Heritable Variation of *ERBB2* and Breast Cancer Risk

Joan P. Breyer,[1] Melinda E. Sanders,[2] David C. Airey,[3] Qiuyin Cai,[1] Brian L. Yaspan,[4]
Peggy A. Schuyler,[5] Qi Dai,[1] Fouad Boulos,[2,7] Maria G. Olivares,[2] Kevin M. Bradley,[1]
Yu-Tang Gao,[8] David L. Page,[2] William D. Dupont,[5] Wei Zheng,[1] and Jeffrey R. Smith[1,6]

Departments of [1]Medicine, [2]Pathology, [3]Pharmacology, [4]Molecular Physiology and Biophysics, and [5]Biostatistics, Vanderbilt-Ingram Cancer Center,
Vanderbilt University School of Medicine; [6]Medical Research Service, VA Tennessee Valley Healthcare System, Nashville, Tennessee;
[7]Department of Pathology and Laboratory Medicine, American University of Beirut Medical Center, Beirut, Lebanon; and
[8]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China

Abstract

Amplification of the epithelial growth factor receptor gene *ERBB2* (*HER2, NEU*) in breast cancer is associated with a poor clinical prognosis. In mammary gland development, this receptor plays a role in ductal and lobuloalveolar differentiation. We conducted a systematic investigation of the role of genetic variation of the *ERBB2* gene in breast cancer risk in a study of 842 histologically confirmed invasive breast cancer cases and 1,108 controls from the Shanghai Breast Cancer Study. We observed that the *ERBB2* gene resides within a locus of high linkage disequilibrium, composed of three major ancestral haplotypes in the study population. These haplotypes are marked by simple tandem repeat and single nucleotide polymorphisms, including the missense variants I655V and P1170A. We observed a risk-modifying effect of a highly polymorphic simple tandem repeat within an evolutionarily conserved region, 4.4 kb upstream from the *ERBB2* transcription start site. Under a dominant genetic model, the age-adjusted odds ratio was 1.74 (95% confidence interval, 1.27-2.37). Its association with breast cancer, and with breast cancer stratified by histology, by histologic grade, and by stage, remained significant after correction for multiple comparisons. In contrast, we observed no association of *ERBB2* single nucleotide polymorphism haplotypes with breast cancer predisposition. (Cancer Epidemiol Biomarkers Prev 2009;18(4):1252–8)

Introduction

The epithelial growth factor receptor gene *ERBB2* is commonly overexpressed in invasive breast cancer, carrying a poor prognosis (1). The receptor serves as an effective therapeutic target of both monoclonal antibodies and small-molecule tyrosine kinase inhibitors (2, 3). A breadth of literature describes somatic alterations of *ERBB2* in breast cancer. Adenocarcinomas of the breast arise in the terminal duct/lobular unit and are histologically classified by growth pattern and grade. *ERBB2* overexpression is observed predominantly among high-grade tumors. Relative to tumors of ductal histologic type (tumors of no special type), tumors of the most common special type, lobular, rarely overexpress *ERBB2* (4-9).

Systematic investigation of a potential role for inherited germline variants of *ERBB2* in breast cancer risk has been conducted using tagging single nucleotide polymorphisms (SNP) and haplotype-based analysis in two prior study populations. These include one by Benusiglio et al. (10, 11) of breast cancer subjects from the population-based Anglian Breast Cancer Study and one by Han et al. (12) within a hospital-based Korean study population. The former evaluated Caucasian subjects at five regional SNPs that capture much of the SNP information content at the locus among HapMap CEU subjects (13). The latter evaluated Korean subjects at six SNPs that were selected as tagging SNPs specifically within the study population. Neither of these studies observed evidence to support a significant association of common *ERBB2* haplotypes with breast cancer risk. The Cancer Genetic Markers of Susceptibility genome-wide association study of breast cancer evaluated a single SNP within *ERBB2*, also without significant association (14).

In this study, we sought to characterize common genetic variation at *ERBB2* within a well-characterized Han Chinese study population, to assess linkage disequilibrium patterns of both single nucleotide and simple tandem repeat (STR) variants, and to systematically evaluate the potential contribution of *ERBB2* germline variation to breast cancer risk. Given the notable somatic role of *ERBB2* in breast cancer, the hypothesis that *ERBB2* germline variation might influence risk of invasive breast cancer is worthy of careful investigation. We also considered the possibility that *ERBB2* variation might modify the specific histologic forms of breast cancer, given the established role of the gene in the differentiation of normal breast tissue (15-17). We tested this hypothesis within the Shanghai

Breast Cancer Study using single allele- and haplotype-based analyses.

