Chromosome 22 array-CGH profiling of breast cancer delimited minimal common regions of genomic imbalances and revealed frequent intra-tumoral genetic heterogeneity

MAGDALENA BENETKIEWICZ[1,4], ARKADIUSZ PIOTROWSKI[1], TERESITA DÍAZ DE STÅHL[1], MICHAL JANKOWSKI[2], DARIUSZ BALA[2], JACEK HOFFMAN[2], EWA SRUTEK[2], RYSZARD LASKOWSKI[2], WOJCIECH ZEGARSKI[2] and JAN P. DUMANSKI[1,3]

[1]Department of Genetics and Pathology, Rudbeck Laboratory, Uppsala University, 751 85 Uppsala, Sweden; [2]Department of Breast Cancer, Clinic of Oncological Surgery, Oncology Center, Collegium Medicum Nicolaus Copernicus University, 85 796 Bydgoszcz, Poland; [3]Department of Genetics, University of Alabama at Birmingham, KAUL 420, 1530 3rd Ave. S, Birmingham, AL 35294, USA

Received March 27, 2006; Accepted June 2, 2006

Abstract. Breast cancer is a common malignancy and the second most frequent cause of death among women. Our aim was to perform DNA copy number profiling of 22q in breast tumors using a methodology which is superior, as compared to the ones applied previously. We studied 83 biopsies from 63 tumors obtained from 60 female patients. A general conclusion is that multiple distinct patterns of genetic aberrations were observed, which included deletion(s) and/or gain(s), ranging in size from affecting the whole chromosome to only a few hundred kb. Overall, the analysis revealed genomic imbalances of 22q in 22% (14 out of 63) of tumors. The predominant profile (11%) was monosomy 22. The smallest identified candidate region, in the vicinity of telomere of 22q, encompasses ~220 kb and was involved in all but one of the tumors with aberrations on chromosome 22. This segment is dense in genes and contains 11 confirmed and one predicted gene. The availability of multiple biopsies from a single tumor provides an excellent opportunity for analysis of possible intra-tumor differences in genetic profiles. In 15 tumors we had access to two or three biopsies derived from the same lesion and these were studied independently. Four out of 15 (26.6%) tumors displayed indications of clonal intra-tumor genotypic differences, which should be viewed as a high number, considering that we studied in detail only a single human chromosome. Our results open up several avenues for continued genetic research of breast cancer.

Correspondence to: Dr M. Benetkiewicz, [4]*Present address*: Unite 509 INSERM, Pathologie Moleculaires des Cancers, Institut Curie Section de Recherche, 26 rue d'Ulm, 75248 Paris cedex 05, France
E-mail: magdalena.benetkiewicz@curie.fr

Key words: genomic array, 22q, telomere, clonal intra-tumor genotypic differences, complex aberrations, biopsies

Introduction

Breast cancer is the most common malignancy among women and the second most common cause of death after lung cancer (1). It is a complex genetic disorder and some aberrations have been correlated with heterogeneous histology and clinical behavior. Breast carcinoma arises from the epithelium of glandular tissue, which includes ducts and lobules. Histologically, this neoplasm can be classified into non-invasive (*in situ*) or invasive ductal and lobular carcinoma. The most common type of breast cancer, invasive ductal carcinoma, develops from ductal carcinoma *in situ* and accounts for 80% of all cases (2). The majority of breast cancer cases are sporadic, while a family history of the disease accounts for 15-20% (3). Less than 1% of all cases are associated with the autosomal dominant or recessive syndromes: Cowden syndrome, Li-Fraumeni syndrome, Peutz-Jeghers syndrome, Bloom syndrome, Werner syndrome and Xeroderma Pigmentosum (4). Individuals recognized with these syndromes or congenital malformations may have a high breast cancer risk. Approximately 5-10% of all cases are attributable to autosomal dominant susceptibility genes: breast cancer susceptibility gene 1 (*BRCA1*) and breast cancer susceptibility gene 2 (*BRCA2*). Mutations in either of these genes account for the majority of families with multiple cases and confer a lifetime risk of breast cancer up to 85% for female *BRCA1* or *BRCA2* mutation carriers (5-8). Many other susceptibility candidate loci including known oncogenes and tumor suppressor genes have also been characterized [breast cancer (OMIM #114480)].

In recent years detailed cytogenetic and molecular investigations of breast cancer have led to the identification of a number of recurrent regions of DNA copy number alteration (9-13). Furthermore, it has been suggested that allelic loss of 22q is a common event in breast carcinoma, with a reported frequency between 11% and 66% (14-16). Previous studies have reported several regions along chromosome 22 showing allelic loss in sporadic breast carcinomas and a candidate tumor suppressor gene region, close to the telo-

Table I. Summary of clinical details and chromosome 22 array-CGH profiles of breast cancer patients.

Clinical data											Chromosome 22 array-CGH data	
						Receptors[f]						
Patient ID[a]	Age at operation	Histology[b]	pTNM[c]	Chemotheraphy pre-operatively[d]	Chemotheraphy post-operatively[e]	ER	PR	HER2	E-cadherin	Array-CGH profile[g]	Aberrant clones[h]	Position on 22q (bp)[i]
001 T	52	d. c.	T2N0	no	CMF	+	+	+	+	dipl.		
002 T	46	d. c.	T2N1M0	D + ADR	ADR + CMF	-	-	+	-	dipl.		
003 T	62	r. c.	n. d.	no	no	-	-	+	+	mono.		
004 T	59	d. c.	T2N1biii	no	AC	-	-	+	+	dipl.		
005 T	46	d. c.	T3N0	no	no	-	+	+	-	dipl.		
007 T	51	a. c.	T1aN0	no	CMF	-	-	-	+	dipl.		
008 T	58	d. c.	T1cN0	no	no	+	+	+	+	gain	Z68323-AL021877	32898948-33388070
										int. del.	AL117256-AC005527;	26333601-28342153;
											AL022334-AL008637;	34214125-AL008637;
											AL022315-Z83845;	36218147-38179916;
											Z83851-AL078613;	40972968-47675305;
											Z94802-AC000050	49219663-49434548
009 T	56	d. c.	T1N1M0	no	AC	+	+	+	+	dipl.		
010 T	63	l. c.	T1cN0	no	no	+	+	-	-	dipl.		
011 T	46	d. c.	T1cN1bii	no	CMF	+	+	-	+	mono.		
012 T	72	d. c.	T2N2	no	ADM	-	-	+	+	dipl.		
013 T	57	mu. b. c.	T2N0	no	CMF	-	-	+	-	dipl.		
014 T	40	l. c.	T2N2	no	CMF	+	+	+	-	dipl.		
014 T1	40	l. c.	T2N3	no	CMF	+	+	+	-	dipl.		
015 T	72	l. c.	T2N1biii	no	yes	+	+	+	-	dipl.		
016 T	62	d. c.	T2N1biii	no	ADR	-	-	+	+	dipl.		
017 T	70	d. c.	T2N1biv	no	CMF	+	+	-	+	dipl.		
017 T1	70	d. c.	T2N1biv	no	CMF	+	+	-	+	dipl.		
018 T	77	d. c.	T1cN0	no	no	-	-	-	+	dipl.		
019 T	64	l. c.	T2N1biv	no	ADR + CMF	+	-	-	-	dipl.		
022 T	67	l. c.	T2N1biii	no	no	+	+	+	-	gain	AC000087-AC000070;	17819985-17947539;
											AL022323-AL049536;	23976814-26143705;
											AL008583-AL078641	37493053- 37810389
										int. del.	AP000531-AC000079;	14686557-17797505;
											AL117256-AL008641;	26333601-32785169;
											Z82196-AL021707;	33387967-37479376;
											AL022353-AC002056	38553570-49495206
023 T	60	d. c.	T2N1b	no	CMF	+	+	+	+	dipl.		
024 T	60	d. c.	pT2N1biii	no	ADR + CMF	-	-	-	-	dipl.		
024 T1	60	d. c.	pT2N1biii		ADR + CMF	-	-	-	-	dipl.		
025 T	28	d. c.	T1N0	no	CMF	+	+	+	+	mon.		
I-025 T	28	d. c.	T1N0	no	CMF	+	+	+	+	dipl.		
II-025 T	28	d. c.	T1N0	no	CMF	+	+	+	+	mono.		
II-025 T1	28	d. c.	T1N0	no	CMF	+	+	+	+	mono.		
027 T	45	d. c.	T1cN1biii	no	CMF	+	+	+	+	dipl.		
028 T	61	d. c.	T2Nbiii	no	AC	+	+	+	+	dipl.		
028 T1	61	d. c.	T2Nbiii	no	AC	+	+	+	+	dipl.		
028 T2	61	d. c.	T2Nbiii	no	AC	+	+	+	+	dipl.		
II-028 T	61	d. c.	T2Nbiii	no	AC	+	+	+	+	dipl.		
II-028 T1	61	d. c.	T2Nbiii	no	AC	+	+	+	+	dipl.		
029 T	68	d. c.	T2N1biii	no	ADR + CMF	+	+	+	+	dipl.		
032 T	44	d. c.	T1bN0	no	no	-	+	+	+	dipl.		
033 T	70	d. c.	T2N1biii	no	CMF	+	+	+	+	dipl.		
035 T	58	d. c.	T2N0	no	no	+	+	+	+	dipl.		
035 T1	58	d. c.	T2N0	no	no	+	+	+	+	dipl.		
036 T	58	d. c.	T2N1biii	no	AC	-	-	+	+	tel. del.	AL022328-AC000036	48840732-49470169
036 T1	58	d. c.	T2N1biii	no	AC	-	-	+	+	dipl.		
036 T2	58	d. c.	T2N1biii	no	AC	-	-	+	+	dipl.		
037 T	61	can. post chemo.	T2N1biii	MVCP + AC	AC	-	-	+	+	cen. del.	AP000525-AP000358	14509865-23444108
037 T1	61	can. post chemo.	T2N1biii	MVCP + AC	AC	-	-	+	+	cen. gain	D87014-AL021153	21249594-25942582

