Genome-wide Search for Loss of Heterozygosity Using Laser Capture Microdissected Tissue of Breast Carcinoma: An Implication for Mutator Phenotype and Breast Cancer Pathogenesis

Chen-Yang Shen,1 Jyh-Cherng Yu, Yen-Li Lo, Chia-Hui Kuo, Chung-Tai Yue, Yuh-Shan Jou, Chiu-Sheng Huang, Jia-Chyi Lung, and Cheng-Wen Wu

Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan [C. Y. S., Y. L. L., C. H. K., J. C. L., C. W. W.]; Department of Surgery, Tri-Service General Hospital, Taipei 100, Taiwan [J. C. Y.]; Department of Pathology, Shin Kong Wu Ho-Su Memorial Hospital, Taipei 111, Taiwan [C. T. Y.]; National Health Research Institutes, Taipei 11529, Taiwan [Y. S. J., C. W. W.]; and Department of Surgery, National Taiwan University Hospital, Taipei 100, Taiwan [C. S. H.]

ABSTRACT

Breast cancer is considered to display a high degree of intratumor heterogeneity, without any obvious morphological and pathological steps to define sequential evolution, and its progression may vary among individual tumors. In an attempt to elucidate these etiological and phenotypic complexities, the present study, based on the fundamental concept that genomic instability is the engine of both tumor progression and tumor heterogeneity, was conducted to test the hypothesis that breast cancer pathogenesis is driven by double-strand break (DSB)-initiated chromosome instability (CIN). The rationale underlying this hypothesis is derived from the clues provided by family breast cancer syndromes, in which susceptibility genes, including p53, ATM, BRCA1 and BRCA2, are involved within the common functional pathway of DSB-related checkpoint/repair. Because genomic deletion caused by DSB is reflected in the genetic mechanism of loss of heterozygosity (LOH), this genome-wide LOH study was conducted, using 100 tumors and 400 microsatellite markers. To minimize the effect of heterogeneity within tumors, the experimental technique of laser capture microdissection was used to ensure that genetic and phenotypic examinations were based on the same tumor cells. Support for our hypothesis comes from the observations that: (a) the extent of DSB-initiated CIN in tumors significantly increased as tumors progressed to poorer grades or later stages; (b) in the sequential steps toward CIN, the loci of p53 and ATM, the key checkpoint genes against DSB, were lost at the earliest stage; and (c) many loci identified to be important in breast tumorigenesis were the genomic sites possibly harboring the genes involved in DSB-related checkpoint/repair (including RAD51, RAD52, and BRCA1) or CIN (including FA-A, FA-D, and WRN), and a higher number of these loci showing LOH was significantly associated with increased level of DSB-initiated CIN (P < 0.0001). Breast cancers are thus considered to be sequentially progressive with CIN. However, CIN might also cause genetic heterogeneity, which was revealed by the findings that LOH at some markers was observed only in the component of ductal carcinoma in situ but not in the invasive component of the same tumors. In addition, some markers were found to preferentially lose at specific tumor grades, implying their contribution to genetic heterogeneity during tumor development. Therefore, this study suggests that breast cancer progression is clonal with regard to CIN, but different breast cancers would present distinct molecular profiles resulting from genetic heterogeneity caused by CIN.

INTRODUCTION

Breast carcinoma is one of the most common cancers in women and is known to arise from a multifactorial process, the effect of reproductive risk factors strongly supporting a hormonal role in its etiology (reviewed in Ref. 1). Early menopause, for example, whether occurring naturally or as a result of oophorectomy, has been shown to significantly reduce the risk. However, these reproductive risk factors have been shown to only moderately increase the risk of developing breast cancer. More frustrating from a cancer prevention viewpoint is that two-thirds of women who develop this disease fit none of the currently identified at-risk groups (2). This has prompted scientists to search for clues at the molecular level that may help in understanding breast tumor pathogenesis.

The greatest insights in our understanding of molecular tumorigenesis have been obtained from the study of family cancer syndromes (3–5). In the case of breast cancer, an inherited component has been suspected for many years because of reports of families with large numbers of affected individuals. In the Li-Fraumeni syndrome, there is a high incidence of premenopausal breast cancer. As a result of the identification of the germline mutations in ~50% of pedigrees, these syndromes were linked to the p53 tumor suppressor gene (6). The documentation of >1000 mutations of the p53, accounting for about 25–30% of all breast cancers tested, has clearly established its role in breast cancer development (7). The inheritance of breast cancer susceptibility in families has also led to the localization of breast cancer susceptibility genes, including BRCA1 localized to chromosome arm 17q and BRCA2 to chromosome arm 13q (reviewed in Ref. 8). Furthermore, the mapping of homozygous deletions on human chromosome 10q23 has resulted in the isolation of a candidate TSG, 2PTEN/MMAC1, that appears to show a considerable frequency of mutation in breast cancer and to be responsible for Cowden disease, an inherited breast and thyroid cancer syndrome (reviewed in Ref. 9). In addition to these highly penetrant genes, other genes, in which mutations might result in a more moderate increase in breast cancer risk, have been predicted; a well-known example is the ataxia telangiectasia (reviewed in Ref. 9). The recent isolation of this gene, designated ATM, supports this hypothesis.

To seek clues to the initiation of breast cancer development, we explored the suggested function of these genes linked to inherited breast cancer syndromes. It is intriguing to find that, except 2PTEN/MMAC1, the products of the other four genes (p53, ATM, BRCA1, and BRCA2) are all directly involved in a common molecular pathway related to cellular responses against DSBs arisen in DNA (reviewed in Refs. 12 and 13). ATM is a component of the cell cycle checkpoint machinery that senses DNA damage and is activated after the formation of DSBs induced by ionizing radiation or other DSB-inducing agents (reviewed in Ref. 11). p53, acting downstream of ATM, is then phosphorylated and consequently triggered to initiate a protective response either by blocking the cell cycle for DNA repair or by

The abbreviations used are: TSG, tumor suppressor gene; DSB, double-strand break; CIN, chromosome instability; LOH, loss of heterozygosity; LCM, laser capture microdissection; DCIS, ductal carcinoma in situ; IDC, invasive ductal carcinoma; CGH, comparative genomic hybridization; MIN, microsatellite instability; NIN, nucleotide instability; PLOH, the proportion of loci showing LOH in a tumor.

