ELSEVIER

Cancer Genetics and Cytogenetics 160 (2005) 141–151

Cancer Genetics and Cytogenetics

Genetic losses in breast cancer: toward an integrated molecular cytogenetic map

Xin Mao[a,b,*], Rifat A. Hamoudi[b], Po Zhao[c], Michael Baudis[d]

[a]*Skin Tumour Unit, St. John's Institute of Dermatology, 4th Floor, South Wing, Block 7, St. Thomas' Hospital, Lambeth Palace Road, London SE1 7EH, UK*

[b]*Cancer Gene Cloning Centre, Haddow Laboratories, Institute of Cancer Research, London, UK*

[c]*Department of Pathology, Chinese General Hospital of PLA, Beijing, China*

[d]*Division of Hematology/Oncology, Department of Pediatrics, University of Florida, Gainesville, FL*

Received 23 August 2004; received in revised form 7 December 2004; accepted 23 December 2004

Abstract Breast cancer is the most common malignant disease in Caucasian women, but is less frequent in Chinese women. The molecular basis for such ethnical difference in disease pathogenesis remains unknown. To address this issue, we performed allelotyping analysis of formalin-fixed, paraffin-embedded samples from 21 Chinese patients with breast cancer using 59 fluorescently tagged oligonucleotide primers amplifying microsatellite loci. Loss of heterozygosity (LOH) was found in all tumor samples. Frequent allelic losses were identified at markers D3S1578 (56%); D7S507 (55%); D1S2766 (50%); D17S789 and D17S946 (43% each); D19S814 (35%); D2S162, D13S158 and D13S296 (33% each); D1S551 and D1S2800 (29% each); D3S1597 and D6S260 (22% each); and D1S1588 (21%). To compare our data to previous reports, we determined the band-specific frequency of chromosomal imbalances in breast cancer karyotypes reported in the Mitelman database, and from the CGH results of cases accessible through the Progenetix website. Furthermore, published LOH analyses of breast cancer cases were compared to our own LOH results, demonstrating the most common chromosomal regions affected by allelic losses. The combined results provide a comprehensive view of genetic losses in breast cancers, indicating the comparability of these different techniques and suggesting the presence of a distinct subset of breast cancers with high-frequency LOH at chromosomes 1 and 2p in Chinese patients.

1. Introduction

Breast cancer is the most frequent neoplasm in women from Western countries, with a cumulative lifetime breast cancer risk of about 1 in 10 [1]. Most cases are sporadic, but familial clustering is observed in ~20% of patients. At least 5–10% of cases appear to be a result of inheritance of an autosomal dominant gene [2]. Breast cancer is less commonly seen in Chinese women [3], with an annual incidence of 25.1 per 100,000 [4]. Although previous studies in Chinese patients have shown a variety of genetic alterations, such as deletion of chromosomes 1, 3p, and 6q [5], allelic losses at 3p24, 7q31, 13q12~q14, 16q24.3, 17p13.1, and 17q11.2~q21 [6–9], and mutations of the *BRCA1* and *BRCA2* genes [10–14], the molecular basis for the ethnic difference in disease pathogenesis remains obscure.

Most malignancies accumulate a series of genetic events that play a role in the development of their malignant phenotype, including loss of tumor suppressor genes (TSGs) and activation of oncogenes. These genetic changes can be detected by different techniques including conventional cytogenetics (Giemsa banding of metaphase spreads after short-term culture), comparative genomic hybridization (CGH), and allelotyping. Several publicly accessible databases and reviews of previously published karyotypes, CGH profiles, and loss of heterozygosity (LOH) reports of a variety of malignancies exist [15–22]; however, none of these has put together the cytogenetic [banding, CGH, multiplex fluorescent in situ hybridization (M-FISH)] and molecular (LOH, genomic and expression microarrays) data of a specific tumor, such as would create a user-friendly map in a single setting. Previously, we constructed the first integrated molecular cytogenetic map for Sézary syndrome via this

* Corresponding author. Tel.: +44-(0)207-188-6253; fax: +44-(0)207-188-6377.

E-mail address: mxmayo@yahoo.co.uk or mxmayo@fsmail.net (X. Mao).

0165-4608/05/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.cancergencyto.2004.12.018

approach, enabling the direct assessment of genetic alterations at chromosomal and molecular levels [23], and this provide a basis for the comparison between different techniques to create integrated molecular cytogenetic maps for different tumors.

The present study was aimed at molecular detection of differences possibly underlying the observed different incidence of breast cancer in China compared to Western countries, then expanding the scope of our integrated molecular cytogenetic cancer maps. We initially conducted allelotyping analysis of formalin-fixed, paraffin-embedded samples from 21 Chinese patients with breast cancer using 59 fluorescently tagged oligonucleotide primers amplifying microsatellite loci. The LOH data from these experiments were combined with published cytogenetic, CGH, and allelotyping data of breast cancer by the use of dedicated karyotype parsing software and conventional literature searching.

2. Materials and methods

2.1. Allelotyping

2.1.1. Specimens and DNA extraction

Breast cancer cases were collected from Chinese patients (Han nationality) who were admitted to the General Hospital of the People's Liberation Army, Beijing, China, from 1987 to 1997. Of these samples, only those well-preserved and clearly marked paraffin blocks with complete clinicopathological data were included in the study from 21 patients (Table 1). Tumor samples were dissected from the paraffin sections of histologically diagnosed breast cancer and the normal control samples were dissected from the paraffin sections of uninvolved breast tissues of the same individual. Dissected samples were incubated in 10 mmol/L Tris HCl (pH 7.5), 1 mmol/L EDTA, 1% sodium dodecylsulfate, and 500 μg/mL proteinase K at 37°C for 72 hours without deparaffinization treatment. The extraction mixture was then heated at 100°C for 10–15 minutes to inactivate the proteinase K, and was then used for polymerase chain reaction (PCR).

Table 1
Clinicopathological findings in 21 sporadic breast cancers

Case no.	Age	Histology/grade	Metastasis
155039	66	IDC/H	—[a]
155176	52	IDC/H	Left lymph
159581	80	IDC/H	—
171323	53	IDC/H	—
172697	32	IDC/H	—
174486	29	IC/L	—
175114	45	IDC/H	—
217101	79	MC/H	Left lymph
223756	70	MC/H	—
228543	27	IDC/H	—
230453	35	IDC/H	Left breast
233303	52	MC/H	—
244100	48	IC/L	—
244338	45	IC/L	—
251935	62	IC/L	—
253747	50	IDC/H	—
278041	43	MC/H	—
280706	40	MC/H	Left lymph
286957	50	MC/H	Left lymph
288291	54	MC/H	—
289873	32	IDC/H	—

Abbreviations: IC/L, intraductal carcinoma/low grade; IDC/H, invasive ductal carcinoma/high grade; MC/H, medullary carcinoma/high grade.

[a] Blanks indicate only local tumor present.

2.1.2. Primers, PCR, and data analysis

Amplification of 59 microsatellite loci distributed throughout the human genome was conducted using fluorescently tagged oligonucleotide primers (Genset, Evry, France) in an Omnigene thermal cycler (Hybaid, Hampshire, UK) (Table 2). The reaction mixture consisted of 1.5 μL of 10× PCR buffer (750 mmol/L Tris HCl, 0.1% Tween 20, 200 mmol/L ammonium sulfate, 15% $MgCl_2$) (Advanced Biotechnology, UK), 1.5 mmol/L $MgCl_2$, 1.5 μL (2 mmol/L each nucleotide) of dNTPs, 0.3 μL of each primer (5 optical density reading [OD]), 0.15 μL (10 mg/mL) of bovine serum albumin, 0.1 μL of *Taq* polymerase, 9.25 μL of water, and 1 μL of DNA. PCR conditions were 35–40 cycles of denaturation at 94°C for 1 minute, annealing at the appropriate temperature (50–60°C) for 1 minute, and extension at 72°C for 1 minute, followed by a final extension for 5 minutes at 72°C.