Materials and Methods

Subjects. The Shanghai Breast Cancer Study has been previously described (18, 19). Briefly, study subjects were female, recruited between August 1996 and March 1998. All subjects were permanent residents of urban Shanghai without a prior history of any cancer and were alive at the time of interview. The study included 1,459 incident breast cancer cases diagnosed at an age between 25 and 64 years in Shanghai during the study period (91% of eligible cases). Cancer diagnoses for all patients were reviewed and confirmed by two senior pathologists of the Shanghai Tumor Hospital, Fudan University. Unaffected controls were randomly selected from the general population using the Shanghai Resident Registry, a population registry containing demographic information for all residents of urban Shanghai. Inclusion criteria for controls were identical to those for cases, with the exception of a breast cancer diagnosis. Controls were frequency matched on age (5-year intervals) to the expected age distribution of the case subjects in a 1:1 ratio. The study included 1,556 control subjects (90.3% of matched eligible controls). Blood samples were collected from each of 1,193 (82%) invasive breast cancer cases and 1,310 (84%) controls for DNA extraction (Gentra Systems). For these subjects, the mean age was 47.5 years for cases and 47.1 years for controls. Sixty-eight percent of the breast cancers were premenopausal. All study participants provided written informed consent under an approved institutional review board protocol.

Histologic Review. H&E-stained slides of 842 of 1,151 invasive breast cancer cases with genotype data were available for study. The slides were independently reviewed by Drs. M.E. Sanders, F. Boulos, M.G. Olivares, and D.L. Page, board-certified pathologists with expertise in breast cancer diagnosis, for detailed evaluation of breast cancer histologic characteristics. Slides were randomly assigned to each pathologist. Invasive breast cancers were graded and subtyped based on the Nottingham modification of the Scarff-Bloom-Richardson grading system (low, intermediate, or high-grade) and histologic features highlighting special types (Table 1; refs. 20-22). Special-type tumors are, by definition, >90% pure in pattern and include classic lobular, tubular, cribriform, medullary, and mucinous cancers. In tumors defined as variants of special types, the defining histologic features comprise 70% to 90% of the lesion and include lobular, tubular, tubular-lobular, tubular-cribriform-lobular, and medullary variants (22, 23). Further, categories of no special type (ductal) with lobular, tubular, cribriform, medullary, mucinous, and micropapillary features are designated when the ''special'' pattern represents less than 70% of the total tumor. The category of no special type (ductal) with lobular features was designated if the tumor had at least two of the following features that are typically associated with lobular carcinoma: (*a*) signet ring cells or lobular cytology, (*b*) single-filing or targetoid growth pattern, or (*c*) a dispersed pattern of infiltration that represented less than 70% of the total tumor. Cases approaching the lower threshold and upper limits for these diagnoses were jointly reviewed and a consensus diagnosis was reached using a multiheaded microscope. Tumors were staged using the tumor-node-metastasis system approved by the American Joint Committee on Cancer (6th edition; ref. 24).

Genotyping. To capture the genetic diversity of *ERBB2* for haplotype estimation, we screened SNPs

Table 1. Characteristics of invasive breast cancers, Shanghai Breast Cancer Study, 1996-1998

Histologic type	*n* (%)	Histologic grade			Median age (y)
		3-5 (low)	6-7 (interm)	8-9 (high)	
All ductal	669* (79.5%)	73	233	354	47
NST	527*	43	167	308	47
NST with lobular features	83	15	44	24	46
NST with tubular features	5	5	0	0	50
NST with mucinous features	12	4	6	2	43
NST with medullary features	12	0	1	11	46
NST with micropapillary features	22	4	9	9	47
NST with cribriform features	8	2	6	0	42
All lobular and tubular	97* (11.5%)	63	30	3	46
Pure lobular	22	22	0	0	46
Lobular variant	32*	2	26	3	49
Pure tubular	5	5	0	0	48
Tubular variant	2	2	0	0	49
Tubular-lobular or tubular-cribriform-lobular variant	36	32	4	0	43.5
All other special types	76 (9.0%)	47	12	17	46
Pure mucinous	25	23	2	0	45
Pure medullary	13	0	0	13	50.5
Medullary variant	5	0	3	2	53
Invasive micropapillary	9	2	6	1	39
Metaplastic carcinoma	2	0	1	1	48.5
Invasive cribriform	17	17	0	0	45
Mixed tubular-cribriform	5	5	0	0	47

