Differential Patterns of Allelic Loss in Estrogen Receptor-Positive Infiltrating Lobular and Ductal Breast Cancer

L. W. M. Loo,[1] C. Ton,[1,2] Y.-W. Wang,[2] D. I. Grove,[2] H. Bouzek,[1] N. Vartanian,[1] M.-G. Lin,[1] X. Yuan,[1] T. L. Lawton,[3] J. R. Daling,[2] K. E. Malone,[2] C. I. Li,[2] L. Hsu,[2] and P. L. Porter[1,2,3]*

[1]Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA
[2]Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, WA
[3]Department of Pathology, University of Washington, Seattle, WA

The two main histological types of infiltrating breast cancer, lobular (ILC) and the more common ductal (IDC) carcinoma are morphologically and clinically distinct. To assess the molecular alterations associated with these breast cancer subtypes, we conducted a whole-genome study of 166 archival estrogen receptor (ER)-positive tumors (89 IDC and 77 ILC) using the Affymetrix GeneChip® Mapping 10K Array to identify sites of loss of heterozygosity (LOH) that either distinguished, or were shared by, the two phenotypes. We found single nucleotide polymorphisms (SNPs) of high-frequency LOH (>50%) common to both ILC and IDC tumors predominately in 11q, 16q, and 17p. Overall, IDC had a slightly higher frequency of LOH events across the genome than ILC (fractional allelic loss = 0.186 and 0.156). By comparing the average frequency of LOH by chromosomal arm, we found IDC tumors with significantly ($P < 0.05$) higher frequency of LOH on 3p, 5q, 8p, 9p, 20p, and 20q than ILC tumors. We identified additional chromosomal arms differentiating the subtypes when tumors were stratified by tumor size, mitotic rate, or DNA content. Of 5,754 informative SNPs (>25% informativity), we identified 78 and 466 individual SNPs with a higher frequency of LOH ($P < 0.05$) in ILC and IDC tumors, respectively. Hierarchical clustering of these 544 SNPs grouped tumors into four major groups based on their patterns of LOH and retention of heterozygosity. LOH in chromosomal arms 8p and 5q was common in higher grade IDC tumors, whereas ILC and low-grade IDC grouped together by virtue of LOH in 16q. Published 2008 Wiley-Liss, Inc.†

INTRODUCTION

Infiltrating lobular carcinoma (ILC) differs both morphologically and clinically from the more common infiltrating ductal carcinoma (IDC) in breast cancer. ILC is more likely to be bilateral, multicentric, and diploid than IDC. Also, ILC tumors have a pattern of metastasis to bone, gastrointestinal tract, and ovary that is different from that of IDC (Harris et al., 1984; Arpino et al., 2004). The widely infiltrative growth of individual and small rows of lobular cancer cells through breast stroma results in a low detection rate by clinical exam or by mammography (Porter et al., 1999). Given that ILC typically has favorable prognostic indicators (Arpino et al., 2004), a more favorable relationship with survival and recurrence would be expected. A survival advantage has been associated with good prognosis in some studies but not in others (Mersin et al., 2003; Li et al., 2003b; Arpino et al., 2004; Cristofanilli et al., 2005), indicating that other, as yet unknown, factors might distinguish ILC from IDC.

Historically, ILC has accounted for between 5 and 14% of breast cancers (Harris, et al., 1992), but recent analyses of Surveillance Epidemiology and End Results (SEER) data indicate that ILC incidence rates have increased rapidly over the past two decades in the United States (Li et al., 2003a), primarily in women over age 50. This differential increase in older women is in keeping with recent studies that have implicated combined (estrogen/progestin) hormone replacement therapy (CHRT) as a risk factor for lobular type breast cancer (Chen et al., 2002; Li et al., 2003b: WHI Steering Committee 2004).

Additional Supporting Information may be found in the online version of this article.

Supported by: NIH; Grant numbers: RO1 CA095717 and RO1 CA085913; DOD Breast Cancer Research Program; Grant numbers: BC045034 and DOD DAMD 17-02-1-0386.

*Correspondence to: Peggy L. Porter, Human Biology and Public Health Sciences, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N, Seattle, WA 98109, USA.
E-mail: pporter@fhcrc.org

Received 7 March 2008; Accepted 11 July 2008

DOI 10.1002/gcc.20610

Published online 21 August 2008 in Wiley InterScience (www.interscience.wiley.com).

Although a few studies have compared molecular alterations between the two subtypes and have generated some information about differences in the frequency and locations of genomic alterations, the small number of ILC tumors and the limited number of loci interrogated in most studies have limited the comprehensive evaluation of genomic differences between ILC and IDC (Gunther et al., 2001; Loo et al., 2004; Roylance et al., 2006).

One of the most common alterations in breast cancer is allelic imbalance or loss of heterozygosity (LOH). An LOH event can be associated with either a gene copy number loss or amplification. LOH has been shown to occur in precancer lesions of the breast and in normal epithelium of women with breast cancer (Deng et al., 1996; O'Connell et al., 1998). The profile of the most common LOH events in breast cancer includes losses in chromosome arms, 1q, 3p, 6q, 11p, 13q, 16q, 17p, 17q, and 18q (Devilee and Cornelisse, 1994), but the contribution that ILC makes to that overall profile is small, given the low number of lobular cancers included in virtually all studies assessing allelic imbalance.

To gain a better understanding of allelic imbalance events inherent in the lobular subtype of breast cancer, we compared LOH profiles of flow cytometrically sorted ER-positive ILC and IDC tumors derived from a population-based sample. Restricting the analysis to ER-positive tumors allowed us to assess similarities and differences in LOH events between the subtypes not confounded by differences due to the ER status of the tumors.

MATERIALS AND METHODS

Patient Population

Participants in this study were drawn from a population-based case–control study designed to examine the relationships of risk factors and histological type of breast cancer. The methods used to select, enroll, and interview participants have been described in detail elsewhere (Li, et al., 2008). The study included ILC (ICD-O histology code 8520/3) and IDC cases (ICD-O histology code 8500/3) 55–74 years of age, ascertained through our population-based cancer registry for 13 western Washington counties and diagnosed between January 2000 and March 2004. This analysis was restricted to the 77 ILC and 89 IDC cases for whom we had sufficient tumor tissue to conduct this work and who had an available blood specimen. All participants consented to tumor- and blood-sample testing.

Pathology Review for Histological Diagnosis/Tumor Characteristics

To verify histological diagnoses for all tumors, pathology reports and histology slides underwent centralized pathology review by the four pathologists (PLP, MGL, XY, TL), using published criteria for ILC and IDC tumors (Page and Anderson, 1987). Histological grade was assessed for IDC tumors using the Nottingham grading scheme (Elston and Ellis, 1991). After agreement by all study pathologists on diagnostic criteria, the histology of each case was assessed by review of all submitted slides by one of two staff pathologists (MGL, XY). A second review of the diagnosis and histological grade of each case was done by the study pathologist (PLP). Discrepancies between the initial and second assessments were resolved by consensus review or submitted for additional review to the study's consulting pathologist (TL). All reviews were conducted independently.

The components of clinical stage: tumor size, lymph node metastasis, and distant metastasis were obtained from SEER and/or original pathology reports.

Immunohistochemistry

To verify the steroid-receptor status of all tumors, immunohistochemistry (IHC) for the expression of estrogen receptor (ER) (ER1D5, AMAC, Inc) (Andersen and Poulson, 1989) and progesterone receptor (PR) (1A6 NovoCastra Lab) (Giri et al., 1988) was done on a representative block for each tumor using standard IHC techniques including antigen retrieval (Hsu et al., 1981; Taylor et al., 1994). Positive and negative controls were included for each antibody.

Flow Cytometric Cell Sorting and DNA Extraction

Tumor cells were collected by flow sorting based on cytokeratin and DNA content, which enriches the sample for tumor cells while excluding contaminating stromal and lymphocytic cells as described previously (Glogovac et al., 1996). Cell populations were classified as diploid/near diploid ($1.0 \leq$ DNA Index (DI) ≤ 1.39) and aneuploid (DI > 1.39 or DI < 1.0). If a tumor contained multiple populations with different ploidy, the most prevalent aneuploid population was assayed. DNA was extracted from the sorted tumor cells (a minimum of 100,000 cells per tumor) using the QIAamp DNA Micro kit (Qiagen, Valencia, CA) with modifications. Briefly,

tumor cells were extracted in Qiagen ATL buffer with the addition of proteinase K (20 μg/μl), and two additional proteinase K applications at 6 and 24 hr, while being mixed at 350 rpm (Eppendorf Thermomixer R, Westbury, NY) at 56°C for a total of ~45 hr. Tumor DNA was then extracted based on the manufacturers' recommendation. We quantified tumor genomic DNA by real-time PCR (Applied Biosystems, Foster City, CA) using two chromosome 2 specific probes at 2p25.3 (29,907–30,162) and 2q31.1 (21,407,882–21,408,181) with normal human female genomic DNA (Promega, Madison, WI) as the reference. Normal genomic DNA from peripheral blood lymphocytes (PBLs) was extracted and purified with the QIAamp DNA Blood Midi Kit (QIAGEN, Valencia, CA), DNA from the tumor cell sample and the blood sample from each subject was aliquoted and stored at -80°C prior to assay.

SNP Array Assay

The GeneChip® Mapping 10K Assay (Affymetrix, Santa Clara, CA), or Map 10K array, was performed as recommended by the manufacturer. Two DNA samples from each woman (tumor and PBL) were assayed on separate arrays. Purified genomic DNA (200–250 ng) was digested with the *XbaI* restriction enzyme, and then amplified by ligation-mediated PCR. As recommended by the manufacturer, multiple PCR reactions were pooled more if necessary for the formalin fixed paraffin embedded (FFPE) samples, to obtain 20 μg of PCR product for labeling and hybridization to the GeneChip® Human Mapping 10K Array Xba131 Version 1.0 (40 tumors) or Xba 142 Version 2.0 (126 tumors). Arrays were scanned with an Affymetrix GeneChip Scanner 3000 7G using GeneChip Operating System (GCOS) 1.3 and analyzed with Affymetrix GeneChip® DNA Analysis Software (GDAS) 3.0. Version 1.0 and Version 2.0 share 9,706 SNPs, which were analyzed here. The genetic map used in the analysis was obtained from GeneChip Mapping 10K library files: Mapping10K_Xba131 or Mapping10K_Xba142.