The PCR products were analyzed on a 29:1 (acrylamide–bis) 4.5% polyacrylamide denaturing gels premix (National Diagnostics, Hull, UK) in 1× Tris–boric acid–EDTA buffer using an ABI 377 automated fluorescent DNA sequencer (Applied Biosystems, Foster City, CA). Two microliters of each PCR reaction was combined with 2 μL blue dye with formamide and 0.5 μL of a Tamra fluorescent size marker (Applied Biosystems). This mix was denatured for 10 minutes at 94°C, after which 1.5 μL was loaded into each well on a prewarmed gel on a 36-cm well-to-read plate. The gel was run for 2.5 hours at 200 W power, 60 A current, 2900 V, scan rate of 2400 scans/hour, and 50°C temperature. While the samples were undergoing electrophoresis, the fluorescence was detected in the laser scanning region using filter set C and was collected and stored using GeneScan software 2.0 (version 2.0; Applied Biosystems). The fluorescent gel data collected during the run were automatically analyzed using GeneScan analysis software (version 2.0.2; Applied Biosystems) at the end of the run. Each fluorescent peak was quantitated in terms of peak height and peak area. The results were then imported into Genotyper (version 1.1.1; Applied Biosystems) for further analysis.

The comparison of the ratios between tumors (T) and their controls (N) was made using two formulas for calculation:

$$(T_1/T_2)/(N_1/N_2) \quad (1)$$

$$(T_2/T_1)/(N_2/N_1) \quad (2)$$

where T_1 and N_1 are the peak height of the smaller allele, and T_2 and N_2 are the peak height of the larger allele. Formula

Table 2
A summary of allelotyping of Chinese breast cancers

Loci	Genetic, cM	Cytogenetic	Informative cases		Cases with LOH	
			no./total	%	no.	%
D1S199	47.7	1p35.3~p35.1	16/17	94	1	6
D1S209	95.9	1p32.3~p31.3	12/17	71	0	0
D1S551	97.9	1p22	12/21	67	4	29
D1S2766	100.4	1p22	10/21	48	5	50
D1S1588	104.5	1p22	19/21	90	4	21
D1S397	185.7	1q25	12/21	57	2	17
D1S2800	256.1	1q41~q42.13	7/17	41	2	29
D2S162	21.3	2p25.1	6/17	35	2	33
D2S112	145.8	2q21.1	14/17	82	1	7
D2S138	191.8	2q31.2~q32.1	4/17	24	0	0
D3S1597	24.1	3p26~p25	9/17	53	2	22
D3S1286	35.8	3p24.3~p24.1	8/17	47	0	0
D3S1578	67.9	3p21.2~p21.1	9/17	53	5	56
D3S1262	207.2	3q26.3~q27	9/17	53	1	11
D4S3039	131.9	4q28.2~q31.1	5/17	29	0	0
D4S1586	146.4	4q28.3~q31.21	9/17	53	0	0
D5S409	109.3	5q14.3~q15	12/17	71	2	17
D5S422	163.9	5q33.1~q34	11/17	65	1	9
D5S394	179.8	5q35.1~q35.2	5/17	29	0	0
D6S260	29.6	6p23~p22.3	9/17	53	2	22
D6S294	78.8	6p12.1~q14.1	15/17	88	1	7
D6S264	179.1	6q27	10/17	59	0	0
D7S507	29.1	7p21.2~p15.3	11/17	65	6	55
D7S492	100.5	7q21.11	4/17	24	0	0
D7S489	101	7q21.11	1/17	6	0	0
D7S495	147	7q31.31~q31.33	9/17	53	1	11
D8S281	122.6	8q23.3~q24.12	4/17	24	0	0
D8S1793	136.5	8q24.13	15/17	88	1	7
D9S274	27.8	9p22.1~p21.1	5/17	29	0	0
D10S249	0	10p15.3	16/17	94	2	13
D10S539	75.4	10q21.1	12/17	71	0	0
D10S574	124.4	10q24.1~q25.1	15/17	88	0	0
D10S187	143.9	10q25.3~q26.11	11/17	65	0	0
D11S902	24.7	11p15.3~p15.2	14/17	82	1	7
D12S1635	66.8	12q11~q13	1/17	6	0	0
D12S105	118.9	12q22	11/17	65	1	9
D13S221	14.6	13q12.13~q13.1	1/17	6	0	0
D13S1229	23.5	13q13.2	9/17	53	1	11
D13S269	58.5	13q21.23~q21.33	11/17	65	4	36
D13S158	86.9	13q22.3~q32.1	14/17	82	5	36
D14S274	53.8	14q22.1~q23.3	5/17	29	0	0
D14S256	86.8	14q24.3~q31.1	3/17	18	0	0
D15S132	44.9	15q21.2~q21.3	17/17	100	2	12
D15S158	84.8	15q25.2~q26.1	11/17	65	1	9
D16S519	19.7	16p13.3~p13.13	9/17	53	1	11
D16S419	65.8	16q12.2~q22.1	11/17	65	0	0
D17S786	18.1	17p13.2~p13.1	14/17	82	6	43
D17S946	61	17q21.1~q21.2	14/17	82	6	43
D17S795	90.2	17q23.3~q24.3	4/17	24	0	0
D18S58	109.1	18q22.3~q23	6/17	35	1	17
D18S70	123.8	18q23	14/17	82	2	14
D19S814	0	19p13.3	17/17	100	6	35
D19S886	0	19p13.3	12/17	71	2	17
D19S883	5.5	19p13.3	13/17	76	2	15
D19S424	10.8	19p13.3	14/17	82	1	7
D20S112	39.3	20p11.22~p11.21	12/17	71	1	8
D20S171	94.4	20q13.13~q13.2	13/17	76	1	8
D21S1260	51.6	21q22.3~qter	6/17	35	1	17
D22S315	16.2	22q11.2	8/17	47	1	13

Abbreviations: LOH, loss of heterozygosity.

(1) was used to calculate the ratio of the smaller allele; formula (2) was used to calculate the ratio of the larger allele. For ratios greater than 1, the reciprocal of the ratio is calculated to give a value between 0.00 and 1.00. A value of $\leqslant$0.25 was assigned as indicative of LOH [24–26].

To exclude the possibility of field effect of uninvolved breast on the determination of LOH, multiple normal samples from different sites including breast tissue of the same individuals were tested with the same microsatellite markers. No LOH or abnormal band shift (microsatellite instability) were detected, indicating that the field effect is insignificant for the present study.

2.2. Literature survey and analysis

2.2.1. Cytogenetics

A total of 1,065 adenocarcinomas of the breast were selected for creation of profiles of genomic losses. Of those, 715 cases analyzed with Giemsa banding were derived from the Mitelman database [21]; 350 cases analyzed with CGH were accessed through the Progenetix database [27]. The karyotype annotations in International System for Cytogenetic Nomenclature (ISCN) format were converted to band-specific aberration status information using dedicated parsing algorithms developed for the Progenetix project. A modified version based on these tools (ISCN2matrix converter) is accessible through the project's Web site (http://www.progenetix.net) [22]. For the transformation of the banding data, only the karyotypes of the main clones were evaluated. Parsing of the karyotypes using a high filter stringency (only completely annotated cases, no unresolved marker chromosomes or questionable bands) resulted in the loss of the majority of cases (321 remaining), with most of the remaining cases showing a low karyotype complexity. In the relaxed analysis method used thereafter (all parsable bands, acceptance of "?" marked annotations), a much richer aberration pattern with clear delineation of hot-spot regions could be observed.