Abbreviation: NST, no special type.
*Missing grading data due to poor histology.
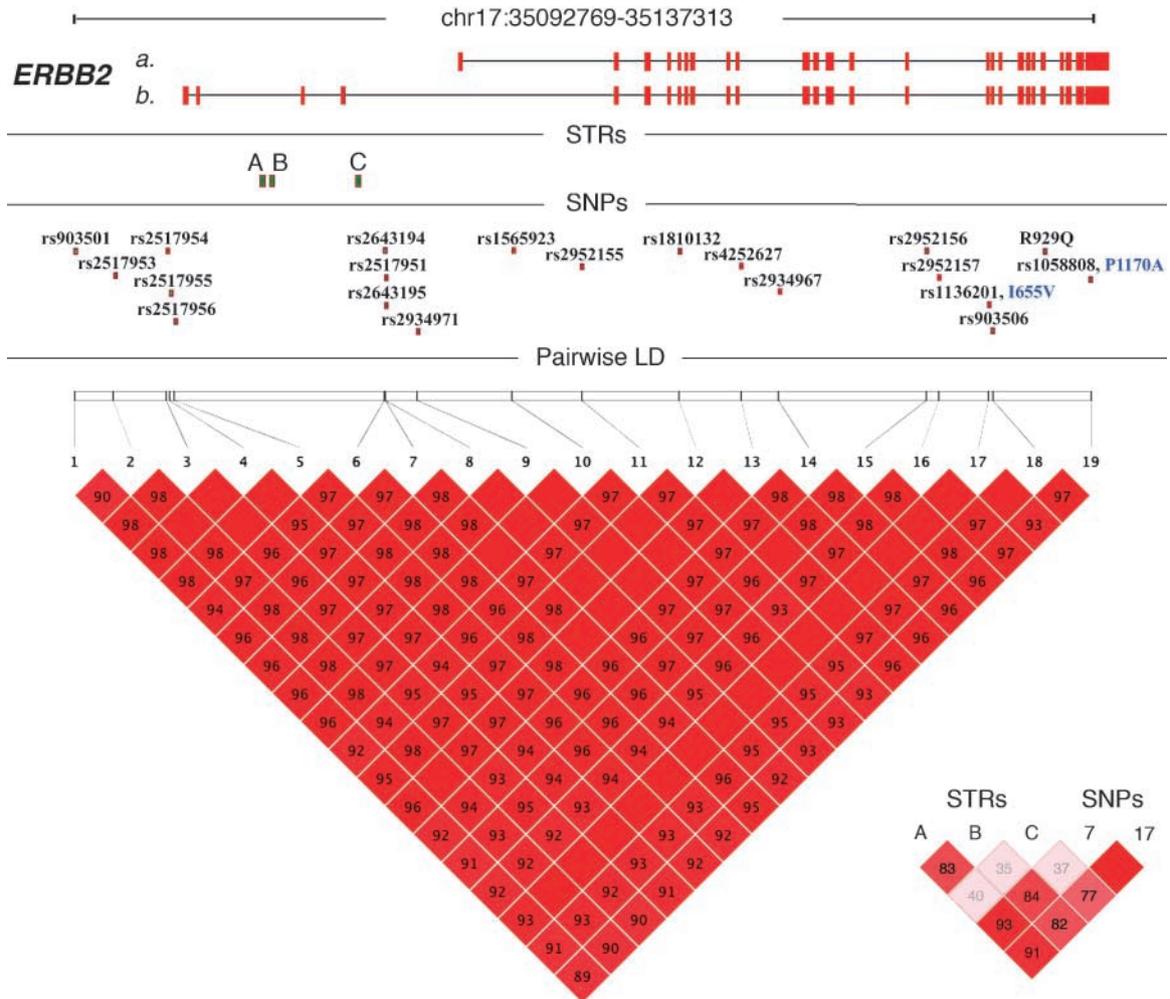
Figure 1. *ERBB2* genetic architecture. A 60-kb interval from chromosome 17q21 is depicted that encompasses the *ERBB2* gene (National Center for Biotechnology Information Build 36.1). The exon-intron structures of transcript *a*, the principal mRNA isoform (NM_004448), and transcript *b*, an alternatively spliced isoform (NM_001005862) are depicted. Positions of three novel polymorphic STRs are shown. Ninety-three database SNPs were screened for polymorphism; those illustrated were commonly polymorphic in the study population. *D′* values are represented by color scale for each pairwise SNP comparison [where the linkage disequilibrium (*LD*) is less than complete; *D′* numerical value is indicated on the matrix]. *Inset,* multiallelic pairwise *D′* among STRs and two tagging SNPs.

within the National Center for Biotechnology Information SNP database for common polymorphism in the study population. Ninety-three annotated SNPs and three STRs across the 28.5-kb *ERBB2* coding region, 17.8 kb of the 5′ flanking sequence, and 23.6 kb of the 3′ flanking sequence (Fig. 1; Supplementary Table S1) were screened in quadruplicate among 22 Han Chinese controls to test for polymorphism. This screening set provided >90% power to detect polymorphism with a minor variant frequency ≥0.05. SNPs rs1801201 (I654V), rs4252633 (W452C), and rs2172826 (P927R) were each predicted to change amino acid residues and were screened further among 356 Han Chinese (half cases, half controls) without an observed polymorphism. An additional SNP (R929Q) was discovered upon sequence confirmation of other polymorphisms but was heterozygous in only one sample.

