Distinct Genomic Profiles in Hereditary Breast Tumors Identified by Array-Based Comparative Genomic Hybridization

Göran Jönsson,[1] Tara L. Naylor,[5] Johan Vallon-Christersson,[1] Johan Staaf,[1] Jia Huang,[5] M. Renee Ward,[5] Joel D. Greshock,[5] Lena Luts,[4] Håkan Olsson,[1] Nazneen Rahman,[6] Michael Stratton,[6] Markus Ringnér,[3] Åke Borg,[1,2] and Barbara L. Weber[5]

[1]Department of Oncology, University Hospital; [2]Lund Strategic Research Center for Stem Cell Biology and Cell Therapy and [3]Department of Theoretical Physics, Lund University; and [4]Department of Pathology, University Hospital, Lund, Sweden; [5]Abramson Family Cancer Research Institute, University of Pennsylvania, Philadelphia, Pennsylvania; and [6]Section of Cancer Genetics, Institute of Cancer Research, Sutton, Surrey, United Kingdom

Abstract

Mutations in *BRCA1* and *BRCA2* account for a significant proportion of hereditary breast cancers. Earlier studies have shown that inherited and sporadic tumors progress along different somatic genetic pathways and that global gene expression profiles distinguish between these groups. To determine whether genomic profiles similarly discriminate among BRCA1, BRCA2, and sporadic tumors, we established DNA copy number profiles using comparative genomic hybridization to BAC-clone microarrays providing <1 Mb resolution. Tumor DNA was obtained from BRCA1 (n = 14) and BRCA2 (n = 12) mutation carriers, as well as sporadic cases (n = 26). Overall, BRCA1 tumors had a higher frequency of copy number alterations than sporadic breast cancers (P = 0.00078). In particular, frequent losses on 4p, 4q, and 5q in BRCA1 tumors and frequent gains on 7p and 17q24 in BRCA2 tumors distinguish these from sporadic tumors. Distinct amplicons at 3q27.1-q27.3 were identified in BRCA1 tumors and at 17q23.3-q24.2 in BRCA2 tumors. A homozygous deletion on 5q12.1 was found in a BRCA1 tumor. Using a set of 169 BAC clones that detect significantly (P < 0.001) different frequencies of copy number changes in inherited and sporadic tumors, these could be discriminated into separate groups using hierarchical clustering. By comparing DNA copy number and RNA expression for genes in these regions, several candidate genes affected by up- or down-regulation were identified. Moreover, using support vector machines, we correctly classified BRCA1 and BRCA2 tumors (P < 0.0000004 and 0.00005, respectively). Further validation may prove this tumor classifier to be useful for selecting familial breast cancer cases for further mutation screening, particularly, as these data can be obtained using archival tissue. (Cancer Res 2005; 65(17): 7612-21)

Introduction

Germ line mutations in the two major breast cancer susceptibility genes, *BRCA1* and *BRCA2*, confer a highly elevated risk of breast and ovarian cancer and account for a significant proportion of inherited breast cancer (1). However, many familial breast cancer cases cannot be attributed to *BRCA1* and *BRCA2*, suggesting a role of additional predisposing genes, although technical limitations and the complexity of BRCA gene regulation and mutation spectrum can probably explain why some disease-causing mutations are missed (2). *BRCA1* and *BRCA2* function as classic tumor suppressor genes with frequent loss of the wild-type allele in tumors of mutation carriers. The BRCA1 protein has been implicated in a broad range of cellular functions, including repair of double-strand breaks by homologous recombination, cell cycle checkpoint control, chromatin remodeling, and transcriptional regulation. The role of BRCA2 is more restricted to DNA recombination and repair processes, being particular important in RAD51 regulation. It is thought that part of the tumor suppressor function of BRCA1 and BRCA2 is attributed to this genome caretaker activity (3).

Earlier studies have suggested different somatic genetic pathways in progression of inherited and sporadic tumors, and several histopathologic and clinical features differ among BRCA1, BRCA2, and sporadic breast cancers. For instance, BRCA1 tumors are often estrogen and progesterone receptor (ER/PR) negative, whereas BRCA2 tumors predominantly are ER and PR positive (4–7). BRCA1 tumors are primarily of high histologic grade and manifest high lymphocyte infiltration and continuous "pushing" tumor margins. BRCA2 tumors are more heterogeneous but have significantly less tubule formation compared with sporadic tumors, which account for an overall higher histologic grade (7). Using global gene expression profiling, we have previously identified a set of genes to distinguish among BRCA1, BRCA2, and sporadic tumors (8). In a parallel fashion, we used conventional metaphase comparative genomic hybridization (CGH) to characterize chromosomal aberrations in hereditary breast tumors (9–11). We found a higher frequency of copy number alterations among BRCA1 and BRCA2 tumors when compared with sporadic cases. BRCA1 tumors harbor frequent loss at 2q, 4p, 4q, 5q, and 12q, whereas BRCA2 tumors are characterized by a higher frequency of 6q and 13q losses as well as gains on 17q22-24 and 20q13. Others have also used metaphase CGH analysis to classify BRCA1 tumors based on patterns of genomic alterations and suggested gain of 3q and loss of 3p and 5q to distinguish BRCA1 from sporadic tumors (12).

Taken together, these findings imply that tumor profiling can be useful as diagnostics tools and encouraged us to extend our previous studies with an array-based CGH technique, providing higher consistency and resolution, to build classifiers based on genomic aberration patterns in tumor tissue. In the familial cancer clinic, candidates for *BRCA1* and *BRCA2* mutation screening are selected based on family history, age at onset, bilateral disease, and cases of ovarian cancers in the family. Because screening of these genes is laborious and costly, more precise criteria for selection and

Requests for reprints: Åke Borg, Department of Oncology, Lund University, SE-221 85 Lund, Sweden, SE 22185. Phone: 46-46-177502; Fax: 46-46-147327; E-mail: ake.borg@onk.lu.se.

doi:10.1158/0008-5472.CAN-05-0570

diagnosis is needed (e.g., a rationale corresponding to the tests for microsatellite instability and mismatch repair protein expression in diagnosis of hereditary nonpolyposis colon cancer; ref. 13). In the present study, we established DNA copy number profiles for BRCA1, BRCA2, and sporadic breast tumors using a microarray containing ~5,000 individual BAC clones providing an average resolution of 1 Mbp. Furthermore, by comparing DNA copy number and gene expression in 11 BRCA1/2 tumors, we found novel genes affected by genomic aberrations in hereditary breast cancer. Our results form the basis for a new classification system of inherited breast cancer and reveal new molecular targets in tumor progression of inherited tumors. The approach of using array CGH in diagnostics may be more advantageous than gene expression profiling because paraffin-embedded material can be used and thus more feasible in the clinical setting.

Materials and Methods

Patients and tumors. Freshly frozen breast tumor tissue was obtained from the Southern Sweden Breast Cancer Group's tissue bank at the Department of Oncology, Lund University Hospital and from the HWP and Delaware Memorial Hospital. Breast cancer patients from which BRCA1 and BRCA2 tumors were derived had been screened for germ line mutations in the *BRCA1* and *BRCA2* genes according to standard techniques (6). From Lund University Hospital, eight frozen BRCA1 tumors were obtained from six patients, including six primary tumors, a local recurrence, and a regional metastasis. Seven primary BRCA2 tumors were obtained from six patients, two of whom were males. In addition, five BRCA1 and five BRCA2 tumors were obtained from paraffin embedded material at the University of Pennsylvania. Except for one deleterious missense BRCA1 mutation, all mutations were of frameshift or nonsense type resulting in protein truncation, most of them located in the large exon 11 of either gene. Age at diagnosis, steroid receptor status, and histologic grade were also recorded (Table 1). Sporadic breast cancer cases constitute an unselected set of 26 tumors from patients without familial history of breast cancer and having variable receptor status and histologic grade (Table 1). The study was approved by the local ethical committees and the Institutional Review Board.