For numerical comparison of the CGH and karyotyping data, correlation analysis and two-sided Kolmogorov–Smirnov test of the band-specific loss percentages was performed using Matlab (Mathworks, Natick, MA) software for Macintosh OS X.

2.2.2. LOH

Two recent reviews have independently compiled and assessed 151 LOH studies of breast cancers published to date, using different approaches [15,18]. Based on these two reviews and the data of the present study, an ideogram was manually constructed to show these chromosomal regions comparatively.

3. Results

3.1. Allelotyping

Our previous study showed that DNA extracted from paraffin sections could not yield products >200 bp upon

PCR [25], so genetic markers with product sizes <200 bp were selected for the present study. Initially, 55 fluorescently labeled oligonucleotide primers amplifying microsatellite loci covering the human genome were used to analyze allelic losses in 17 Chinese breast cancer samples. An additional four primer pairs amplifying microsatellite loci flanking the region of *BCL10* at 1p22 [28] and a locus at 1q25 were tested in 21 Chinese breast cancer samples. All these tumor samples were found to have LOH on at least one locus. Frequent allelic losses (>20%) were identified at markers D3S1578 (56%); D7S507 (55%); D1S2766 (50%); D17S789 and D17S946 (43% each); D19S814 (35%); D2S162, D13S158 and D13S296 (33% each); D1S551 and D1S2800 (29% each); D3S1597 and D6S260 (22% each); and D1S1588 (21%) (Table 2). A low level of LOH (<20%) was also noted at several other loci throughout the genome (Table 2). Although patients with high-grade invasive ductal carcinoma and medullary carcinoma seemed to have slightly more LOH than those with low-grade intraductal carcinoma (Table 3), there was no significantly statistical difference among these patients.

3.2. Literature surveys and analysis

3.2.1. Conventional cytogenetics

As noted, breast cancer is not only one of the most important but also one of the most extensively studied malignancies in Western nations, due to its high incidence and prevalence in Caucasian women [1]. The question of why there is ethnic difference in disease occurrence between Chinese women and their Western counterparts could be addressed by comparing genetic changes identified in breast cancers from a high-incidence area with that from China, a low-incidence area.

Since the first karyotype of breast cancer was described by Toews et al. [29] in 1968, there has been a steady increase of the number of cytogenetic report on breast cancer in the literature. According to the Mitelman database [21], by early 2004 there had been 126 publications describing 959 karyotypes of 940 cases of breast tumors. For the present study, we selected 715 cases of breast adenocarcinoma for the assessment of overall chromosome aberrations using Karyotype2matrix software. As Fig. 1 shows, there was a dominant pattern of chromosomal losses affecting each individual chromosome, particularly 1p, 3p, 6q, 8p, 13, 16q, 17p, 18, 19, 22, and X. Gains of 1q, 7, and 8q were also common cytogenetic abnormalities.

3.2.2. CGH

Starting with reports by Kallioniemi and du Manoir over 10 years [30,31], there have been hundreds of publications applying comparative genomic hybridization to virtually every type of human neoplasia. Since December 2000, the Progenetix database project has attempted to collect published CGH results and make them accessible to the scientific community in a format suitable for data-mining procedures [32]. Recently, data obtained using other techniques (banding, array CGH) have also been included. Of the 9,710 cases from 358 publications, CGH data from 350 adenocarcinomas of the breast were selected. To compare overall CGH profiling of these breast cancers to the conventional cytogenetic changes found, ISCN2matrix conversion software was used to analyze these cases. As Fig. 2 shows, DNA copy number gains on 1q, 8q, 11q13, 16p, 17q, and 20q were the dominant genetic alterations, occurring in >15% of the cases. In addition, however, DNA copy number losses were prominent for 8p, 11q23q24, 13q, 16q, and 17p, occurring in >10% of cases each.

3.2.3. Cytogenetic data comparison

From visual inspection of the aberration ideograms generated using Progenetix software packages, differences between the karyotyping and CGH-based data became apparent, with the CGH data showing more distinct regional gains and generally fewer, but regionally pronounced deletions.

In the statistical comparison, the band-specific deletion vectors of the karyotyping and CGH data showed clear correlation in a two-sided Kolmogorov–Smirnov test ($P < 0.001$ for confidence interval CI = 95.5%), as well as a strong similarity using correlation analysis ($P < 0.001$).

3.2.4. LOH

Since the first reports by Ali et al. [33] and Lundberg et al. [34] in 1987, there have been >150 published studies describing LOH in >15,000 breast cancers, and there have also been primary allelotypes in >4,300 cases in the literature [35]. Using integrated chromosome-specific maps of polymorphic markers, Osborne and Hamshere [15] identified 26 chromosome regions as common deletion regions (CDRs) in breast cancers from these individual LOH studies (Fig. 3). More recently, through a likelihood-based approach, Miller et al. [18] found, from these studies, 19 chromosomal regions most likely containing LOH with logarithmic odds (LOD) scores of >3.50 (Fig. 3).

There is a concordance between these two reviews with 11 consistent LOH regions (1p, 2q, 3p, 6q, 7q, 8p, 9p, 13q, 16q, 17p, and 11q). However, some discrepancy is also seen, because 10 CDRs (4q, 5q, 8q, 10q, 11, 14q, 18p, 19p, 21q, and 22q) appear in only one or the other list (Fig. 3). Furthermore, Miller et al. [18] observed that the 7q31.2 region containing *FRA7G* had the highest LOD score of LOH (17.76) among the 19 chromosomal regions most likely harboring a breast TSG, and this is far higher than that of 13q14.11 region (11.06) containing *BRCA2*; however, the 17q21.23 region containing *BRCA1* had a LOD score < 3.50 and it was therefore excluded from their list [18]. In contrast, Osborne and Hamshere noted that there were two CDRs on 13q: one, 13q12.3, containing *BRCA2* and 13q14.2 containing *RB1*, and another on 17q21.23 (*BRCA1*) [15]. Overall, these LOH regions lie within regions of

Table 3
Chromosomal distribution of LOH in 17 Chinese breast cancers

Loci	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
D1S199	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−
D1S551	−	+	−	+	−	−	−	−	+	−	+	−	−	−	−	−	−
D1S552	−	−	+	+	−	+	−	−	−	−	−	−	+	−	−	−	−
D1S553	−	+	+	+	−	−	+	−	−	−	−	−	−	−	−	−	−
D1S554	−	+	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−
D1S555	−	−	+	−	−	−	−	+	−	−	−	−	−	−	−	−	−
D1S556	−	−	−	−	+	+	−	−	−	−	−	−	−	−	−	−	−
D1S557	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−
D1S558	−	−	−	−	−	−	−	+	−	−	−	−	−	+	−	−	−
D1S559	−	−	−	−	−	−	+	−	+	+	−	−	+	−	−	−	−
D1S560	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−
D1S561	+	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−
D1S562	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−
D1S563	−	−	−	−	−	−	+	−	−	−	−	−	+	−	−	−	−
D1S564	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−
D1S565	−	−	+	+	+	+	−	−	−	−	−	−	+	−	+	−	−
D1S566	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−
D1S567	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−
D1S568	−	−	−	−	−	+	+	−	−	−	−	−	−	−	−	−	−
D1S569	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−
D1S570	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−
D1S571	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−
D1S572	−	−	−	−	+	−	−	−	+	+	−	−	−	−	−	+	−
D1S573	−	+	−	−	−	−	−	+	−	+	−	−	+	−	+	−	−
D1S574	−	−	−	−	−	−	−	+	+	−	−	−	−	−	−	−	−
D1S575	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−
D1S576	−	−	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−
D1S577	−	−	−	−	−	−	+	+	+	−	+	−	+	−	+	−	−
D1S578	−	−	+	−	−	−	+	−	+	−	−	+	−	−	+	+	−
D1S579	−	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−
D1S580	−	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	+
D1S581	+	+	+	−	−	−	−	+	+	−	−	−	−	+	−	−	−
D1S582	−	−	+	−	−	−	+	−	−	−	−	−	−	−	−	−	−
D1S583	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−
D1S584	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−	−	−
D1S585	−	−	−	−	−	−	−	−	−	−	−	−	−	+	−	−	−
D1S586	−	−	+	−	−	−	−	−	−	−	−	−	−	−	−	−	−
D1S587	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−
D1S588	−	−	−	−	−	−	−	+	−	−	−	−	−	−	−	−	−