Table I. Continued.

Clinical data											Chromosome 22 array-CGH data	
Patient ID[a]	Age at operation	Histology[b]	pTNM[c]	Chemotheraphy pre-operatively[d]	Chemotheraphy post-operatively[e]	Receptors[f] ER	PR	HER2	E-cadherin	Array-CGH profile[g]	Aberrant clones[h]	Position on 22q (bp)[i]
038 T	75	d. c.	T2N1bi	no	CMF	-	-	+	+	gain	AL096701-Z99716	30121563-40798159
										ter. del.	AL110122-AC002056	46622060-49495206
039T	40	can. post chemo.	T2N1biii	D + ADR	CMF	+	+	-	+	dipl.		
039 T1	40	can. post chemo.	T2N1biii	D + ADR	CMF	+	+	-	+	dipl.		
040 T	70	l. c.	T1cN0	no	no	+	-	-	-	dipl.		
041 T	45	d. c.	T2N0	no	CMF	-	-	+	+	dipl.		
042 T	48	l. c.	T2N1biii	no	CMF	+	+	-	-	dipl.		
043 T	77	d. c.	T2N1biii	no	no	+	+	+	+	dipl.		
044 T	77	d. c.	T2N1biii	no	ADR + CMF	+	+	+	+	dipl.		
044 T1	77	d. c.	T2N1biii	no	ADR + CMF	+	+	+	+	dipl.		
044 T2	77	d. c.	T2N1biii	no	ADR + CMF	+	+	+	+	dipl.		
045 T	72	d. c.	T1cN0	no	CMF	-	-	-	+	dipl.		
047 T	47	d. c.	T2N0	no	CMF	+	+	-	+	dipl.		
048 T	60	a. c.	T2N0	no	CMF	-	-	+	+	dipl.		
048 T1	60	a. c.	T2N0	no	CMF	-	-	+	+	dipl.		
049 T	73	l. c.	T2N1biii	no	CMF	+	+	+	-	dipl.		
050 T	63	d. c.	T1bN0	no	no	+	+	+	+	dipl.		
051 T	41	d. c.	T2N1	no	CMF	-	-	-	+	ter. del.	AL021707-AC002055	37348024-49534710
053 T	65	d. c.	T2N1biii	no	no	+	+	+	+	dipl.		
054 T	53	l. c.	T1cN0	no	no	+	+	+	-	mono.		
055 T	63	d. c.	pT2N1biii	no	CMF	+	+	+	+	dipl.		
056 T	65	d. c.	pT2N0	no	no	+	-	+	+	dipl.		
056 T1	65	d. c.	pT2N0	no	no	+	-	+	+	mono.		
057 T	52	d. c.	T2N0	no	no	+	+	+	+	dipl.		
058 T	67	d. c.	T2N0	no	AC	-	+	+	+	dipl.		
059 T	48	l. c.	T2N1biii	no	CMF	+	+	-	-	dipl.		
060 T	68	l. c.	T2 N1biv	no	CMF	+	+	+	-	dipl.		
061 T	69	m. c.	n. d.	no	no	n. d.	n. d.	n.d.	n. d.	dipl.		
062 T	67	can. post chemo.	T3Nx	MVCP	CMF	-	-	-	+	dipl.		
062 T1	67	can. post chemo.	T3Nx	MVCP	CMF	-	-	-	+	dipl.		
062 T2	67	can. post chemo.	T3Nx	MVCP	CMF	-	-	-	+	dipl.		
063 T	48	d. c.	T1cN0	no	CMF	+	+	+	+	dipl.		
064 T	57	l. c.	T2 N1biii	no	ADR + CMF	+	+	+	-	dipl.		
065 T	57	l. c.	T2N0	no	no	+	-	-	-	dipl.		
065 T1	57	l. c.	T2N0	no	no	+	-	-	-	mono.		
065 T2	57	l. c.	T2N0	no	no	+	-	-	-	dipl.		
066 T	64	d. c.	T2N0	no	AC	-	-	+	+	dipl.		
067 T	56	d. c.	T1cN0	no	CMF	+	-	-	+	cen. del.	AP000525-AC000092;	14509865-17759840;
											D87012-AC000102	20626387-21833044
068 T	62	l. c.	T2N1biii	no	ADR + CMF + ADM	+	+	-	-	dipl.		
069 T	68	l. c.	T2N1biii	no	ADR + CMF	-	-	+	-	dipl.		

[a]T, T1 or T2 at the end of the case name identifies the different biopsies taken within a tumor. I and II at the beginning of the case name indicates additional tumor samples obtained from the same patient. [b]d. c., ductal carcinoma; r. c., recurrent carcinoma; a. c., apocrine carcinoma; l.c., lobular carcinoma; mu. b. c., mucinous bifocal carcinoma; can. post chemo., cancer post chemotherapy; m. c., metastatic carcinoma. [c]pTNM, pathological tumor-node-metastasis staging classification; n. d., no data. [d]D, Docetaxel; ADR, adriamycin; MVCP, metotrexat + vincristin + cyclophosphamide + prednisone; AC, doxorubicin+cyclophosphamide. [e]CMF, cyclophosphamide + driamycin methotrexate + 5-fluouracil; ADM, doxorubicin. [f]ER, estrogen receptor; PR, progesterone receptor; HER2, human epidermal growth factor receptor 2; E-cadherin, epithelial cadherin 1; +, positive; -, negative. [g]dipl., diploid; mono., monosomy; int. del., interstitial deletion; tel. del., telomeric deletion; cen. del., centromeric deletion; cen. gain, centromeric gain; ter. del., terminal deletion. [h]Copy number changes affecting two or more consecutive clones showing the same changes are indicated. [i]Position of aberrated clones on 22q according to http://www.ncbi.nlm.nih.gov.

meric end of 22q has been determined (17,18). Additionally, allelic loss at 22q13 has been shown to correlate with post-operative recurrence (19). However, the target tumor suppressor gene(s) residing on this chromosome, involved in breast cancer development has not yet been identified. The aim of this study was to perform high resolution gene copy number profiling of chromosome 22 aberrations in breast tumors using a methodological approach which is superior to others applied previously. We used a full-coverage tiling path chromosome 22 genomic microarray with an average resolution of 75 kb (20), which allows the precise detection of minute DNA copy number aberrations.