BAC array platforms. BAC clones included in the "1 Mb–array" platform were collected from several sources, including ~3,600 fluorescence *in situ* hybridization and sequence-tagged site (STS)–mapped BAC clones obtained from the RPCI-11 and CalTech A and B libraries, and ~1,400 from the end-sequenced collection at The Institute for Genomic Research. Genome coverage was assessed by the UCSC human genome assembly Build 34 (http://www.genome.ucsc.edu), revealing an average resolution of ~1 Mbp, with no gaps >2 Mbp. BAC DNA was amplified using degenerate oligonucleotide PCR (DOP) primers. Arrays were constructed with at least two replicates per clone on each slide using a Molecular Dynamics Gen3 spotter and a spotting solution of 50% DMSO (14). Subsequently, microarrays with tiling coverage and high resolution of chromosome 5 (>2,600 clones) were produced from BAC clones included in the 32K set of CHORI BACPAC Resources (http://bacpac.chori.org/genomicRearrays.php). Here, DOP-PCR products were obtained from 6 ng BAC DNA template and purified using filter based 96-well plates (PALL), dried, and resuspended in 50% DMSO to a concentration of 500 to 1,000 ng/μL. Arrays were printed using a MicroGrid II spotter (Biorobotics, Cambridge, MA) as described in detail elsewhere.[7]

DNA extraction and array comparative genomic hybridization hybridization. Genomic DNA was extracted from frozen tissue sections using a proteinase K treatment followed by phenol chloroform purification. DNA extraction from paraffin-embedded tumors was done using a xylene treatment before proteinase K digestion and phenol chloroform purification. For all samples, 1 μg of genomic DNA was labeled according to published protocols (14) using a random labeling kit (Invitrogen Life Technologies, Carlsbad, CA). Tumor DNA and male reference DNA was differentially labeled, pooled, mixed with human COT-1 DNA, dried down, and resuspended in a formamide-based buffer (14). The hybridization reactions were applied to arrays and the arrays were incubated under coverslips for 48 to 72 hours at 37°C. Slides were washed according to published protocols (15) and scanned using an Axon 4000A scanner (Axon Instruments, Wheatherford, TX). All experiments were done with "dye-swap" to account for variations in dye labeling efficiency and fluorescence emission and all BACs are spotted on the array in duplicate; thus, each data point is represented by four replicates.

RNA extraction and gene expression analysis. RNA was extracted from freshly frozen tumor tissue using Trizol reagent (Invitrogen Life Technologies) followed by RNeasy Midi kit (Qiagen, Chatsworth, CA). RNA quality was assessed using the RNA 6000 Nano LabChip Kit for Agilent 2100 bioanalyzer (Agilent Technologies, Palo Alto, CA) and concentration was determined using a NanoDrop Spectrophotometer (NanoDrop Technologies, Wilmington, DE). Arrays were produced by the SWEGENE DNA Microarray Resource Centre, Department of Oncology at Lund University. Human Genome Oligo Set Version 2.1 (containing 21,329 70-mer probes) and Human Genome Oligo Set Version 2.1 Upgrade (containing 5,462 70-mer probes) were obtained from Operon Biotechnologies, Inc. (Huntsville, AL). cDNA synthesis, labeling, and cleanup were overall done according to manufacturers' instructions using Pronto! Plus System 6 (Corning, Inc., Corning, NY) with 5 μg of total RNA as starting material. As reference RNA a commercial pool of cell lines (Stratagene Universal Reference, La Jolla, CA) was used. Arrays were scanned at two wavelengths using an Agilent G2505A DNA microarray scanner (Agilent Technologies).

Data analysis. Identification of individual spots on scanned arrays was done with Gene Pix Pro 4.0 (Axon Instruments, Union City, CA), and the quantified data matrix was loaded into Bio Array Software Environment BASE (16). Background-correction of Cy3 and Cy5 intensities was calculated using the median feature and median local background intensities provided in the quantified data matrix. Within arrays, intensity ratios for individual probes were calculated as background corrected intensity of tumor sample divided by background corrected intensity of reference sample. For the gene expression data spots for which background-corrected Cy3 or Cy5 intensities were <1 or >65,000 were removed from further analysis. In addition, spots that had been flagged during image analysis, either by the GenePix software or after manual inspection, were removed. Data within individual arrays were then adjusted by an implementation of the intensity-dependent normalization method based on a lowess fit, as described by Yang et al. (17). A filter was applied on remaining spots excluding all spots having signal-to-noise ratio of <2 and intensity less than background intensity. Next, data were mean centered using an iterative process so that all row-wise and column-wise mean $\log_2$ ratio values are close to 0.

For the CGH data, a signal-to-noise filter of ≥ 2 for the tumor channel and ≥ 1.5 for the reference channel was applied to the data and spots that failed to pass these criterions were excluded from further analysis and regarded as missing values. Average intensity ratios were calculated for spots present in both dye-swap hybridizations after filters had been applied and used in subsequent analysis. Furthermore, BAC clones with >10 missing values across the 52 tumors were excluded from further analysis (81% presence required). The filtered data was, for each array separately, centralized to a median ratio of unity. All filtering and normalization was done in BASE (16). The CGH-Plotter software, as an R (http://www.r-project.org) implementation in BASE, was used to identify regions of gains and losses (18). In CGH-Plotter, a moving median sliding window of three clones and a constant variable value of 15 were applied. Cutoff ratios for gains and losses were set to 1.15 and 0.87, respectively, corresponding to $\log_2$ (ratio) of ± 0.2. All clones were designated gained, lost, or not changed giving us a ternary scale. Using the values given from CGH-Plotter, we calculated the percentage of altered clones, the number of altered regions in each tumor, and the mean value within each group. We used a two-sided Mann-Whitney test to calculate significant differences in copy number frequency between

[7] Jönsson et al., submitted for publication 2005.

Table 1. Clinical and tumor characteristics of breast cancer cases

Sample ID	Tumor type	F/P	Age at diagnosis	ER status*	Histologic grade	Mutation Nt change	Mutation aa change
BRCA1 mutation positive							
Ca 11808	Primary tumor	F	39	−	NA	3829delT	Stop 1263
Ca 12224 †	Primary tumor	F	39	−	3	1806C>T	Q563X
Ca 12530 †	Local recurrence	F	39	−	3	1806C>T	Q563X
Ca 13494	Primary tumor	F	62	−	3	3438G>T	E1107X
Ca 13928 ‡	Primary tumor	F	53	−	3	1806C>T	Q563X
Ca 13996 ‡	Lymph node metastasis	F	53	−	NA	1806C>T	Q563X
Ca 14007	Primary tumor	F	34	−	3	3171ins5	Stop 1025
Ca 15504	Primary tumor	F	74	−	3	3040T>A	L1013X
F1070-1	Primary tumor	P	32	−	NA	3053T>G	Y978X
F2039-1	Primary tumor	P	51	NA	NA	3347delAG	Stop 1084
Ca 9002	Primary tumor	F	39	−	NA	1235G>A	W372X
F700-9	Primary tumor	P	39	NA	NA	4286delTG	Stop 1402
F27-40	Primary tumor	P	50	−	NA	309T>G	C64G
F205-7	Primary tumor	P	45	NA	NA	5382insC	Stop 1829
BRCA2 mutation positive							
Ca 9167	Primary tumor, male	F	69	++	NA	6503delTT	Stop 2098
Ca 10588	Primary tumor	F	75	+++	NA	4486delG	Stop 1477
Ca 11506§	Primary tumor	F	59	+++	2	6293C>G	S2022X
Ca 13816	Primary tumor	F	83	+	3	3058A>T	K944X
Ca 14616	Primary tumor, male	F	86	+++	3	4486delG	Stop 1477
Ca 14767	Primary tumor	F	58	+++	2	4486delG	Stop 1477
Ca 15243§	Primary tumor	F	63	++	2	6293C>G	S2022X
F572	Primary tumor	P	39	+	3	4392delTT	Stop 1401
F644-302	Primary tumor	P	54	+	NA	7985G>A	W2586X
F644-301	Primary tumor	P	45	+	3	7985G>A	W2586X
F714-301	Primary tumor	P	36	+	3	6503delTT	Stop 2098
F795	Primary tumor	P	39	+	3	8219delT	Stop 2672
Sporadic							
TB004	Primary tumor	F	41	+	3	—	—
TB007	Primary tumor	F	36	+	NA	—	—
TB017	Primary tumor	F	34	−	3	—	—
TB033	Primary tumor	F	62	+	3	—	—
TB038	Primary tumor	F	52	+	2	—	—
TB071	Primary tumor	F	94	+	2	—	—
TB275	Primary tumor	F	97	−	3	—	—
TB277	Primary tumor	F	45	−	3	—	—
TB313	Primary tumor	F	85	−	3	—	—
TB346	Primary tumor	F	47	+	2	—	—
TB369	Primary tumor	F	36	+	2	—	—
TB316	Primary tumor	F	43	−	3	—	—
TB482	Primary tumor	F	57	NA	3	—	—
TB088	Primary tumor	F	49	+	3	—	—
TB064	Primary tumor	F	67	+	NA	—	—
TB268	Primary tumor	F	38	−	3	—	—
TB318	Primary tumor	F	68	+	2	—	—
TB352	Primary tumor	F	87	+	2	—	—
TB348	Primary tumor	F	80	+	2	—	—
TB063	Primary tumor	F	51	NA	NA	—	—
TB001	Primary tumor	F	51	−	3	—	—
TB046	Primary tumor	F	52	NA	NA	—	—
TB026	Primary tumor	F	70	+	2	—	—
TB016	Primary tumor	F	74	+	2	—	—
TB344	Primary tumor	F	38	+	3	—	—
TB476	Primary tumor	F	74	+	2	—	—