2 The abbreviations used are: TSG, tumor suppressor gene; DSB, double-strand break; CIN, chromosome instability; LOH, loss of heterozygosity; LCM, laser capture microdissection; DCIS, ductal carcinoma in situ; IDC, invasive ductal carcinoma; CGH, comparative genomic hybridization; MIN, microsatellite instability; NIN, nucleotide instability; PLOH, the proportion of loci showing LOH in a tumor.

Received 12/17/99; accepted 5/16/00.

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

1 To whom requests for reprints should be addressed, at Institute of Biomedical Sciences, Academia Sinica, Taipei 11529, Taiwan. Fax: 886-2-2782-3047; E-mail: bncys@ccvax.sinica.edu.tw.
inducing apoptosis in genetically damaged cells (reviewed in Refs. 14 and 15). Alternatively, the proteins encoded by BRCA1 or BRCA2 can work in concert with a protein called Rad51, which is known to help repair DSBs in the DNA double helix (reviewed in Refs. 16 and 17). Given the central role of DSB-induced chromosome breakage to cause genomic deletion or nonhomologous recombination (18, 19), we postulated that breast cancer etiology is specifically linked to DSB-initiated CIN, a defining characteristic of mutator phenotype of cancer cells (20). Because genomic deletion or nonhomologous recombination are reflected in a genetic mechanism called LOH (21), this genome-wide LOH study was conducted to test our hypothesis. We sought to determine whether breast cancer progression is associated with an increased level of genomic instability characterized by a specific mutator phenotype manifested as DSB-initiated CIN.

LOH is common to all human solid tumors and allows the expressivity of recessive loss-of-function mutations in TSGs (21). Therefore, the detection of recurrent LOH in a chromosomal region is now considered to be critical evidence for the localization of TSGs. However, two methodological limitations always affect valid assessment of LOH results. Because detection of LOH is based on the comparison of tumor cells and corresponding normal tissues to identify genetic deletion that occurs specifically in tumor tissues, it is essential to be able to obtain pure tumor cells to provide the homogeneous materials required for reliable analysis. However, the infiltrating nature of most invasive solid tumors, including breast cancer, leads to contamination of the surrounding normal tissue and creates technical problems in LOH analysis. To overcome this problem, the labor-intensive, inaccurate, and time-consuming process of tissue microdissection must be used, which originally involved manual or micromanipulator-guided scraping of the area of interest of a thin tissue section. The recently developed technique of LCM has greatly facilitated this procedure and now provides an efficient and reliable one-step method for obtaining pure populations of cells from specific microscopic regions of tissue sections under direct visualization (22). LCM was used in the present study to obtain pure tumor cells. Secondly, because the magnitude of the genomic instability manifested as LOH varies among different forms of tumors and between the same tumor in different study populations, it would not be precise to define the loci of frequent LOH simply on the basis of the magnitude of an increased frequency used in previous studies. To overcome this problem, the present study used a genome-wide search for LOH and identification of loci showing a high LOH frequency based on comprehensive alteration profiles. On the basis of these improved methods, this study aimed at dissecting out the contribution of putative genomic loci involved in breast cancer progression and elucidating the genetic mechanism underlying breast tumorigenesis.

MATERIALS AND METHODS

Study Population and Tumor Tissue. The study of this genome-wide search for LOH is part of an ongoing cooperative study aimed at understanding the causes of breast cancer in Taiwan, which is characterized by low incidence (23), early tumor onset (24), reproductive hormone dependency (25), and novel genomic alterations (24, 26). The study subjects (100 patients) were a subset of women randomly selected from this large, ongoing hospital-based breast cancer cohort. Their ages ranged from 35 to 85 years, and five of the patients had a family history of breast cancer (mothers or sisters). These tumors were diagnosed with incident, histologically confirmed hyperplasia of breast epithelium (6 cases), DCIS (5 cases), IDC (86 cases), or lobular carcinomas (3 cases) of the breast. Invasive carcinomas were further graded using the system proposed by Bloom and Richardson (27) and staged using the Tumor-Node-Metastasis classification (28). Institutional review board-approved informed consent was obtained from each patient prior to tissue collection. None of the patients were receiving neoadjuvant treatment at the time of primary surgery. Peripheral blood was also collected from each patient on the day of surgery, and WBCs were isolated by centrifugation. All specimens were maintained at −80°C until subsequent analysis.

LCM. To ensure that tissue samples assayed consisted of >90% tumor cells, LCM was performed on routinely immunostained slides using a PixCell laser capture microscope (Arcturus Engineering, Mountain View, CA) as described previously (22) with minor modification. Briefly, the stained, dehydrated tissue section was overlaid with a thermoplastic film mounted on an optically transparent cap. The visually selected areas (tumor cells) were bound to the membrane by short, low-energy laser pulses, resulting in focal melting of the polymer. On an average, the LCM shot performed to obtain tumor cells for this genome-wide LOH for one single tumor was ~2,000, which yielded 20,000 tumor cells. The cells were immersed in 50–100 μl of digestion buffer, containing 0.2% Triton X-100, 10 mm Tris-HCl (pH 8.0), 1 mm EDTA, 400 μg/ml proteinase K, and 1% Tween 20, and digested at 55°C overnight. After digestion, the enzyme was heat inactivated (95°C for 10 min), and the extract was used directly for alleleotyping PCR (LOH detection). Efforts were made to optimize LCM conditions, including improving capture efficiency by overnight pretreatment of the slide with xylene and ensuring successful PCR amplification by avoiding fragmentation of DNA in the captured tissue by the use of a modified protocol for tissue fixation. Genomic DNA extracted from the corresponding WBC samples using conventional proteinase K-phenol/chloroform extraction served as the normal control for LOH analysis.