We genotyped SNPs by single nucleotide primer extension and fluorescence polarization in a 384-well format. Reaction processing entailed three steps: a 4.4-μL PCR reaction, addition of 4 μL of an exonuclease I (New England Biolabs) and calf intestinal alkaline phosphatase (Promega) reagent mix to degrade unincorporated primer and dephosphorylate dNTPs, and a final addition of 4 μL of an Acyclopol and Acycloterminator reagent mix for the primer extension reaction (AcycloPrime FP SNP Detection System, Perkin-Elmer; ref. 25). Each PCR mixture included 0.1 unit AmpliTaq Gold DNA polymerase, 1× Buffer II (Applied Biosystems), 2.5 mmol/L $MgCl_2$, 0.25 mmol/L dNTPs, 335 nmol/L of each primer, and 2 ng DNA template. We detected incorporation of R110- and TAMRA-labeled terminators by fluorescence polarization on a Molecular Devices/LJL Analyst HT. Primers are provided in Supplementary Table S1. The

completion rate for the sought genotypes was 96%. CEPH control individual 1347-02 was included on all genotyping assay plates, with a study-wide 98.7% genotype concordance rate.

STRs were genotyped by 5′-dye–labeled fluorescent amplimers, detected on an ABI PRISM 3700 (Applied Biosystems). Primers (Supplementary Table S1) were designed using a tailing strategy to promote full non-templated nucleotide addition by AmpliTaq Gold DNA polymerase (Applied Biosystems), providing unambiguous detection of alleles separated by 1 bp (26). PCR conditions were as described above. Allele fragment size estimation was accomplished using the internal size standard Genescan 400HD ROX and the local Southern algorithm of GENESCAN software. Editing of alleles was done in GENOTYPER (Applied Biosystems).

Statistical Analysis. Characterization of haplotypes and linkage disequilibrium was conducted among a pilot subset of subjects of the Shanghai Breast Cancer Study that included 178 cases and 178 controls. Pairwise linkage disequilibrium was estimated by D' for SNPs (using Haploview; ref. 27) and by both D' and multi-allelic D' for STRs (using MIDAS; ref. 28). Tagging SNPs were selected based on a pairwise r^2 threshold of 0.8 and a minor allele frequency threshold of 0.05. When multiple SNPs were assigned as tagging SNPs for a particular bin, the SNP with most robust assay perfor-

mance was selected for that bin. Tagging SNPs [rs2517951 and rs1136201 (formerly rs1801200)] and STRs were genotyped among all cases and controls. Each of the SNPs and STRs was in Hardy-Weinberg equilibrium by likelihood ratio test. Subject diplotypes were estimated by the Bayesian method implemented in PHASE version 2.1 (29-31). Diplotypes of highest probability ≥0.9 were used for subsequent tests of association.

The χ^2 test statistic was used to evaluate differences in allele and haplotype frequency of case and control groups. Alleles or haplotypes with an overall frequency of <0.04 were grouped for analysis. Permutation testing was used to assess significance. Nominally significant risk or protective variants identified in contingency tests were subsequently modeled by logistic regression conditioned on age to estimate odds ratios (OR) and 95% confidence intervals (95% CI; Intercooled Stata 10, Stata Corporation). Mitotic rate differences were evaluated using Wilcoxon rank-sum tests. Unless specifically noted, P values are unadjusted for multiple comparisons. All P values are calculated with respect to two-sided alternative hypotheses.

Results

All cases included in the study were confirmed to have invasive breast cancer on histologic review. Seventy-nine

Table 2. Common *ERBB2* alleles, haplotypes, and invasive breast cancer risk, Shanghai Breast Cancer Study, 1996 to 1998

Allele or haplotype					Cases, n (%)	Controls, n (%)	P	
STR A	STR B	STR C	rs2517951	rs1136201				
			T (A1170 tag)		678 (40.5)	963 (39.7)	0.601	
				G (V655)	210 (12.5)	304 (12.7)	0.837	
135					897 (58.3)	1270 (57.9)	0.832	
136					205 (13.3)	314 (14.3)	0.396	Overall χ^2 = 6.8, P = 0.098
138					416 (27.0)	555 (25.3)	0.256	
	346				774 (52.2)	1095 (51.9)	0.826	
	348				91 (6.1)	124 (5.9)	0.746	Overall χ^2 = 2.5, P = 0.661
	350				289 (19.5)	394 (18.7)	0.536	
	354				131 (8.8)	219 (10.4)	0.139	
		113			100 (6.4)	149 (6.7)	0.634	
		115			114 (7.3)	162 (7.3)	0.922	
		117			95 (6.0)	116 (5.3)	0.289	
		119			102 (6.5)	94 (4.3)	0.002	
		121			113 (7.2)	166 (7.5)	0.717	
		123			131 (8.3)	153 (6.9)	0.111	Overall χ^2 = 17.1, P = 0.107
		125			143 (9.1)	244 (11.0)	0.054	
		127			174 (11.1)	249 (11.3)	0.849	
		129			156 (9.9)	219 (9.9)	0.983	
		131			138 (8.8)	194 (8.8)	1.000	
		133			92 (5.9)	138 (6.2)	0.620	
			T	G	205 (12.4)	301 (13.0)	0.723	
			T	A	468 (28.2)	641 (27.1)	0.455	Overall χ^2 = 3.3, P = 0.405
			C	A	982 (59.2)	1420 (60.1)	0.576	
136	354		T	G	90 (6.8)	145 (7.8)	0.333	
138	350		T	A	232 (17.6)	312 (16.7)	0.516	Overall χ^2 = 1.3, P = 0.742
135	346		C	A	670 (50.8)	949 (50.8)	0.994	
136	354	113	T	G	29 (3.0)	61 (5.0)	0.074	
135	346	115	C	A	46 (5.4)	51 (4.2)	0.210	
135	346	125	C	A	47 (5.5)	76 (6.2)	0.491	
135	346	127	C	A	59 (6.9)	84 (6.9)	0.993	Overall χ^2 = 1.1, P = 0.911
135	346	129	C	A	57 (6.6)	79 (6.5)	0.896	
135	346	131	C	A	50 (5.8)	80 (6.6)	0.488	
135	346	133	C	A	42 (4.9)	62 (5.1)	0.848	