To ensure that the DNA from FFPE samples was compatible with the Map 10K array, we tested DNA from paired fresh and FFPE samples of four human cell lines (X-chromosome aneusomy cell lines, Coriell Cell Repositories, Camden NJ) on the array. The FFPE samples were processed using the same procedures as those used for tumor samples including embedding in paraffin. There was high concordance for the genotype calls (≥99%) between the matched FFPE and frozen samples and excellent reproducibility in four samples done in duplicate on different days (data not shown).

Short Tandem Repeat (STR) Validation Assays

To validate the array findings, we assayed for LOH using microsatellite markers (short tandem repeats or STRs) for multiple locations of a subset of the 166 cases. Seventy tumor samples were assayed with fluorescent primers (Operon Technologies, Alameda, CA) and standard PCR reactions, for the following STR markers in these chromosomal locations: 2q34.35 (*D17S1322* and *D17S932*), 16q22.1 (*D16S767* and *D17S932*), and 17q21 (*D2S2319* and *D16S767*). Forty-three tumor samples were assayed at 6q25.1 (*D16S496*). DNA from PBLs served as the constitutive normal genomic DNA for comparison with tumor sample. Primer sequence, map location (in Mb), and additional information associated with the STR markers can be obtained at National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/).

Statistical Analysis

An LOH event at a specified SNP for a tumor was defined as a SNP with a heterozygous (AB) genotype in the constitutive blood DNA and homozygous (AA or BB) genotype in the tumor. Genome-wide fractional allelic loss (FAL) was calculated by taking the ratio of the number of SNPs exhibiting LOH and the total number of informative SNPs in the whole genome for each tumor. FAL was also calculated for each chromosomal arm by using the SNPs on that arm only. Two-sample t statistics were used to assess the significance of the mean differences of FAL between ductal and lobular subtypes and following stratification by the categorized DI (diploid/near diploid versus aneuploid), mitotic rate (low versus intermediate or high), and tumor size T1a/T1b (≤2.0 cm), T1c (2.1–5.0 cm), and T2/T3, (>5.0 cm). The two-sample t statistics were also used to assess the association of FAL with DI and mitotic rate stratified by ILC and IDC status. The F-test was used to assess FAL in relation to tumor size stratified by ILC and IDC status.

The association of LOH at individual SNPs with lobular and ductal subtypes was also examined using a likelihood-ratio test statistic based on the logistic regression model. SNPs with informativity less than 25% across tumors were excluded to ensure an adequate sample size. The validity of this test statistic for our data was examined by randomly permuting the ILC and IDC status of these

tumors while maintaining the same informativity and frequency of LOH for each SNP. The *P* values were uniformly distributed, indicating that the test statistic is adequate in testing the differences of LOH frequencies at each SNP between ILC and IDC tumors with our data (results not shown). Genesis (http://genome.tugraz.at) was used for hierarchical clustering to group tumors based on the SNPs that show significant differences in LOH between ILC and IDC at level 0.05.

P values in all results are two-sided and considered statistically significant if they were less than 0.05 level. For individual SNP and subtype association, the proportion of SNPs of statistical significance at level 0.05 was compared to the expected proportion of false positives when no association exists between LOH and breast cancer subtypes. Because this study will be used to generate hypotheses for future studies, multiple comparison adjustment was not explicitly performed. This minimized the number of false negative results. All analyses were performed using an OpenSource R package which can be downloaded at http://www.r-project.org/.

The software dCHIP was used to obtain copy number (CN) information for each SNP (Lin et al., 2004). The Circular Binary Segmentation (CBS) method (Olshen et al., 2004) was used to smoothen the original copy number data into segments of similar copy number. CBS is a modified binary segmentation procedure to find two change-points at a time by considering the spliced segment as a circle. The R package *DNAcopy* (http://bioconductor.org/) was used for this procedure with the recommended default settings, with one exception, alpha was set as 0.05. Segmented copy number was categorized as "Gain" ($CN \geq 3.0$), "No Change" ($1.0 < CN < 3.0$), and "Loss" ($CN \leq 1.0$), assuming diploid $CN = 2.0$.

RESULTS

Clinical Characteristics of the Lobular and Ductal Tumors

One hundred and sixty-six (77 ER-positive ILC and 89 ER-positive IDC) women with ER-positive breast cancer tumors in the parent population study had DNA yield and quality sufficient for array testing. Within the group of Map10K-tested tumors, the ILC tumors were more likely than the IDC tumors to be larger, to exhibit a lower mitotic rate, and to have diploid/near diploid DNA content (Table 1). IDC and ILC tumors in the tested group did not differ significantly with respect to stage of disease, lymph node status, histological grade, or PR status. To determine the relevance of the results reported here to the general population of women with breast cancer, we compared the characteristics of the tested and untested IDC and ILC groups. The tested IDC tumors were more likely to have a higher stage ($P = 0.02$), higher histological grade ($P = 0.02$), and higher mitotic rate ($P = 0.01$) than the untested IDC group. There were no statistically significant differences between the tested and untested ILC tumors.

Reproducibility and Accuracy of LOH Identification from Archival Samples by Map 10K Array

To test the accuracy of LOH calls by the Map 10K array for the FFPE tumor samples, we performed STR genotyping at loci in or near four genes: 70 tumor/control pairs at *BRCA1* (17q21), *CDH1* (E-cadherin; 16q22.1), and *BARD1* (2q34.35), and 43 tumor/control pairs at *ESR1* (6q25.1). The LOH calls for each STR were compared on a case-by-case basis to infer calls by the Map 10K based on the flanking SNPs. The concordances in LOH calls by Map 10K array and STR were: *BRCA1* (*D17S1322*) 89%, (*D17S932*) 85%; *CDH1* (*D16S767*) 100%, (*D17S932*) 100%; *BARD1* (*D2S2319*) 100%, (*D16S767*) 100%; and *ESR1* (*D16S496*) 96%. The strong concordance of the LOH calls from the Map 10K array with an alternative and established method for determining regions of LOH supported our confidence in the identification of LOH with the Map 10K array.

We examined the SNP call rate as a measure for the performance of the Map 10K Array for the 166 archival ER-positive breast tumors. The call rate was defined as the percentage of SNPs given a genotype call ("AB," "AA," or "BB") on the array and was generated automatically for each array by the GDAS software. Similar to a previous report (Thompson et al., 2005), the SNP call rates in this study were comparatively lower for the FFPE samples (median call rate 86%; range 60–99%) than the constitutive normal PBL samples (median call rate 97%; range 82–100%). This was likely due to the compromised quality and quantity of DNA from the FFPE samples compared to the frozen, non-FFPE, constitutive normal PBL DNA.

Sites of High Frequency LOH in ER-Positive ILC and IDC tumors

SNPs showing high frequency of LOH (informativity $> 25\%$ and with LOH frequency $\geq 50\%$) were identified in ILC (165 SNPs) and in IDC (200 SNPs) (Supplemental Data 1). Table 2 shows

TABLE 1. Clinical and Tumor Characteristics Among ER-Positive IDC and ILC Carcinomas Not Tested or Tested by the Map10K Assay

	Not-MAP 10K tested ER+ cases					MAP10K tested ER+ cases				
	IDC ($n = 234$)		ILC ($n = 141$)		*P*-value[a]	IDC ($n = 89$)		ILC ($n = 77$)		*P*-value[a]
Tumor marker	*n*	%	*n*	%	IDC vs. ILC	*n*	%	*n*	%	IDC vs. ILC
AJCC stage										
I	152	66	68	48		44	49	33	43	
II	71	31	58	41		42	47	38	49	
III/IV	9	4	15	11	<0.0001	3	3	6	8	0.23
Tumor size (cm)										
T1a/T1b	167	77	75	56		57	66	38	50	
T1c	44	20	47	35		27	31	28	37	
T2/T3	7	3	13	10	<0.0001	2	2	10	13	0.008
Lymph node status										
Negative	165	72	90	65		55	63	42	57	
Positive	64	28	48	35	0.17	33	38	32	43	0.46
Histologic grade										
Low	114	49	68	48		32	36	26	34	
Intermediate	93	40	70	50		39	44	47	61	
High	27	12	3	2	0.19	18	20	4	5	0.21
Mitotic rate										
Low	181	77	134	95		55	62	71	92	
Intermediate	34	15	6	4		21	24	5	7	
High	19	8	1	1	<0.0001	13	15	1	1	<0.0001
Ploidy										
Diploid/near diploid	86	58	74	68		52	58	62	81	
Aneuploid	63	42	35	32	0.1	37	42	15	20	0.003
PR[b]										
Positive	204	88	114	82		84	94	71	93	
Negative	28	12	25	18	0.12	5	6	5	7	0.8

[a]chi square *P*-value.
[b]Allred score <3, negative; ≥3, positive.

the cytogenetic band locations of the SNPs with high frequency LOH for ILC and IDC tumors. We observed SNPs with high frequency LOH that were common to both histological subtypes (SNPs in 11q, 16q, and 17p) and as has been reported previously, we observed these regions of common LOH events were associated with copy number "loss" and "no change" (Carter et al., 1994; Tomlinson et al., 1996; Roylance et al., 2006). There were also SNPs for which LOH occurred predominantly and frequently in ILC (6q) or in IDC (8p, 13q). These results indicate that ILC and IDC ER-positive tumors share common sites of high frequency LOH in addition to LOH events occurring primarily in one histological subtype and not the other.

Differential LOH Frequencies Between ER-Positive ILC and IDC Tumors

To characterize further the differences in the frequency and location of LOH events between IDC and ILC, we calculated the average genome-wide FAL for both tumor types. We found IDC had only slightly more LOH events than ILC tumors (FAL = 0.186 vs. 0.156, respectively). The distribution of FAL across the genome for the two subtypes differs in shape: the tumors in the upper quartile range have fewer LOH events in ILC than IDC (FAL = 0.204 and 0.499, respectively), although the distributions have nearly identical medians (FAL = 0.152 and 0.151 for ILC and IDC, respectively) (Fig. 1A).