Abbreviations: −, normal; +, loss of heterozygosity (LOH).

Sample numbers 1 to 17 are, respectively: 155039, 155176, 171323, 172697, 174486, 175114, 217101, 223756, 228543, 230453, 233303, 244100, 244338, 251935, 253747, 278041, 286957.

chromosomal losses detected by G-banded karyotyping and CGH (Figs. 1 and 2).

Comparison of our allelotyping results with the reviews shows that there is a consistency in the three most frequent LOH regions (3p, 17p, and 19p; >50%) and in four slightly different regions (1p, 7p, 13q, and 17q) (Fig. 3). In addition, the LOH data from Chinese tumor samples showed a similar peak distribution, with generally lower frequency compared to chromosomal losses from cases collected in the Mitelman database (banding analysis) or the Progenetix project (CGH). For the bands 1p22 (21–50%), 1q25 (17%), 1q41q42 (29%) and 2p25.1 (33%), however, an exceptionally high LOH frequency was observed (Fig. 4).

4. Discussion

To assess differences in the genomic deletion patterns as a possible explanation for the difference in breast cancer incidence observed in China compared to Western countries, we performed allelotyping analysis in breast adenocarcinomas from Chinese patients and compared the results to previous LOH and cytogenetic studies. Frequent allelic losses at 1p, 1q, 2p, 3p, 6p, 7p, 13q, 16q, 17q, and 19p were identified in >20% of cases, which is consistent with previous cytogenetic and molecular studies of Chinese breast cancers [5–9]. In addition, these LOH results are also generally in line with previous LOH and cytogenetic studies in Western countries [15,18,21,22], suggesting the presence of common

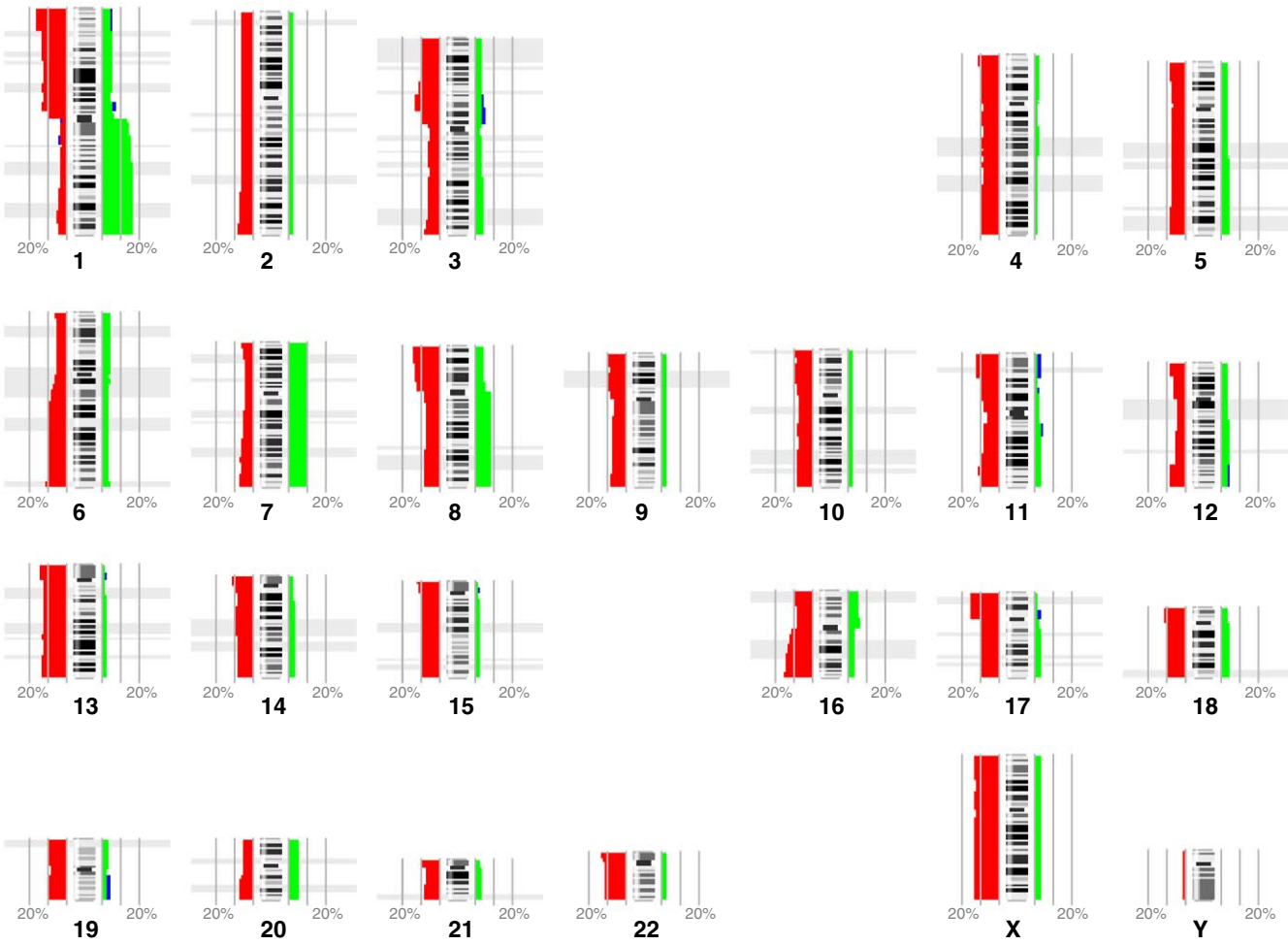

Fig. 1. A summary of karyotypic changes of 715 breast adenocarcinomas automatically retrieved from the Mitelman database [21] using ISCN2matrix software [22]. A dominant pattern of chromosomal losses involving individual chromosomes (bar on the left), and gains of 1q, 7, and 8q (bar on the right) are clearly visible. Shaded background marks regions for which allelotyping was attempted.

molecular defects contributing to the pathogenesis of this malignancy from both high- and low-incidence areas.