Allelotyping PCR. DNA from the microdissected tumor specimens and blood samples was used for alleleotyping PCR using fluorescence-labeled primers from Applied Biosystems PRISM Linkage Mapping Set-MD10 (PE Biosystems, Foster City, CA), comprising 400 polymorphic microsatellite markers, with an average spacing of 9.2 cM and average heterozygosity of 0.79. The markers were selected from the Genethon linkage map (29), on the basis of chromosomal location and heterozygosity. These markers were arranged in 28 panels, facilitating simultaneous amplification of multiple primer pairs (about four markers/per PCR assay) and significantly reducing the efforts required to perform the arduous task of genome-wide screening. PCR amplification was carried out using DNA (of tumor and blood), 0.4 unit of Taq polymerase, 0.2 mM deoxynucleotides, and 2.5 mM MgCl2 in a total reaction volume of 10 μl. The PCR conditions used were 95°C for 12 min to activate the Taq polymerase, followed by 45 cycles of denaturation of 94°C for 15 s, annealing at 55°C for 15 s, and extension at 72°C for 30 s. The final elongation was at 72°C for 10 min. PCRs were run in a GeneAmp PCR 9600 thermocycler (PE Biosystems, Foster City, CA). The reaction conditions sometimes yielded unsatisfactory amplification for certain markers and were modified, and if necessary, these particular markers were run singly. PCR products were electrophoresed on a 377 ABI PRISM sequencer, and the fluorescent signals from the different sized alleles were recorded and analyzed using GENOTYPER version 2.1 and GENESCAN version 3.1 software.

Definition of Allelic Loss (LOH). For a given informative marker, we used our previous definition (24, 30) that the marker is considered to display LOH when a 3-fold or greater difference is seen in the relative allele intensity of the different sized alleles were recorded and analyzed using GENOTYPER version 2.1 and GENESCAN version 3.1 software.

Data Analysis. The statistical methods used in previous genome-wide screening studies (CGH studies), which involved examination of genomic changes at the chromosome level, were not applicable to our study. More importantly, the characteristics of the 400 markers used in the present study, including a very intense distribution throughout the whole genome and regular spacing among markers, provided a unique opportunity to precisely measure genomic instability at the level of individual genomic loci (markers). This allowed us to develop a novel index to estimate the level of genomic abnormality of individual tumors. The index of the PLOH was a measure of the proportion (prevalence) of acquired allelic loss (LOH) in a breast cancer tumor or at specific stages/grades of a breast tumor and was obtained by dividing the total number of markers showing LOH in a tumor by the number of informative (heterozygous) markers.

To obtain evidence for the role of genomic instability in breast cancer progression, the average PLOHs of tumors of different grades or stages were correlated to tumor grade or tumor stage. The difference in the average PLOH between tumors with different pathological indexes was examined using Student t test. In addition, linear regression was used to evaluate the trend of an
increasingly higher average PLOH as tumor progressed to poorer grade or later stage.

It was of particular interest to determine whether tumor progression was specifically associated with LOH at particular genomic loci. Given the fact that most cancers are clonal (31, 32), implying breast cancer development will generally progress from low grade/early stage to high grade/later stage (reviewed in Refs. 33 and 34), we applied a well-known strategy to establish a model for breast tumorigenesis. The strategy used, which has been successfully applied in defining the genetic events underlying the progression of colon cancer (35), is that a prevalent genetic alteration that is consistently seen, regardless of histological stage or grade, is considered an early event in the tumorigenesis, and an alteration that is associated with a histologically and clinically advanced stage or grade is interpreted as a late event. However, in contrast, it has also been suggested that, although clonal, cancers are highly heterogeneous (31, 36). More importantly, a recent study argued against the concept of sequential progression and proposed an alternative model, i.e., different breast tumor grades may have distinct molecular origins (37). Thus, to account for the complexity of these issues, including the possibilities that: (a) either genomic instability per se or allelic loss at specific loci are the driving forces in tumor progression; and (b) molecular heterogeneity of individual cancer cells might be an inevitable consequence of genetic instability (31), we used two different approaches to select putative markers for the sites of TSGs from the 400 markers used in the evaluation. The first approach was to select 20 markers (the top 5%), the LOH of which showed the highest degree of association with an increase of genomic instability, from the top 50 markers most frequently targeted by LOH. This was determined from the difference in average PLOH in tumors with or without LOH at these particular markers (evaluated by \( t \) test). The identification of markers significantly associated with an increase in PLOH may provide clues about the sites of putative TSGs implicated in genomic instability leading to tumor formation. The second approach selected markers showing a significantly high frequency of LOH in tumors of different pathological grades, and these markers were identified if their LOH frequencies were among the top 5% of the 400 markers in tumors of specific grades. Recurrent loss at particular genomic loci detected by markers showing high LOH frequency is always considered to suggest the sites of putative TSGs (21), but genetic changes that are very prevalent within a particular subgroup may be missed or their significance diluted when all breast tumors are analyzed together (37). Thus, this approach to select grade-specific markers would avoid this problem, and the loss of the sites identified might either confer cells with selective growth advantage during clonal evolution or lead to genetic divergence or heterogeneity characterizing specific grades of tumor. We did not select markers based on tumor staging, because tumor stage is often a combined clinical and pathological process and is not necessarily related to tumor cell differentiation (27, 28).

RESULTS

Detection of LOH. Using 400 microsatellite markers and paired DNA from blood and laser capture microdissected tumor tissues, this genome-wide search for allelic loss was carried out on 100 breast tumors in Taiwan. Allelic loss was detected by LOH, which was measured using standardized fluorescence-based methodology and apparatus. Two examples of our use of LCM and allelotype PCR to detect LOH are illustrated in Fig. 1. Particular attention was paid to the definition of marker heterozygosity. Because of the nature of short tandem repeat (microsatellite) markers on electrophoresis, it is usual to see, in addition to the main bands, minor “shadow bands,” which are thought to result from slipped strand mispairing at tandem dinucleotide repeats (38). Sometimes, the band of the shorter allele can be especially intense because it contains both the main band of this allele and the major shadow band derived from the longer allele. If the size difference of the two alleles is small, this artifact might become obvious and might confound the definition of heterozygosity status. Because all of our study subjects were Taiwanese (Chinese), a relatively genetically homogenous population as compared with the Caucasian population (on which the Genethon microsatellite markers are based), a relatively higher proportion of individuals possess parental alleles with very similar allelic sizes. Accordingly, this artifact caused by “shadow band confounding” would be expected to be more common. To avoid this problem, in addition to using the usual definition of heterozygosity based on a \( \geq 2 \)-bp difference in allelic length, we also adopted a more conservative criterion that only a \( \geq 4 \)-bp difference in allelic length between two alleles would be considered as heterozygous. On thorough comparison of the results using either of the two above criteria, only minor differences were found. Therefore, the study results were presented based on the routinely used criterion, i.e., a \( \geq 2 \)-bp-difference between two alleles was considered to represent heterozygosity (informative), and the heterozygous markers served as the target to assess LOH status.