We compared the average FAL for ILC and IDC for each chromosomal arm. Several chromosomal arms showed a relatively high frequency of LOH events (>25% FAL) for both ILC and IDC (6q, 11q, 16q, 17p, and 22q), consistent with what was observed above in the genomic distribution of individual high-frequency LOH SNPs, but we also observed that IDC tumors have a significantly higher frequency ($P < 0.05$) of FAL than ILC on 3p, 5q, 8p, 9p, 20p, 20q (Fig. 1B). Several chromosomal arms had a higher frequency of FAL for ILC (1q, 6q, 16p and q, and X) but the differences were

TABLE 2. SNPs with High Frequency LOH (≥50% Fractional Allelic Loss)

		ILC	IDC	
Chromosomal arm	Cytogenetic band	Number of SNPs with high frequency LOH in the cytoband	Number of SNPs with high frequency LOH in the cytoband	Genes 100 kb up or downstream of SNPs with high frequency LOH
1p	1p36.12		**1**	C1QB, EPHA8, ZBTB40, C1QA, C1QC
3p	3p26.1[a]	**1**	**1**	GRM7
	3p14.2		**2**	CADPS
	3p14.1		**1**	MAGI1
3q	3q21.3	**1**		UROC1, AB113648, BC042038, BC103896, C3orf22, CHST13
	3q29[a]	**1**	**1**	FGF12, C3orf59
5q	5q23.2	**1**		
6p	6p12.3	**1**	**1**	OPN5, FLJ41841
6q	6q12		**1**	PTP4A1
	6q13	**2**		BAI3, RIMS1
	6q14.1	**1**	**2**	HTR1B
	6q16.1	**1**		
	6q16.3	**1**		
	6q21	**2**	**1**	SEC63, OSTM1, LOC442247
	6q22.2	**1**		DCBLD1, ROS1
	6q22.31	**2**		HSF2, SERINC1
	6q23.3	**1**		HBS1L, ALDH8A1
	6q24.3	**1**		STXBP5, BC107737
	6q25.1	**1**	**1**	C6orf71, PPP1R14C
	6q25.2	**3**		VIP
	6q26	**1**	**2**	PARK2, PACRG, QKI
8p	**8p23.3**		**1**	MYOM2
	8p23.2		**3**	
	8p23.1		**7**	ANGPT2, MCPH1, DEFA4, DEFA6, DEFA1, DEFB1, PPP1R3B, THEX1, AY203962, AB073660, XKR6, AJ307469, C8orf16, C8orf15
	8p22[a]	**1**	**11**	SGCZ, TUSC3, MSR1, BC098128, FGF20, EFHA2, VPS37A, MTMR7, AY176665, PSD3
	8p21.3		**2**	C8orf58, BIN3, KIAA1967, SORBS3, PPP3CC, PDLIM2
	8p21.2		**5**	DOCK5, EBF2, AK130123, AY700779, AK130123
	8p21.1		**4**	ESCO2, CCDC25, SCARA3, PBK, ELP3, ZNF395, PNOC, RC74, FLJ21616, EXTL3
	8p12		**1**	WRN, PURG
8q	8q11.23	**1**		OPRK1
9q	**9q31.1**	**1**		PPP3R2, GRIN3A
10q	10q21.1	**1**		PCDH15
11q	11q14.1	**1**		DLG2
	11q14.2		**1**	
	11q14.3	**1**		
	11q21[a]	**3**	**2**	FAM76B, CEP57, MTMR2, MAML2, AY358248
	11q22.1[a]	**2**	**4**	CNTN5, PGR, AK057372, AD031, KIAA1377, ANGPTL5
	11q22.2	**1**		MMP7, MMP8, MMP20, MMP27
	11q22.3[a]	**7**	**6**	AB231766, MMP13, DCUN1D5, AB082528, AB231766, BC037496, BX538093, DDI1, PDGFD
	11q23.1[a]	**4**	**4**	ARHGAP20, BTG4, POU2AF1, FLJ46266, AK096925, LAYN, BTG4, FLJ46266
	11q23.2		**3**	DRD2, TTC12, AK125909, ANKK1, ZBTB16
	11q23.3		**2**	BACE1, DSCAML1, PCSK7, DKFZp547C195, CEP164, ARHGEF12
	11q24.1	**3**	**5**	SORL1, BRCC2, ASAM, OR4D5, OR10S1, PMP22CD, LOH11CR2A, OR10G7, OR10G8

(Continued)

TABLE 2. SNPs with High Frequency LOH (≥50% Fractional Allelic Loss) (Continued)

Chromosomal arm	Cytogenetic band	ILC: Number of SNPs with high frequency LOH in the cytoband	IDC: Number of SNPs with high frequency LOH in the cytoband	Genes 100 kb up or downstream of SNPs with high frequency LOH
	11q24.2	**3**	**3**	PKNOX2, AK055281, FEZ1, MGC39545, CHEK1, ACRV1, STT3A, KIRREL3
	11q25	**1**	**3**	HNT, OPCML
12p	12p13.31	**1**		PZP
	12p11.22	**1**		AF113698, MLSTD1
12q	12q24.21	**1**		
	12q24.33[a]	**1**	**1**	KIAA1944
13q	13q12.12[a]	**1**	**1**	SPATA13
	13q14.11		**1**	
	13q21.1		**2**	PCDH17
	13q21.32		**2**	
	13q31.1	**1**		
	13q31.2		**1**	
	13q31.3		**1**	
	13q33.1		**1**	FGF14
14q	**14q31.3**	**1**		GALC,GPR65
15q	15q21.2		**1**	
16p	16p13.3	**1**		
16q	**16q12.1**[a]	**11**	**11**	ZNF423, CYLD, SALL1
	16q12.2[a]	**13**	**13**	CHD9, FTS, RBL2, BC003583, IRX3, IRX6
	16q13[a]	**8**	**8**	MT1E, MT1G, MT1X, MT1A, MT1B, MT1F, MT1H, MT2A, MT3, MTM, MT1JP, MT1M, NUP93,C16orf50, GPR56, KATNB1, GPR114, GPR97, CNGB1, C16orf50, KIFC3, BC082977, KATNB1
	16q21[a]	**20**	**20**	CDH8, CDH11
	16q22.1[a]	**2**	**2**	LOC497190, MGC34761, BC063633, LOC348174, WWP2, COG4, FUK, SF3B3, ST3GAL2
	16q22.3[a]	**6**	**6**	AY358233,AP1G1,PHLPPL,LOC55565, KIAA0174,DHODH,LOC55565, PMFBP1,HP,DHX38,TXNL4B,HPR
	16q23.1[a]	**19**	**19**	ADAMTS18, WDR59, FA2H, ZNRF1, KIAA1576, CLEC3A, WWOX, MAF
	16q23.2[a]	**1**	**1**	
	16q23.3[a]	**3**	**3**	CDH13, HSBP1
	16q24.1[a]	**5**	**5**	KIAA0703, KIAA1609, ZDHHC7, KIAA0513, FAM92B, AK127184, IRF8
	16q24.2[a]	**2**	**2**	
	16q24.3[a]	**1**	**1**	AK023048, AK092264, ANKRD11, AK125549
17p	17p13.3[a]	**1**	**1**	CARKL, P2RX5, CR624034, TMEM93, CTNS, ITGAE, TAX1BP3, HSA277841, GSG2
	17p13.2		**1**	ZZEF1, AK074369, AK128386, ATP2A3, CYB5D2
	17p13.1[a]	**3**	**8**	MYH10, CCDC42, FLJ35773, WDR16, STX8, GAS7, C17orf48, MYH2, LOC388335, MYH3,SCO1, FLJ45455
	17p12[a]	**3**	**3**	AK092009, MYOCD, KIAA0672, ELAC2
18p	18p11.32		**1**	
20p	20p13[a]	**1**	**1**	DEFB125, BC048429, DEFB128, DEFB126, DEFB129, DEFB127, DEFB32
22q	22q12.1	**2**	**1**	SEZ6L, AB051436, AK172861, KREMEN1, HS747E2A
	22q12.3[a]	**3**	**2**	RAXLX
Xp	Xp11.3	**1**		

Cytobands in bold contain one or more statistically significant SNP(s) that have differential LOH frequencies that differentiate ILC from IDC.
[a]Cytobands containing some (or all) common SNPs with high frequency LOH in both ILC and IDC.