Materials and methods

Clinical material and chromosome 22 genomic microarray. The study included 83 tumor biopsies from 63 distinct tumor lesions that were obtained from 60 female patients. Single or multiple biopsies were collected from each tumor. From 15 of the tumors, either two or three independent biopsies were collected, followed by independent DNA isolation and genomic profiling. From two patients (ID 025 and 028), two separate distinct tumors were studied, which were located within the same breast. Details of all studied patients are summarized in Table I. Tumor tissues and corresponding surrounding healthy tissue were sampled during operation and stored at -70˚C prior to DNA isolation. Matched peripheral blood was also collected. Samples were obtained from patients treated at the Oncology Centre, Bydgoszcz, Poland. All tumors were classified and graded by experienced pathologist according to the TNM criteria. The clinical samples were characterized immunohistochemically for expression of estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2) and epithelial cadherin 1 (E-cadherin) (Table I). The expression of HER2 receptor was determined by immunohistochemical assay performed on consecutive paraffin sections using standardized automated procedures with Dako Autostainer (HerceptTest™; Carpinteria, CA). Immunohistochemical staining was prepared in accordance to scoring guidelines (possible score: 0 to 3+). The cases, in which overexpression of HER2 receptor occurred, were investigated for *HER2* gene amplification using FISH analysis (details not shown). For the determination of ER, PR and E-cadherin expression, the mouse monoclonal and rabbit polyclonal anti-human receptor antibodies (EnVision™+ Single Reagents; Carpinteria, CA) were used. High molecular weight DNA was isolated from both tumor and peripheral blood using standard methods (21). A full-coverage chromosome 22 genomic array was used for DNA copy number analysis (20). The array covers the whole long arm of chromosome 22, with genomic DNA clones derived from the minimal tiling path and provides an average resolution of 75 kb. It contains a set of 460 chromosome 22 measurement points derived from chromosome 22, as well as X chromosome controls and loci derived from other chromosomes (20), all printed in triplicate. The identity of each clone was verified and quality controlled by STS-PCR (using at least one specific pair of primers) as well as EcoRI cleavage. Each batch of printed slides was validated using DNA derived from a male subject with a 7.4 Mb constitutional deletion on chromosome 22. The deletion was previously identified by cytogenetics studies and confirmed by restriction fragment length polymorphism and simple sequence repeat polymorphism analysis as well as by FISH (22). A complete list of the set of clones included in the array and methods used for preparation of DNA from genomic clones are displayed at http://puffer.genpat. uu.se/chrom_22_array/chrom22.htm.

Hybridizations, scanning and image analysis. Protocols used for test and reference DNA labeling, hybridization and post-hybridization processing have been previously described (20) and are accessible at http://puffer.genpat.uu.se/chrom_22_array/chrom22.htm. A pool of peripheral blood derived DNA, from 10 normal females, was used as control for all performed hybridizations. In brief, two micrograms of test and reference DNA were differentially labeled by random priming using Cy3-dCTP (PA53021, GE Healthcare, Piscataway, NJ) or Cy5-dCTP (PA55021, GE Healthcare). These were then mixed with 100 μg of Cot-1 DNA (Roche, Basel, Switzerland) and hybridized to the array. Image acquisition was done using the GenePix 4000B scanner (Axon Instruments, Inc., Union City, CA). Analysis of hybridization intensity was carried out using the GenePixPro 6 image analysis software (Axon Instruments). Ratio of means for the fluorescence intensities between test and reference DNA was calculated, as well as average and standard deviation for each clone. Clones displaying a standard deviation >5% of the average, between a minimum of two replica spots, were discarded from further analysis. The average ratio from the non-chromosome 22 autosomal controls was used in the normalization of data in each hybridization experiment. Autosomal control clones displaying a normalized average fluorescence ratio indicative for copy number alteration (R>1.2 or R<0.8) were not used in the normalization. The Average Normalized Inter-Locus Fluorescence Ratio (ANILFR), representing the normalized ratio for successfully scored loci from a certain continuous region or regions on the array, was calculated in order to assess the average fluorescence values for a given number of clones as well as the inter-locus variation.

Results

Chromosome 22 gene copy number profiling. We performed DNA copy number profiling of chromosome 22 aberrations in a series of 83 breast tumor biopsies derived from 60 patients (Table I). In order to investigate the presence of genetic heterogeneity within a single tumor, two or three biopsies were collected from 12 patients and independently profiled for copy number changes. Furthermore, for selected patients, DNA derived from corresponding surrounding healthy breast tissue (9 cases), metastatic tumor tissue (3 cases) and peripheral blood (12 cases) were also analyzed. The clinical details of patients included in the study, together with the summary of results from array-CGH profiling of tumor biopsies, are summarized in Table I and Fig. 1. A general conclusion is that multiple distinct patterns of genetic aberrations were observed, which included several types of deletion(s) and/or gain(s), ranging in size from affecting the whole chromosome to regional aberrations encompassing only a few hundred kb. Overall, the analysis revealed genomic imbalances of 22q in

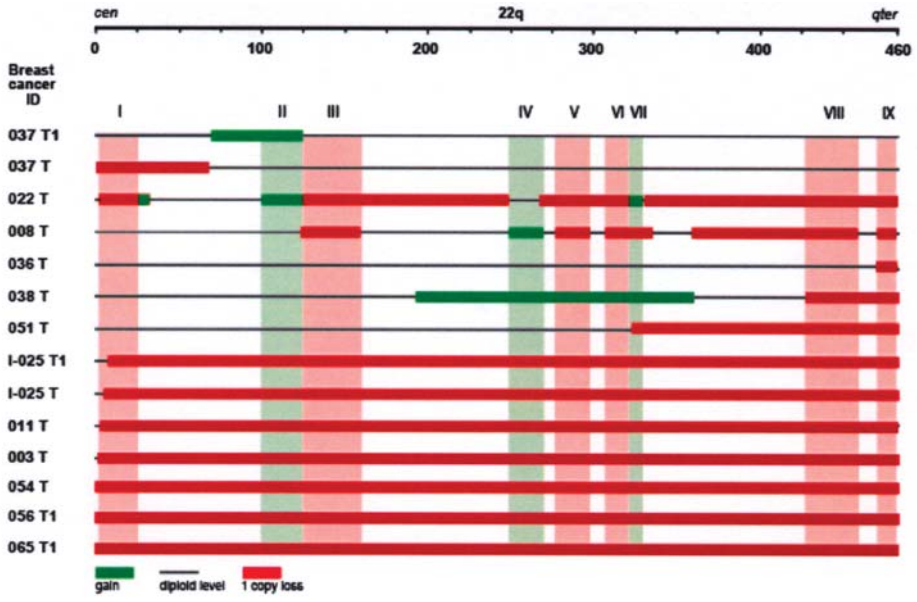


Figure 1. Schematic view of chromosome 22 DNA copy number imbalances detected in 14 breast cancer samples. Tumor samples (case number indicated on the left side), with their corresponding genomic aberration(s) displayed by the designated colors are summarized. Gains are indicated with green bars and deletions with red bars. Regions with normal allelic/copy number are drawn by black lines. Minimum overlapping regions of deletions (I, III, V, VI, VIII and IX, red shadow) and gains (II, IV and VII, green shadow) are displayed on the scale of 460 array data points derived from chromosome 22 on the top of the figure. Two independent biopsies were studied for tumor I-025 (T and T1, see Fig. 4 and Table I) and 037 T (T and T1, see Table I, array CGH profile not shown).