Abbreviations: F, freshly frozen tissue; P, paraffin-embedded formalin-fixed tissue; NA, not available; Nt, nucleotide; aa, amino acids.
*ER status in fmol/mg protein: <25 (−), 25-100 (+), 100-300 (++), >300 (+++) fmol/mg protein.
†Ca 12530 is a local recurrence of Ca 12224.
‡Ca 13996 is a lymph node metastasis of Ca 13928.
§Ca 11506 and Ca 15243 are bilateral primary tumors derived from the same patient.

groups. To identify clones discriminating the three tumor groups, we did a Kruskal-Wallis test (19) as implemented in the Statistics Toolbox in Matlab (Mathworks, Natick, MA), on the ternary data. To determine the correlation coefficients, we used a Pearson correlation estimate using the results obtained from the CGH-Plotter software. For the Kruskal-Wallis test, only primary tumors were included in the analysis. Hierarchical clustering was done in BASE using a bottom-up approach with center of mass linkage and Pearson correlation coefficient distance measure.

Correlation of gene expression and copy number data was determined as follows. After normalization and quality filters had been applied, we selected oligonucleotide probes (henceforth referred to as reporter) for which expression data was present in 80% of 11 tumor samples. The data set was then filtered for variance across all hybridizations for each reporter individually, removing reporters having a $\log_2$ ratio SD of <0.5. To map reporters to genomic location of corresponding transcript target, each reporter sequence was blasted against Genbank sequences in Unigene build 176. We used 100% sequence similarity as cutoff criteria for associating a reporter with accession sequences. The ACID database (http://bioinfo.thep.lu.se/acid.html; ref. 20) was used to extract full gene information for accession numbers associated to each reporter through the blast analysis (Unigene build 176). Each gene was then mapped to a chromosomal position based on its gene symbol trough the UCSC Genome Browser (Hg16). For each BAC clone present in the BAC data set, we extended its stop position to cover the full base pair range between its adjacent BAC neighbors. The new BAC mapping was then connected to the previously mapped oligonucleotide reporters, creating reporter-BAC pairs. Standard Pearson correlation (henceforth referred to as correlation) for each reporter-BAC pair, in total 7,410, was then calculated to find pairs where gene expression and genomic copy number behaved concordantly. Next, sample assignments for the copy number profiles were randomly permuted and correlations recalculated keeping the expression data and reporter-BAC pair mapping intact. This procedure was repeated 10,000 times. The permutations allowed us to calculate P values for the different correlation bins (each bin represent a correlation range of 0.05). A P value cutoff of 0.01 corresponded to a correlation cut off of 0.75; that is, by chance, one would expect 1% of all pairs to have a correlation of $\geq$0.75. To determine whether the gene is "overexpressed" or "underexpressed," we used an SD method based on using a sliding window through an M-A plot. The method uses a defined window size for calculation of mean and SD for spots within the window. Each spot is assigned a window and is then compared with mean ratio and SD to see how many SD the spot is from the mean (21).

For classification of BRCA and sporadic breast tumors, we used supervised support vector machines in a leave-one-out iterative validation approach included in the TIGR MultiExperimentViewer (TMEV) software (22, 23). In the classification, whole genome copy number profiles obtained from the array CGH experiment were used. In the leave-one-out procedure, one tumor is left out and the remaining samples are used to construct a classifier that is validated on the left-out sample. The procedure is iterated for each sample, such that every sample gets a classification. A linear kernel with default values for the variables, except the diagonal factor set to 4, was used within the TMEV software (22). The significance of classification results was calculated using Fisher's exact test on 2 $\times$ 2 contingency tables using R (http://www.r-project.org).

Results

Array-based CGH was used in genomic profiling of 14 BRCA1, 12 BRCA2 mutation-positive tumors, and 26 sporadic breast tumors. The arrays comprised $\sim$5,000 individual BAC clones evenly distributed over the genome, corresponding to an average resolution of $\sim$1 Mbp. Gains, losses, and high level amplifications were readily detected as shown by visualization in overview plots (Fig. 1*A-B*). The sizes of regions manifesting gain or loss were determined using the distance between the two closest, not altered clones and ranged from 1 Mbp to whole chromosome gains/losses. Overall, we found a higher frequency of copy number changes in BRCA1 (47.8 $\pm$ 12.7% altered clones) than in BRCA2 (34.9 $\pm$ 10.5% altered clones, P = 0.029) and sporadic tumors (28.1 $\pm$ 14.7% altered clones, P = 0.00078). The difference between BRCA2 and sporadic tumors was not significant (P = 0.18). In addition, we found a higher number of distinct altered regions in BRCA1 (91.1 $\pm$ 20.8) than in BRCA2 (79.4 $\pm$ 45.7, P = 0.064) and sporadic tumors (72.6 $\pm$ 32.1, P = 0.091). There was no apparent association between type or location of the *BRCA1* or *BRCA2* mutations and copy number changes, although it should be pointed out that all mutations but