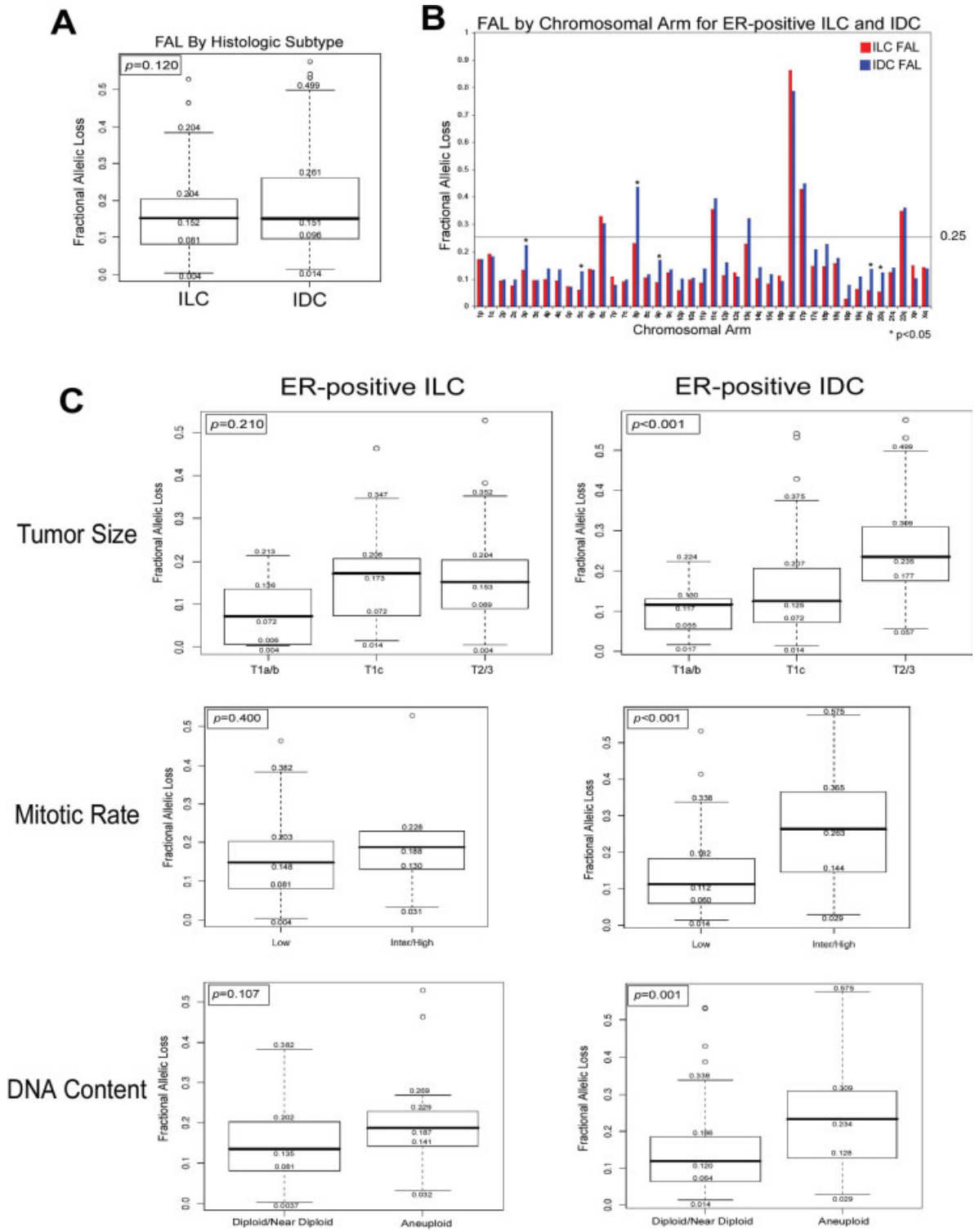

Figure 1. Average LOH Frequencies across the genome or by chromosomal arm in ER-positive ILC and IDC tumors. (A) A box plot of fractional allelic loss calculated for all informative SNPs in the genome (FAL; ratio of the # LOH and # informative SNPs) for tumors grouped by histological subtype. (B) Average FAL by chromosomal arm for ER-positive ILC (red bars) and IDC (blue bars) tumors. The line represents 25% FAL. (C) Box plots of FAL across the genome for each ER-positive ILC and IDC tumors stratified by tumor size, mitotic rate, and DNA content. The median FAL for ILC and IDC tumors are shown as the bold line in the box with the 1st and 3rd quartiles as the lower and upper bars of the box. Two-sample *t* test was used to calculate the *P* values for genome-wide and chromosomal-arm FAL comparisons. The F-test was used to calculate the *P* values for genome-wide FAL for tumors stratified by tumor size.

not significant. Taken together, these results indicate that significant differences in LOH frequency between ILC and IDC tumors can be detected in specific chromosomal arms.

We used categorized copy number data (see Materials and Methods) to assess genomic alteration events associated with the allelic loss on the six chromosomal arms with statistically significant differences in LOH frequencies between ILC and IDC. We observed the majority of LOH events to be associated with no copy number change (89 and 82%), for ILC and IDC respectively, with a lower percentage associated with copy number loss (11 and 17%), and dramatically lower with copy number gain (0.14 and 0.53%) for these six chromosomal arms, suggesting that the mechanism for the majority of observed LOH, associated with no copy number change, on these six chromosomal arms is the result of an initial deletion of one allele followed by reduplication of the remaining allele.

Differential LOH Frequency in Relation to Tumor Size, Mitotic Rate, or DNA Content

Three clinical characteristics, tumor size, mitotic rate, and DNA content, distinguished ILC and IDC tumors (Table 1). We stratified the tumors based on these three clinical characteristics and compared both genome-wide and chromosomal arm-specific FAL for ILC and IDC tumors.

We compared the whole genome FAL for ILC and IDC tumors grouped in AJCC tumor size categories (T1a/T1b, T1c, and T2/T3) and found a strong correlation between increased genome-wide FAL and tumor size ($P < 0.001$) for the three size categories in IDC tumors (Fig. 1C). This correlation was not observed in ILC tumors ($P = 0.210$). The average FAL for individual chromosomal arms were determined for the tumor size groups to grossly map where there may be similarities or differences in the quantities and locations of these LOH events. We observed no significant differences in FAL at the chromosomal arms for ILC and IDC in the T1a/T1b size group. However, there were significant differences in three chromosomal arms (7p, 8p, 16q) in the T1c group and in nine chromosomal arms (3p, 5q, 8p, 9p, 13q, 15q, 18p, 20p, 20q) in the T2/T3 group (Table 3). Overall these results indicate divergence of frequency and location of genomic alterations between the subtypes with increasing tumor size.

ILC tumors generally demonstrate low mitotic activity. For this study, 92% of the ILC tumors were in this category. We observed a significant difference ($P < 0.001$) in the genome-wide FAL between low versus inter/high mitotic rate IDC tumors, whereas ILC tumors stratified by mitotic rate did not show this difference ($P = 0.400$) (Fig. 1C). We observed significant differences in FAL between ILC and IDC for chromosome arms 1p, 7p, 8p, and 12q in the low mitotic rate group, and in 4q, 5q, 7q, 16q, 17q, 19p, 20p, and 20q in the high mitotic rate group (Table 3).

DNA index is a quantitative measure of DNA content and was determined for all tumors during the flow sorting step. For the 77 ILC and 89 IDC tumors, 62 and 52 were diploid/near diploid and 15 and 37 were aneuploid, respectively. When tumors of each histological subtype were stratified by ploidy, we observed the aneuploid tumors of both ILC and IDC had a higher median whole genome FAL than the diploid/near diploid tumors in that subtype (Fig. 1C). This difference between ploidy groups was statistically significant for IDC tumors ($P = 0.001$), but not significant for ILC tumors ($P = 0.107$). Two chromosomal arms (7p and 12q) showed significant differences ($P < 0.05$) in FAL between diploid ILC and IDC. Five chromosomal arms (8p, 13q, 16q, 20p, and 20q) showed significant difference ($P < 0.05$) in FAL between aneuploid ILC and IDC (Table 3).

Identification of differences in LOH frequency and locations between ILC and IDC tumors measured by both whole genome and chromosomal arm FAL is most dramatic when comparing tumors with increased tumor size, mitotic rate, and ploidy.

Identification of Loci with Differential LOH Frequencies for ER-Positive ILC and IDC Tumors

In addition to the observed differences between ILC and IDC in overall or chromosome-arm specific FAL, we identified specific SNPs that showed differential LOH frequencies in ILC and IDC. We found that 9.5% ($n = 544$) of SNPs with >25% informativity had statistically significant ($P < 0.05$) differences in LOH frequencies between ILC and IDC (Supplemental Data 2). This percentage of SNPs is much greater than the 5% expected if none of the SNPs is associated with subtype. These SNPs are distributed throughout the genome (Fig. 2A) and of the 544 SNPs, 466 (86%) had a higher frequency of LOH in IDC, and 78 SNPs (14%) had a higher frequency of LOH in ILC (Table 4). The LOH frequencies for individual SNPs in each subgroup are plotted along the genome in Figure 2B to illustrate the differences in frequencies for the two subgroups at these SNPs.

TABLE 3. Relationship of Tumor Characteristics and Difference in FAL Frequency for ILC and IDC Tumors by Chromosomal Arm[a]

	Chromosomal arm	1p	1q	2p	2q	3p	3q	4p	4q	5p	5q	6p	6q	7p	7q	8p	8q	9p	9q	10p	10q
	No stratification (ILC=77 ; IDC=89)					0.04					0.02					<0.01		0.03			
Tumor size	T1 a/b (ILC= 6; IDC=12)																				
	T1c (ILC=32; IDC=48)													0.05		0.04					
	T2 & T3 (ILC=39; IDC=29)					0.01					0.03					<0.01		0.03			
Mitotic rate	Mitotic rate low (ILC=71 ; IDC=55)	0.03												<0.01		0.01					
	Mitotic rate Inter/High (ILC=6 ; IDC=34)								0.05		0.01				<0.01						
DNA content	Diploid/near diploid (ILC=62 ; IDC=52)													0.03							
	Aneupioid (ILC=15; IDC=37)															<0.01					

	Chromosomal arm	11p	11q	12p	12q	13q	14q	15q	16p	16q	17p	17q	18p	18q	19p	19q	20p	20q	21q	22q	Xp	Xq
	No stratification (ILC=77 ; IDC=89)																0.01	0.01				
Tumor size	T1 a/b (ILC= 6; IDC=12)																					
	T1c (ILC=32; IDC=48)									0.01												
	T2 & T3 (ILC=39; IDC=29)					0.02		0.04					0.02				0.02	0.01				
Mitotic rate	Mitotic rate low (ILC=71 ; IDC=55)				0.05																	
	Mitotic rate Inter/High (ILC=6 ; IDC=34)									0.01		<0.01			0.03		0.03	0.01				
DNA content	Diploid/near diploid (ILC=62 ; IDC=52)				0.01																	
	Aneupioid (ILC=15; IDC=37)					0.03				0.01							0.01	0.02				

[a]Chromosomal Arme with statistically significant ($P < 0.05$; Two-sample *t*-test for Mitotic rates and DNA content and the *F* -test for tumor size comparisons) differences in FAL frequencies for ILC and IDC tumors. Pink and blue fill indicates a significantly higher frequency of FAL in ILC and IDC tumors, respectively, on the chromosomal arm.

