11:1327–1332.
- Paik S, Shak S, Tang G, Kim C, Baker J, Cronin M, Bachner FL, Walker MG, Watson D, Park T, Hiller W, Fisher ER, Wickerham DL, Bryant J, Wolmark N. 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351:2817–2826.
- Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, Fluge O, Pergamenschikov A, Williams C, Zhu SX, Lonning PE, Borresen-Dale AL, Brown PO, Botstein D. 2000. Molecular portraits of human breast tumours. *Nature* 406:747–752.
- Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. 1999. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* 23:41–46.
- Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, Tibshirani R, Botstein D, Borresen-Dale AL, Brown PO. 2002. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* 99:12963–12968.
- Rampaul RS, Pinder SE, Nicholson RI, Gullick WJ, Robertson JF, Ellis IO. 2005. Clinical value of epidermal growth factor receptor expression in primary breast cancer. *Adv Anat Pathol* 12:271–273.
- Richard F, Pacyna-Gengelbach M, Schluns K, Fleige B, Winzer KJ, Szymas J, Dietel M, Petersen I, Schwendel A. 2000. Patterns of chromosomal imbalances in invasive breast cancer. *Int J Cancer* 89:305–310.
- Richardson AL, Wang ZC, De Nicolo A, Lu X, Brown M, Miron A, Liao X, Iglehart JD, Livingston DM, Ganesan S. 2006. X chromosomal abnormalities in basal-like human breast cancer. *Cancer Cell* 9:121–132.
- Ross JS, Fletcher JA. 1998. The HER-2/neu oncogene in breast cancer: Prognostic factor, predictive factor, and target for therapy. *Stem Cells* 16:413–428.
- Schuler GD. 1997. Pieces of the puzzle: Expressed sequence tags and the catalog of human genes. *J Mol Med* 75:694–698.
- Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE, Borresen-Dale AL. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* 98:10869–10874.
- Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A, Deng S, Johnsen H, Pesich R, Geisler S, Demeter J, Perou CM, Lonning PE, Brown PO, Borresen-Dale AL, Botstein D. 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci USA* 100:8418–8423.
- Subramaniam DS, Isaacs C. 2005. Utilizing prognostic and predictive factors in breast cancer. *Curr Treat Options Oncol* 6:147–159.
- Tanner MM, Tirkkonen M, Kallioniemi A, Holli K, Collins C, Kowbel D, Gray JW, Kallioniemi OP, Isola J. 1995. Amplification of chromosomal region 20q13 in invasive breast cancer: Prognostic implications. *Clin Cancer Res* 1:1455–1461.
- Tirkkonen M, Johannsson O, Agnarsson BA, Olsson H, Ingvarsson S, Karhu R, Tanner M, Isola J, Barkardottir RB, Borg A, Kallioniemi OP. 1997. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. *Cancer Res* 57:1222–1227.
- Tusher VG, Tibshirani R, Chu G. 2001. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98:5116–5121.
- van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415:530–536.
- Wang ZC, Lin M, Wei LJ, Li C, Miron A, Lodeiro G, Harris L, Ramaswamy S, Tanenbaum DM, Meyerson M, Iglehart JD, Richardson A. 2004. Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancers. *Cancer Res* 64:64–71.
- Yeon CH, Pegram MD. 2005. Anti-erbB-2 antibody trastuzumab in the treatment of HER2-amplified breast cancer. *Invest New Drugs* 23:391–409.