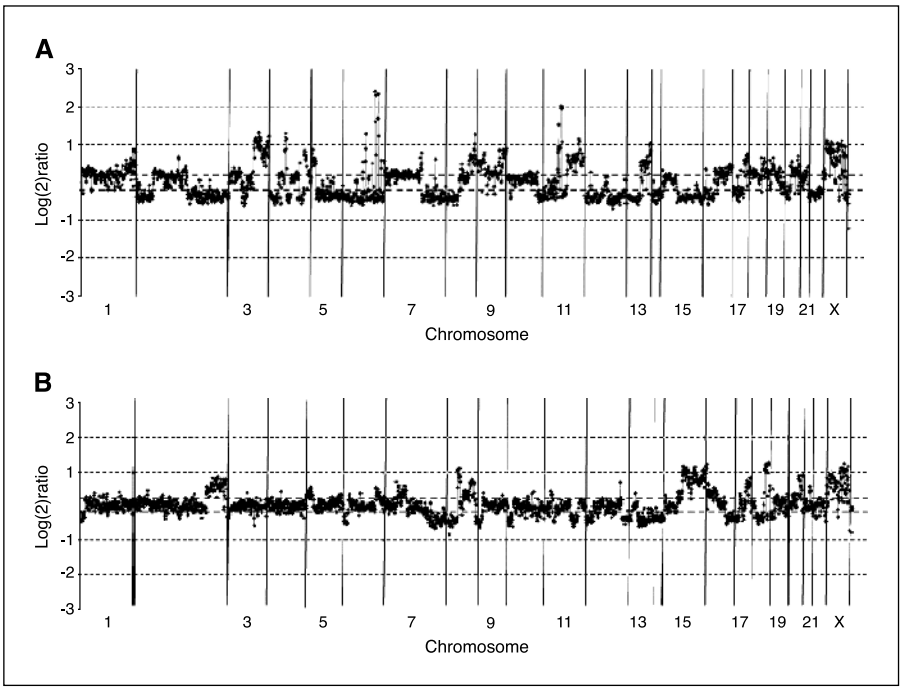

Figure 1. *A,* genome-wide copy number profile for a BRCA1 tumor (Ca 13996). Characteristic alterations for BRCA1 tumors are losses on 4p, 4q, and 5q and gains on 3q. *B,* genome-wide copy number profile for a BRCA2 tumor (Ca 10588). Characteristic for BRCA2 tumors are amplifications on 17q22-24 and 20q13, also deletions on 11q and 13q are frequent events.

one were of protein truncating type and clustered within the central regions of the genes. There was, however, a concordance in genomic profiles of tumors occurring in the same patient (see below). The two male BRCA2 tumors had an overall genomic profile similar to BRCA2 tumors from female patients (see below).

Regions with gains and amplifications. Overlapping regions with frequent (>75% of samples) gains in BRCA1 tumors included 1q42.12-q42.13, 3q26.32-q26.33, 3q27.1-q27.33, 7q36.1-q36.3, 8q24.23-q24.3, 10p15.3, and 10p15.1-p14; each region comprising 0.6 to 6.7 Mbp genomic DNA and multiple candidate genes. Some of these regions such as 1q and 8q also showed recurrent gains (>50% of samples) in sporadic breast tumors. Ten BRCA1 tumors showed gain on 3q27.1-q27.33, comprising ~3.5 Mbp and including >50 genes. Another region with frequent (83%) gain in BRCA1 tumors was found at 7q36.1-q36.3, including the cyclin-dependent kinase 5 (*CDK5*) gene. Narrow amplifications were detected on 6q, including the *MYB* oncogene and *HBS1L* gene. In BRCA2 tumors, overlapping regions with frequent (>75%) gains included 1q32.1-q41, 8q21.13, 8q22.1-q24.3, 17q23.3-q24.2, 17q25.1-qter, and 20q12-q13.12. The 17q23.3-q24.2 and 20q13.13 amplicons were shown to be most specific for BRCA2 tumors. A narrow amplification peak on 12q14.2-q21.1 including the *MDM2* oncogene was revealed in one BRCA2 tumor. In addition, corroborating earlier findings, BRCA2 tumors exhibited recurrent gains on 20q, including high frequency of gain on 20q13.13. Among other frequently altered genes in breast cancer, the *ERBB2* gene on 17q12 was amplified in seven (27%) of sporadic tumors but in none of the BRCA1 or BRCA2 tumors. The *CCND1* gene on 11q13.3 was amplified in only one of the sporadic tumors, whereas another sporadic tumor (TB088) manifested discrete amplification of the *EMSY* gene, located 6 Mbp distal to *CCND1*.

Regions with deletions. Regions showing recurrent loss in BRCA1 tumors (>75% of samples) included 4p15.32-p14, 4q31.3, 4q32.1-q34, 5q11.2-q23.3, 8pter-p12, 13q13.3-q21.32, 15q12-13.1, 15q15.3-q21.1, and 17p13.2-p12. Some of these regions (4q, 8p, and 13q) were commonly deleted (35-61%) also in sporadic tumors, although not to the same extent as in the inherited breast cancers. In addition to frequent heterozygous losses on chromosome 5q, a putative homozygous deletion on 5q12.1 was detected in a primary tumor and a local recurrence obtained from a BRCA1 mutation carrier. This deletion comprised two adjacent BAC clones present on the 1 Mb array resulting in $\log_2$ (ratios) below -1.5. The size of the deleted region was estimated to 2.8 Mbp and included nine known genes. Analysis of these samples using a BAC array with complete coverage and high-resolution (in average 50 kbp) of chromosome 5, narrowed down this homozygous deletion to a 700-kbp region including only three genes (*PART*, *FKSG52*, and *DEPDC1B*) and the 3′ flanking regions of *PDE4D* (Fig. 2*A*). Primers for STSs mapped to the putative homozygous deletion were used for PCR analysis, which confirmed the homozygote loss in both cases (Fig. 2*B*). In 72% of the BRCA1 tumors investigated, a heterozygous loss of 5q12.1 was present, suggesting that this region harbors genes of importance for BRCA1 tumor development. Regions showing recurrent loss (>75%) in BRCA2 tumors included 8p23.3-p21.2, 11q14.3-q21, 11q24.2-q25, 13q13.3-q14.13, 16q22.2-23.2, and 17p13.1-p12. Of these regions, losses on 11q were the most specific alterations for BRCA2 tumors. The region comprising *BRCA1* and adjacent clones on 17q21.31 was deleted in six BRCA1, three BRCA2, and two sporadic tumors, whereas the *BRCA2* gene on 13q13.1 was deleted in nine BRCA1, eight BRCA2, and 12 sporadic tumors.

Concordance in genomic profiles of tumors arising in predisposed individuals. In two BRCA1 germ line mutation carriers, samples were available from both the primary tumor and a local recurrence or a lymph node metastasis, respectively. As expected, the recurrent /metastatic tumors had genomic profiles very similar to their primary tumors (correlation coefficient, 0.84-0.93). More intriguingly, two tumors from a BRCA2 germ line mutation carrier also showed a striking concordance with regard to their genomic profiles (correlation coefficient, 0.79). Both tumors were considered as primary cancers by histopathologic criteria, derived from contralateral breasts and diagnosed 5 years apart, the first (Ca 11506) in the right and the second (Ca 15243) in the left breast, both including ductal carcinoma *in situ* (DCIS) and invasive ductal cancer.

Correlation of gene expression and DNA copy number. To determine the effect of DNA alterations on gene expression, we did genome wide expression analysis on four BRCA2- and seven BRCA1-associated breast tumors from which array CGH results had been obtained. By matching BAC and oligonucleotide probes (in total, 7,410 BAC-oligonucleotide probe pairs), a significant correlation between genomic alterations and gene expression was found for 746 genes ($P < 0.01$), including several relevant candidate genes such as *RAD50*, *HBS1L*, *BCAS4*, *CDK5*, and *TM4SF1*. Moreover, 30 of the 746 genes were mapped to the chromosomal regions (see below) discriminating BRCA1, BRCA2, and sporadic tumors (Fig. 3).