LOH at the Levels of Markers, Chromosomal Arms, and Individual Tumors. As expected, the proportion of heterozygosity varied among the 400 markers studied, the mean being 68.7% and the median 71.7%. A total of 379 markers showed \( >50 \% \) heterozygosity, yielding a stable estimation of LOH frequency in the present study. The mean percentage of LOH among the markers was 17.5% (SD, 7.08\%), with the results statistically fitting a normal distribution when examined by a normal probability plot. Sixty-three markers displayed \( \geq 25 \% \) LOH, the highest values being seen with \( D17S921 \) on 17p12 (48.4\%), \( D16S503 \) on 16q22.1 (42.0\%), \( D17S831 \) on 17p13.3 (41.6\%), \( D17S938 \) on 17p13.2 (37.7\%), \( D11S4175 \) on 11q22.1 (34.4\%), and \( D15S249 \) on 1q32.1 (33.9\%). Some of these loci are known to harbor TSGs; these include p53 (17p12–13), E-cadherin (16q22.1), and ATM (11q22–23), all of which have been suggested to play a critical role in breast tumorigenesis (reviewed in Refs. 34 and 39; reported in Ref. 40). In contrast, 55 markers showed \( <10 \% \) LOH among the informative cases.

The frequencies of LOH on individual chromosomal arms can be also estimated by dividing the number of markers showing LOH on a given chromosomal arm by the number of total informative markers on that arm. The highest frequency of LOH for all arms was seen on chromosome 17p (35\%), which was mainly attributable to LOH at the locus of p53. The arms with the highest frequency of LOH (>20%) were 1q, 4p, 8p, 8q, 11q, 13q, 16q, 17p, 17q, and 22q. This chromosomal arm deletion profile is similar to that obtained using CGH reported for breast carcinoma in Western populations (reviewed in Ref. 41), implying that similar tumorigenic mechanisms might underlie the causes of breast cancer in Taiwan, although it is unique in that its incidence is the lowest in the world. Nevertheless, one notable exception is that the loss of 4p, which has rarely been reported in previous genome-wide searches in breast cancer and may suggest the site of ethnic-specific cancer susceptibility genes. It is also possible that the relatively low resolution of the method of CGH, used previously to detect a small deletion at 4p, might be the explanation.

To estimate the magnitude of genomic instability in individual tumors, the PLOH was calculated on the basis of comprehensive deletion profiles of 400 markers distributed throughout the entire genome. The values for the PLOH ranged from 2 to 66.3%, with a mean of 17.6\% (SD, 13.0\%) and the median of 14.6\%. The reason for the higher mean value was attributable to extremely high frequencies of PLOH (>40%) in some tumors. After careful evaluation of their deletion profiles, these tumors were found to show evidence of aneuploidy, in which an extensive loss of all informative markers on particular chromosomal arms or even of the whole chromosome was found. We therefore purposely differentiated the six cases with putative aneuploidy from the other tumors, with LOH profiles consistent with the pattern of interstitial deletion. Although aneuploidy and interstitial deletion are both classified as CIN (20), the mechanisms underlying these two genetic events are obviously different. Aneuploidy in cancer results from a defective mitotic machinery resulting in chromosome missegregation (reviewed in Refs. 20 and 42).
was more obvious in tumors displaying very few (75%) tubule higher individual grade indices (Fig. 2). An increase in average PLOH poorer tumor differentiation, as reflected by a higher tumor grade or higher average PLOH was consistently found to be associated with clinical outcomes. In the interstitial deletion tumors analyzed, a difference in average PLOH among tumors with different pathological initiated CIN and manifests as a mutator phenotype, we compared the our hypothesis that breast cancer progression is associated with DSB-present study.

whereas interstitial deletion is caused by the formation of DSBs in the DNA (reviewed in Refs. 18–20), the latter being the main target of the present study.

**Genomic Instability during Breast Cancer Progression.** To test our hypothesis that breast cancer progression is associated with DSB-initiated CIN and manifests as a mutator phenotype, we compared the difference in average PLOH among tumors with different pathological and clinical outcomes. In the interstitial deletion tumors analyzed, a higher average PLOH was consistently found to be associated with poorer tumor differentiation, as reflected by a higher tumor grade or higher individual grade indices (Fig. 2). An increase in average PLOH was more obvious in tumors displaying very few (<75%) tubule formation and substantial nuclear pleomorphism or in tumor cells with higher-than-moderate mitotic counts (P < 0.05 as compared with tumors displaying >75% of tubular architecture, little nuclear pleomorphism, or very few mitotic counts, respectively). Overall, a significantly increasing trend toward DSB-initiated CIN was seen as IDC progressed to poorer grades or latter stages (P < 0.05 assessed by linear regression analysis). In contrast, lobular carcinomas displayed a genetically stable phenotype, in which the average PLOH was close to that for well-differentiated IDC (Fig. 2).