Table 3. *ERBB2* STR C and invasive breast cancer risk, Shanghai Breast Cancer Study, 1996 to 1998

Stratification (cases/controls)	119 Allele carrier				125 Allele carrier			
	Frequency		OR* (95% CI)	*P*	Frequency		OR* (95% CI)	*P*
	Cases (%)	Controls (%)			Cases (%)	Controls (%)		
All (785/1108)	13.4	8.3	1.74 (1.27-2.37)	0.001	17.6	20.6	0.78 (0.61-1.00)	0.051
Histology								
Lobular-tubular (93/1108)	10.8	8.3	1.47 (0.73-2.96)	0.278	8.6	20.6	0.34 (0.15-0.74)	0.007
Ductal (619/1108)	13.4	8.3	1.81 (1.30-2.51)	<0.001	18.7	20.6	0.83 (0.64-1.08)	0.174
Combined histologic grade								
3-5 (158/1108)	13.9	8.3	1.90 (1.13-3.18)	0.015	15.3	20.6	0.68 (0.42-1.10)	0.119
8-9 (344/1108)	14.3	8.3	1.89 (1.29-2.77)	0.001	19.5	20.6	0.87 (0.63-1.19)	0.382
Stage								
TNM 0, I, or IIa (422/1108)	11.8	8.3	1.51 (1.04-2.19)	0.031	18.7	20.6	0.88 (0.66-1.18)	0.388
TNM IIb, III, or IV (267/1108)	15.7	8.3	2.09 (1.40-3.12)	<0.001	15.4	20.6	0.69 (0.48-1.00)	0.050

*Dominant genetic model; noncarriers as the reference group.

percent of tumors were no-special-type (ductal) carcinomas (Table 1). Eleven percent of tumors were categorized as tubular or lobular of special or variant types, with good prognoses. Other special types represented 9% of tumors. Using the Nottingham combined grading system (20, 21), the lobular and tubular special-type tumors consisted of 65% low grade, 31% intermediate grade, and 3% high grade. The anticipated converse trend was observed for no-special-type tumors: 11% low-grade, 35% intermediate-grade, and 54% high-grade tumors. Mitotic activity in all lobular, tubular, and their good prognosis variant carcinomas as a group averaged 2 mitoses per 10 high-power fields. In contrast, all no-special-type carcinomas averaged 12 mitoses per 10 high-power fields. The remaining special-type carcinomas as a group averaged 4 mitoses per 10 high-power fields, predominantly as a result of the medullary and medullary variant carcinomas. Relative to the no-special-type (ductal) carcinomas, the mitotic rate of tubular and lobular tumors of special or variant types, and the mitotic rate of the remaining special-type groups, was significantly lower ($P < 0.00005$ for each comparison).

To characterize ancestral versions of the *ERBB2* gene within the study population, we sought to identify haplotype-defining polymorphisms and disequilibrium architecture. We found that 19 of 93 screened database SNPs were commonly polymorphic in the study population (Fig. 1), including two altering evolutionarily conserved amino acids: rs1136201 (I655V, minor allele frequency 0.13) and rs1058808 (P1170A, minor allele frequency 0.40). We observed strong disequilibrium among SNPs spanning the locus, with a minimum observed D' value of 0.89 among controls. Eighteen of the SNPs seemed to mark the same two ancestral DNA fragments in the study population (all pairwise r^2 values ≥0.8). We selected rs2517951 as the tagging SNP for this linkage disequilibrium bin due to its robust assay, effectively also distinguishing rs1058808 (P1170A) alleles among subjects. Only rs1136201 (I655V) provided independent information (maximum pairwise r^2 value 0.19 with other SNPs) and was also selected as a tagging SNP. Together, these tagging SNPs identify three common haplotypes among study subjects, marking versions of the receptor with residue combinations: I655-P1170, I655-A1170, and V655-A1170 (frequencies of 0.60, 0.13, and 0.27, respectively).