Regions discriminating BRCA1, BRCA2, and sporadic breast tumors. In addition to the descriptive representation of the individual gains and losses for BRCA1 and BRCA2 tumors, we sought to identify chromosomal regions discriminating hereditary from sporadic breast tumors. Of the ~5,000 BAC clones on our arrays, 3,591 passed our missing value requirements in these experiments. First, we used CGH-Plotter to assign a discrete state (gained, lost, or unchanged) to each copy number estimate for these 3,591 clones. Second, unsupervised hierarchical clustering using 3,591 BAC clones was conducted where a trend of subgrouping was apparent (Fig. 4*A*). Next, we used a Kruskal-Wallis test and supervised model to identify BAC clones that discriminate among the three tumor groups, including only primary tumors in the analysis. We found 169 significant BAC clones ($P < 0.001$), whereas only four clones would be expected by chance alone. Several regions confirmed findings from other studies using techniques with lower resolution, including large regions on chromosomes 4p, 4q, 5q, and 12q we previously showed as being frequently deleted in BRCA1 tumors. In addition, deletions on chromosome 15q were distinct features of BRCA1 tumors as well as gains on chromosome 3q and 7q. Moreover, gains on 7p and 17q24 and losses on 8p were found to be more prevalent in BRCA2 cases. Using hierarchical clustering, we could discriminate BRCA1, BRCA2, and sporadic tumors based on the 169 BAC clones (Fig. 4*B*). In the cluster analysis, 14 of 14 BRCA1 tumors were tightly clustered and separated from sporadic cases. BRCA2 tumors showed a somewhat higher similarity with the sporadic breast tumors but still displayed a distinct genomic profile of their own. Eight of 12 BRCA2 tumors were grouped in a subcluster whereas one outlier case clustered together with the sporadic cases and three showed high similarity to the BRCA1 tumors. In addition, two sporadic tumors (TB088 and TB026) showed high similarity to BRCA1 tumors.

Leave-one-out iterative classification method. We used a leave-one-out iterative classification approach to determine

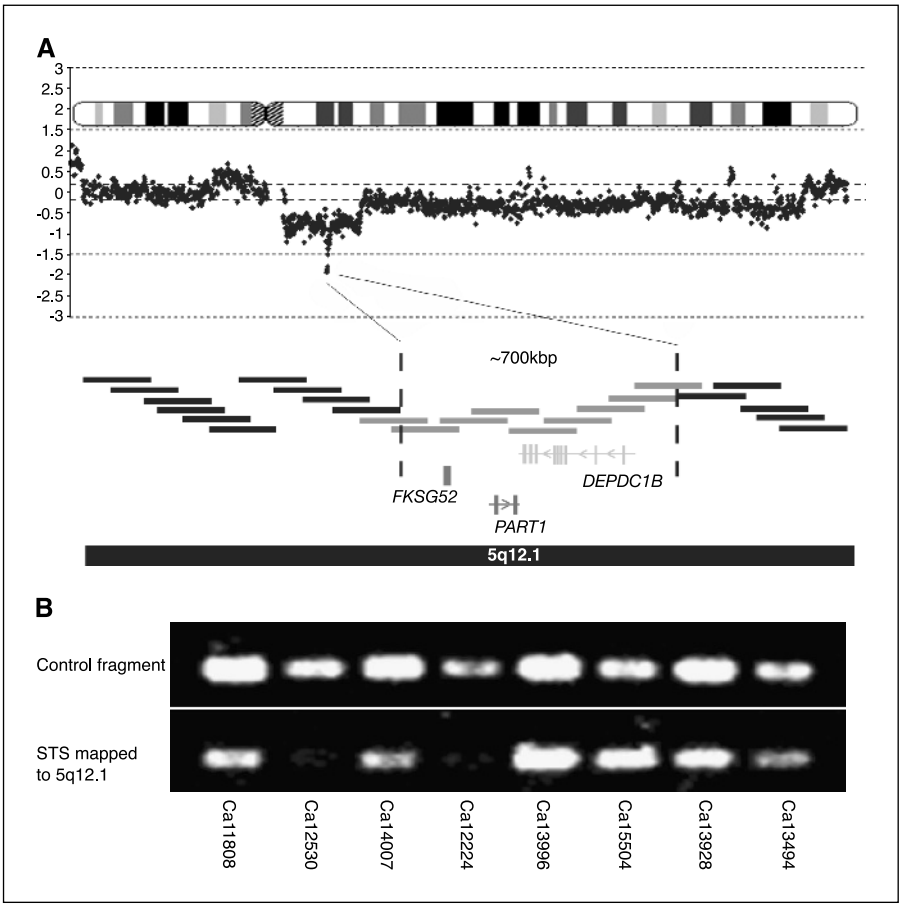

Figure 2. *A,* homozygous deletion on chromosome 5q detected in a BRCA1 tumor (Ca 12530). Most of the q arm of chromosome 5 is hemizygously-deleted and a small region on 5q12.1 was found to be homozygously deleted. Using the 1-Mb resolution arrays, a 2.8-Mbp region including nine genes were identified. Using the 50-kb resolution array we narrowed down the region to 700-kbp deletion including *PART1* (prostate-specific and androgen-regulated transcript 1), *PDE4D* (cyclic AMP–specific phosphodiesterase variant), *FKSG52* (unknown function), and *DEPDC1B* (or *XPT1,* encoding a Rho GTPase activation protein). *B,* PCR analysis verified the 5q12.1 homozygous deletion in a primary tumor (Ca 12224) and its local recurrence (Ca 12530). A marker for a fragment within a nondeleted region was used as reference to an STS marker mapped to the putative homozygous deletion.

whether the genomic copy number profiles of the 50 primary breast tumor samples accurately identified them as positive or negative for *BRCA1* mutations or positive or negative for *BRCA2* mutations (Table 2). Eleven of 12 samples with *BRCA1* mutation were correctly identified in the BRCA1 classification. In addition, one BRCA2 tumor and three sporadic tumors were also classified as BRCA1 tumors. All sporadic breast tumors classified as BRCA1 tumors were ER negative, had no distinct amplification at the *ERBB2* locus, and were of histologic grade 3. Moreover, 9 of 12 samples with BRCA2 mutation and 34 of 38 without BRCA2 mutation were correctly identified in the BRCA2 classification. All four sporadic tumors misclassified as BRCA2 tumors were ER positive, two were of histologic grade two and two samples of grade 3.

Discussion

Array-based CGH, as well as its predecessor metaphase CGH, has the advantage of revealing DNA copy number changes throughout the entire genome. In contrary to the latter, array-based CGH provides high resolution limited only by the number of probes available on the array. As CGH techniques continue to develop, their use in cancer studies will undoubtedly give new insight in cancer genetics, development, and etiology as already exemplified by published work (15, 24, 25). We here present the first array-based CGH study investigating genomic profiles of breast cancers arising in women with *BRCA1* and *BRCA2* mutations. Three previous studies based on metaphase CGH show similar results including a higher frequency of copy number alterations in inherited compared with sporadic breast cancers (9, 11, 12). A *BRCA1*-specific genome profile was suggested, in which regions on chromosomes 3 and 5 were included (12). Descriptive studies have shown recurrent gains and losses on chromosomes 3 and 5 indicating that alterations in these regions are important for BRCA1 tumor progression (9).

Here, we report genomic profiles for BRCA1 and BRCA2 tumors using DNA arrays comprised of ~5,000 individual BAC clones providing an average resolution of 0.9 Mb. Genomic gains, heterozygous losses, high-level amplifications, and homozygous deletions were mapped. Our results confirm that BRCA1 and BRCA2 tumors exhibit a higher frequency of copy number alterations, although mainly shown here for the BRCA1 tumors compared with sporadic breast tumors with regard to percentage of altered clones. Although the number of discrete chromosomal regions altered was higher in hereditary than in sporadic tumors, this difference was not significant, indicating the presence of large alterations which signifies the gross chromosomal instability and aggressive phenotype generally displayed by BRCA1 and BRCA2 tumors. Both gene products are, among a variety of functions, important key players in homologous recombination mediated DNA repair; thus, loss of function increase genomic instability (3). Moreover, these tumors may evolve along distinct genomic programs as suggested by the occurrence of specific recurrent chromosomal alterations. For instance, the majority of all BRCA1 tumors showed losses on chromosomes 4p, 4q, and 5q, often accompanied by copy number gains or amplifications on 3q23-qter. The *PIK3CA* oncogene is one of several candidate genes in the

relatively large 3q region. Interestingly, a potential role of *CDK5* in the 7q36.1 amplicon is corroborated by gene expression analysis, where one of the BRCA1 tumors with 7q36-qter amplification was shown to have an elevated expression of *CDK5* in comparison with a large number of other tumors (data not shown).