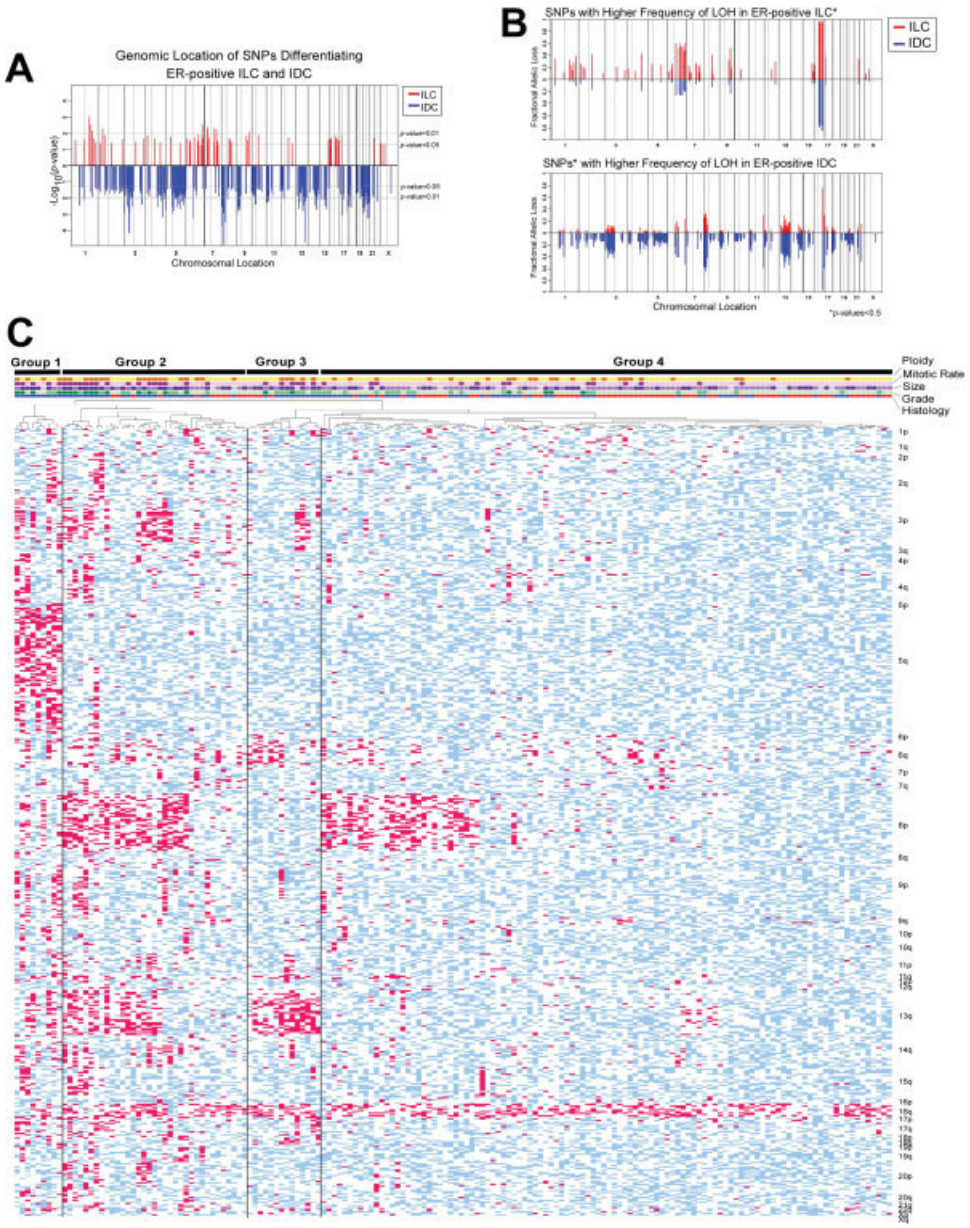

Figure 2. Genomic Distribution and Patterns of LOH Differentiates ER-positive ILC and IDC Tumors. (A) The genomic distribution of all SNPs whose LOH frequency differs significantly between IDC and ILC (P values < 0.05) (Likelihood Ratio Test). The y-axis is the $-\text{Log}_{10}(P$ value) reflecting the P values of SNPs with significant differences in the frequency of LOH between ILC and IDC tumors. The red bars above the x-axis represent SNPs with a significantly higher frequency of LOH in ILC than IDC and the blue bars below the x-axis represent SNPs with a significantly higher frequency of LOH in IDC than ILC. Chromosomal locations are indicated. (B) Whole genome plots of the FAL for SNPs ($P < 0.05$) illustrating the relative differences in FAL frequency at each. The top plot represents the differential SNPs with a higher frequency of LOH in ILC and the bottom plot represents the differential SNPs with a higher frequency of LOH in IDC The red bars represent the FAL of ILC tumors and the blue bars represent the FAL of IDC tumors. (C) Hierarchical clustering of the differential SNPs ($P < 0.05$). Tumors are grouped into four major groups based on their patterns of LOH and retention of heterozygosity (ROH). Tumor characteristics such as histological subtype (red is ILC and blue is IDC), grade (dark to light green represent high to low grade, respectively), size (dark to light purple represent T2/3 to T1a/b, respectively), mitotic rate (dark and light violet represent high/intermediate to low, respectively), and ploidy (aneuploid orange and diploid yellow) are indicated above the dendrogram. A pink block in the clustermap represents LOH, a blue block represents ROH; a white block represents a non-informative SNP.

Hierarchical clustering was performed using the 544 SNPs showing differential LOH to evaluate common patterns of LOH events and their association with tumor characteristics such as histology, grade, tumor size, mitotic rate, and ploidy (Fig. 2C). The dendrogram shows that the 166 tumors form four major groups. Tumors in Group 1 are exclusively IDC, usually large (>2.0 cm), and have LOH primarily in 3p, 5q, and 13q. Group 2 consists primarily of IDC tumors (67%), the majority of

TABLE 4. Cytoband Location of SNPs[a] with Differential Frequencies of LOH for ER-Positive ILC and IDC Tumors

	Higher frequency of LOH in ILC		Higher frequency of LOH in IDC	
Chromosomal arm	Cytogenetic band	Genes 100 kb up or downstream of SNPs with differential LOH frequencies	Cytogenetic band	Genes 100 kb up or downstream of SNPs with differential LOH frequencies
1p	1p34.3, 1p13.2	CSMD2, C1orf94, ADORA3, ATP5F1, RAP1A, WDR77, C1orf162	1p31.3, 1p31.1, 1p22.1, 1p13.3	AK124028, GNG12, ST6GALNAC3, ABCA4, ARHGAP29, PSRC1, AK127086, AK127745, MYBPHL, CELSR2, PSMA5, SORT1
1q	1q23.1, 1q23.2, 1q25.1, 1q31.11q32.1, 1q41, 1q42.12	OR10K2, CRP, APCS, TNR, IPO9, NAV1, BX648460, SUSD4, C1orf65, ENAH	1q25.2, 1q31.1, 1q41, 1q42.2	PAPPA2, C1orf136, DISC1
2p	2p24.1, 2p21, 2p16.1		2p25.1, 2p24.1, 2p16.1, 2p14, 2p12	RSAD2, RNF144, LOC129607, AK096196, LRRTM4
2q	2q14.3, 2q36.1	CNTNAP5, FARSLB, MOGAT1	2q11.2, 2q12.1, 2q14.3, 2q22.1, 2q23.1, 2q24.1, 2q24.3, 2q33.1, 2q33.3, 2q34, 2q36.2, 2q36.3, 2q37.1	NPAS2, CNTNAP5, LRP1B, EPC2, MBD5, GALNT13, COBLL1, HECW2, STK17B, HECW2, MPP4, ALS2, AK131512, ALS2CR7, KIAA0971, MDH1B, KIAA0971, CPS1, LANCL1, DOCK10, FLJ20701, DNER, AK128274, SP110, SP140,
3p			3p26.3, 3p26.2, 3p26.1, 3p25.2, 3p24.3, 3p24.2, 3p24.1, 3p23, 3p22.3, 3p22.1, 3p21.31, 3p14.3, 3p14.2, 3p14.1, 3p13, 3p12.3, 3p12.1	SUMF1, AK058041, AK094424, ATG7, VGLL4, TIMP4, RPL15, UBE2E1, NKIRAS1, NR1D2, RPL15, NKIRAS1, MYRIP, SEC22L3, VIPR1, LYZL4, LARS2, WNT5A, CAST1, FHIT, PTPRG, SUCLG2, FAM19A1, LOC401072, ROBO1, IGSF4D,
3q	3q13.12, 3q26.33	CD47,IFT57, FXR1, DNAJC19	3q24, 3q25.2, 3q26.1	PLSCR1, P2RY1, ARL14 , PPM1L, PPM1L,
4p	4p15.33	BC036758	4p15.33, 4p15.32, 4p15.1, 4p14	APBB2
4q	4q21.21, 4q28.3	BX648337, RASGEF1B	4q24, 4q25, 4q26, 4q28.1, 4q28.3, 4q31.21, 4q31.3, 4q32.1, 4q33, 4q34.2, 4q34.3, 4q35.1, 4q35.2	BANK1, TACR3, HADHSC, LEF1, CYP2U1, MGC26963, SPRY1,IL15, AF494508, USP38, GAB1, PET112L, DCHS2, GUCY1A3, AK172805, C4orf18, FLJ11155, NEIL3, AGA, NEIL3, FAT, MTNR1A
5p	5p13.1,	OXCT1	5p15.31, 5p13.1	ADCY2, FLJ40243, C7
5q	5q23.2		5q11.2, 5q12.1, 5q12.3, 5q13.1 5q13.3, 5q14.1, 5q14.2, 5q14.3, 5q15, 5q21.1 5q21.2, 5q21.3, 5q22.1, 5q22.2, 5q23.1, 5q23.2, 5q23.3, 5q31.1, 5q31.2, 5q31.3, 5q32, 5q33.2, 5q33.3, 5q34, 5q35.1, 5q35.2	KCTD9, ARL15, AJ973643, ESM1, GZMK, FLJ37927, LOC493869, GZMA, UNG2, R7BP, ADAMTS6, AK130467, FLJ13611, SGTB, BC047475, TRIM23, PPWD1, FKSG14, AK057601, LOC375449, FLJ35779, LOC441087, AK128395, TBCA, OTP, MSH3, RASGRF2, XRCC4, MGC23909, ARTS-1, LRAP, LNPEP, AF119888, SLCO6A1, SLCO4C1, MAN2A1, CAMK4, WDR36, APC, AK125966, MCC, TSSK1, YTHDC2, TNFAIP8, DMXL1, SNCAIP, CSNK1G3, FLJ36090, AK093561, LMNB1, MARCH3, AK026965, H2AFY, PCDHGB7, PCDHGA10, PCDHGB5, PCDHGB2, PCDHGA2, PCDHGC3, PCDHGA12, PCDHGB4, TAF7, PCDHGA8, PCDHGA1, PCDHGA11, PCDHGA3, PCDHGA4, PCDHGA5, PCDHGA6, PCDHGA7, PCDHGA9, PCDHGB1, PCDHGB3, PCDHGB6, PCDHGC4, PCDHGC5, SLC25A2, HB-1, POU4F3, TCERG1, POU4F3, TCERG1, AK172724, CYFIP2, MGC27121, AK128840, ADRA1B, TTC1, GABRG2, MAT2B, HMMR, WWC1, SLIT3, FLJ20364, DOCK

(Continued)

TABLE 4. Cytoband Location of SNPs[a] with Differential Frequencies of LOH for ER-Positive ILC and IDC Tumors (Continued)