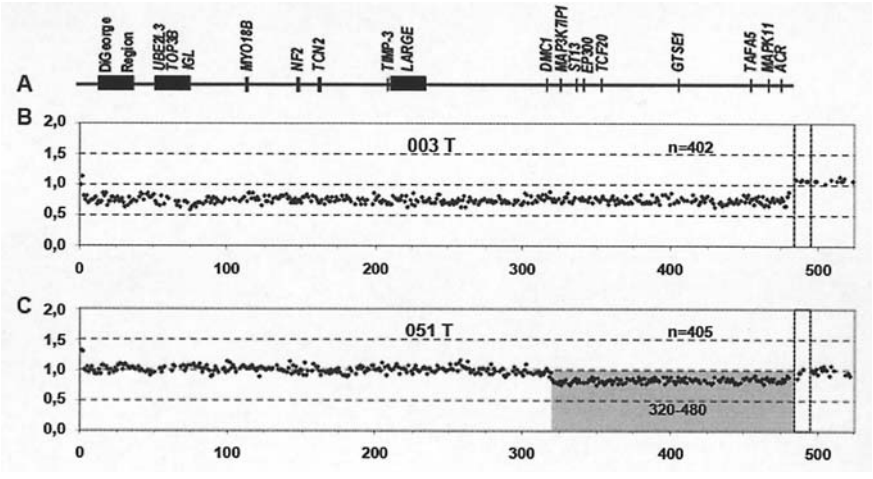


Figure 2. Detection of tumor-specific chromosome 22 hemizygous deletions in breast cancer identified by array-CGH. The X-axis displays chromosome 22 data points, ordered from centromere (left hand side) to telomere (right hand side) while the Y-axis denotes the average normalized fluorescence ratio. Dots between the vertical broken lines indicate control loci derived from chromosome X. Autosomal controls derived from chromosomes other than chromosome 22 are plotted to the right hand side of the X controls. The n-value denotes the number of positively scored 22q-derived measurement points loci in each experiment. The grey box highlights the clones presenting partial gene copy number aberrations; with the ID of the clones indicated. (A) Schematic map of selected chromosome 22 loci/genes. (B) Breast cancer DNA from female patient 003 T presents a profile consistent with monosomy 22 (Average Normalized Inter-Locus Fluorescent Ratio, ANILFR 0.75 ± SD 0.06). Values for the X controls and the non-chromosome 22 controls are at the expected diploid level (ANILFR 1.08 ± SD 0.01 and 1.07 ± SD 0.03, respectively). (C) A 12-Mb tumor-specific terminal deletion detected in case 051 T, spanning from clone AL021707 to the telomere (position 37348024-37479376 bp on 22q). Chromosome 22 clones scored as deleted exhibit a fluorescence ratio of 0.83 ± SD 0.04.

21.6% of patients (13 out of 60) or 22.2% of studied tumor lesions (14 out of 63).

The predominant array-CGH profile, detected in 7 out of 63 tumors (11%), was a single copy loss encompassing all measurement points on the chromosome 22 array, consistent with monosomy 22, presumably due to a mitotic non-disjunction event. One representative case from this category 003 T is shown in Fig. 2B. The level of normalized fluorescence ratio for the majority of chromosome 22-derived measurement points was comparable to the level of one copy loss. Furthermore, we identified three cases 036 T, 037 T and 051 T, presenting partial tumor-specific chromosome 22 deletions as the only detectable aberration (Table I and Fig. 1). A profile revealing the presence of a 12 Mb telomeric deletion in female case 051 T is illustrated in Fig. 2C. The level of fluorescence ratio for chromosome 22-

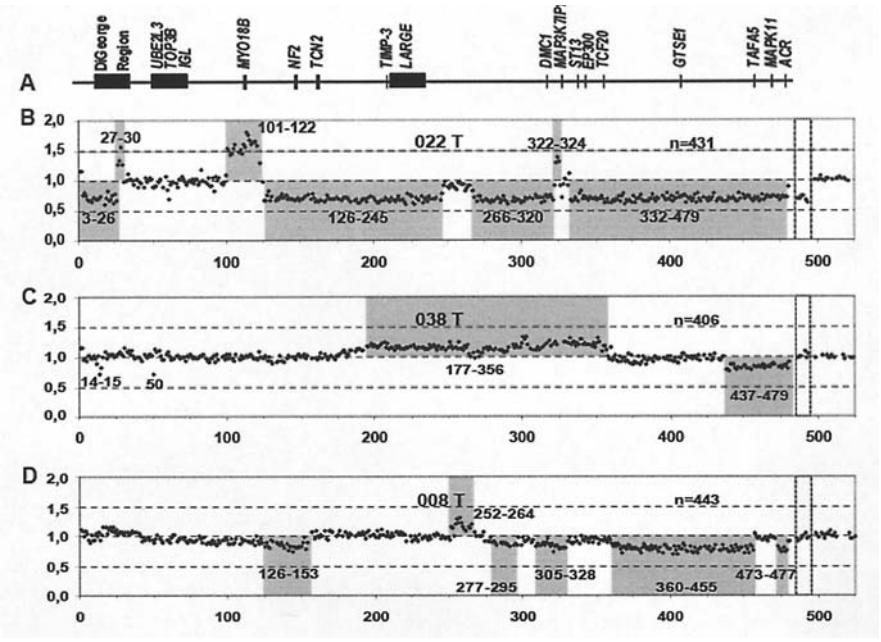


Figure 3. Complex chromosome 22 array-CGH profiles of three breast cancer samples, demonstrate the co-existence of deletions and gains. The general outline of this figure follows the layout of Fig. 2. The grey boxes highlight the position of deletions and gains; with the ID of the clones indicated. (A) Schematic drawing of selected chromosome 22 loci/genes. (B) An interstitial deletion on chromosome 22q together with partial gains in case 022 T. This tumor displays four regional gains, the first one of ~130 kb, encompassing clones ID 27-30 (AC000087-AC000070; 17819985-17947539 bp), at the centromeric part of 22q with an ANILFR of 1.35 ± SD 0.13; the second of ~2 Mb, affecting 21 consecutive clones (ID 101-122; AL022323-AL049536; 23976814-26143705 bp) and consistent with up to 3 gene copies (ANILFR 1.56 ± SD 0.12) and the third of ~310 kb (ID 322-324; AL008583-AL078641; 37493053- 37810389 bp) with the fluorescence ratio 1.32 ± SD 0.05. This sample also displays partial deletions along the whole long arm of chromosome 22, which display a fluorescent ratio consistent with haploid level (ANILFR 0.69 ± SD 0.05). Notably, the chromosome X values (ANILFR 0.69 ± SD 0.04), in this female patient, are lower than expected and consistent with single DNA copy level. (C) Co-occurrence of deletion and gain in tumor case 038 T. The identified ~2.8 Mb deletion on the telomeric end of 22q, encompassed 43 clones (ID 437-479; AL110122-AC002056; 46622060-49495206 bp, ANILFR 0.8 ± SD 0.04). The ANILFR value is not consistent with the deletion being present in all cells from the sample studied, suggesting heterogeneity of this sample. This tumor also displays a regional deletion in the vicinity of the centromere ~190 kb, affecting two consecutive clones on the array (ID 14-15; AC008132-AC008103; 17071632-17306230 bp), with a fluorescence ratio consistent with the haploid level (0.80 ± SD 0.05). A single clone ID 50 (AC018751; 20223407-20391344 bp) with a fluorescent ratio consistent with haploid level is also indicated. A regional gain of ~10.6 Mb, affecting clones ID 177-356 (AL096701-Z99716; 30121563-40798159 bp) identified in this sample is also indicated. Chromosome 22 data points scored at the diploid level exhibit a fluorescence ratio similar to the value of X controls and non-chromosome 22 controls (ANILFR 1.03 ± SD 0.09, 1.08 ± SD 0.04 and 0.99 ± SD 0.03, respectively). (D) Several interstitial deletions on chromosome 22q together with a partial gain in case 008 T. The smallest regional deletion of ~220 kb, encompassing five consecutive clones ID 473-477 (Z94802-AC000050; 49219663-49434548 bp, ANILFR 0.77 ± SD 0.04) was present at the terminal end of 22q. This tumor displays also four additional regional deletions, the first one of ~2Mb at the centromeric part of 22q (clones ID 126-153; AL117256-AC005527; 26333601-28342153 bp), the second of ~1 Mb, affecting 18 consecutive clones (ID 277-295; AL022334-AL008637; 34214125-AL008637 bp), the third of ~2 Mb (clones ID 305-328; AL022315-Z83845; 36218147-38179916 bp, ANILFR 0.84 ± SD 0.05) and the fourth of ~7 Mb (clones ID 360-455; Z83851-AL078613; 40972968-47675305 bp, ANILFR of 0.79 ± SD 0.05). All these deletions display ANILFR values not consistent with the deletions being present in all cells from the sample studied, suggesting heterogeneity of this sample. The profile also displays a regional gain of ~490 kb, affecting clones ID 252-264 (Z68323-AL021877; 32898948-33388070 bp), with the fluorescence ratio ANILFR 1.26 ± SD 0.04. The chromosome X and non-22q autosomal controls are as expected at the diploid level.