We additionally identified three novel polymorphic STRs (denoted A, B, and C in Fig. 1) within the 5′ alternative coding region/promoter. STR A and B alleles had high pairwise D' values with tagging SNPs at *ERBB2* (range 0.86-0.96 and 0.72-0.95, respectively, for alleles of > 0.04 frequency). Pairwise multiallelic D' values between STRs and tagging SNPs are presented to the bottom right of Fig. 1. In contrast to STRs A and B, STR C was considerably more polymorphic (respective heterozygosities: 0.58, 0.68, and 0.93). A total of 28 alleles were observed for STR C in the study population, 11 with a frequency of >0.04. The major alleles of STR C had pairwise D' values ranging from 0.04 to 1.0 with alleles of the tagging SNPs.

Select alleles of each of the STRs marked ancestral SNP haplotypes reasonably well. For example, allele 135 (bp) of STR A had a pairwise r^2 value of 0.85 with tagging SNP rs2517951, and allele 136 had a pairwise r^2 value of 0.74 with tagging SNP rs1136201. Allele 346 of STR B had a pairwise r^2 value of 0.62 with tagging SNP rs2517951, and allele 354 had a pairwise r^2 value of 0.46 with tagging SNP rs1136201. In contrast, all pairwise r^2 values between alleles of STR C and tagging SNPs were <0.06, with the exception of allele 113 and rs1136201 (I655V; $r^2 = 0.30$). STR C is embedded within an evolutionarily conserved sequence, raising the possibility of a direct role in gene function. We used the STRs and tagging SNPs to capture genetic diversity for subsequent tests of association between *ERBB2* and breast cancer.

We conducted tests of association with breast cancer among the histologically confirmed invasive breast cancer cases and controls by contingency tests, presented in Table 2. Among single allele tests of association, an allele of STR C was nominally significant, with an excess of the 119 allele among cases ($P = 0.002$). The association was also evident when comparing the superset of cases, including both those with and without available histologic slides ($n = 1,151$) to controls ($P = 0.0016$; data not shown). The frequency of V655 was 0.125 among histologically confirmed cases, and 0.127 among controls, yielding no evidence for association with invasive breast cancer. The frequency of V655 among all cases (additionally including those without histologic slides) was also not significantly different than among controls ($P = 0.60$; data not shown). We also evaluated possible evidence for the association of *ERBB2* haplotypes with

breast cancer risk; these included haplotypes marked by tagging SNPs, marked additionally by STRs A and B, and finally additionally including the highly polymorphic STR C. No common haplotype yielded significant evidence of association with breast cancer risk (Table 2).

We further estimated the effect size and strength of potential association between STR C alleles 119 and 125 under a dominant logistic regression model adjusting for age (Table 3). Recessive and additive models were not evaluated because only one case was homozygous for allele 119 and only four were homozygous for allele 125. Allele 119 carried an OR of 1.74 (95% CI, 1.27-2.37; $P = 0.001$, Bonferroni corrected for the 47 tests of Tables 2 and 3 to $P = 0.046$), whereas allele 125 carried an OR of 0.78 (95% CI, 0.61-1.00; $P = 0.051$). Upon stratification for the major histologic subtypes, allele 119 carried an OR of 1.81 (95% CI, 1.30-2.51; $P < 0.001$, corrected to $P < 0.046$) for ductal breast cancer, whereas allele 125 carried an OR of 0.34 (95% CI, 0.15-0.74; $P = 0.007$, corrected to $P = 0.28$) for lobular and tubular breast cancer (Table 3). The OR estimate of allele 119 with breast cancer was similar for both low- and high-grade cancer, but greater for advanced-stage cancer (OR, 2.09; 95% CI, 1.40-3.12, $P < 0.001$, corrected to $P < 0.046$). Tests of STR C allele 119 that were stratified by ductal histology, by high combined histologic grade, and by advanced stage each remained significant upon correction for multiple comparisons.

Discussion

ERBB2 plays a key role in the differentiation of mammary tissue (15-17). The ERBB2 receptor hetero-dimerizes with EGFR, a receptor with a prominent role in ductal morphogenesis during adolescence. It also hetero-dimerizes with ERBB3 and ERBB4, receptors with a prominent role in lobuloalveolar morphogenesis. Somatic alteration of *ERBB2* expression is well established in breast cancer, with clinical utility for stratification of breast cancer patients. Overexpression of *ERBB2* is observed in approximately one third of breast cancers and is a prognostic indicator of decreased overall and disease-free survival (1). Lobular carcinomas, however, rarely manifest *ERBB2* amplification and overexpression (4, 5). Better overall and disease-free survival has been suggested to be associated with both lobular and tubular histology (22, 32). Given the established role for *ERBB2* in breast cancer progression and in normal breast differentiation, our investigation tested the hypothesis that heritable variation of *ERBB2* influences risk of breast cancer and of its major histologic types.