To further elucidate the effect of copy number changes to gene expression, we did gene expression analysis on 11 BRCA1 and BRCA2 tumors and found 746 genes for which DNA copy number and RNA expression were significantly correlated ($P < 0.01$). Interestingly, *TM4SF1* located on 3q25.1, a region known to be commonly amplified in BRCA1 tumors, showed a very high correlation with gene expression data. *TM4SF1* is a member of the transmembrane 4 superfamily, also known as the tetraspanin family. These are putative regulators of cancer cell motility and metastasis (26) which indicates that *TM4SF1* is a candidate oncogene for BRCA1 tumors. Furthermore, several regions on 5q are frequently deleted in BRCA1 tumors, including 5q23.3 and *RAD50*, another gene for which gene and transcript copy numbers were significant correlated. RAD50 interacts with BRCA1 and is part of the "BRCA1-associated genome surveillance complex" that serves as a sensor of abnormal DNA structures (27). Moreover, BRCA2 tumors have a high frequency of DNA copy number gains on 20q13, which was confirmed in the present study suggesting a role of *BCAS4* as one of the target genes. Recently, *BCAS4* was found to be amplified and overexpressed in 9 of 13 breast cancer cell lines, suggesting an oncogenic function in sporadic breast cancer as well (28).

Intriguingly, we found a striking concordance regarding genomic profiles for two different tumors obtained from one patient carrying a *BRCA2* mutation. These tumors were regarded as

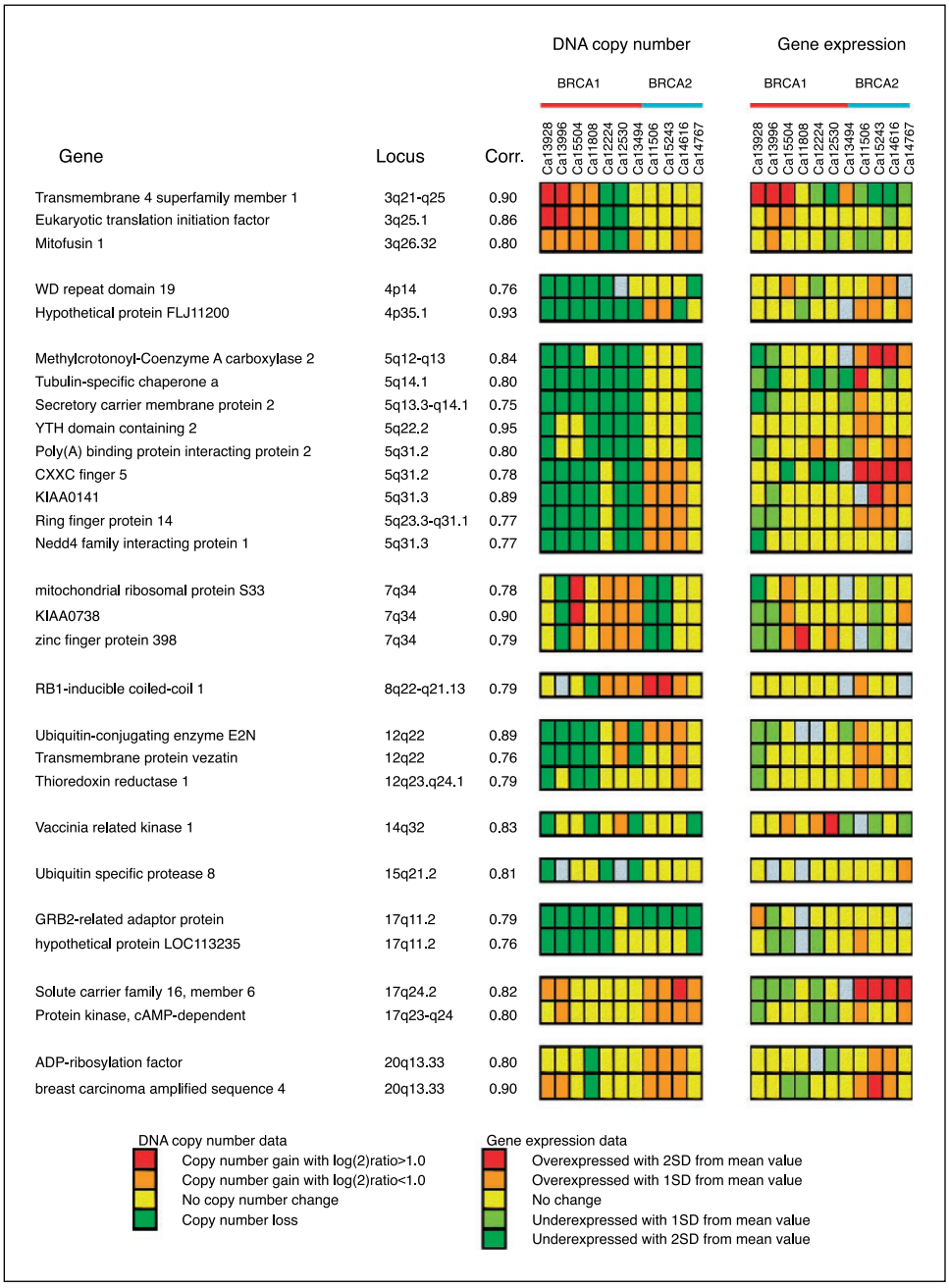

Figure 3. Genes with significant correlation with regards to DNA copy number changes and RNA expression and that are located in regions discriminating BRCA1, BRCA2, and sporadic breast tumors. Also noted are the correlation coefficients (*Corr.*) for each gene-BAC clone pair.