	Higher frequency of LOH in ILC		Higher frequency of LOH in IDC	
Chromosomal arm	Cytogenetic band	Genes 100 kb up or downstream of SNPs with differential LOH frequencies	Cytogenetic band	Genes 100 kb up or downstream of SNPs with differential LOH frequencies
6p	6p24.1, 6p21.2, 6p21.1	PHACTR1, MDGA1, AK126965, AK127725, C6orf129, SLC35B2, AF090903, HSP90AB1, NFKBIE, MGC33600, SPATS1, CDC5L, AARSL	6p25.3	GMDS
6q	6q14.1, 6q21, 6q22.2, 6q23.3, 6q25.2, 6q25.3, 6q26	HTR1B, SEC63, OSTM1, DCBLD1, ROS1, HBS1L, ALDH8A1, SYNE1, MYCT1, RSHL2, TAGAP, PNLDC1, ACAT2, MAS1, MRPL18, TCP1, MAP3K4, AK094629	6q12, 6q13, 6q15, 6q16.3, 6q26	CD109, MAP3K7, PARK2
7p	7p15.2, 7p15.1, 7p14.3, 7p14.2, 7p14.1	HIBADH, CREB5, CPVL, KIAA0241, LOC441208, PTHB1, FLJ22313, SEPT7	7p21.3, 7p21.2	
7q	7q21.13, 7q21.3, 7q31.1, 7q31.33	PFTK1, PON2, PON1, PON3, FOXP2, POT1, GPR37, GRM8	7q22.3, 7q31.1, 7q33	AK127860, PIK3CG, PRKAR2B, EXOC4, AKR1B1, SLC35B4, FLJ32786
8p			8p23.3, 8p23.2, 8p23.1, 8p22, 8p21.3, 8p21.2, 8p21.1, 8p12	MYOM2, ANGPT2, MCPH1, AF217970, AK126737, AY461701, PPP1R3B, THEX1, AY203962, PINX1, SOX7, RP1L1, AB073660, XKR6, AJ307469, C8orf16, C8orf15, AJ291678, AK131319, BC080558, C8orf13, BLK, SGCZ, BC098128, FGF20, EFHA2, AY176665, PSD3, SLC18A1, LPL, ATP6V1B2, LZTS1, C8orf58, PEBP4, BIN3, KIAA1967, SORBS3, EGR3, PPP3CC, PDLIM2, FLJ14107, DOCK5, AK130123, AY700779, EBF2, ELP3, ZNF395, PNOC, RC74, FLJ21616, EXTL3, WRN, PURG, FUT10, CR596972
8q	8q21.11, 8q21.13, 8q21.3	TPD52, TMEM64, EFCBP1	8q12.1, 8q12.2, 8q12.3, 8q13.2, 8q21.3, 8q23.1, 8q24.22	IMPAD1, RAB2, CHD7, ZFPM2, ADCY8, ST3GAL1
9p			9p24.2, 9p24.1, 9p23, 9p22.2, 9p21.3, 9p21.2, 9p21.1, 9p13.3, 9p13.1	SMARCA2, CR595322, GLIS3, TPD52L3, UHRF2, JMJD2C, MPDZ, SH3GL2, MLLT3, KIAA1797, IFNA5, IFNA21, IFNA14, IFNA10, IFNA16, IFNW1, IFNA7, IFNA4, IFNA17, MTAP, ELAVL2, TUSC1, LRRN6C, AF370382, AY358408, C9orf127, NPR2, OR2S2, C9orf128, SPAG8, HINT2, SHB
9q	9q21.33, 9q31.1, 9q31.2	NTRK2, TMEFF1, LOC347273, PRG-3, PPP3R2, GRIN3A	9q21.2, 9q21.31	GNAQ
10p			10p15.3, 10p15.1, 10p14, 10p13, 10p12.33, 10p12.31, 10p12.1, 10p11.23, 10p11.22	ZMYND11, DQ335455, DIP2C, PFKFB3, AK128185, CCDC3, CAMK1D, SLC39A12, MRC1, PLXDC2, KIAA1217, SVIL, ARHGAP12

(Continued)

TABLE 4. Cytoband Location of SNPs[a] with Differential Frequencies of LOH for ER-Positive ILC and IDC Tumors (Continued)

	Higher frequency of LOH in ILC		Higher frequency of LOH in IDC	
Chromosomal arm	Cytogenetic band	Genes 100 kb up or downstream of SNPs with differential LOH frequencies	Cytogenetic band	Genes 100 kb up or downstream of SNPs with differential LOH frequencies
10q	10q21.1	C10orf70, IPMK	10q11.21, 10q21.1, 10q21.3, 10q22.3, 10q23.31	RET, PCDH15, RAI17, LIPA, IFIT2, IFIT3, CH25H,
11p			11p15.4, 11p15.3, 11p15.2, 11p15.1, 11p14.3, 11p11.2	RRM1, STIM1, TRIM6, TRIM6-TRIM34, TRIM34, UBQLN3, MGC20470, SYT9, STK33, LOC58486, MICAL2, DKK3, LOC387755, AB231748, MRGPRX2, PHF21A, PTPRJ
11q			11q11, 11q25	OR5M10, OR5M8, HNT, OPCML
12p			12p13.32, 12p13.2, 12p13.1, 12p12.2	CCND2, C12orf5, FGF23, FGF6, BCL2L14, GRIN2B, SLCO1B3, LST-3TM12, SLCO1C1
12q	12q15, 12q23.3	CPM, HCFC2, NFYB	12q14.1, 12q21.1	
13q			13q12.11, 13q12.3, 13q14.11, 13q14.3, 13q21.1, 13q21.32, 13q21.33, 13q22.3, 13q31.1, 13q31.2, 13q31.3, 13q32.1 13q32.31 ,3q33.1	CRYL1, BC071810, ZDHHC20, FLJ14834, MGC40178, AY116215, AY116216, FLJ40919, EPST11, DLEU7, FLJ30707, GUCY1B2, LECT1, PCDH8, PCDH17, PCDH9, DACH1, GPC6, DOCK9, PCCA, BC001077, FGF14
14q			14q11.2, 14q22.2, 14q23.1, 14q23.2, 14q24.1, 14q24.2, 14q24.3, 14q31.1, 14q31.3, 14q32.11, 14q32.13, 14q32.2, 14q32.33	FLJ10357, HNRPC, BC031469, CR617382, CR618390, ZNF219, RNASE7, BC035680, BC063432, MGC40069, PLEKHC1, DDHD1, FLJ46156, PPP2R5E, RHOJ, GALNTL1, EXDL2, ERH, SLC39A9, SIPA1L1, RGS6, AF113687, AF130112, DPF3, WDR21A, TTLL5, TGFB3, MGC16028, NRXN3, DIO2, TSHR, GTF2A1, AK130927, CHES1, BC069658, SERPINA9, SERPINA11, SERPINA5, SERPINA3, SERPINA4, SERPINA12, DICER1, CLMN, FLJ45244, C14orf68, SLC25A29, WDR25, CR607410, YY1, WARS, AK124507, ZFYVE21, KNS2, XRCC3, PPP1R13B
15q	15q26.1	AK124722, SV2B, SLCO3A1		
16p	16p13.2, 16p13.13	A2BP1, NUBP1, EMP2, FLJ32871		
16q	16q12.1, 16q12.2, 16q13, 16q22.3, 16q23.1	ZNF423, BC003583, BC003583, IRX3, IRX6, C16orf50, GPR56, KATNB1, GPR114, GPR97, AY358233, AP1G1, PHLPPL, LOC55565, MAF	16q23.1	KIAA1576
17p			17p13.1	GAS7, FLJ45455
17q	17q11.1	WSB1, KSR1	17q11.2, 17q12, 17q21.2, 17q25.3	SSH2, RAB11FIP4, ACCN1, GAS2L2, AP2B1, C17orf50, MMP28, RASL10B, TCF2, KRT25A, KRT25D, KRT12, KRT10, KRT24, KRT25C, KRT25B, FLJ21865, LOC146713, raptor
18p			18p11.31, 18p11.22, 18p11.21	AK126828

(Continued)

TABLE 4. Cytoband Location of SNPs[a] with Differential Frequencies of LOH for ER-Positive ILC and IDC Tumors (Continued)

	Higher frequency of LOH in ILC		Higher frequency of LOH in IDC	
Chromosomal arm	Cytogenetic band	Genes 100 kb up or downstream of SNPs with differential LOH frequencies	Cytogenetic band	Genes 100 kb up or downstream of SNPs with differential LOH frequencies
18q			18q11.2, 18q12.1	DSG1, DSC1
19p			19p13.3, 19p13.12	ZNRF4, PGLYRP2, AK024488, AK131404, BC030281, BRD4, AKAP8, AKAP8L
19q			19q13.12, 19q13.2, 19q13.31, 19q13.33, 19q13.42, 19q13.43	AF289566, ZNF565, AK090827, COX7A1, ZNF146, KIAA1559, ZNF420, AK098695, BC026081, ZNF585A, AF090927, EIF3S12, AK094707, AK126566, LGALS7, ACTN4, ECH1, LGALS4, MAP4K1, CAPN12, ZNF180, ZNF229, KLK4, KLK3, KLK2, KLK15, KLK1, BC033237, BC066878, GPR32, CLEC11A, SHANK1, ACPT, MGC13170, KLK11, KLK7, KLK6, KLK5, AY551001, AY646152, KLK13, KLK12, KLK14, KLK10, KLK8, KLK9, ATPBD3, UNQ3033, PRKCG, CACNG8, CACNG7, CACNG6, LILRB1, LILRA2, FLJ00060, LILRB4, LAIR2, LILRA1, ZNF272, AURKC, LOC390980, ZNF264, USP29, IM3, ZNF543
20p			20p12.3, 20p12.2, 20p12.1, 20p11.23, 20p11.21	BMP2, PLCB1, AK126221, PLCB1, JAG1, BTBD3, C20orf38, AK074473, AK075271, SLC24A3, THBD, CD93, GGTLA4, ENTPD6, CR600960, CR936765, PYGB, C20orf22
20q			20q11.23, 20q12, 20q13.13, 20q13.2, 20q13.31, 20q13.32	C20orf132, RPN2, GHRH, SRC, MANBAL, PTPRT, PTPN1, C20orf175, ZFP64, AK096426, RAE1, RNPC1, SPO11, AK128005,
21q	21q22.3	AF289552, UMODL1, TFF3, ABCG1, U34919	21q22.11, 21q22.13	KRTAP19-1, KRTAP19-4, KRTAP19-5, KRTAP19-7, AB096950, KRTAP6-2, AB096956, KRTAP20-2, AB096958, KRTAP19-2, KRTAP22-1, KRTAP20-1, KRTAP6-1, KRTAP6-3, KRTAP19-3, KRTAP19-6, TIAM1, DSCR6, TTC3, HLCS, PIGP
22q	Xp22.31, Xp11.4		22q11.21	LZTR1, SLC7A4, DKFZp434K191, BC063553, P2RXL1, CR600536, LOC400891, THAP7, MGC16703, FLJ42953
Xq			Xq22.1	BEXL1, TCEAL5, BEX2, TCEAL8

SNPs whose LOH frequency differs significantly between IDC and ILC (P-values < 0.05 based on the likelihood ratio test).

which were of high or intermediate grade. Tumors in this group had LOH in 3p, 8p, 13q, and 16q. Group 3 is predominantly IDC tumors (67%) of intermediate grade with several large (>2.0 cm) ILC diploid/near diploid tumors. The tumors in this group had LOH in 3p, 6q, 13q, and 16q. Group 4 consists predominantly of ILC and low-grade IDC tumors, together accounting for 75% of the tumors in this group. Tumors in this group had an overall lower frequency of LOH in these 544 SNPs but virtually all showed LOH in 16q and a subset additionally had LOH in 8p.