derived measurement points included in this deleted region was, however, higher than expected for the haploid level. This suggests that this deletion is not present in all tumor cells within this tumor biopsy. Alternatively, this biopsy might contain a considerable percentage of normal diploid stromal cells. Analysis of paired blood DNA from cases 003 T, 011 T, 051 T and the corresponding surrounding healthy tissue from cases 003 T revealed no detectable DNA copy number imbalances (not shown), which confirmed that the observed deletions were tumor-specific.

Complex chromosome 22 array-CGH profiles - co-existence of deletions and gains. Three cases were identified with very complex array-CGH profiles, in which partial deletions of the chromosome were combined with independent peaks of gain (Fig. 3B-D). In case 022 T, three different loci, affected by regional gain were identified, encompassing four, twenty-one and three consecutive clones on the array (Fig. 3B). The size of these regional gains was ~130 kb, ~2 Mb and ~310 kb, respectively and the maximum average fluorescence ratio observed was of ~1.8 (ID 114; AL035044), corresponding to, on average, >3 copies of this locus in each tumor cell. It should be stressed that each of these consecutive and independent measurement points along the chromosome function as an internal control for other clones within the locus affected by the aberration. In addition, four interstitial deletions were observed in this sample. Non-chromosome 22 autosomal controls were at the diploid level (ANILFR 1.01 ± SD 0.02). However, chromosome X control clones suggested one copy loss of the measurement points located on this chromosome (ANILFR 0.69 ± SD 0.04). Case 038 T presented a noticeable low-copy number gain of ~10.6 Mb in size (Fig. 3C). A small terminal deletion of ~2.8 Mb, extending from ID 437 to ID 479 (AL110122-AC002056) was also observed in this case. The level of fluorescence ratio for measurement points within this deleted region was, however,

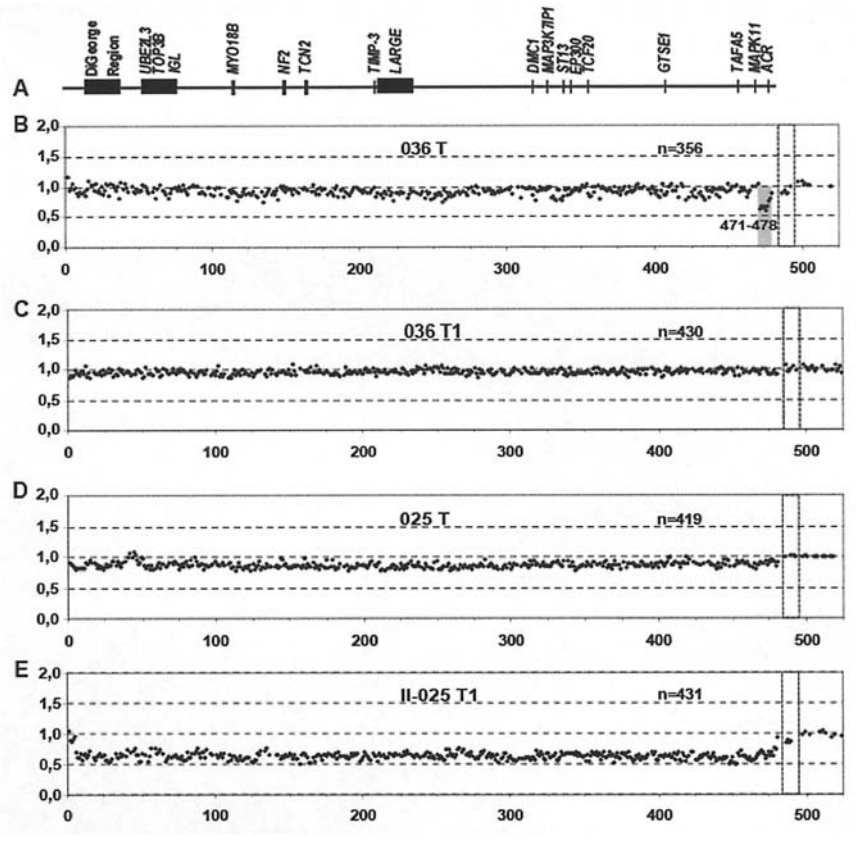


Figure 4. Identification of intra-tumor genetic heterogeneity in breast tumors. The general outline of this figure follows Fig. 2. (A) Selected chromosome 22 loci/genes. (B and C) Array-CGH profiles of two different biopsies derived from a single tumor in case 036. (B) Sample 036 T presented a regional deletion of ~300 kb encompassing seven consecutive clones (AL096767-AC000036; 49153145-49470169 bp) at the telomeric end of 22q (ANILFR 0.67 ± SD 0.07). (C) Sample 036 T1 derived from a different biopsy within the same tumor presented a diploid profile. The values of chromosome 22-derived loci were similar to that of chromosome X and non-chromosome 22 controls (ANILFR 0.96 ± SD 0.04, 0.96 ± SD 0.05 and 1.00 ± SD 0.04 respectively). (D and E) Array-CGH profiles of two different tumors located within the same breast and operated at the same time from patient 025. (D) Detection of monosomy in sample 025 T. The ANILFR values of chromosome 22-derived loci (0.87 ± SD 0.04) are in between haploid and diploid level and consistent with a mixed population of tumor and normal stromal cells. Alternatively, this profile represents a mixture of tumor cells with different genetic profiles. ANILFR values for the chromosome X and autosomal controls (1.07 ± SD 0.05 and 1.08 ± SD 0.04, respectively) are, as expected, at the diploid level. (E) Sample II-025 T1 presented similar findings encompassing almost the entire 22q and consistent with monosomy 22. The fluorescence ratio for clones in the immediate vicinity of the centromere and telomere of 22q are at the diploid level, most likely due to copy number variation within these sequences, which are rich in repeats and segmental duplications. Similar results have been observed earlier upon analyses of chromosome 22 and chromosome 1 in the context of tumor-related studies (28). The chromosome 22 data points scored as deleted display a fluorescence ratio consistent with the haploid DNA level (ANILFR 0.63 ± SD 0.05).

higher than expected for the haploid level. Clones ID 14-15 presented a fluorescence ratio consistent with hemizygous deletion, however this locus has been previously shown to be polymorphic (23). Tumor sample 008 T also displayed a complex pattern of aberrations, with at least five different loci affected by interstitial deletions in the proximal and terminal half of chromosome 22 (Fig. 3D). Additionally, a regional low copy number gain of ~490 kb, affecting clones ID 252-264 (Z68323-AL021877) was also identified in this sample. The analysis of paired blood DNA and the corresponding surrounding healthy breast tissue from these cases revealed no detectable DNA copy number imbalances (not shown). The large complexity of co-existing aberrations in the latter three cases calls for caution in interpreting the biological significance of these results. However, one overall conclusion can be drawn from analysis described so far. All but one of the tumors (biopsies 037 T and 037 T1) (Fig. 1), i.e. 13 out of 63 (20%), that displayed aberrations on chromosome 22, involved the most telomeric 22q region, encompassing clones ID 473-477. This indicates a relatively well defined locus of 220 kb that should be studied further.