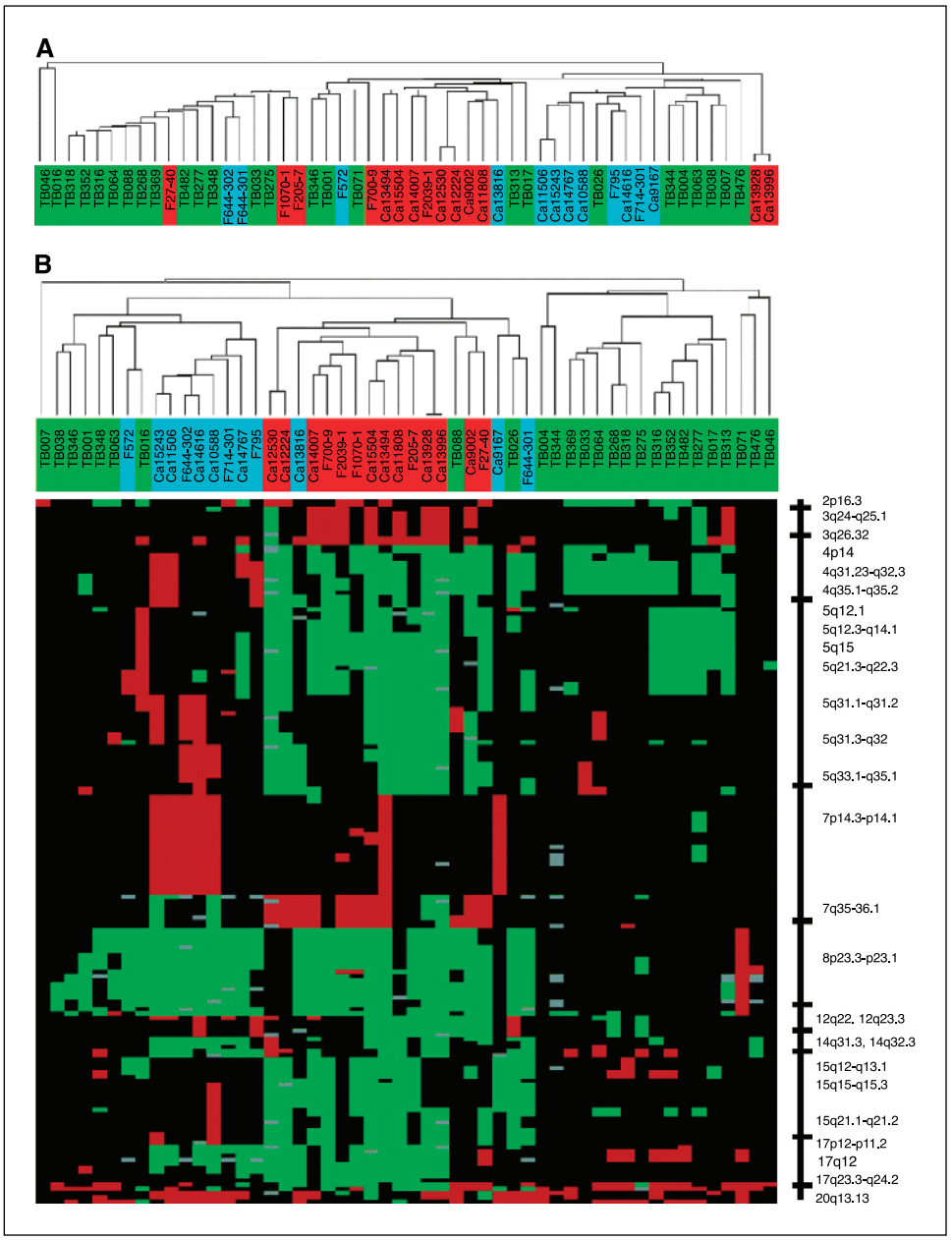

Figure 4. *A,* unsupervised hierarchical clustering of BRCA1 (*red*), BRCA2 (*blue*), and sporadic breast tumors (*green*) using 3,591 BAC clones. Samples are numbered according to Table 1. *B,* hierarchical clustering using 169 BAC clones discriminating BRCA1, BRCA2, and sporadic tumors. BAC clones are ordered according to their genomic position. Label for cytogenetic bands for specific regions (*right*).

independent primary cancers as determined by histopathologic variables and arose in opposite breasts with a considerable time interval. Both tumors were invasive and included extensive DCIS, further arguing for an independent origin and against a metastatic spread. Recently, a study using metaphase CGH analysis of bilateral sporadic breast cancer revealed widely different alterations in the contralateral tumors, indicating that most bilateral cancers indeed have different origin (29). Although we cannot exclude the

Table 2. Classification of BRCA1- and BRCA2-associated breast cancers using support vector machines and leave-one-out classification

Classification	No. samples analyzed	No. misclassified samples	*P*
BRCA1 mutation positive versus *BRCA1* mutation negative	50	5 (*BRCA1* mutation positive 1; *BRCA1* mutation negative 4)	<0.0000004
BRCA2 mutation positive versus *BRCA2* mutation negative	50	7 (*BRCA2* mutation positive 3; *BRCA2* mutation negative 4)	<0.00005

possibility that a contralateral breast cancer arise via hematogenic metastasis, our results indicate that a predisposing BRCA gene mutation and other constitutional factors, possibly in combination with a specific progenitor cell type of origin, may influence tumor development and genomic alterations along distinct genetic pathways or 'programs'.

Heterozygous 5q deletions have previously been found to be common in BRCA1 tumors indicating the presence of one or several tumor suppressor genes (9). Moreover, CGH-targeted linkage analysis revealed a possible BRCA1 modifier locus on chromosome on 5q33-q34 (30). In our set of 12 independent primary tumors from *BRCA1* mutation carriers, we found nine heterozygous and one homozygous deletion on 5q12.1. This novel homozygous deletion was found in both the primary tumor and its local recurrence. To determine its exact boundaries, we used a BAC array with complete coverage of chromosome 5, which allowed narrowing the region to approximately 700 kbp. The region fully comprises two known genes, *PART1* that regulated by androgens in prostate cancer cells (31) and *DEPDC1B* that encodes a Rho GTPase activation protein and a putative tumor suppressor protein. The region is situated in the proximity of a proposed familial prostate cancer susceptibility locus (32) and also includes a single exon gene (*FKSG52* or *AF336879*) of unknown function and the 3′ flanking region of *PDE4D*, which encodes a phosphodiesterase associated with ischemic stroke (33). The role of *PART1*, *FKSG52*, *PDE4D*, or *DEPDC1B* in BRCA1 tumor development remains to be determined.

Using hierarchical clustering methods to group tumors based on ternary data for the 169 discriminating BAC clones, all BRCA1 tumors were tightly clustered and separated from sporadic cases. BRCA2 tumors showed a somewhat higher similarity with the sporadic tumors but still displayed a distinct genomic profile of their own. This is a remarkable finding in light of the inclusion of DNA extracts from both fresh-frozen and formalin-fixed tumors. Eight of 12 BRCA2 tumors were tightly grouped in a subcluster, whereas one BRCA2 outlier clustered together with the sporadic cases and three outlier cases clustered with BRCA1 tumors. This was not due to inclusion of male BRCA2 breast cancers, which had an overall similar genomic profile as female tumors, as shown earlier (10). It might, however, reflect that BRCA2 tumors constitute a more heterogeneous group, as also shown by a number of different histopathologic and clinical variables (7). It could also be a result of the relatively small number of samples in the present study. In addition, two sporadic cases seemed highly similar to the BRCA1 tumors with regard to their genomic profiles. Sporadic tumors with methylation of the *BRCA1* gene promoter have been shown to display an expression profile similar to BRCA1-mutated tumors (8), and further analysis will reveal if *BRCA1* methylation can influence also the genomic profiles of tumors. On the other hand, BRCA-like sporadic tumors could also represent missed BRCA1 germ line mutation carriers. Interestingly, one of the sporadic tumors (TB088) in the BRCA1 cluster exhibited amplification of the *EMSY* oncogene, which encodes a novel BRCA2-binding protein that suppress the chromatin remodeling function of BRCA2 (34). It is tempting to speculate in the role of *EMSY* activation to promote a tumor development similar to inherited tumors, but this will clearly need further studies to prove.

To further elucidate the biological mechanism behind the distinct copy number profiles for BRCA1/2-associated breast cancers, we extracted 30 of 746 significant genes correlating in gene expression and copy number data mapped to the discriminating regions pinpointing genes that might be important in the distinct tumor progression patterns in BRCA1 and BRCA2 breast cancers (Fig. 3). In addition to hierarchical clustering methods, a supervised leave-one-out iterative classification of the samples was conducted based on whole genome copy number data. Using this model, only one BRCA1 breast tumor was misclassified. Interestingly, this tumor was obtained from the single missense mutation ($Cys^{64}Gly$) carrier included in the study, whereas all other BRCA tumors were from nonsense or frameshift mutation carriers. Although this mutation is known to be deleterious and disease-associated, transcript with $Cys^{64}Gly$ may escape nonsense-mediated decay, become translated, and retain some of the many functions of wild-type BRCA1 protein or acquiring new ones. In addition, the three sporadic breast tumors that were classified in the BRCA1 group were all ER and ERBB2 negative and of histologic grade 3 thus resembling tumors of basal-like phenotype that are similar to BRCA1 breast tumors with regard to gene expression profiles (35). In the BRCA2 classification model, 9 of 12 BRCA2 tumors were correctly classified indicating that these tumors constitute a more heterogeneous group than BRCA1 tumors. In addition, four sporadic tumors were classified as BRCA2-associated tumors. Nonetheless, even if the results of the present study must await confirmation in a larger and independent tumor material, our overall findings strongly suggest that BRCA1 and BRCA2 tumors can be classified based on copy number alterations. In addition, using CGH arrays with higher resolution, such as tiling DNA microarrays with complete coverage of the human genome, the regions discriminating the different tumor groups will be refined (36). Such a BRCA tumor classifier, applicable to formalin-fixed, paraffin-embedded tumor specimens, would be a valuable tool for test laboratories and oncogenetic clinics. It may also become of value for selection of "BRCAness" tumors for targeted treatment (37, 38).

Acknowledgments

Received 2/18/2005; revised 6/11/2005; accepted 6/24/2005.

