Int J Mol Epidemiol Genet 2013;4(4):235-249
www.ijmeg.org /ISSN:1948-1756/IJMEG1310004

Original Article

Epidermal growth factor receptor (*EGFR*) polymorphisms and breast cancer among Hispanic and non-Hispanic white women: the Breast Cancer Health Disparities Study

Avonne E Connor[1], Richard N Baumgartner[1], Kathy B Baumgartner[1], Christina M Pinkston[1], Esther M John[2,3], Gabriela Torres-Mejía[4], Lisa M Hines[5], Anna R Giuliano[6], Roger K Wolff[7], Martha L Slattery[7]

[1]Department of Epidemiology & Population Health, School of Public Health and Information Sciences, University of Louisville, Louisville, KY, USA; [2]Cancer Prevention Institute of California, Fremont, CA; [3]Division of Epidemiology, Department of Health Research and Policy, and Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA; [4]Instituto Nacional de Salud Pública, Centro de Investigación en Salud Poblacional, Cuernavaca, Morelos, México; [5]Department of Biology, University of Colorado Springs, Colorado Springs, CO, USA; [6]H.Lee Moffit Cancer Center & Research Institute, Tampa, FL, USA; [7]Department of Internal Medicine, University of Utah, Salt Lake City, UT, USA

Received October 24, 2013; Accepted November 8, 2013; Epub November 28, 2013; Published December 15, 2013

Abstract: The epidermal growth factor receptor (*EGFR*), a member of the ErbB family of receptor tyrosine kinases, functions in cellular processes essential to the development of cancer. Overexpression of *EGFR* in primary breast tumors has been linked with poor prognosis. We investigated the associations between 34 *EGFR* tagging SNPs and breast cancer risk and breast cancer-specific mortality in 4,703 Hispanic and 3,030 non-Hispanic white women from the Breast Cancer Health Disparities Study. We evaluated associations with risk of breast cancer defined by estrogen/progesterone receptor (ER/PR) tumor phenotype. Only one association remained statistically significant after adjusting for multiple comparisons. Rs2075112$_{GA/AA}$ was associated with reduced risk for ER-/PR+ tumor phenotype (odds ratio (OR), 0.34; 95% confidence interval (CI) 0.18-0.63, p adj=0.01). All additional results were significant prior to adjustment for multiple comparisons. Two of the *EGFR* polymorphisms were associated with breast cancer risk in the overall study population (rs11770531$_{TT}$: OR, 0.56, 95% CI 0.37-0.84; and rs2293348$_{AA}$: OR, 1.20, 95% CI 1.04-1.38) and two polymorphisms were associated with risk among Hispanics: rs6954351$_{AA}$: OR, 2.50, 95% CI 1.32-4.76; and rs845558$_{GA/AA}$: OR, 1.15, 95% CI 1.01-1.30. With regard to breast cancer-specific mortality, we found positive associations with rs6978771$_{TT}$ hazard ratio (HR), 1.68; 95% CI 1.11-2.56; rs9642391$_{CC}$ HR, 1.64; 95% CI 1.04-2.58; rs4947979$_{AG/GG}$ HR, 1.36; 95% CI 1.03-1.79; and rs845552$_{GG}$ HR, 1.62; 95% CI 1.05-2.49. Our findings provide additional insight for the role of *EGFR* in breast cancer development and prognosis. Further research is needed to elucidate *EGFR*'s contribution to ethnic disparities in breast cancer.

Keywords: Breast cancer, Hispanic, epidermal growth factor receptor, polymorphisms, tumor phenotype

Introduction

The epidermal growth factor receptor (*EGFR*) is a member of the ErbB family of receptor tyrosine kinases, and is expressed in epithelial and mesenchymal tissues and tissues of neuronal origin. This gene plays an important role in the processes of the normal cell, which includes differentiation, proliferation, and development [1]. *EGFR* also functions with various cellular processes essential to the development of cancer, including cell division, angiogenesis, migration, and inhibition of apoptosis [2]. There are six known direct binding ligands for *EGFR*, which include EGF, transforming growth factor, amphiregulin, betacellulin, epiregulin, and heparin-binding EGF [3]. Receptor dimerization is initiated by ligand binding, and subsequently activates signaling pathways by the triggering of the cytosolic kinase domain of the receptor

tyrosine kinase, leading to cross-autophosphorylation of the receptors [3]. These *EGFR*-signaling pathways, such as the pathway that leads to suppression of apoptosis through phosphatidylinositol 3-kinase and subsequent Akt activation, have been recognized to be supportive of the development and progression of cancer [3].