**Selection of Markers Based on Their Association with Genomic Instability.** On the basis of the first selection approach, a total of 20 markers (the top 5%) were selected from the 400 markers, because they were frequently targeted by LOH (overall LOH frequency ≥25% and >50% of heterozygosity), and more importantly, their LOH status showed the most significant association with an increased level of PLOH, correlating to an increase of >15% of DSB-initiated CIN (Fig. 3). LOH in these markers was expected to be either the causes or the consequences of the increasing level of genomic instability during tumor progression found in previous section. Thus, it is notable that, among these four loci, 15q15.1, 2q34, 13q32.3, and 3p24.3, reflected by LOH at four markers (D13S1007, D2S325, D13S138, and D3S1266), are the genomic sites in which RAD51, Ku80, DNA ligase IV, and FA-D are located. These genes are key players in DSB-related repair or CIN (20, 43).

**Genetic Heterogeneity.** The previous findings provide evidence in support of a clonal proliferation model driven by DSB-initiated CIN during breast tumor progression. However, it is suggested that, although most cancers are clonal, they are yet highly heterogeneous, i.e., nonclonal with regard to the karyotype and phenotype of individual cancer cells. This heterogeneity results in genetic divergence during the clonal evolution of breast cancer (44, 45). More importantly, genomic instability (including CIN) is considered to contribute not only to tumor progression but also to tumor heterogeneity (20, 31, 36). In the light of the findings that DSB-initiated CIN has been shown to be associated with tumor progression in the present study (see above), it becomes necessary to examine whether, and at what stages, genetic heterogeneity in our tumors existed and acted as the other determinant adding genetic complexity in breast cancer development. To examine this less definite issue in the present series of patients, we took advantage of having six synchronous tumors in which both DCIS and IDC components were simultaneously available for allelotyping analysis (Fig. 4). Using 10 microsatellite markers, we examined LOH on nine chromosomal arms (6q, 8p, 13q, 16q, 17q, 19p, 20q, and 22q) in DNA microdissected from both the DCIS and IDC components. At least three of these tumors (case nos. 467, 541, and 557) supported the clonal evolution hypothesis, simultaneously showing LOH at the same markers. However, two tumors (case nos. 467 and 557) revealed a specific LOH profile, in which LOH at a given locus was only detected in the DCIS component, but this marker remained intact in the IDC component, supporting the presence of genetic heterogeneity (Fig. 4). These results suggest that genetic heterogeneity was common in these breast cancers, and that heterogeneity of breast cancer might occur at the earliest detectable grade/stage, i.e., DCIS, of progression.

**Selection of Markers Based on LOH Frequency.** Given the previously demonstrated presence of genetic heterogeneity in the tumors of this series, and more importantly, given the fact that the markers showing recurrent LOH suggest the sites of TSGs, we selected grade-specific markers that showed high LOH frequencies in tumors of specific grades (the second approach of marker selection; Table 1). Notably, several of the genomic loci identified, including 17p12–13, 11q22.1, and 16q22, were commonly found in both lower grade tumors and higher grade tumors, suggesting that the genes located in these loci, e.g., p53 at 17p12–13, ATM at 11q22.1, E-cadherin at 16q22 (reviewed in Refs. 34 and 39), might contribute to an early event during progression. DCIS was not included in this table, because the limited number of cases of DCIS included in this study might result in an imprecise estimation of LOH frequency.


**Construction of a Breast Tumorigenic Model.** Considering the individual contributions of clonal evolution, genetic heterogeneity, and DSB-initiated CIN (genomic instability), which might occur simultaneously during breast cancer pathogenesis, 10 patterns profiling the genetic alterations of LOH frequency of a particular marker could be proposed (Fig. 5B). The changes in patterns of A1, A2, and A3 are consistent with the conventional concept of clonal evolution, in which tumor progression occurs via a sequential pathway driven by genetic alterations occurring at early (well differentiated; A1), intermediate (moderately differentiated; A2), and late (poorly differentiated; A3) stages. In contrast, if LOH of a marker was preferentially seen at specific grades/stages, putative genes located around this marker may be involved in grade/stage-specific heterogeneity. Thus, the patterns of A4, A5, and A3 were considered to represent such genetic alterations leading to heterogeneity of well-differentiated, moderately differentiated, and poorly differentiated tumors. Finally, if the contribution of genomic instability is also taken into consideration, in which DSB-initiated CIN would become more obvious as the tumor progresses, the patterns of A1 to A5 would become B1 to B5, respectively. The timing of LOH of individual markers belonging to these 10 patterns can be represented in a breast tumorigenic model, as shown in Fig. 5A.

The markers selected on the basis of two selection approaches were then classified into 1 of 10 patterns according to their LOH profiles in tumors of different grades (Fig. 5B). Interestingly, the loci for ATM (11q22) and p53 (17p12–13), the key checkpoint genes of DSB, were lost at the earliest stage. In addition, as mentioned earlier, several loci identified (specified in Fig. 5B) were those possibly harboring genes participating in DSB-related checkpoint/repair. More interestingly, these loci were found to be lost at different grades, yielding a clue to suggest that accumulation of multiple genetic defects associated with DSB-initiated CIN may be required for the continuous decay of chromosomal integrity. To test this possibility, we performed a multivariate regression analysis to correlate the extent of CIN in a tumor (measured by PLOH) with the LOH status of eight loci that probably harbor CIN-related genes (including 17q12–13 for p53, 11q22 for ATM, 16q24.3 for FA-A, 15q14–15 for RAD51, 12p12 for RAD52, 17q21 for BRCA1, 14q32 for XRCC3, and 2q34 for Ku80). A consistently significant association of increased CIN in tumors with a higher number of LOH at these CIN-related loci was found, showing, on average, a 7% increase in PLOH for each additional LOH at these loci ($P < 0.0001$).

With regard to genetic heterogeneity, it is notable that no DSB-related genes were identified in well-differentiated tumors, which showed the lowest level of CIN (Fig. 5B). Interestingly, two loci (3p22.1 and 9q22.33) that contributed to the well-differentiated state were the sites at which hMLH1 and XPA are located. Defects of these two genes would result in MIN and NIN, which is consistent with the suggestion that MIN/NIN tumors display a low level of CIN (20).