As a preface to this goal, we identified common variation of *ERBB2* to enable comprehensive tests of association with breast cancer. We used a tagging SNP approach, specifically selecting tagging SNPs among a large set directly tested for polymorphism in the study population. We screened more than twice as many candidate SNPs for polymorphism at this locus than were included in build 36/phase II of the HapMap. We also developed STR markers within the promoter region to capture additional genetic diversity within the study population. The most notable overall observation is the lack of evidence to support a significant association between *ERBB2* SNPs and breast cancer initiation, despite the wealth of information supporting its role in breast cancer progression. These results are concordant with those of Han et al. (12) and Benusiglio et al. (10).

A relatively large number of prior studies have evaluated the role of the I655V variant in breast cancer risk, with inconsistent results. The first of these studies was an investigation within a relatively small subset of the Shanghai Breast Cancer Study, observing a significant association of the minor valine allele with breast cancer risk (33). This association did not replicate in our investigation of a much larger number of cases. We do not believe that these results are an artifact of genotyping assay interference by an adjacent SNP. Frank et al. (34) observed that a rare A to G variant (rs1801201, I654V) might confound genotyping of rs1136201 (I655V), two variants only three bases apart. Because the rare V654 occurred only when adjacent to V655, an I654/V654 heterozygote might falsely genotype as an I655/I655 homozygote. However, rs1801201 was not polymorphic among our screening set of 372 study subjects; thus, it is not likely to have been a significant source of confounding in our study.

No prior study has included an investigation of STR polymorphism of *ERBB2*. Information content of STRs A and B seemed to be similar (although not identical) to that of the SNP haplotypes. We have previously observed that multiallelic STRs can very efficiently tag SNP haplotypes and may also be functional candidates in disease (35). STR C was among the most polymorphic markers that we have observed. Alleles of this STR were not highly correlated with the tested SNPs by pairwise r^2. We cannot fully exclude a potential role for allele 119 of STR C in breast cancer risk, after correction of significance for multiple comparisons. The association of allele 119 with breast cancer risk was strongest among those with aggressive (high grade or stage) disease, a facet closely intertwined with the histologic types. The potential association of STR C alleles 119 and 125 alternatively with major histologic types of breast cancer (ductal and lobular-tubular, respectively) is intriguing. It is conceivable that the half-helical turn of these *ERBB2* alleles, which separate evolutionarily conserved flanking sequences, modifies risk for specific types of breast cancer. Our observed association of *ERBB2* STR C may be worthy of independent evaluation to distinguish a true association from a multiple-comparisons artifact. Given the concordance of our study and the two prior haplotype-based evaluations of *ERBB2*, it seems unlikely that common population single nucleotide variants at the gene play a significant role as modifiers of breast cancer risk.

In summary, we characterized common genetic variation at *ERBB2* within a well-characterized Han Chinese study population, assessed linkage disequilibrium patterns of both SNPs and STRs, and systematically evaluated the potential contribution of *ERBB2* germline variation to breast cancer risk. We observed no association of common SNP variants or haplotypes of *ERBB2* with breast cancer risk. We did, however, observe a risk-modifying effect of a highly polymorphic STR within an evolutionarily conserved region of the gene.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