The role of *EGFR* in breast cancer etiology is of considerable interest [4]. Overexpression of *EGFR* in primary breast tumors has been linked with poor prognosis [5] and 30-52% of triple negative (estrogen receptor negative (ER-), progesterone receptor negative (PR-), Her2/neu negative) breast cancer overexpress *EGFR* [6]. Mutations in *EGFR* also have been documented in triple negative breast cancer [7]. A recent study conducted by Jacot and colleagues identified the possibility of geographic and ethnic variations in the frequency of these specific *EGFR* mutations [8]. Single nucleotide polymorphisms (SNPs) account for the majority of human genetic variation [9] and some research has shown that *EGFR* SNPs may regulate protein expression [10] and could potentially change gene expression [11, 12]. No epidemiological studies to date have examined the direct relationships between *EGFR* polymorphisms and risk of breast cancer by tumor phenotypes, or considered these associations among women with Hispanic ethnicity, as Hispanic women with breast cancer more frequently have ER- or triple negative tumors than non-Hispanic whites [13, 14]. Furthermore, the functional significance for many of the genetic variants in the *EGFR* gene and the potential for interethnic differences of these SNPs have yet to be completely illuminated [15]. Additionally, there have not been any epidemiological studies that examined the direct relationship between *EGFR* polymorphisms and breast cancer mortality.

Several recent studies have investigated the relationship between *EGFR* polymorphisms and breast cancer risk [2, 9, 12, 16]. One study that examined the effect of rs11568315 found that women with the SS (short/short) genotypes were almost two times more likely (OR, 1.86; 95% CI, 1.02-4.67) to develop breast cancer compared to women with the LL (long/long) genotypes. Additionally, these women were over three times more likely (OR, 3.36; 95% CI, 1.04-10.91) to develop breast cancer before the age of 55 years [2]. Kallel et al. found no association with rs11543848; however, the homozygous GG genotype was more prevalent among breast cancer cases with lymph node metastasis and high grade tumors [16]. Choura et al. did not find significant associations between rs17337451 or rs17290699 with breast cancer risk; however, their results demonstrated that the T allele of rs1140476 was associated with increased breast cancer risk [9].

We hypothesized that the *EGFR* gene would be associated with risk of breast cancer and breast cancer-specific mortality in our sample of Hispanic and non-Hispanic white (NHW) women from the Breast Cancer Health Disparities Study. We also evaluated the association between *EGFR* polymorphisms and breast cancer risk according to ER/PR tumor phenotypes and investigated effect modification by self-reported ethnicity, percentage of Native American ancestry, and menopausal status.

Materials and methods

The Breast Cancer Health Disparities Study (BCHDS) includes participants from three population-based case-control studies: the 4-Corner's Breast Cancer Study (4-CBCS), the San Francisco Bay Area Breast Cancer Study (SFBCS), and the Mexico Breast Cancer Study (MBCS) [17]. All participants signed informed written consent prior to participation, completed an interview, and had a blood or mouth sample available for DNA extraction; the study was approved by the Institutional Review Board for Human Subjects at each institution.

The 4-CBCS participants were Hispanic and NHW women between 25 and 79 years of age with a histological confirmed diagnosis of *in situ* (n=341) or invasive (n=1492) cancer between October 1999 and May 2004; controls were selected from the target populations of cases living in Arizona, Colorado, New Mexico, and Utah and were frequency matched to cases on ethnicity and 5-year age distribution [18]. Participants from the MBCS were Hispanic and between 28 and 74 years of age, living in one of three states, Monterrey, Veracruz and Mexico City, for the past five years. Eligible cases in Mexico were women diagnosed with either a new histologically confirmed *in situ* or invasive breast cancer between January 2004 and December 2007 at 12 participating hospi-