Nonetheless, there is obvious discrepancy between this and previous LOH studies, in which the frequent allelic losses at 1p22, 7p21, 13q22, and 17q21.1 detected in Chinese breast cancer samples do not appear among the 26 CDRs described by Osborne and Hamshere [15] or the 19 chromosomal regions with statistically significant LOD score for LOH reported by Miller et al. [18]. This difference may be due to technical variations, given that there are 10 different CDRs reported only by Osborne and Hamshere [15] or only by Miller et al. [18], respectively, despite both of them analyzing the same literature. Furthermore, the majority of previous LOH studies have used either silver or radioactive gels [36], and there might be discrepancies in the interpretation of the respective results. In previous comparative allelotyping studies using these three types of gels, we have observed a high degree of consistency in the determination of LOH in different tumors [24–26,37]. In addition, low-density microsatellite markers used in the present study and heterogeneity in non-Chinese breast cancer cases compiled by Osborne and Hamshere [15] and Miller et al. [18] may contribute to such a difference. It is nonetheless likely, however, that this inconsistency results also from population differences due to genetic variations that may lead to ethnic difference in disease pathogenesis [38].

Deletion distribution in both CGH and banding cases is similar, with very high statistical correlation of distribution of deletions between the data collected with both techniques. The karyotype data, however, apparently show a number of randomly distributed deletions, leading to a higher (on average) but much less pronounced deletion profile compared to the CGH data. The LOH of our present cases actually has the lowest frequencies (with the exception of high frequency of LOH on chromosomes 1 and 2p). There are several possible explanations for such contrasting

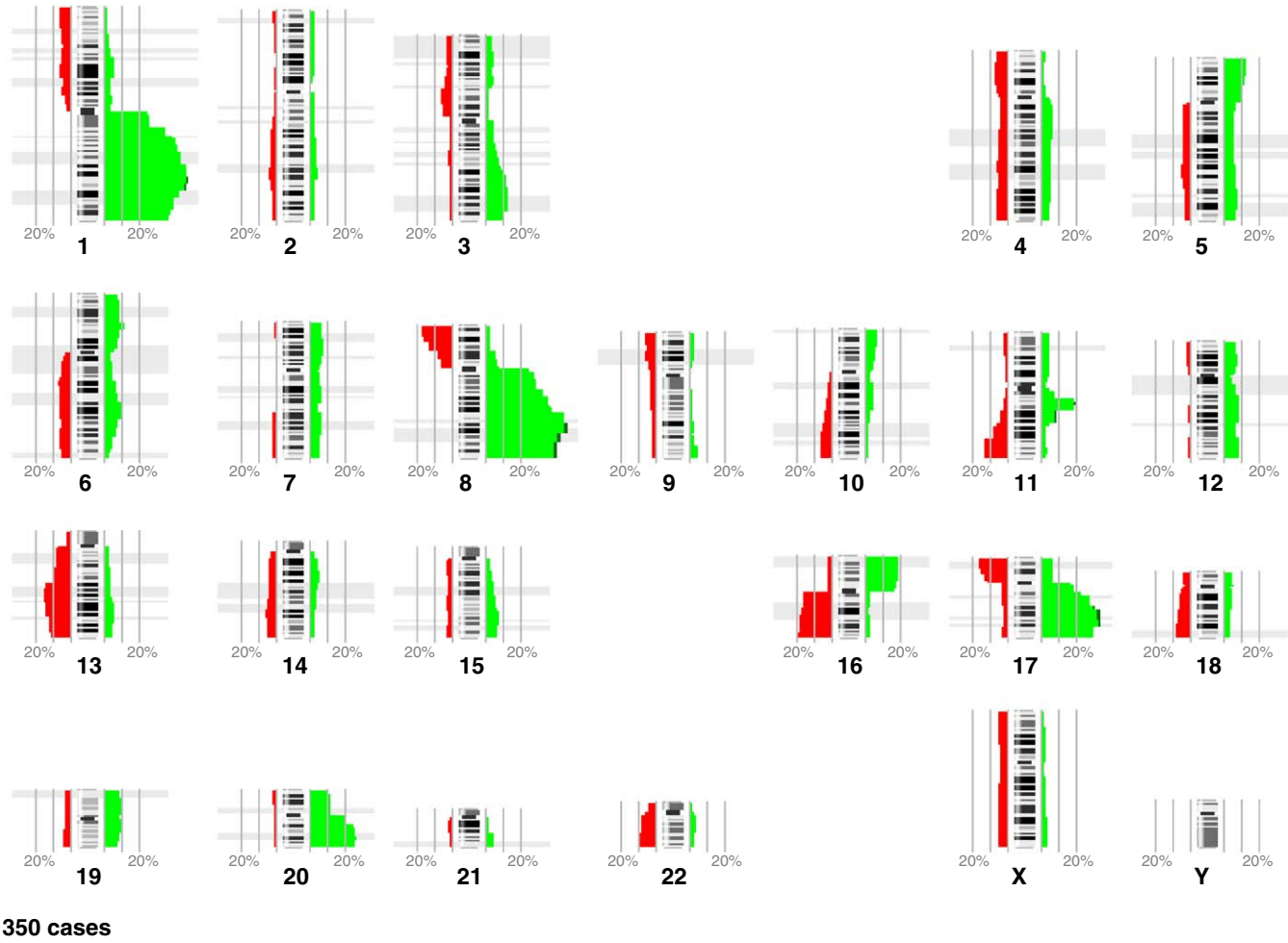

Fig. 2. A summary of CGH profiling of 350 cases of breast cancer from the Progenetix data collection of published CGH cases worldwide and retrieved using ISCN2matrix conversion software [22], revealing the most common DNA copy number changes as being gains of 1q, 8q, 11q13, 16p, 17q, and 20 (bar on the right) and losses of 8p, 11q23q24, 13q, 16q, and 17p (bar on the left), which are clearly comparable with the banding profile of breast adenocarcinomas shown in Fig. 1.

findings among the different techniques. The CGH results could be explained by the absence of a subset of aberrations due to low tumor cell number in the specimen, or to conservative thresholds, or both. In contrast, reasons for the LOH data being low could be the different background or a skewing of the tumor type, stage, and grade; there was a small sample number for the Chinese cases, with different cases used for the different LOH loci, whereas CGH and banding data are essentially complete for each case. Nevertheless, in the present study, the Chinese samples with high frequency of LOH on chromosomes 1 and 2p may represent a generous subset of breast cancers with different genetic–pathological features, in that similar findings have also been described in other Asian breast cancer cases [39–41]. We suggest that further large-scale comparative study using high-density oligonucleotide array-based single nucleotide polymorphism analysis [42] is necessary to clarify this hypothesis.

A second intriguing part of the present study was the construction of integrated molecular cytogenesis maps for breast cancers, based on extensive literature search with the assistance of Karyotype2matrix and ISCN2matrix software. The rationale for producing such maps is twofold. First, breast cancer is one of the well-characterized malignancies in the world, due to its high incidence in the West. There are large numbers of cytogenetic and molecular studies of breast cancers in the literature. To date, however, those reports have not been combined to evaluate the consistency of results derived by a large number of observers using different molecular and cytogenesis techniques. We have previously drawn an integrated molecular cytogenetic map for Sézary syndrome, combining published G-banded karyotypes with our CGH and M-FISH data. This approach allowed us to directly and easily compare chromosomal aberrations in this skin lymphoma at cellular and molecular levels, and showed that the pattern of cytogenetic changes