**DISCUSSION**

Cancer is believed to result from a series of genetic alterations leading to the progressive disordering of the normal mechanisms controlling growth, differentiation, cell death, or genomic instability (35). The complete understanding of the etiology of breast cancer or any other tumors will require studies that comprehensively evaluate both the genes participating in tumorigenesis and tumors of different pathological and clinical stages to reflect the sequential steps occurring during breast cancer progression. This understanding is likely to emerge slowly, because research is only beginning to move from single-gene studies to multigenic or genome-wide studies. Colorectal cancer is one of the most successful models for studying tumorigenic progression, because carcinomas appear to evolve through a series of morphologically and histologically recognizable steps representing different stages of cancer development (35). In contrast, many breast cancers display a high degree of intratumor heterogeneity in terms of histological appearance and phenotypic character and do not exhibit easily observable pathological manifestations to define sequential evolution. In an attempt to elucidate these etiological and phenotypic complexities, the present study proposed a tumor progression model to explore breast tumorigenesis, based on the fundamental concept postulated by Nowell (31) in 1976, that genomic instability is the engine of both tumor progression and tumor heterogeneity. In light of the clues provided by familial breast cancer syndromes (reviewed in Refs. 3, 4, and 5), this genome-wide study was carried out to examine whether breast cancer progression is manifested as a mutator phenotype and whether breast tumorigenesis is driven by a particular form of genomic instability, i.e., DSB-initiated CIN. Unlike previous genome-wide studies (reviewed in Ref. 41), in which CGH was usually used to detect genomic alterations at the chromosomal level, the present study, based on genome-wide LOH of highly intensive microsatellite markers, was able to address this hypothesis more comprehensively, providing a more concise insight into breast tumorigenesis driven by genomic instability. Furthermore, the effects of heterogeneity within a tumor were minimized by ensuring that genetic and phenotypic examinations were carried out on the same tumor cells using the newly invented method of LCM, which allows a more precise evaluation of specific associations between genetic alteration and pathological manifestation.

Using a PCR-based method to detect LOH may be confounded by genomic amplification, which might yield a false-positive loss of markers. To exclude the possibility that the loci defined to show high LOH frequency were attributable to amplification, we examined whether these genomic regions are those known to be frequently amplified and to harbor known breast cancer oncogenes. The onco-
genes implicated in breast cancer development are MYC (at 8q24.12–
24.13), cyclin D1 (at 11q13–14), and ERBB2 (at 17q11–12; reviewed in Ref. 46). As a result, LOH at these loci must be interpreted with caution. However, none of the loci identified to show high LOH frequency in the present study were the “frequently amplified” loci. Thus, possible artifacts attributable to amplification was minor.

Genomic instability exists at two distinct levels, at the nucleotide sequence level, resulting in base substitutions (NIN) or deletions or insertions of a few nucleotides (MIN), and at the chromosome level, resulting in the loss or gain of whole chromosomes or large portions thereof (CIN; Ref. 20). In breast cancer, MIN is seen in only a small subset (<10%) of tumors (34), and there is little evidence of mutation hotspots to support a significant etiological role of NIN. In contrast, aneuploidy, an abnormal number of chromosomes representing the prototype of CIN, is relatively common in breast cancer. On the basis of genome-wide detection of allelic loss (LOH) and the exclusion of tumors presenting aneuploidy, we specifically defined the other form of CIN, i.e., DSB-initiated CIN, as a contributing determinant to breast cancer progression. Support for our hypothesis comes from the observations that: (a) DSB-initiated CIN, measured by PLOH, significantly increased as the tumors progressed to a poorer grade or later stages; (b) in sequential steps toward CIN, the loci of p53 and ATM, the key checkpoint genes against DSB, were lost at the earliest stage; (c) several loci identified as associated with breast tumorigenesis are genomic sites where the genes involved in DSB-related checkpoint/repair are probably located, and tumors possessing a higher number of LOH loci at these sites displayed a greater level of CIN. Furthermore, on the basis of known interaction between genes implicated in the same tumorigenic mechanism, we are also able to verify this proposed model of tumor progression. Our result shows that sequential loss of the locus for p53 (17p12–13) occurred in well-differentiated tumors, followed by loss of the locus for BRCA1 (17q21.31) in moderately differentiated tumors. A recent model of mice homozygous for mutations in both the p53 and Brca1 genes is concordant with this finding. In this model, p53−/− (a defective checkpoint) is the initial event in tumor progression to acquire an increased growth potential and genomic instability. The subsequent loss of BRCA1 function, which plays a role in multiple repair pathways, would be expected to accelerate the accumulation of additional mutations leading to malignant transformation (47). Etiologically, this model was considered to be a more plausible mechanism for explaining the relationship between p53 and BRCA1 in breast cancer formation, because it does not envisage a growth-arrest stage attributable to loss of BRCA1 (17, 47, 48). Our model certainly does not exclude the contribution of aneuploidy to breast tumorigenesis. Rather, it shows that several tumorigenic loci probably harbor the genes involved in either chromosomal stability (e.g., WRN, FA-A, and FA-D) or DSB-related checkpoint/repair. These findings suggest that CIN affecting the integrity of both chromosomal number and chromosomal structure is of particular importance for breast cancer development. Consequently, an increased mutation rate (mutator phenotype) attributable to CIN, probably in combination with selection and clonal expansion, provides the basic requirement for breast tumor progression.