Table 1. Description of *EGFR* polymorphisms

EGFR SNPs	Coordinate	Region	Major/Minor Allele[1]	non-Hispanic Whites			Hispanics			Proportion Missing
				Major allele freq.	Minor allele freq.	FDR adjusted HWE *p* value	Major allele freq.	Minor allele freq.	FDR adjusted HWE *p* value	
rs6978771	55140296	INTRON	C/T	0.75	0.25	0.62	0.61	0.39	0.95	0.00023753
rs11487218	55141540	INTRON	T/C	0.66	0.35	0.93	0.80	0.20	0.08	0.000475059
rs10225877	55150822	INTRON	A/T	0.81	0.19	0.96	0.86	0.14	0.30	0.00023753
rs12718945	55192963	INTRON	G/T	0.55	0.45	0.98	0.67	0.33	0.05	0.000950119
rs2075112	55219611	INTRON	G/A	0.58	0.42	0.93	0.59	0.41	0.62	0
rs6944906	55251953	INTRON	A/G	0.59	0.41	0.96	0.56	0.44	0.98	0.00023753
rs17586365	55140786	INTRON	G/A	0.86	0.14	0.96	0.90	0.10	0.93	0
rs1344307	55137888	INTRON	A/G	0.80	0.20	0.96	0.89	0.11	0.59	0
rs9642391	55245364	INTRON	G/C	0.72	0.28	0.84	0.65	0.35	0.83	0
rs4947971	55160995	INTRON	C/T	0.69	0.31	1.00	0.76	0.24	0.70	0
rs12671550	55173675	INTRON	C/G	0.69	0.31	0.93	0.60	0.40	0.20	0.00023753
rs1558544	55228053	INTRON	T/A	0.72	0.28	0.96	0.85	0.15	0.72	0.00023753
rs1140475	55266417	CODING	C/T	0.88	0.12	0.96	0.89	0.11	0.86	0
rs6593205	55168692	INTRON	G/A	0.60	0.40	0.98	0.63	0.37	0.61	0.00023753
rs17151957	55200512	INTRON	G/A	0.77	0.23	0.89	0.68	0.32	0.48	0.00023753
rs6970262	55259763	INTRON	G/A	0.62	0.39	0.92	0.72	0.28	0.56	0
rs884419	55276280	INTERGENIC	C/T	0.90	0.10	0.96	0.86	0.14	0.50	0.00023753
rs763317	55095197	INTRON	G/A	0.52	0.48	0.96	0.72	0.28	0.35	0
rs723527	55134872	INTRON	G/A	0.59	0.41	0.86	0.67	0.33	0.10	0.004038005
rs12535536	55154381	INTRON	A/G	0.70	0.30	0.96	0.85	0.15	0.91	0.00023753
rs917880	55162011	INTRON	C/T	0.53	0.47	0.96	0.58	0.42	0.88	0.00023753
rs11977660	55162336	INTRON	T/C	0.53	0.47	0.98	0.53	0.47	0.29	0
rs6954351	55171190	INTRON	G/A	0.86	0.14	0.86	0.92	0.08	0.67	0.000712589
rs2330951	55174342	INTRON	A/C	0.75	0.25	0.66	0.75	0.26	0.23	0
rs759160	55181442	INTRON	A/G	0.76	0.24	0.93	0.83	0.17	0.41	0.00023753
rs4947979	55195625	INTRON	A/G	0.79	0.21	0.86	0.85	0.15	0.95	0.00023753
rs11770531	55220905	INTRON	C/T	0.86	0.14	0.52	0.91	0.09	0.30	0
rs3752651	55229543	INTRON	T/C	0.80	0.20	0.96	0.85	0.15	0.44	0.00023753
rs9692301	55243754	INTRON	A/G	0.69	0.31	0.98	0.64	0.36	0.60	0.000475059
rs845552	55245507	INTRON	A/G	0.52	0.48	0.96	0.54	0.46	0.52	0.059144893
rs845558	55247588	INTRON	G/A	0.58	0.42	0.94	0.57	0.43	0.61	0.012589074
rs2472520	55265940	INTRON	C/G	0.57	0.44	1.00	0.54	0.46	0.66	0
rs2293348	55266757	INTRON	G/A	0.69	0.31	0.93	0.63	0.37	0.19	0.003562945
rs2280653	55276094	INTERGENIC	T/C	0.84	0.16	0.96	0.81	0.19	0.68	0.00023753

[1]Major/minor allele reported for NHW population; minor allele frequency and Hardy-Weinberg Equilibrium (HWE) based on control population.

tals from three main health care systems; controls were randomly selected from the catchment area of the 12 participating hospitals using a probabilistic multi-stage design [19]. The SFBCS included Hispanic and NHW women aged 35 to 79 years from the San Francisco Bay Area diagnosed with a first primary histologically confirmed invasive breast cancer between April 1995 and April 2002; controls were identified by random-digit dialing and frequency-matched to cases based on the expected race/ethnicity and 5-year age distribution [20, 21].

Data harmonization

Interview data were harmonized across the three studies [17]. The present analyses considered adjusting for body mass index (BMI, kg/m^2) calculated as self-reported weight during the referent year (or more distantly recalled weight if referent year weight was not available or measured weight if neither were available) divided by measured height squared, parity (number of live births and stillborn pregnancies), self-reported ethnicity in the U.S. studies (all women in Mexico were considered Hispanic), and highest level of education. The referent year was defined as the calendar year prior to diagnosis for cases or selection into the study for controls.