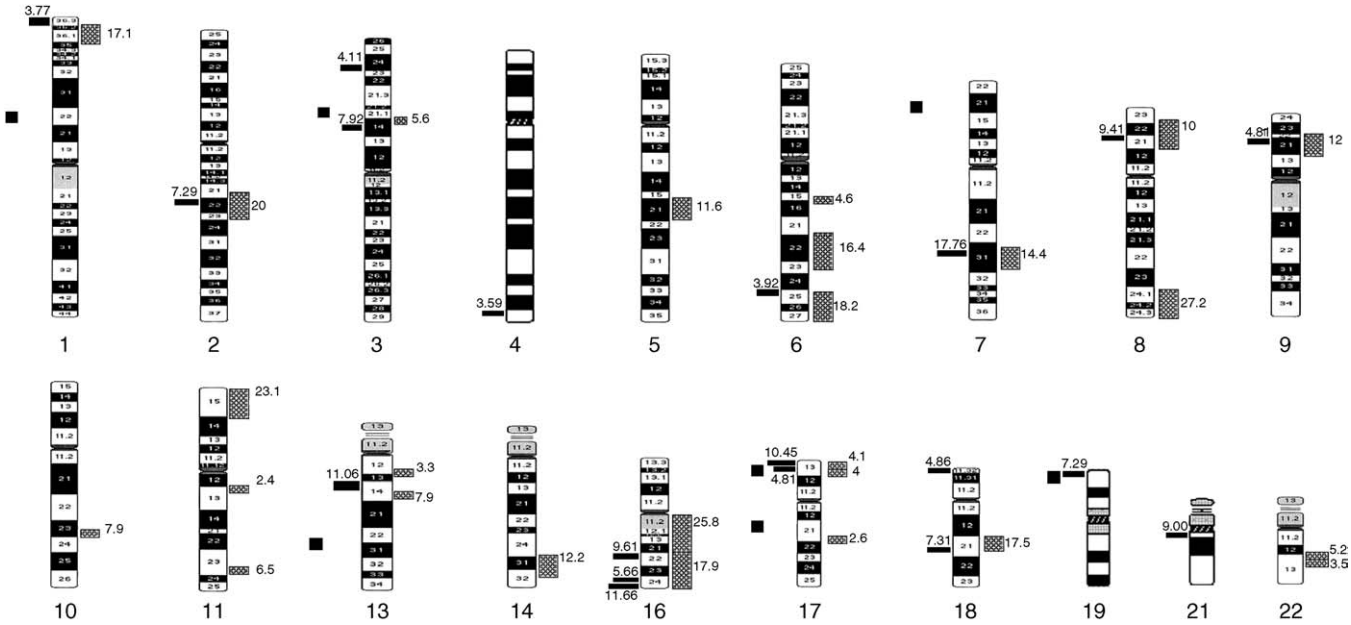

Fig. 3. Illustration of chromosomal regions containing the most frequent LOH in breast cancers. Patterned boxes on the right depict chromosomal deletion regions reported by Osborne and Hamshere [15] and numbers show the size (Mb) of each CDR [36]. The horizontal black bars the left demonstrate the most likely regions of allelic losses described by Miller et al. [18] and the numbers above and below indicates the LOD score of likelihood of LOH. Black squares on the left represent most common LOH (>50%) identified in the present study.

was consistent with the CGH profile [23]. This provides a basis for the systematic comparison between different techniques to create integrated molecular cytogenetic maps for different tumors in terms of large sample size. In the present study, we expanded the maps to cover LOH study, and the results further support our notion that integrated molecular cytogenetic maps could be used for rapid visual assessment of chromosomal and molecular abnormalities in human malignancies.

In the light of Knudson's two-hit hypothesis [43], the first event in tumor formation is a mutation in a specific cancer gene, which in hereditary cancers is present in the germ-line. The second hit in both hereditary and nonhereditary cases would be a somatic mutation leading to functional inactivation of the second copy of the gene [42,43]. This second hit frequently results from a large deletion of chromosomal material as a result of aberrant mitotic recombination or nondisjunction, although several other mechanisms have been suggested [44–46]. This model suggests that frequent or above-background rates of loss of DNA or of LOH at a specific chromosomal locus in a tumor may signify the presence of a TSG in the region of DNA that is lost [47]. Despite the thousands of LOH studies of a variety of cancers [35], no TSG has been identified solely through LOH analysis, which raises concern about the validity and efficiency of allelotyping as a technique for finding TSGs in cancer genetic research [48]. The present study, however, revealed a high degree of similarity between conventional cytogenetic and LOH maps of breast cancers, suggesting that LOH represents part of a spectrum of genetic instabilities, in that numerical chromosomal losses reflect chromosome instabilities in disease pathogenesis, and allelotyping remains a valid technique for the assessment of genetic losses in malignancies and the identification of supporting evidence for TSG.

In the present study, a CGH map showed that the dominant pattern of genetic alterations in a breast cancer cell line is DNA copy number gains, not losses, which is supported by our recent observation in breast cancer cell line. In addition, the similarity of banding and CGH data summary profiles in describing regions of nonrandom genomic deletions should counteract the view on solid tumor cytogenetics as being technically biased, due to the difficulty of obtaining high-quality metaphase chromosome preparations for banding analysis [49].

Newer array- or matrix-CGH (a-CGH) methods have been applied to genomic cancer research [50–52]. We have used this novel technique to analyze gene copy number changes in primary cutaneous lymphomas [53–56]. One question that a-CGH could address would be high-resolution mapping of losses, thereby providing the basis for selection of polymorphisms for allelotyping. For the future, the development of robust, disease-specific analysis approaches based on genomic array technologies [57] is to be expected. It remains to be seen whether a-CGH analysis results will support the more pronounced regional copy number changes observed with CGH compared to banding analysis approaches in terms of a large number cases, as we found in the present study. In addition, to date there have been >1,000 reports on expression microarray analysis of a variety of human malignancies [35]. A recent study has also compiled and analyzed