Grant support: Swedish Cancer Society, the Swedish Research Council, the Berta Kamprad Foundation, the Gunnar Nilsson Cancer Foundation, the Franke and Margareta Bergqvist Foundation, the Lund University Hospital Foundations, the King Gustav V:s Jubilee Foundation, the Ingabritt and Arne Lundberg Foundation, the Swedish Foundation for Strategic Research, the Knut and Alice Wallenberg Foundation via the SWEGENE program (Å. Borg), the Breast Cancer Research Foundation, QVC/FFNY and the Abramson Family Cancer Research Institute (B.L. Weber), and the American Cancer Society (G. Jönsson).

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 Kazutoyo Osoegawa and Pieter de Jong (BACPAC Resources, Children's Hospital Oakland Research Institute, Oakland, CA; http://bacpac.chori.org/genomicRearrays.php) for obtaining the BAC clone DNA aliquots from the 32K set.

References

1. Ford D, Easton DF, Stratton M, et al. Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. The Breast Cancer Linkage Consortium. Am J Hum Genet 1998;62:676–89.
2. Nathanson KL, Wooster R, Weber BL. Breast cancer genetics: what we know and what we need. Nat Med 2001;7:552–6.
3. Venkitaraman AR. Cancer susceptibility and the functions of BRCA1 and BRCA2. Cell 2002;108:171–82.
4. Loman N, Johannsson O, Bendahl PO, Borg A, Ferno M, Olsson H. Steroid receptors in hereditary breast carcinomas associated with BRCA1 or BRCA2

mutations or unknown susceptibility genes. Cancer 1998;83:310–9.
5. Johannsson OT, Idvall I, Anderson C, et al. Tumour biological features of BRCA1-induced breast and ovarian cancer. Eur J Cancer 1997;33:362–71.
6. Loman N, Johannsson O, Kristoffersson U, Olsson H, Borg A. Family history of breast and ovarian cancers and BRCA1 and BRCA2 mutations in a population-based series of early-onset breast cancer. J Natl Cancer Inst 2001;93:1215–23.
7. Lakhani SR, Gusterson BA, Jacquemier J, et al. The pathology of familial breast cancer: histological features of cancers in families not attributable to mutations in BRCA1 or BRCA2. Clin Cancer Res 2000; 6:782–9.
8. Hedenfalk I, Duggan D, Chen Y, et al. Gene-expression profiles in hereditary breast cancer. N Engl J Med 2001; 344:539–48.
9. Tirkkonen M, Johannsson O, Agnarsson BA, et al. Distinct somatic genetic changes associated with tumor progression in carriers of BRCA1 and BRCA2 germ-line mutations. Cancer Res 1997;57:1222–7.
10. Tirkkonen M, Kainu T, Loman N, et al. Somatic genetic alterations in BRCA2-associated and sporadic male breast cancer. Genes Chromosomes Cancer 1999; 24:56–61.
11. Kainu T, Juo SH, Desper R, et al. Somatic deletions in hereditary breast cancers implicate 13q21 as a putative novel breast cancer susceptibility locus. Proc Natl Acad Sci U S A 2000;97:9603–8.
12. Wessels LF, van Welsem T, Hart AA, van't Veer LJ, Reinders MJ, Nederlof PM. Molecular classification of breast carcinomas by comparative genomic hybridization: a specific somatic genetic profile for BRCA1 tumors. Cancer Res 2002;62:7110–7.
13. Umar A, Risinger JI, Hawk ET, Barrett JC. Testing guidelines for hereditary non-polyposis colorectal cancer. Nat Rev Cancer 2004;4:153–8.
14. Greshock J, Naylor TL, Margolin A, et al. 1-Mb resolution array-based comparative genomic hybridization using a BAC clone set optimized for cancer gene analysis. Genome Res 2004;14:179–87.
15. Snijders AM, Nowak N, Segraves R, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. Nat Genet 2001;29:263–4.
16. Saal LH, Troein C, Vallon-Christersson J, Gruvberger S, Borg A, Peterson C. BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data. Genome Biol 2002;3: SOFTWARE0003.
17. Yang YH, Dudoit S, Luu P, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res 2002;30:e15.
18. Autio R, Hautaniemi S, Kauraniemi P, et al. CGH-Plotter: MATLAB toolbox for CGH-data analysis. Bioinformatics 2003;19:1714–5.
19. Kruskal WH. WW. Use of ranks in one-criterion variance analysis. J Am Stat Assoc 1952;47:583–621.
20. Ringner M, Veerla S, Andersson S, Staaf J, Hakkinen J. ACID: a database for microarray clone information. Bioinformatics 2004;20:2305–6.
21. Yang IV, Chen E, Hasseman JP, et al. Within the fold: assessing differential expression measures and reproducibility in microarray assays. Genome Biol 2002;3: research0062.
22. Saeed AI, Sharov V, White J, et al. TM4: a free, open-source system for microarray data management and analysis. Biotechniques 2003;34:374–8.
23. Brown MP, Grundy WN, Lin D, et al. Knowledge-based analysis of microarray gene expression data by using support vector machines. Proc Natl Acad Sci U S A 2000;97:262–7.
24. Heidenblad M, Schoenmakers EF, Jonson T, et al. Genome-wide array-based comparative genomic hybridization reveals multiple amplification targets and novel homozygous deletions in pancreatic carcinoma cell lines. Cancer Res 2004;64:3052–9.
25. Paris PL, Andaya A, Fridlyand J, et al. Whole genome scanning identifies genotypes associated with recurrence and metastasis in prostate tumors. Hum Mol Genet 2004;13:1303–13.
26. Kao YR, Shih JY, Wen WC, et al. Tumor-associated antigen L6 and the invasion of human lung cancer cells. Clin Cancer Res 2003;9:2807–16.
27. Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J. BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. Genes Dev 2000;14:927–39.
28. Barlund M, Monni O, Weaver JD, et al. Cloning of BCAS3 (17q23) and BCAS4 (20q13) genes that undergo amplification, overexpression, and fusion in breast cancer. Genes Chromosomes Cancer 2002;35:311–7.
29. Teixeira MR, Ribeiro FR, Torres L, et al. Assessment of clonal relationships in ipsilateral and bilateral multiple breast carcinomas by comparative genomic hybridisation and hierarchical clustering analysis. Br J Cancer 2004;91:775–82.
30. Nathanson KL, Shugart YY, Omaruddin R, et al. CGH-targeted linkage analysis reveals a possible BRCA1 modifier locus on chromosome 5q. Hum Mol Genet 2002;11:1327–32.
31. Lin B, White JT, Ferguson C, et al. PART-1: a novel human prostate-specific, androgen-regulated gene that maps to chromosome 5q12. Cancer Res 2000;60: 858–63.
32. Wiklund F, Gillanders EM, Albertus JA, et al. Genome-wide scan of Swedish families with hereditary prostate cancer: suggestive evidence of linkage at 5q11.2 and 19p13.3. Prostate 2003;57:290–7.
33. Gretarsdottir S, Thorleifsson G, Reynisdottir ST, et al. The gene encoding phosphodiesterase 4D confers risk of ischemic stroke. Nat Genet 2003;35:131–8.
34. Hughes-Davies L, Huntsman D, Ruas M, et al. EMSY links the BRCA2 pathway to sporadic breast and ovarian cancer. Cell 2003;115:523–35.
35. Sorlie T, Tibshirani R, Parker J, et al. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc Natl Acad Sci U S A 2003;100:8418–23.
36. Ishkanian AS, Malloff CA, Watson SK, et al. A tiling resolution DNA microarray with complete coverage of the human genome. Nat Genet 2004;36:299–303.
37. Bryant HE, Schultz N, Thomas HD, et al. Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. Nature 2005;434:913–7.
38. Farmer H, McCabe N, Lord CJ, et al. Targeting the DNA repair defect in BRCA mutant cells as a therapeutic strategy. Nature 2005;434:913–21.