References

1. Abd El-Rehim DM, Pinder SE, Paish CE, et al. Expression and co-expression of the members of the epidermal growth factor receptor (EGFR) family in invasive breast carcinoma. Br J Cancer 2004;91: 1532–42.
2. Kirschbaum MH, Yarden Y. The ErbB/HER family of receptor tyrosine kinases: a potential target for chemoprevention of epithelial neoplasms. J Cell Biochem 2000;77:52–60.
3. Ross JS, Fletcher JA, Linette GP, et al. The Her-2/neu gene and protein in breast cancer 2003: biomarker and target of therapy. Oncologist 2003;8:307–25.
4. Bilous M, Ades C, Armes J, et al. Predicting the HER2 status of breast cancer from basic histopathology data: an analysis of 1500 breast cancers as part of the HER2000 International Study. Breast 2003;12: 92–8.
5. Hoff ER, Tubbs RR, Myles JL, Procop GW. HER2/neu amplification in breast cancer: stratification by tumor type and grade. Am J Clin Pathol 2002;117:916–21.
6. Janocko LE, Brown KA, Smith CA, et al. Distinctive patterns of Her-2/neu, c-myc, and cyclin D1 gene amplification by fluorescence *in situ* hybridization in primary human breast cancers. Cytometry 2001;46:136–49.
7. Kommoss F, Colley M, Hart CE, Franklin WA. *In situ* distribution of oncogene products and growth factor receptors in breast carcinoma: c-erbB-2 oncoprotein, EGFr, and PDGFr-β-subunit. Mol Cell Probes 1990;4:11–23.
8. Ramachandra S, Machin L, Ashley S, Monaghan P, Gusterson BA. Immunohistochemical distribution of c-erbB-2 in *in situ* breast carcinoma—a detailed morphological analysis. J Pathol 1990;161:7–14.
9. Sastre-Garau X, Jouve M, Asselain B, et al. Infiltrating lobular carcinoma of the breast. Clinicopathologic analysis of 975 cases with reference to data on conservative therapy and metastatic patterns. Cancer 1996;77:113–20.
10. Benusiglio PR, Lesueur F, Luccarini C, et al. Common ERBB2 polymorphisms and risk of breast cancer in a white British population: a case-control study. Breast Cancer Res 2005;7:R204–9.
11. Benusiglio PR, Pharoah PD, Smith PL, et al. HapMap-based study of the 17q21 ERBB2 amplicon in susceptibility to breast cancer. Br J Cancer 2006;95:1689–95.
12. Han W, Kang D, Lee JE, et al. A haplotype analysis of HER-2 gene polymorphisms: association with breast cancer risk, HER-2 protein expression in the tumor, and disease recurrence in Korea. Clin Cancer Res 2005;11:4775–8.
13. Frazer KA, Ballinger DG, Cox DR, et al. A second generation human haplotype map of over 3.1 million SNPs. Nature 2007;449:851–61.
14. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 2007;39:870–4.
15. Casalini P, Iorio MV, Galmozzi E, Menard S. Role of HER receptors family in development and differentiation. J Cell Physiol 2004;200: 343–50.
16. Stern DF. ErbBs in mammary development. Exp Cell Res 2003;284: 89–98.
17. Troyer KL, Lee DC. Regulation of mouse mammary gland development and tumorigenesis by the ERBB signaling network. J Mammary Gland Biol Neoplasia 2001;6:7–21.
18. Gao YT, Shu XO, Dai Q, et al. Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. Int J Cancer 2000;87:295–300.
19. Zheng W, Gao YT, Shu XO, et al. Population-based case-control study of CYP11A gene polymorphism and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2004;13:709–14.
20. Elston CW, Ellis IO. 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 1991;19: 403–10.
21. Harris JR. Diseases of the breast. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2004.
22. Page DL. Special types of invasive breast cancer, with clinical implications. Am J Surg Pathol 2003;27:832–5.
23. Page DL, Anderson TJ. Diagnostic histopathology of the breast. Edinburgh: Churchill Livingstone; 1987.
24. Singletary SE, Allred C, Ashley P, et al. Revision of the American Joint Committee on Cancer staging system for breast cancer. J Clin Oncol 2002;20:3628–36.
25. Chen X, Levine L, Kwok PY. Fluorescence polarization in homogeneous nucleic acid analysis. Genome Res 1999;9:492–8.
26. Brownstein MJ, Carpten JD, Smith JR. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 1996;20: 1004–6, 1008–10.
27. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of linkage disequilibrium and haplotype maps. Bioinformatics 2005;21:263–5.
28. Gaunt TR, Rodriguez S, Zapata C, Day IN. MIDAS: software for analysis and visualisation of interallelic disequilibrium between multiallelic markers. BMC Bioinformatics 2006;7:227.
29. Marchini J, Cutler D, Patterson N, et al. A comparison of phasing algorithms for trios and unrelated individuals. Am J Hum Genet 2006;78:437–50.
30. Stephens M, Donnelly P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet 2003;73:1162–9.
31. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet 2001;68:978–89.
32. Cristofanilli M, Gonzalez-Angulo A, Sneige N, et al. Invasive lobular carcinoma classic type: response to primary chemotherapy and survival outcomes. J Clin Oncol 2005;23:41–8.
33. Xie D, Shu XO, Deng Z, et al. Population-based, case-control study of HER2 genetic polymorphism and breast cancer risk. J Natl Cancer Inst 2000;92:412–7.
34. Frank B, Hemminki K, Burwinkel B. A bias in genotyping the ERBB2 (HER2) Ile655Val variant. Carcinogenesis 2005;26:1649.
35. Yaspan BL, Breyer JP, Cai Q, et al. Haplotype analysis of CYP11A1 identifies promoter variants associated with breast cancer risk. Cancer Res 2007;67:5673–82.

Heritable Variation of *ERBB2* and Breast Cancer Risk

Joan P. Breyer, Melinda E. Sanders, David C. Airey, et al.

Updated version	Access the most recent version of this article at: http://cebp.aacrjournals.org/content/18/4/1252
Supplementary Material	Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2009/03/31/1055-9965.EPI-08-1202.DC1

Cited articles	This article cites 33 articles, 7 of which you can access for free at: http://cebp.aacrjournals.org/content/18/4/1252.full#ref-list-1
Citing articles	This article has been cited by 1 HighWire-hosted articles. Access the articles at: http://cebp.aacrjournals.org/content/18/4/1252.full#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, use this link http://cebp.aacrjournals.org/content/18/4/1252. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.