Evidence for clonal intra-tumor genotypic differences for aberrations affecting 22q. The availability of multiple biopsies from a single tumor provides an excellent opportunity for the analysis of possible intra-tumor differences in genetic profiles. In 15 independent tumor lesions we had access to two or three biopsies derived from the same tumor and these were independently profiled for DNA copy number changes. Out of these 15 tumors, where at least two histopathologically controlled biopsies were analyzed, four tumors displayed differences in genetic profiles on chromosome 22 (036 T, 037 T, 056 T and 065 T) (Table I and Figs. 1 and 4). Furthermore, from two cases (028 and 025) we collected two, respectively three, samples derived from independent tumors growing within the same breast (Table I). Array-CGH profiles for samples 036 T and T1, illustrating intra-tumor differences, and for patient 025 where two independent tumor lesions were studied are displayed in Fig. 4B-E. Profiling of female case 036 T revealed a minute interstitial hemizygous deletion of ~300 kb, encompassing seven clones (ID 471-478; AL022328-AC000036) on the array and mapping to the telomeric end of 22q. Two additional surgical components (036 T1 and 036 T2),

obtained from the same tumor, demonstrated no detectable aberrations for all chromosome 22-derived measurement points (Table I and Fig. 4). It should be emphasized that the deletion in case 036 T overlaps with the small region of allelic loss detected at the terminal end of 22q in case 008 T (Fig. 3). The analysis of the corresponding surrounding healthy breast tissue from case 036 revealed no detectable DNA copy number imbalances (not shown). Another intriguing case is tumor 037, where distinct profiles of aberrations within the proximal half of 22q were identified in two distinct biopsies. Sample 037 T presented a centromeric hemizygous deletion of ~6.4 Mb, affecting 62 clones (ID 1-62; AP000525-AP000358) on the array. Profiling of the surgical sample obtained from the second biopsy disclosed a noticeable gain of ~4.7 Mb in size, affecting numerous consecutive clones (ID 70-120; D87014-AL021153) (Table I and Fig. 1) (profiles not shown). The latter region is also overlapping with a regional gain detected in case 022 T (Fig. 3B). It should be also noted that the centromeric breakpoint of gain in sample 037 T1 clearly starts after the terminal breakpoint of regional deletion identified in sample 037 T. In two additional cases (056 T and 065 T) similar results were obtained with regard to presence of findings compatible with monosomy 22 in one biopsy and no detectable aberrations in the remaining samples from the same tumor (Table I) (profiles not shown). Thus, in summary, four out of 15 (26.6%) tumors displayed indications of clonal intra-tumor genotypic differences, which should be viewed as a high number, considering that we studied in detail only a single human chromosome.

Minimal common regions of genomic imbalances on chromosome 22q in breast cancer. Detection of common segments of allelic loss or gain on a chromosome is usually considered as indicative for the presence of genes involved in the tumorigenic process. In this study, the distribution of overlapping regions of deletion across the chromosome 22 allowed to delimit at least six distinct minimal common areas of loss (I, III, V, VI, VIII and IX; Fig. 1). However, several of these regions are delimited based on only a few tumors, which often displayed a complex pattern of aberrations. Furthermore, we studied in detail ~1% of the genome and we therefore lack a broader perspective. Nevertheless, the most commonly observed aberration was a deletion of ~220 kb, affecting five consecutive clones ID 473-477 (Z94802-AC000050), and mapping to a gene rich region at the telomere of 22q; candidate region IX. This locus is determined by the telomeric deletion detected in case 008 T (Fig. 3D), which is also encompassed by the heterozygously deleted region identified in sample 022 T, 036 T, 038 T and 051 T, in addition to cases with monosomy 22 (Figs. 2-4 and Table I). A schematic view of the genes and clones included in this minimum overlapping region is presented in supplementary Fig. 1 (http://puffer.genpat.uu.se/publications/supplement/breastcancer/Fig_5.pdf). We have also performed expression analysis *in silico* of these genes using information from Oncomine-Cancer profiling database, http://www.oncomine.org, which revealed that three genes (*SCO2, ECGF1* and *MGC16635*) from this region were found to be under-expressed in different types or stages of breast carcinoma (supplementary Table I; http://puffer.genpat.uu.se/publications/supplement/breastcancer/Table_2.pdf). This makes them plausible candidates for tumor suppressors and warrant further studies. The distribution of the overlapping regions affected by regional gain identified candidate regions along the chromosome (regions II, IV and VII, Fig. 1). Using publicly available databases, we assessed the content of genes with possible implication in cancer within these segments of chromosome 22 affected by overlapping genomic imbalances (genome build 35.1, http://www.ncbi.nlm.nih.gov/) and identified several known genes. We also applied our chromosome 22 genomic clone-based array to profile blood derived DNA from a series of 45 normal individuals and further confirmed that these loci did not vary in DNA copy number (data not shown). This analysis therefore indicates that these specific variations are not frequent DNA copy number polymorphisms (CNP) in normal population.

Discussion

Human chromosome 22 is very rich in cancer related genes. A recent global review on genes that have been shown to be mutated in various malignancies, places chromosome 22 in the top position in the human genome, when number of cancer-related genes versus all genes located on this autosome are considered (24). However, there is also a consensus that chromosome 22 harbors a number of additional, not yet characterized genes important for various forms of human cancer. The main reason for this assumption is that a large number of tumors display specific genetic aberrations of this autosome and do not show evidence of mutations in the already known cancer genes (http://www.ncbi.nlm.nih.gov/entrez/). Breast cancer is among the neoplasms, which have been previously shown to display allelic loss of chromosome 22. The reported frequencies of 22q-associated deletions display great divergence, ranging from 11 to 66% (14-16). In this respect, our comprehensive and high-resolution analysis revealed genomic imbalances of 22q in 21.6% (13 out of 60) of patients or 22% (14 out of 63) of studied tumors. We identified different patterns of aberrations including hemizygous deletions and/or gains of 22q, which were: monosomy, terminal and centromeric deletions, interstitial gains, as well as complex patterns, in which interstitial deletions were combined with low gene copy number gains (Fig. 1 and Table I). However, the most commonly observed aberration, detected in 7 out of 63 tumors (11%), was a single copy loss encompassing all measurement points on the array, consistent with monosomy 22. The latter finding suggests that the involvement of human chromosome 22 in the development/progression of breast cancer should not be viewed from a single gene perspective. Monosomy 22 is most likely the result of mitotic non-disjunction event during divisions of breast cancer cells. The haploinsufficiency for more than 600 known genes from 22q might be an event which as such introduces imbalance in the normal function of the genes encoded from this chromosome. It may further predispose the affected cell to gain further genetic or epigenetic mutations in many other genes residing on the remaining copy of chromosome 22.