DISCUSSION

In this study of ER-positive breast cancer, we used the Mapping 10K SNP microarray to assess the similarities and differences of allelic imbalance in ILC and IDC tumors across the genome. This study was unique in that the participants were drawn from a population-based setting, and tumors

were flow cytometrically sorted to provide a genomic profile with little or no contamination from normal cells. To our knowledge, this is the largest set of lobular breast cancers assessed for LOH by genome-wide SNP analysis.

Frequent Sites of LOH Common to Both Subtypes of ER-Positive Tumors

We identified genomic regions with a high frequency of LOH (≥50% FAL) in ILC and IDC tumors and show that there are several regions with an overall high frequency of LOH regardless of subtype (11q, 16q, and 17p) (Table 2). Previous studies have reported allelic imbalance in regions on these chromosomal arms and suggested that they are components of early events in breast cancer formation (Fujii et al., 1996; Ando et al., 2000). Because the majority of these reports focused on IDC tumors, and our study consists of a near equal number of IDC and ILC tumors, our data both complement and add to what was previously reported and suggest that these are events that are common to ER-positive breast cancers regardless of histological subtype.

We identified several sites of high-frequency LOH containing genes (within 100 kb up or downstream of the SNP) that have been implicated in tumorigenesis, including those involved in metastasis, *GPR56* (16q12.2) (Xu et al., 2006); DNA damage/integrity, *CHEK1* (11q24.2) (Petermann et al., 2006); and tumor suppression, *WWOX* (16q23.1) (Ramos and Aldaz, 2006) and *RBl2* (16q12.2) (Jackson and Pereira-Smith, 2006) (Table 2). These findings are preliminary and further studies will be pursued to confirm the role of these genes in breast cancer

It is apparent from our study and many others that 16q loss is a common and likely early event in ER-positive ILC and IDC breast cancer. In a high-resolution assessment of gain and loss on 16q using array CGH virtually all of the 30 ILC tumors tested showed loss of 16q distal to approximately 57.5 Mb (Roylance et al., 2006). Loss of the 16q arm was reported in 100% of 24 tumors containing both ILC and lobular carcinoma in situ (LCIS) (Hwang et al., 2004). Low-grade IDC also showed frequent 16q loss and only subsets of high grade IDC exhibited loss at 16q (Roylance et al., 1999, 2006).

Aside from the morphological differences between ILC and IDC, one of the most established molecular distinctions is the expression of the E-cadherin protein (Berx et al., 1995). In this study, we failed to observe a significant difference in the frequency of LOH for the SNPs neighboring to the *CDH1* gene (16q22.1) between ILC and IDC tumors (94 and 81% upstream and 91 and 90% downstream of the *CDH1* gene for ILC and IDC, respectively). The loss of E-cadherin protein expression in ILC tumors has been linked to LOH, mutation, hypermethylation, and transcriptional regulation (Korkola et al., 2003; Zhao et al., 2004). Thus our data confirm earlier observations that loss of E-cadherin expression in ILC cannot be attributed solely to allelic loss, but must be the outcome of additional events that contribute to the loss of protein expression in ILC specifically.

Differential LOH Frequencies for ER-Positive ILC and IDC Tumors

In an earlier study using array CGH, we found that ER-positive ILC had a slightly higher overall frequency of genomic alterations (copy number gains and losses) than IDC tumors (Loo et al., 2004). In this study, we observe an overall lower frequency of LOH in ILC than IDC tumors. However, a small number of ER positive ILC tumors ($n = 15$) were included in the previous study, which likely accounts for the difference in the results.

Larger sample size in this study also allowed us to stratify the two histological subtypes by tumor size, mitotic rate, or DNA content and compare whole genome FAL quantities of ILC and IDC tumors. We observed a statistically significant trend for IDC tumors to have higher overall FAL with increasing tumor size, mitotic rate, and DNA content, whereas ILC tumors did not (Fig. 1). These results indicate a fundamental difference in the dynamics and accumulation of LOH events between ILC and IDC, especially in large, highly proliferative tumors with a high DNA content.

When ILC and IDC were compared for FAL per chromosomal arm, IDC tumors had significantly more FAL than ILC tumors in 3p, 5q, 8p, 9p, 20p, and 20q (Fig. 1). The most discriminating difference in FAL was on 8p. IDC tumors had significantly higher frequencies of LOH at the SNPs on 8p with and without stratification of tumors. Loss on 8p was identified in earlier studies using CGH and microsatellite markers in IDC tumors and proposed this location to be a potential site of tumor suppressor genes (Anbazhagan et al., 1998). We identified several differential SNPs in 8p neighboring genes (within 100 kb up or downstream) involved in DNA-damage response and integrity including *BRIT1/MCPH1* (8p23.1) involved in regulating the expression of *CHK1* and *BRCA1* and a proposed tumor suppressor in the ATM/ATR path-

way (Xu et al., 2004; Rai et al., 2006), and *WRN* (8p12), a key regulator of double-strand break repair (Chen et al., 2003), and *PINX1* (8p23), a putative telomerase inhibitor shown to have reduced expression in gastric cancer (Banik and Counter, 2004; Kondo et al., 2005). Perturbation of the protein expression, often associated with allelic imbalance, of one or more of these genes in IDC may contribute to the initiation and maintenance of increased numbers of genomic alterations associated with aneuploidy. This might be one explanation for the lower levels of aneuploidy observed in ILC tumors.

Similar to 8p, 5q showed more frequent LOH in IDC than ILC, both with and without stratification into size and mitotic-rate categories. Other studies have also identified genomic alterations on 5q to be associated primarily with ER-negative IDC breast tumors (Loo et al., 2004; Wang et al., 2004). The IDC tumors in Group 1 with LOH at 5q could represent a small subset of ER-positive IDC tumors with genomic profiles that are more similar to ER-negative IDC tumors. These tumors also showed little or no LOH on 8p (Fig. 2C). Two genes involved in DNA damage response and integrity, *XRCC4* (5q14.2) (Lee et al., 2003) and *APC* (5q22.2) (Jaiswal et al., 2005) neighbor individual SNPs with significantly different frequency of LOH by histological subtype and may contribute to genomic instability in this subset of IDC tumors.

The majority of ILC and low-grade IDC tumors (Fig. 2C) appeared to have little or no LOH at the 544 distinguishing SNPs in 5q and 8p and a marked decrease in LOH events on other chromosomes. These data support an earlier observation that ILC and low-grade IDC tumors have fewer genomic alteration events, with the exception of allelic loss at 16q, which occurs at a higher frequency in ILC and low-grade IDC tumors (Buerger et al., 1999). We observed a higher percentage of LOH across 16q in ILC (86%) and low-grade IDC tumors (82%) compared to (77%) for inter/high grade IDC. In addition, similar to what we observed with Group 4 (Fig. 2C), Fridlyand et al. (2006) observed high levels of 16q loss in a subset of ER-positive IDC tumors with low levels of whole genome copy number changes based on hierarchical clustering. They observed this subset to be associated with best patient outcome.

We believe that the reported genomic alteration events in IDC and ILC tumors reflect sites of allelic imbalance that are either shared or differentiate these two ER-positive histological subtypes of breast cancer. Although there are limitations to our study, including the moderate resolution of the array (newer versions of the array provide a more complete coverage with over 900K SNPs) and modest informativity (~34%) of the SNPs on the array, our results help further to define the genotype of lobular breast cancer. These results can be used in future studies to help understand possible molecular events underlying the increased incidence of ILC and the relationship of ILC and use of combined hormone replacement therapy.

ACKNOWLEDGMENTS

We thank Stephanie Stafford, Kelly Wirtala, and Ann Yoder of the Porter Lab and Jeff Delrow and Cassandra Sather of the Genomics Shared Resource, FHCRC for technical assistance.

REFERENCES

Anbazhagan R, Fujii H, Gabrielson E. 1998. Allelic loss of chromosomal arm 8p in breast cancer progression. Am J Pathol 152:815–819.

Andersen J, Poulson H. 1989. Immunohistochemical estrogen receptor determination in paraffin embedded tissue. Cancer 64:1901–1908.

Ando Y, Iwase H, Ichihara S, Toyoshima S, Nakamura T, Yamashita H, Toyama T, Omoto Y, Karamatsu S, Mitsuyama S, Fujii Y, Kobayashi S. 2000. Loss of heterozygosity and microsatellite instability in ductal carcinoma in situ of the breast. Cancer Lett 156:207–214.

Arpino G, Bardou V, Clark G, Elledge R. 2004. Infiltrating lobular carcinoma of the breast: Tumor characteristics and clinical outcome. Breast Cancer Res Epub 6:R149–R156.

Banik S, Counter C. 2004. Characterization of interactions between PinX1 and human telomerase subunits hTERT and hTR. J Biol Chem 279:51745–51748.

Berx G, Cleton-Jansen A, Nollet F, de Leeuw W, van de Vijver M, Cornelisse C, van Roy F. 1995. E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers. EMBO 14:6107–6115.