One cautionary note should be raised. We were aware that LOH at a few markers cannot be taken as sufficient evidence for the involvement of the genes in tumorigenesis, and the demonstration of specific mutation in putative genes would be required. However, the causal link between some of these suggested genes, such as p53, ATM, and BRCA1, and breast cancer has been confirmed (reviewed in Refs. 8, 11, and 17). The increased risk of developing breast cancer in familial cancer syndromes, such as xeroderma pigmentosum and Muir-Torre syndrome (4), lends additional support to the contribution to breast cancer development of certain other genes, such as XPA and hMLH1. On the other hand, to the best of our knowledge, with the exception of p53, the probability of finding somatic mutation in these checkpoint/repair (caretaker) genes (e.g., RAD51, RAD52, FA-A, FA-D, XPA, and hMLH1) has been shown to be extremely low. Because LOH only represents the specific indicator of the “one-hit” needed to inactivate TSGs, these genes do not seem to fulfill Knudson’s “two-hit” criterion for a TSG in cancer formation. However, recent evidence, which suggests two nonmutually exclusive possibilities, provides support for a tumorigenic role of these genes defined solely by LOH in breast cancer. The first is that a growing number of common TSGs, including p53 and ATM, have been found to exhibit the haplo-insufficiency phenotype (49, 50), which implies that homozygous inactivating mutations and complete loss of function are not necessary to cause defective tumor suppressor function. Thus, a half normal level of the gene product, resulting from allelic loss or LOH per se, would be able to produce the phenotypic manifestation (51). It is still possible that other (epigenetic) mechanisms of inactivation abrogating the function of these genes are implicated in breast tumorigenesis. For example, hypermethylation of the BRCA1 promoter region has been found in breast cancer (52), supporting the role of BRCA1 in sporadic breast cancer development (53). Therefore, our identification of multiple LOH loci possibly targeting DSB-related genes should not be considered to merely reflect the sites of putative TSGs. Instead, these findings can be further regarded as clues suggesting that breast cancer progression is driven by the accumulation of multiple genetic alterations to inactivate the genes participating in the common defensive mechanism against DSB. In accordance with our findings, a recent breast cancer study involving the detection of LOH at the loci of five DSB-repairing genes (BRCA1, BRCA2, RAD51, RAD52, and RAD54) showed that the probability of simultaneous LOH at more than two of these loci was much higher than expected (54) and underlines the additional importance of finding LOH at these DSB-related loci. More importantly, these findings, including our own, suggest that these proteins associated with DSB-related checkpoint/repair work together to maintain chromosomal stability. As a result, a decrease in the amount of any one of these DSB-related genes as a result of LOH may alter the stoichiometric relationship between them. This inference can be supported by molecular findings that certain of these proteins are
found in the same DNA-repairing complex or pathway (reviewed in Refs. 12, 13, and 16). Consequently, disruption of any one component of this DNA-repairing complex or pathway may hinder DNA repair and result in the accumulation of unrepaired DSB, subsequently leading to CIN without cell cycle checkpoint and to tumor formation.

The most inconsistent findings regarding the pathogenesis of breast cancer are those related to whether breast cancer progression follows a process of dedifferentiation, from well differentiated (grade I) to poorly differentiated (grade III) over time. The current grading system for breast cancer is based on semiquantitative indices that assess the degree of tubule formation, nuclear pleomorphism, and mitotic activity (27) and assumes a sequential change from lower grade to higher grade tumors. Supporting this concept, epidemiological observations based on a randomized screening trial involving a large cohort of women indicate that the natural history of breast carcinoma follows a progressive disease model, by which regular breast screening to prevent deterioration of the malignancy grade of tumors is recommended (33). If this dedifferentiation is to happen, one would see a group of lower grade cells adjacent to a group of higher grade cells in the same tumor population. However, it is not unusual for some tumors to always present with a poorly differentiated phenotype without a well-differentiated counterpart. Furthermore, a recent study that showed distinct chromosomal abnormality profiles in well-differentiated and poorly differentiated breast tumors (the most striking difference being the loss of 16q, seen in 65% of grade I tumors but only 16% in grade III tumors) concludes that different breast tumor grades may have distinct molecular origins and pathogenesis (37). The findings of the present study, which are based on a comprehensive and more precise picture of genome-wide LOH profiles, only confirmed that grade III tumors are not directly progressed from grade I tumors. However, the results indicate that the loss of 16q markers was an early event, because the LOH frequencies were relatively high in tumors of all grades. We also do not agree with the suggestion that different breast cancer grades have totally distinct origins. More comprehensively, our model suggests that breast cancers are sequentially progressive with regard to CIN. However, CIN might also cause genetic heterogeneity (31), which was revealed by the findings that LOH at some markers was observed only in the DCIS component but not in the IDC component of the same tumors. In addition, some markers were found to preferentially lose at specific tumor grades, suggesting their contribution to genetic heterogeneity during tumor development. Therefore, breast cancer progression follows a clonal evolution model driven by CIN, but different tumors would present distinct molecular profiles caused by genetic heterogeneity resulting from CIN. On the basis of this model, the development of a grade I tumor is the consequence of genetic heterogeneity initially driven by DSB-initiated CIN, and if the genomic loci targeted by LOH happen to be those triggering MIN or NIN, tumor cells would progress via a molecular pathway less affected by CIN and consequently become well-differentiated tumors, generally presenting a relatively mild clinical progression. This model is more biologically plausible and is consistent with colorectal cancers with MIN, which are always associated with genetic heterogeneity (31), which was revealed by the findings that LOH at some markers was observed only in the DCIS component but not in the IDC component of the same tumors. In addition, some markers were found to preferentially lose at specific tumor grades, suggesting their contribution to genetic heterogeneity during tumor development. Therefore, breast cancer progression follows a clonal evolution model driven by CIN, but different tumors would present distinct molecular profiles caused by genetic heterogeneity resulting from CIN. On the basis of this model, the development of a grade I tumor is the consequence of genetic heterogeneity initially driven by DSB-initiated CIN, and if the genomic loci targeted by LOH happen to be those triggering MIN or NIN, tumor cells would progress via a molecular pathway less affected by CIN and consequently become well-differentiated tumors, generally presenting a relatively mild clinical progression. This model is more biologically plausible and is consistent with colorectal cancers with MIN, which are always associated with
Interestingly, the locus (5q31.3) at which Tcf1 is located is supported by the finding that disruption of this molecular pathway could result in breast cancer through its targeted transactivation of growth-promoting genes (59). The possibility that p27 could also be a direct growth suppressor via a mechanism mediated by E-cadherin escape from growth-control signals (57). Alternatively, p27 could also play a role in the initiation of cell proliferation by allowing cells to progress into the cell cycle (58).

The known function of E-cadherin as a cell-cell adhesion molecule contributing to the suppression of tumor metastasis has been studied extensively (56). The high frequency of LOH of the selected markers suggests sites of putative TSGs playing an initiating role in tumor progression. In addition to those associated with DSB, one of these markers shown in boxes represent genomic loci possibly harboring genes involved in DSB-related checkpoint/repair or CIN; these include 11q22 for ATM, 12p12 for RAD52, 16q24.3 for FA-A, 17q12-13 for p53, 13q32 for DNA ligase IV, 14q24 for RAD51B, 17q21 for BRCA1, and 15q14–15 for RAD51.