Array-CGH is a powerful methodology, which has not yet made a major impact on the genetic research in cancer, mainly due to lack of studies using high-resolution arrays covering the whole human genome. Array-CGH allows studying very large

genomic segments with a high sensitivity in a single experiment. It also permits an independent comparison of tumor and constitutional DNA against a normal unrelated reference DNA, which allows discriminating between aberrations that are tumor-specific and those which might be present in the normal surrounding tissues (25). Furthermore, careful comparison of fluorescence ratios from different experiments allows a rough assessment whether a certain observed genetic aberration is present in all, or in a fraction, of cells within the studied tumor sample. Previous studies of breast cancer-associated deletions using low-resolution methodology failed to define a specific gene, or a region small enough, which would allow testing a reasonable number of candidate genes. This is likely due to the application of sub-optimal technology, such as RFLP- or micro-satellite-based markers. The candidate region IX from this study, encompassing ~220 kb (clone ID 473-477; Z94802-AC000050) (Fig. 1), is defined by the smallest tumor-specific interstitial deletion detected in case 008 T. It should be stressed here that all but one of the tumors, i.e. 13 out of 63 (20%), with aberrations on chromosome 22, involved clones ID 473-477. In one of the tumors (036 T) a minute deletion encompassing this locus was the only copy number aberration detected on 22q (ID 471-478; AL022328-AC000036, Fig. 4B) and present in only one of the two biopsies studied from this tumor lesion. The latter suggests that the gene(s) located there might be related to progression of breast cancer, rather than early predisposing events. Analysis of the corresponding surrounding healthy tissue derived DNA from this case confirmed that the observed allelic loss was also tumor-specific. Further support for the presence of a tumor suppressor gene at this particular location have resulted from the analysis of Wilms tumors samples (26). We recently uncovered, using the same chromosome 22 genomic array, a candidate Wilms tumor locus at the telomeric end of 22q (ID 471-480; AL022328-AC002055), which is entirely in agreement with the candidate region IX presented here (26). The concordance of results from breast cancer and Wilms tumor samples might indicate that a gene(s) located there might play a more general role in tumor development. This ~220 kb segment of 22q is very dense in genes and contains at least 11 genes (in order from centromere to telomere of 22q: encoding hypothetical proteins BC002942 and 384D8_6, *SCO*, *ECGF1* genes, encoding hypothetical protein BC009980, *CHKB-CPT1B, CPT1B, CHKB, MAPK8IP2, ARSA* and *SHANK3* genes) and one predicted gene (*LOC440836*, presumably encoding a protein similar to MGC52679) (http://www.ncbi.nlm.nih.gov/mapview/). Based on the review of available literature and results stored in public databases, a number of these genes present plausible candidates for tumor suppressors, e.g. genes involved in purine/pyrimidine metabolism and different signaling pathways. Interestingly, it has been hypothesized that loss-of-function mutations in *ECGF1* (endothelial cell growth factor 1, platelet-derived) leads to impaired replication or/and maintenance of mtDNA (27). *ECGF1* has also been annotated as a gene with possible implication in cancer according to Atlas Chromosomes in Cancer (http://www.infobiogen.fr/services/chromcancer/). A second important result of our study was the detection of frequent intra-tumoral clonal variation in gene copy number profiles. The number we report (~27%) should be viewed as very high, considering that only a fraction of the genome is profiled in detail. Our findings indicate that a considerable variation should be expected when the whole genome is being profiled in a similar way as performed here for 22q. This is also the next logical step to follow-up our results.

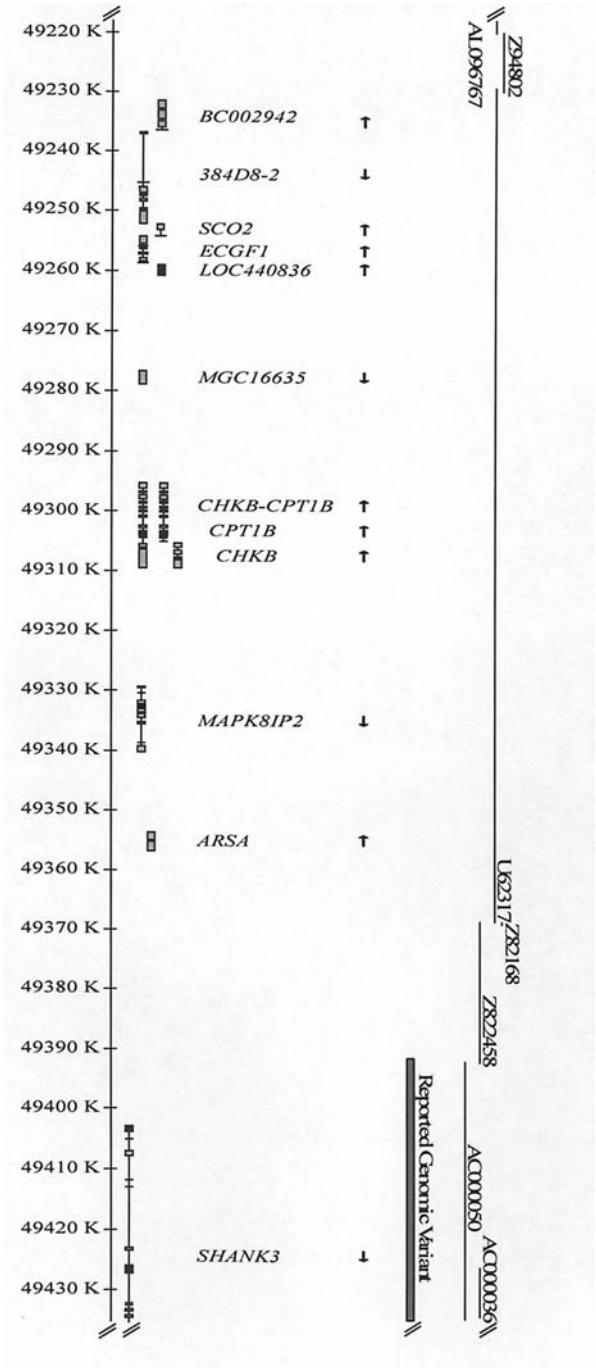
Acknowledgements

This work was supported by grants from the US Army Medical Research and Materiel Command (award no. W81XWH-04-1-0269), the Wallenberg Consortium North, the Swedish Cancer Foundation, the Swedish Research Council and Uppsala University to J.P.D. We thank Dr Kiran Mantripragada for his valuable assistance in printing slides.

References

1. Greenlee RT, Murray T, Bolden S and Wingo PA: Cancer statistics, 2000. CA Cancer J Clin 50: 7-33, 2000.
2. Vogelstein B and Kinzler KW: The genetic basis of human cancer. Vol. section II. McGraw-Hill Professional, pp403-415, 2002.
3. Nathanson KL, Wooster R and Weber BL: Breast cancer genetics: what we know and what we need. Nat Med 7: 552-556, 2001.
4. Lindor NM and Greene MH: The concise handbook of family cancer syndromes. Mayo Familial Cancer Program. J Natl Cancer Inst 90: 1039-1071, 1998.
5. Ford D, Easton DF, Bishop DT, Narod SA and Goldgar DE: Risks of cancer in BRCA1-mutation carriers. Breast Cancer Linkage Consortium. Lancet 343: 692-695, 1994.
6. 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 62: 676-689, 1998.
7. Brose MS, Rebbeck TR, Calzone KA, Stopfer JE, Nathanson KL and Weber BL: Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program. J Natl Cancer Inst 94: 1365-1372, 2002.
8. Easton DF, Bishop DT, Ford D and Crockford GP: Genetic linkage analysis in familial breast and ovarian cancer: results from 214 families. The Breast Cancer Linkage Consortium. Am J Hum Genet 52: 678-701, 1993.
9. Forozan F, Mahlamaki EH, Monni O, *et al*: Comparative genomic hybridization analysis of 38 breast cancer cell lines: a basis for interpreting complementary DNA microarray data. Cancer Res 60: 4519-4525, 2000.
10. Reyal F, Stransky N, Bernard-Pierrot I, *et al*: Visualizing chromosomes as transcriptome correlation maps: evidence of chromosomal domains containing co-expressed genes - a study of 130 invasive ductal breast carcinomas. Cancer Res 65: 1376-1383, 2005.
11. Pollack JR, Perou CM, Alizadeh AA, *et al*: Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 23: 41-46, 1999.
12. Nessling M, Richter K, Schwaenen C, *et al*: Candidate genes in breast cancer revealed by microarray-based comparative genomic hybridization of archived tissue. Cancer Res 65: 439-447, 2005.
13. Mao X, Hamoudi RA, Zhao P and Baudis M: Genetic losses in breast cancer: toward an integrated molecular cytogenetic map. Cancer Genet Cytogenet 160: 141-151, 2005.
14. Bieche I and Lidereau R: Genetic alterations in breast cancer. Genes Chromosomes Cancer 14: 227-251, 1995.
15. Sato T, Tanigami A, Yamakawa K, *et al*: Allelotype of breast cancer: cumulative allele losses promote tumor progression in primary breast cancer. Cancer Res 50: 7184-7189, 1990.
16. Allione F, Eisinger F, Parc P, Noguchi T, Sobol H and Birnbaum D: Loss of heterozygosity at loci from chromosome arm 22Q in human sporadic breast carcinomas. Int J Cancer 75: 181-186, 1998.
17. Castells A, Gusella JF, Ramesh V and Rustgi AK: A region of deletion on chromosome 22q13 is common to human breast and colorectal cancers. Cancer Res 60: 2836-2839, 2000.