Buerger H, Otterbach F, Simon R, Schafer K, Poremba C, Diallo R, Brinkschmidt C, Dockhorn-Dworniczak B, Boecker W. 1999. Different genetic pathways in the evolution of invasive breast cancer are associated with distinct morphological subtypes. J Pathol 189:521–526.

Carter S, Negrini M, Baffa R, Gillum D, Rosenberg A, Schwartz G, Croce C. 1994. Loss of heterozygosity at 11q22-q23 in breast cancer. Cancer Res 54:6270–6274.

Chen C, Weiss N, Newcomb P, Barlow W, White E. 2002. Hormone replacement therapy in relation to breast cancer. JAMA 287:734–741.

Chen L, Huang S, Lee L, Davalos A, Schiestl R, Campisi J, Oshima J. 2003. WRN, the protein deficient in Werner syndrome, plays a critical structural role in optimizing DNA repair. Aging Cell 2:191–199.

Cristofanilli M, Gonzalez-Angulo A, Sneige N, Kau S, Broglio K, Theriault R, Valero V, Buzdar A, Kuerer H, Buccholz T, Hortobagyi G. 2005. Invasive lobular carcinoma classic type: Response to primary chemotherapy and survival outcomes. J Clin Oncol 23:41–48.

Deng G, Lu Y, Zlotnikov G, Thor A, Smith H. 1996. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. Science 274:2057–2059.

Devilee P, Cornelisse C. 1994. Somatic genetic changes in human breast cancer. Biochim Biophys Acta 1198:113–130.

Elston C, Ellis I. 1991. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: Experience from a large study with long-term follow-up. Histopathology 19:403–410.

Fridlyand J, Snijders A, Ylstra B, Li H, Olshen A, Segraves R, Dairkee S, Tokuyasu T, Ljung B, Jain A, McLennan J, Ziegler J, Chin K, Devries S, Feiler H, Gray J, Waldman F, Pinkel D, Albertson D. 2006. Breast tumor copy number aberration phenotypes and genomic instability. BMC Cancer 6:96–103.

Fujii H, Szumel R, Marsh C, Zhou W, Gabrielson E. 1996. Genetic progression, histological grade, and allelic loss in ductal carcinoma in situ of the breast. Cancer Res 56:5260–5265.

Giri D, Goepel J, Rogers K. 1988. Immunohistological demonstration of progesterone receptor in breast carcinomas: Correlation with radioligand binding assays and oestrogen receptor immunohistology. J Clin Pathol 41:444–447.

Glogovac J, Porter P, Banker D, Rabinovich P. 1996. Cytokeratin labeling of breast cancer cells extracted from paraffin embedded tissue for bivariate flow cytometric analysis. Cytometry 24:260–267.

Gunther K, Merkelbach-Bruse S, Amo-Takyi BK, Handt S, Schroder W, Tietze L. 2001. Differences in genetic alterations between primary lobular and ductal breast cancers detected by comparative genomic hybridization. J Pathol 193:40–47.

Harris J, Lippman M, Veronesi U, Willit W. 1992. Breast Cancer (1). N Engl J Med 327:319–328.

Harris M, Howell A, Chrissohou M, Swindell RI, Hudson M, Sellwood R. 1984. A comparison of the metastatic pattern of infiltrating lobular carcinoma and infiltrating duct carcinoma of the breast. Br J Cancer 50:23–30.

Hsu S, Raine L, Fanger H. 1981. Use of biotin-avidin-peroxidase conplex (ABC) in immunoperoxidase techniques: A comparison between ABC and unlabeled antibody techniques. Am J Clin Pathol 75:816–821.

Hwang E, Nynte S, Chen Y, Moore D, Devries S, Korkola J, LJ E, Waldman F. 2004. Clonality of lobular carcinoma in situ and synchromus invasive lobular carcinoma. Cancer 100:2562–2572.

Jackson J, Pereira-Smith O. 2006. Primary and compensatory roles for RB family members at cell cycle gene promoters that are deacetylated and downregulated in doxorubicin-induced senescence of breast cancer cells. Mol Cell Biol 26:2501–2510.

Jaiswal A, Balusu R, Narayan S. 2005. Involvement of adenomatous polyposis coli in colorectal tumorigenesis. Front Biosci 10:1118–1134.

Kondo T, Oue N, Mitani Y, Kuniyasu H, Noguchi T, Kuraoka K, Nakayama H, Yasui W. 2005. Loss of heterozygosity and histone hypoacetylation of the PINX1 gene are associated with reduced expression in gastric carcinoma. Oncogene 24:157–164.

Korkola J, DeVries S, Fridlyand J, Hwang E, Estep A, Chen Y, Chew K, Dairkee S, Jensen R, Waldman F. 2003. Differentiation of lobular versus ductal breast carcinomas by expression microarray analysis. Cancer Res 63:7167–7175.

Lee J, Yannone S, Chen D, Povirk L. 2003. Requirement for XRCC4 and DNA ligase IV in alignment-based gap filling for nonhomologous DNA end joining in vitro. Cancer Res 63:22–24.

Li C, Anderson B, Daling J, Moe R. 2003a. Trends in incidence rates of invasive lobular and ductal breast carcinoma. JAMA 289:1421–1424.

Li C, Malone K, Porter P, Weiss N, Tang M, Cushing-Haugen K, Daling J. 2003b. Relationship between long durations and different regimens of hormone therapy and risk of breast cancer. JAMA 289:3254–3263.

Li C, Malone K, Porter P, Lawton T, Voigt L, Cushing-Haugen K, Lin M, Yuan X, Daling J. 2008. Relationship between menopausal hormone therapy and risk of ductal, lobular, and ductal-lobular breast carcinomas. Canc Epidemiol Biomarkers Prev 17:43–50.

Lin M, Wei L, Sellers W, Lieberfarb M, Wong W, Li C. 2004. dChipSNP: Significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics 20:1233–1240.

Loo L, Grove D, Neal C, Williams E, Cousens L, Schubert E, Holcomb I, Massa H, Glogovac J, Li C, Malone K, Daling J, Delrow J, Trask B, Hsu L, Porter P. 2004. Array comparative genomic hybridization analysis of genomic alterations in breast cancer subtypes. Cancer Res 64:8541–8549.

Mersin H, Yildirim E, Gulben K, Berberoglu U. 2003. Is invasive lobular carcinoma different from invasive ductal carcinoma. Eur J Surg Oncol 29:390–395.

O'Connell P, Pekkel V, Fuqua S, Osborne C, Clark G, Allred D. 1998. Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. J Natl Cancer Inst 90:697–703.

Olshen A, Venkatraman E, Lucito R, Wigler M. 2004. Circular binary segmentation for the analysis of array-based DNA copy number data. Biostatistics 5:557–572.

Page D, Anderson T. 1987. Diagnostic Histopathology of the Breast. Edinburgh: Churchill Livingston.

Petermann E, Maya-Mendoza A, Zachos G, Gillespie D, Jackson D, Caldecott K. 2006. Chk1 requirement for high global rates of replication fork progression during normal vertebrate S phase. Mol Cell Biol 26:3319–3326.

Porter P, El-Bastawissi A, Mandelson M, Lin M, Khalid N, Watney E, Cousens L, White D, Taplin S, White J. 1999. Breast tumor characteristics as predictors of mammographic detection: Comparison of Interval-and screen-detected cancers. J Natl Cancer Inst 91:2020–2028.

Rai R, Dai H, Multani A, Li K, Chin K, Gray J, Lahad J, Liang J, Mills G, Meric-Bernstam F, Lin S. 2006. BRIT1 regulates early DNA damage response, chromosomal integrity, and cancer. Cancer Cell 10:145–157.

Ramos D, Aldaz C. 2006. WWOX, a chromosomal fragile site gene and its role in cancer. Adv Exp Med Biol 587:149–159.

Roylance R, Gorman P, Harris W, Leibmann R, Barnes D, Hanby A, Sheer D. 1999. Comparative genomic hybridization of breast tumors stratified by histological grade reveals new insights into the biological progression of breast cancer. Cancer Res 59:1433–1436.

Roylance R, Gorman P, Papior T, Wan Y, Ives M, Watson J, Collins C, Wortham N, Langford C, Fiegler H, Carter N, Gillett C, Sasieni P, Pinder S, Hanby A, Tomlinson I. 2006. A comprehensive study of chromosome 16q in invasive ductal and lobular breast carcinoma using array CGH. Oncogene 25:6544–6553.

Taylor C, Shi S-R, Chaiwun B, Young L, Imam S, Cote R. 1994. Strategies for improving the immunohistochemical staining of various intranuclear prognostic markers in formalin-paraffin sections: Androgen receptor, estrogen receptor, progesterone receptor, p53 protein, proliferating cell nuclear antigen and Ki-67 antigen revealed by antigen retrieval techniques. Hum Pathol 25:263–270.

Thompson E, Herbert S, Forrest S, Campbell I. 2005. Whole genome SNP arrays using DNA derived from formalin-fixed, paraffin-embedded ovarian tumor tissue. Hum Mutat 26:384–389.

Tomlinson I, Nicolai H, Solomon E, Bodmer W. 1996. The frequency and mechanism of loss of heterozygosity on chromosome 11q in breast cancer. J Pathol 180:38–43.

Wang Z, Lin M, Wei L-J, Li CI, Miron A, Lodeiro G, Harris L, Ramaswa y, S, Tanenbaum D, Meyerson M, Iglehart J, Richardson A. 2004. Loss of heterozygosity and its correlation with expression profiles in subclasses of invasive breast cancer. Cancer Res 64:64–71.

Women's Health Initiative Steering Committee. 2004. Effects of conjugated equine estrogen in postmenopausal women with hysterectomy: The Women's Health Initiative randomized controlled trial. JAMA 291:1701–1712.

Xu L, Begum S, Hearn J, Hynes R. 2006. GPR56, an atypical G protein-coupled receptor, binds tissue transglutaminase, TG2, and inhibits melanoma tumor growth and metastasis. Proc Natl Acad Sci USA 103:9023–9028.

Xu X, Lee J, Stern D. 2004. Microcephalin is a DNA damage response protein involved in regulation of CHK1 and BRCA1. J Biol Chem 279:34091–34094.

Zhao H, Langerod A, Ji Y, Nowels K, Nesland J, Tibshirani R, Bukholm I, Karesen R, Botstein D, Borresen-Dale A, Jeffrey S. 2004. Different gene expression patterns in invasive lobular and ductal carcinomas of the breast. Mol Biol Cell 15:2523–2536.