The genomic instability responsible for the development of familial gastric cancer has been found between loss of E-cadherin expression and tumor progression. In addition to those associated with DSB, one of these markers shown in boxes represent genomic loci possibly harboring genes involved in DSB-related checkpoint/repair or CIN; these include 11q22 for ATM, 12p12 for RAD52, 16q24.3 for FA-A, 17q12-13 for p53, 13q32 for DNA ligase IV, 14q24 for RAD51B, 17q21 for BRCA1, and 15q14–15 for RAD51.

In conclusion, this study demonstrates that DSB-initiated CIN is the major driving force of breast cancer progression, and this mutator phenotype also contributes to breast cancer progression in our model. If the proposed DSB-initiated CIN model correctly describes breast cancer pathogenesis, it would be interesting to ask why breast epithelium cells are particularly susceptible to defective DSB-related checkpoint/repair, because this system is believed to be essential in all cells. This tissue specificity may be explained by the possibility that breast cells that have lost DSB-related checkpoint/repair have a growth advantage with respect to DSB-checkpoint/repair-proficient cells. This may be imposed by a selective micro-environment, and in the breast, we expect an increased exposure to estrogen may be one cause. Recently, the role of reproductive hormones in breast tumorigenesis has been greatly expanded to suggest that estrogen might be a complete carcinogen directly causing genetic alteration and tumor initiation (61). We have recently provided epidemiological evidence to suggest that breast cancer can be initiated by estrogen via exposure to estrogen metabolites, including catechol estrogen and catechols estrogen quinones (25). These metabolites cause oxidative DNA damage and depurination (abasic sites), possibly resulting in clustered sites of base damage in both DNA strands, and attempts to correct such damage by standard base excision repair can further lead to the formation of DSBs (62). This may shed light on why DSB-related checkpoint/repair is particularly important in breast tumorigenesis.

In conclusion, this study demonstrates that DSB-initiated CIN is the major driving force of breast cancer progression, and this mutator phenotype also contributes to tumor heterogeneity. Because the issue of the involvement of CIN in the development of breast cancer has seldom been addressed so extensively, these results, based on genome-wide survey, will serve as an essential foundation for a better understanding of the molecular mechanisms that occur during breast tumorigenesis.

Fig. 5. Breast cancer progression model (A) and the genomic loci (and markers) involved in specific steps during progression (the lower part of B) along with representative LOH profiles of selected markers in tumors of different grades (the upper part of B). In B, well indicates well-differentiated tumors, moder., moderately differentiated tumors, and poor, poorly differentiated tumor, whereas the numbers on the X axis represent the proportion of LOH of the selected markers. In the lower part of B, the markers shown in boxes represent genomic loci possibly harboring genes involved in DSB-related checkpoint/repair or CIN; these include 11q22 for ATM, 12p12 for RAD52, 16q24.3 for FA-A, 17q12-13 for p53, 13q32 for DNA ligase IV, 14q24 for RAD51B, 17q21 for BRCA1, and 15q14–15 for RAD51.

was also defined as one of the important loci contributing to breast cancer progression in our model.

a better prognostic phenotype (reviewed in Ref. 55). Furthermore, the high number of genetic alterations found to be associated with genetic heterogeneity in poorly differentiated tumor cells (i.e., the loci of the patterns of A3 and B3 in Fig. 5B) has also led to the suggestion that, as DSB-initiated CIN reaches a maximum (the highest PLOH), breast cancer would progress to be a highly heterogeneous disease with diverse molecular profiles. Consequently, as suggested previously, every breast tumor has a different set of genomic deletions (40).

We were particularly interested in genomic loci showing a significantly high frequency of LOH in well-differentiated tumors, because they suggest sites of putative TSGs playing an initiating role in tumor progression. In addition to those associated with DSB, one of these regions, 16q22.1, has been suggested to reflect the involvement of E-cadherin (reviewed in Refs. 34 and 39). In breast cancer, an association has been found between loss of E-cadherin expression and advanced stage, poor tumor grade, or unfavorable prognosis (reviewed in Ref. 56). The known function of E-cadherin as a cell-cell adhesion molecule contributing to the suppression of tumor metastasis does not seem to favor its initiating role during tumor progression. However, a recent demonstration that the E-cadherin germline mutation is responsible for the development of familial gastric cancer raises the possibility that disruption of normal cell-cell adhesion could also play a role in the initiation of cell proliferation by allowing escape from growth-control signals (57). Alternatively, E-cadherin could also be a direct growth suppressor via a mechanism mediated by p27Kip1 (58).

Furthermore, the cytoplasmic domain of E-cadherin can modulate the Wnt signaling pathway by inhibiting the availability of -catenin-mediating Wnt signaling (60). The possibility that disruption of this molecular pathway could result in breast cancer progression is supported by the finding that Tcf1 (the other effect modifier of the -catenin pathway) mice developed adenomas in the mammary glands (60). Interestingly, the locus (5q31.3) at which Tcf1 is located was also defined as one of the important loci contributing to breast cancer progression in our model.
Pathogenesis Implication for Mutator Phenotype and Breast Cancer Capture Microdissected Tissue of Breast Carcinoma: An
Genome-wide Search for Loss of Heterozygosity Using Laser Capture Microdissected Tissue of Breast Carcinoma: An
Implication for Mutator Phenotype and Breast Cancer Pathogenesis

Chen-Yang Shen, Jyh-Cherng Yu, Yen-Li Lo, et al.

Cancer Res 2000;60:3884-3892.

Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/60/14/3884

Cited articles
This article cites 57 articles, 26 of which you can access for free at:
http://cancerres.aacrjournals.org/content/60/14/3884.full.html#ref-list-1

Citing articles
This article has been cited by 36 HighWire-hosted articles. Access the articles at:
/content/60/14/3884.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.