18. Iida A, Kurose K, Isobe R, *et al*: Mapping of a new target region of allelic loss to a 2-cM interval at 22q13.1 in primary breast cancer. Genes Chromosomes Cancer 21: 108-112, 1998.
19. Hirano A, Emi M, Tsuneizumi M, *et al*: Allelic losses of loci at 3p25.1, 8p22, 13q12, 17p13.3, and 22q13 correlate with post-operative recurrence in breast cancer. Clin Cancer Res 7: 876-882, 2001.
20. Buckley PG, Mantripragada KK, Benetkiewicz M, *et al*: A full-coverage, high-resolution human chromosome 22 genomic microarray for clinical and research applications. Hum Mol Genet 11: 3221-3229, 2002.
21. Sambrook J, Fritsch E and Maniatis T: Molecular clonning; a Laboratory Manual. Cold Spring Harbor Laboratory Press, New York, 1989.
22. Bruder CE, Ichimura K, Blennow E, *et al*: Severe phenotype of neurofibromatosis type 2 in a patient with a 7.4-MB constitutional deletion on chromosome 22: possible localization of a neurofibromatosis type 2 modifier gene? Genes Chromosomes Cancer 25: 184-190, 1999.
23. De Ståhl TD, Hansson C, De Bustos C, *et al*: High-resolution array-CGH profiling of germline and tumor-specific copy number alterations on chromosome 22 in patients affected with schwannomas. Hum Genet 118: 35-44, 2005.
24. Futreal PA, Coin L, Marshall M, *et al*: A census of human cancer genes. Nat Rev Cancer 4: 177-183, 2004.
25. Mantripragada KK, Buckley PG, De Stahl TD and Dumanski JP: Genomic microarrays in the spotlight. Trends Genet 20: 87-94, 2004.
26. Benetkiewicz M, De Stahl TD, Göndör A, *et al*: Identification of limited regions of genetic aberrations in patients affected with Wilms tumor using a tiling-path chromosome 22 array. Int J Cancer 119: 571-578, 2006.
27. Nishino I, Spinazzola A and Hirano M: Thymidine phosphorylase gene mutations in MNGIE, a human mitochondrial disorder. Science 283: 689-692, 1999.
28. Buckley PG, Jarbo C, Menzel U, *et al*: Comprehensive DNA copy number profiling of meningioma using a chromosome 1 tiling path microarray identifies novel candidate tumor suppressor loci. Cancer Res 65: 2653-2661, 2005.



Supplementary Figure 1. Schematic view of the minimum overlapping region affected by deletion, ~220 kb, affecting five consecutive clones on the array ID 473-477 (Z94802-AC000050) and mapping to the telomere of 22q. The scale denotes the position on the chromosome according to NCBI (*Homo sapiens* Genome, build 35.1, http://www.ncbi.nlm.nih.gov). Genes are drawn to scale and exons represented with filled boxes. The direction of gene transcription is indicated by arrows for each gene (down arrow, positive strand, up arrow: negative strand). The grey bar highlights a CNP, at the telomeric end of the chromosome at 22q13.33, encompassing *SHANK3* locus (gain detected in two individuals) (1). This CNP maps to position 49396359-49493490 bp on 22q, which corresponds to ID 477-479 in our array (AC000050, AC000036 and A002056).

1. Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW and Lee C: Detection of large-scale variation in the human genome. Nat Genet 36: 949-951, 2004.

Supplementary Table I. Genes included in the ~220 kb minimum overlapping 22q deleted region in breast carcinoma.

Genes	Definition	Pathway[a]	Expression Organ/tissue[b]	Differential expression[c,d] Underexpressed	Overexpressed	Refs.
BC002942	Hypothetical protein BC002942	Not defined				
384D8-2	Hypothetical protein 384D8_2	Not defined				
SCO2	SCO cytochrome oxidase deficient homolog 2 (yeast)	Not defined	Heart, liver, brain, kidney, striatum	In ER positive breast cancer		(1)
ECGF1	Endothelial cell growth factor 1 (platelet-derived)	Purine metabolism	Pancreas, brain, prostate, lung, kidney, glands	In sporadic vs breast cancer positive for *BRCA1* mutation		(2)
		Pyrimidine metabolism		In ER positive breast cancer		(3)
				In primary breast cancer vs normal tissue		(4)
LOC440836	Similar to MGC52679 protein	Not defined				
MGC16635	Hypothetical protein BC009980	Not defined		In ER positive breast cancer		(1,3)
				In lymph node negative breast cancer		(2)
CPT1B		Fatty acid metabolism	Heart, striatum		In ER positive breast cancer	(1)
		Adipocytokine signaling pathway			In PR positive breast cancer	(5)
					In non-metastatic breast cancer	(1,3)
					In *E-cadherin* positive breast carcinoma	(6)
CHKB	Choline kinase beta	Glycine, serine and threonine metabolism			In ER positive breast cancer	(1)
		Glycerophospholipid metabolism			In PR positive breast cancer	(5)
					In non-metastatic breast cancer	(3)
					In *E-cadherin* positive breast carcinoma	(6)
MAPK8IP2	Mitogen-activated protein kinase 8 interacting protein 2	MAPK signaling pathway	Brain, pancreas, glands		In ER positive breast cancer	(1)
				In *BRCA1* vs *BRCA2* positive breast cancer)		(2)
				In *TP53* positive breast cancer		(7)
				In metastatic breast cancer		(2)
				In ER positive breast cancer		(3)
					In breast cancer positive for *BRCA2* mutation	(2)
					In breast cancer vs normal breast tissue	(7,8)
ARSA	Arylsulfatase A	Glycosphingolipid metabolism	Liver, brain, kidney	In higher grade breast carcinoma		(5,8)
				In ER positive breast cancer		(9)
				In lymph node positive breast cancer		(10)
					In higher stage breast cancer	(11)
					In *E-cadherin* positive breast cancer	(6)
					In breast cancer vs normal breast tissue	(8)
					In higher grade breast cancer	(6)
SHANK3	SH3 and multiple ankyrin repeat domains 3	Not defined	Brain	In ER positive breast cancer		(3)
				In PR positive breast cancer		(6)
				In metastatic breast cancer		(12)
					In metastatic breast cancer	(3)
					In LC vs DC breast cancer	(7)

[a]According to Kyoto Encyclopedia of Genes and Genomes database: http://www.genome.jp/kegg/. [b]According to GenAtlas database, http://www.dsi.univ-paris5.fr/genatlas. [c]According to Oncomine-Cancer profiling database: http://www.oncomine.org/main/index.jsp. [d]ER, estrogen receptor; PR, progesterone receptor; *E-cadherin*, epithelial cadherin 1; *BRCA1*, breast cancer 1; *BRCA2*, breast cancer 2; *TP53*, tumor protein p53; LC, lobular carcinoma; DC, ductal carcinoma.