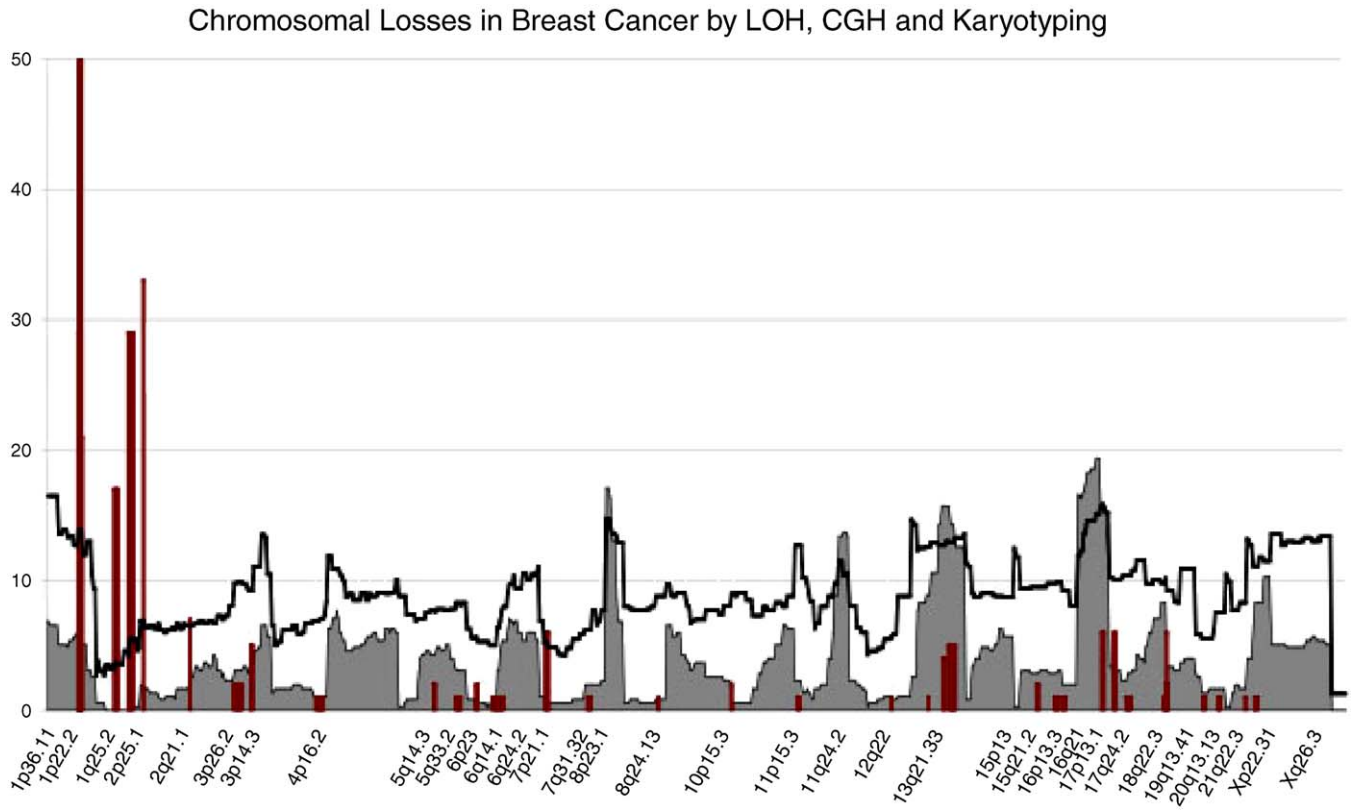

Fig. 4. A comparison of regional LOH in Chinese breast carcinoma cases (bar) with chromosomal losses from cases collected from the Mitelman database [21] (banding, black curve line) and the Progenetix project [22] (CGH, shaded area). Although the absolute values differ between the CGH and the banding data, both datasets show a similar distribution of hot-spot regions. The LOH data from Chinese tumor samples generated in the present study also showed a similar peak distribution with general lower frequency; however, for the bands 1p22 (21–50%), 1q25 (17%), 1q41q42 (29%), and 2p25.1 (33%), an exceptionally high LOH frequency was seen.

all these data revealing common and distinct gene expression patterns and clusters of signaling pathways in different types of cancers [58]. It would be worth knowing whether such global gene expression changes accompany genomic and chromosomal alterations in a specific tumor type by the combination of the integrated molecular cytogenetic maps with the published data of gene expression microarrays with dedicated software.

References

[1] Wingo PA, Tong T, Bolden S. Cancer statistics, 1995. CA Cancer J Clin 1995;45:8–30. [Erratum in: CA Cancer J Clin 1995;45:127–8].

[2] Stratton MR, Wooster R. Hereditary predisposition to breast cancer. Curr Opin Genet Dev 1996;6:93–7.

[3] Ursin G, Bernstein L, Pike MC. Breast cancer. Cancer Surv 19-1994;20:241–64.

[4] Jin F, Devesa SS, Zheng W, Blot WJ, Fraumeni JF Jr, Gao YT. Cancer incidence trends in urban Shanghai, 1972–1989. Int J Cancer 1993; 53:764–70.

[5] Lu YJ, Xiao S, Yan YS, Fu SB, Liu QZ, Li P. Direct chromosome analysis of 50 primary breast carcinomas. Cancer Genet Cytogenet 1993;69:91–9.

[6] Deng GR, He LW, Lin BY. Loss of heterozygosity at different chromosomes in patients with breast cancer. [In Chinese; English abstract]. Chin J Med 1994;74:31–4.

[7] Tseng SL, Yu IC, Yue CT, Chang SF, Chang TM, Wu CW, Shen CY. Allelic loss at *BRCA1*, *BRCA2*, and adjacent loci in relation to *TP53* abnormality in breast cancer. Genes Chromosomes Cancer 1997; 20:377–82.

[8] Cui J, Shen F, Jiang F, Wang Y, Bian J, Shen Z. Loss of heterozygosity and microsatellite instability in the region including *BRCA1* of breast cancer in Chinese. [In Chinese; English abstract]. Chin J Med Genet 1998;15:348–50.

[9] Ju W, Wang J, Li B, Li Z. An epidemiology and molecular genetic study on breast cancer susceptibility. Chin Med Sci J 2000;15:231–7.

[10] Tang NL, Choy KW, Pang CP, Yeo W, Johnson PJ. Prevalence of breast cancer predisposition gene mutations in Chinese women and guidelines for genetic testing. Clin Chim Acta 2001;313:179–85.

[11] Zhi X, Szabo C, Chopin S, Suter N, Wang QS, Ostrander EA, Sinilnikova OM, Lenoir GM, Goldgar D, Shi YR. *BRCA1* and BRCA2 sequence variants in Chinese breast cancer families. Hum Mutat 2002;20:474.

[12] Hu Z, Wu J, Liu CH, Lu JS, Luo JM, Han QX, Shen ZZ, Shao ZM. The analysis of *BRCA1* mutations in eastern Chinese patients with early onset breast cancer and affected relatives. Hum Mutat 2003;22:104.

[13] Suter NM, Ray RM, Hu YW, Lin MG, Porter P, Gao DL, Zaucha RE, Iwasaki LM, Sabacan LP, Langlois MC, Thomas DB, Ostrander EA. *BRCA1* and *BRCA2* mutations in women from Shanghai China. Cancer Epidemiol Biomarkers Prev 2004;13:181–9.

[14] Zhou YZ, Sun Q, Lin SQ, Wang J, Liu B, Li JX, Zhou YD, Ye J, Han H, Fang FD. Germline mutations in the *BRCA1* and *BRCA2* genes from breast cancer families in China Han people. [In Chinese; English abstract]. Zhonghua Yi Xue Za Zhi 2004;84:294–8.

[15] Osborne RJ, Hamshere MG. A genome-wide map showing common regions of loss of heterozygosity/allelic imbalance in breast cancer. Cancer Res 2000;60:3706–12.

[16] Struski S, Doco-Fenzy M, Cornillet-Lefebvre P. Compilation of published comparative genomic hybridization studies. Cancer Genet Cytogenet 2002;135:63–90.

[17] Knuutila S, Aalto Y, Autio K. Online access to CGH data of DNA sequence copy number changes. Am J Pathol 2000;157:689. [Erratum in: Am J Pathol 2000;157:1413].

[18] Miller BJ, Wang D, Krahe R, Wright FA. Pooled analysis of loss of heterozygosity in breast cancer: a genome scan provides comparative evidence for multiple tumor suppressors and identifies novel candidate regions. Am J Hum Genet 2003;73:748–67.

[19] Atlas of Genetics and Cytogenetics in Oncology and Haematology. Available at: http://www.infobiogen.fr/services/chromcancer/index.html. Accessed May 2, 2004.

[20] Cancer Genome Anatomy Project (CGAP), Available at: http://cgap.nci.nih.gov/. Accessed May 2, 2004.

[21] Mitelman F, Johansson B, Mertens F, editors. Mitelman database of chromosome aberrations in cancer [Internet]. Accessed May 2, 2004. Available at: http://cgap.nci.nih.gov/Chromosomes/Mitelman.

[22] Baudis M. Progenetix: cytogenetic abnormalities in human cancer [Internet]. Online database (2000–2004). Available at: http://www.progenetix.net/. Accessed May 2, 2004.

[23] Mao X, Lillington D, Czepulkowski B, Russell-Jones R, Young B, Whittaker S. Molecular cytogenetic characterization of Sezary syndrome. Genes Chromosomes Cancer 2003;36:250–60.

[24] Mao X, Barfoot R, Hamoudi RA, Noble M. Allelotyping of an oligodendrocyte-type-2 astrocyte lineage derived from a human glioblastoma multiforme. J Neurooncol 1998;40:243–50.