Genetic data

DNA was extracted from either whole blood (n=7286) or mouthwash (n=637) samples. Whole Genome Amplification (WGA) was applied to the mouthwash-derived samples prior to genotyping. A tagSNP approach was used to characterize variation across candidate genes. TagSNPs were selected as follows: linkage disequilibrium (LD) blocks were defined using a Caucasian LD map and an r^2=0.8; minor allele frequency (MAF) >0.1; range=-1500 bps from the initiation codon to +1500 bps from the termination codon; and 1 SNP/LD bin. A total of 104 Ancestral Informative Markers (AIMs) was used to distinguish European and Native American ancestry in the study population [17]. All markers were genotyped using a multiplexed bead array assay format based on GoldenGate chemistry (Illumina, San Diego, California). A genotyping call rate of 99.93% was attained (99.65% for WGA samples). We included 132 internal replicates that were blinded representing 1.6% of the sample set. The duplicate concordance rate was 99.996% as determined by 193,297 matching genotypes among sample pairs [17].

In the current analysis, we examined 34 *EGFR* polymorphisms: rs6978771, rs72352, rs1148-7218, rs12535536, rs10225877, rs917880, rs12718945, rs11977660, rs2075112, rs695-4351, rs6944906, rs2330951, rs17586365, rs759160, rs1344307, rs4947979, rs96423-91, rs11770531, rs4947971, rs3752651, rs1-2671550, rs9692301, rs1558544, rs845552, rs1140475, rs845558, rs6593205, rs24725-20, rs17151957, rs2293348, rs6970262, rs2280653, rs884419, and rs763317. **Table 1** describes the *EGFR* polymorphisms in detail, including the minor allele frequencies (MAF) and adjusted Hardy-Weinberg equilibrium (HWE) *p* values. Online supplement 1 describes the LD between all 34 *EGFR* polymorphisms by self-reported ethnicity.

Tumor characteristics and survival data

Information on ER and PR status was obtained from the cancer registries in New Mexico, Utah, Colorado, Arizona, and California for 979 (68%) NHW cases and 958 (75%) Hispanic cases. These data were not available for the MBCS.

Survival status was available for the New Mexico, Utah, Colorado, Arizona, and California study centers. Each center's respective cancer registry provided information on date of death or last follow-up (month and year). Survival (in months) was calculated as the difference between diagnosis date and date of death or last follow-up. The cause of death was classified as breast cancer if either the primary or contributing cause of death noted on the death certificate was breast cancer. Survival data were not available for the MBCS.

Statistical methods

STRUCTURE was used to compute individual ancestry assuming two founding populations [22, 23] and each study participant was classified by level of percent Native American ancestry. The following strata for percentage of genetic ancestry were created using cut-points based on the distribution of genetic ancestry in the control population: 0-28%, 29-70%, and 71-100%. The groups were categorized in this manner to ensure sufficient power to assess

Table 2. Characteristics of study population, stratified by ethnicity and case-control status, the Breast Cancer Health Disparities Study *(n=7,733)*

	Non-Hispanic Whites (n=3,030)				Hispanics (n=4,703)				
	Cases		Controls		Cases		Controls		
	No.	%	No.	%	No.	%	No.	%	*p* value[a]
Total Subjects	*1431*		*1599*		*2093*		*2610*		
Study Site									
4-CBCS	1177	82.3	1335	83.5	579	27.7	736	28.2	<0.001
MBCS	-		-		816	39.0	994	38.1	
SFBCS	254	17.8	264	16.5	698	33.4	880	33.7	
Age (years)									
<40	87	6.1	117	7.3	198	9.5	313	12.0	<0.001
40-49	401	28.0	409	25.6	708	33.8	834	32.0	
50-59	403	28.2	410	25.6	614	29.3	758	29.0	
60-69	340	23.8	356	22.3	425	20.3	530	20.3	
70+	200	14.0	307	19.2	148	7.1	175	6.7	
Percentage of Native American Ancestry									
≤0.28	1420	99.2	1591	99.5	276	13.2	280	10.7	<0.001
0.28-0.70	7	0.5	7	0.4	1373	65.6	1697	65.0	
>0.70	4	0.3	1	0.1	444	21.2	633	24.3	
Menopausal Status									
Premenopausal	475	34.1	494	31.5	831	41.2	1027	40.7	<0.001
Postmenopausal	919	66.9	1075	68.5	1186	58.8	1499	59.3	
Estrogen/Progesterone Receptor Status									
ER+/PR+	674	68.5	---		594	62.1	---		0.002
ER+/PR-	116	11.8	---		114	11.9	---		
ER-/PR+	15	1.5	---		28	2.8	---		
ER-/PR-	179	18.3	---		223	23.2	---		