[25] Mao X, Barfoot R, Hamoudi RA, Easton DF, Flanagan AM, Stratton MR. Allelotyping analysis of uterine leiomyomas. Cancer Genet Cytogenet 1999;114:89–95.

[26] Mao X, Lillington D, Scarisbrick J, Mitchell T, Czepulkowski B, Russell-Jones R, Young B, Whittaker SJ. Molecular cytogenetic analysis of cutaneous T-cell lymphomas: Identification of common genetic alterations in Sezary syndrome and mycosis fungoides. Br J Dermatol 2002;147:464–75.

[27] Baudis M, Cleary ML. Progenetix.net: an online repository for molecular cytogenetic aberration data. Bioinformatics 2001;17:1228–9.

[28] Willis TG, Jadayel DM, Du MQ, Peng H, Perry AR, Abdul-Rauf M, Price H, Karran L, Majekodunmi O, Wlodarska I, Pan L, Crook T, Hamoudi R, Isaacson PG, Dyer MJ. Bcl10 is involved in t(1;14)(p22;q32) of MALT B cell lymphoma and mutated in multiple tumor types. Cell 1999;96:35–45.

[29] Toews HA, Katayama KP, Masukawa T, Lewison EF. Chromosomes of benign and malignant lesions of the breast. Cancer 1968;22:1296–307.

[30] Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 1992;258:818–21.

[31] du Manoir S, Speicher MR, Joos S, Schröck E, Popp S, Dohner H, Kovacs G, Robert-Nicoud M, Lichter P, Cremer T. Detection of complete and partial chromosome gains and losses by comparative genomic in situ hybridization. Hum Genet 1993;90:590–610.

[32] Baudis M., Cleary M.L.. Progenetix: cytogenetic abnormalities in human cancer. Available at: http://www.progenetix.net/. Accessed May 2, 2004.

[33] Ali IU, Lidereau R, Theillet C, Callahan R. Reduction to homozygosity of genes on chromosome 11 in human breast neoplasia. Science 1987;238:185–8.

[34] Lundberg C, Skoog L, Cavenee WK, Nordenskjold M. Loss of heterozygosity in human ductal breast tumors indicates a recessive mutation on chromosome 13. Proc Natl Acad Sci U S A 1987;84:2372–6.

[35] National Library of Medicine. PubMed. U.S. Institutes of Health. Available at: http://www.ncbi.nih.gov/pubmed/. Accessed May 2, 2004.

[36] Ellsworth RE, Ellsworth DL, Lubert SM, Hooke J, Somiari RI, Shriver CD. High-throughput loss of heterozygosity mapping in 26 commonly deleted regions in breast cancer. Cancer Epidemiol Biomarkers Prev 2003;12:915–9.

[37] Mao X, Jones TA, Tomlinson I, Rowan AJ, Fedorova LI, Zelenin AV, Mao JI, Gutowski NJ, Noble M, Sheer D. Genetic aberrations in glioblastoma multiforme: translocation of chromosome 10 in an O-2A derived cell line. Br J Cancer 1999;79:724–31.

[38] Krieger N, van den Eeden SK, Zava D, Okamoto A. Race/ethnicity, social class, and prevalence of breast cancer prognostic biomarkers: a study of white, black, and Asian women in the San Francisco Bay area. Ethn Dis 1997;7:137–49.

[39] Matsumoto S, Emi M, Kasumi F, Nakamura Y. Genetic alterations and DNA-based diagnosis in breast cancer. [In Japanese; English abstract]. Nippon Geka Gakkai Zasshi 1996;97:375–80.

[40] Takami S, Kawasome C, Kinoshita M, Koyama H, Noguchi S. Chromosomal instability detected by fluorescence in situ hybridization in Japanese breast cancer patients. Clin Chim Acta 2001;308:127–31.

[41] Kang D. Genetic polymorphisms and cancer susceptibility of breast cancer in Korean women. J Biochem Mol Biol 2003;36:28–34.

[42] Wong KK, Tsang YT, Shen J, Cheng RS, Chang YM, Man TK, Lau CC. Allelic imbalance analysis by high-density single-nucleotide polymorphic allele (SNP) array with whole genome amplified DNA. Nucleic Acids Res 2004;32:e69.

[43] Knudson AG. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 1971;68:820–3.

[44] Cavenee WK, Dryja TP, Phillips RA, Benedict WF, Godbout R, Gallie BL, Murphree AL, Strong LC, White RL. Expression of recessive alleles by chromosomal mechanisms in retinoblastoma. Nature 1983;305:779–84.

[45] Huynh H, Alpert L, Pollak M. Silencing of the mammary-derived growth inhibitor (MDGI) gene in breast neoplasms is associated with epigenetic changes. Cancer Res 1996;56:4865–70.

[46] Cappione AJ, French BL, Skuse GR. A potential role for NF1 mRNA editing in the pathogenesis of NF1 tumors. Am J Hum Genet 1997;60:305–12.

[47] Chen LC, Kurisu W, Ljung BM, Goldman ES, Moore D, Smith HS. Heterogeneity for allelic loss in human breast cancer. J Natl Cancer Inst 1992;84:506–10.

[48] Tomlinson IP, Lambros MB, Roylance RR. Loss of heterozygosity analysis: practically and conceptually flawed? Genes Chromosomes Cancer 2002;34:349–53.

[49] Mitelman F, Johansson B, Mertens F. Fusion genes and rearranged genes as a linear function of chromosome aberrations in cancer. Nat Genet 2004;36:331–4.

[50] Solinas-Toldo S, Lampel S, Stilgenbauer S, Nickolenko J, Benner A, Dohner H, Cremer T, Lichter P. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosomes Cancer 1997;20:399–407.

[51] Pinkel D, Segraves R, Sudar D, Clark S, Poole I, Kowbel D, Collins C, Kuo WL, Chen C, Zhai Y, Dairkee SH, Ljung BM, Gray JW, Albertson DG. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 1998;20:207–11.

[52] Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 1999;23:41–6.

[53] Mao X, Lillington D, Child F, Russell-Jones R, Young B, Whittaker S. Comparative genomic hybridization analysis of primary cutaneous B-cell lymphomas: Identification of common genomic alterations in disease pathogenesis. Genes Chromosomes Cancer 2002;35:144–55.

[54] Mao X, Orchard G, Lillington D, Russell-Jones R, Young B, Whittaker S. Amplification and overexpression of JUNB is associated with primary cutaneous T-cell lymphomas. Blood 2003;101:1513–9.
[55] Mao X, Lillington D, Russell-Jones R, Young B, Whittaker S. Genetic alterations in primary cutaneous CD30+ anaplastic large cell lymphoma. Genes Chromosomes Cancer 2003;37:176–85.
[56] Mao X, Onadim Z, Price EA, Child F, Lillington DM, Russell-Jones R, Young BD, Whittaker S. Genomic alterations in blastic natural killer/extranodal natural killer-like T-cell lymphoma with cutaneous involvement. J Invest Dermatol 2003;121:618–27.
[57] Schwaenen C, Nessling M, Wessendorf S, Salvi T, Wrobel G, Radlwimmer B, Kestler HA, Haslinger C, Stilgenbauer S, Dohner H, Bentz M, Lichter P. Automated array-based genomic profiling in chronic lymphocytic leukemia: development of a clinical tool and discovery of recurrent genomic alterations. Proc Natl Acad Sci U S A 2004;101:1039–44.
[58] Segal E, Friedman N, Koller D, Regev A. A module map showing conditional activity of expression modules in cancer. Nat Genet 2004;36:1090–8.