Missing information: Menopausal status n=227; eligible cases missing estrogen receptor status n=765. [a]Ethnic group comparison, regardless of case-control status. *p* values from chi-square tests.

associations. When used as an adjusting variable to assess confounding, genetic ancestry was modeled as a continuous variable. Descriptive statistics were calculated for all covariates and t-tests and chi-square tests were used to assess differences between groups. The homozygous wildtypes for each polymorphism were used as the referent categories. Using co-dominant models, genotype associations for all *EGFR* SNPs were estimated as odds ratios (ORs) with 95% confidence intervals (CIs) by unconditional logistic regression with adjustments for age and study center. Based on initial assessment of the co-dominant associations, dominant and recessive models were also examined. Potential confounders included BMI, menopausal status, menopausal hormone therapy (HT) use, physical activity, caloric intake per day, and smoking status (ever or never). These covariates were included in multivariable models if their univariate *P* values were ≤0.20 and if they changed the point estimate for the main effects of the *EGFR* genotypes by ≥10% for SNPs that were found to be statistically significant prior to multiple comparisons [24]. However, there was no evidence of confounding and the models were adjusted for age, study site, and percentage of Native American ancestry. Interactions between *EGFR* variants with ethnicity, genetic ancestry, and menopausal status were assessed using the likelihood-ratio test comparing the model including an interaction term with a reduced model without the term. For survival analyses, hazard ratios (HR) and 95% CIs were derived using multivariable Cox proportional hazard models and were adjusted for SEER disease stage at diagnosis, age, genetic ancestry, and study center. Stratified analyses were also conducted for survival analyses to determine if there was evidence of effect modification by genetic ancestry or by ethnicity.

Women were classified as either premenopausal or postmenopausal based on self-reported responses to questions on menstrual history. Women who reported menstruation during the referent year were classified as premenopausal. The classification for postmenopausal women was established by using criteria provided by each individual study. If women were taking (HT) and still having periods and were at or above the 95th percentile of age for ethnicity of those who reported having a natural menopause among their study site, they were classified as postmenopausal. This age was 58 for NHW and 56 for Hispanics in the 4-CBCS, age 54 in the MBCS, and 55 for NHW and 56 for Hispanics in the SFBCS.

Multinomial logistic regression models were constructed to evaluate the associations between *EGFR* genotypes and breast cancer risk by ER/PR status [25, 26]. Results were adjusted for multiple comparisons taking into account tagSNPs within the gene using the step-down Bonferroni correction (i.e., Holm's method) based on the effective number of independent SNPs as determined using the SNP spectral decomposition method proposed by Nyholt [27] and modified by Li and Ji [28]. The interaction *p* values, based on 1-df likelihood-ratio tests, were adjusted using the step-down Bonferroni correction or the Holm's test [29]. We considered an adjusted *p* value of 0.10 or less as potentially important for main effects and a Holm's *p* value of 0.15 or less for interactions. All data analyses were performed using SAS version 9.3 (SAS Institute, Cary NC).

Results

The distributions of the demographic and major risk factors for breast cancer in the Breast Cancer Health Disparities Study have been previously reported [17, 30]. A total of 7,733 breast cancer cases and controls were included in analyses that evaluated breast cancer risk; 1,943 cases were available for analyses by ER/PR tumor status. **Table 2** describes the distribution of selected variables of importance to the present analysis.

Table 3 describes the significant associations (p<0.05), prior to adjustment for multiple comparisons, between the *EGFR* polymorphisms and breast cancer overall and by ethnicity. Two of the polymorphisms were associated with risk (rs11770531$_{TT}$: OR, 0.56; 95% CI 0.37-0.84, Wald p=0.01; and rs2293348$_{AA}$: OR, 1.20; 95% CI 1.04-1.38, p trend=0.12). Among Hispanic women, the AG/GG genotypes versus the AA genotype of rs6944906 were associated with decreased risk (OR, 0.87; 95% CI 0.77-0.99) and positive associations were found for rs6954351 AA genotype (OR, 2.49; 95% CI 1.31-4.72, Wald p=0.01), the rs845558 GA/AA genotypes (OR, 1.15; 95% CI 1.01-1.30, Wald p=0.03), the CC genotype of rs3752651 (OR, 1.51; 95% CI 1.03-2.20, Wald p=0.03), and the AA genotype of rs2293348 (OR, 1.24; 95% CI 1.03-1.48, p trend=0.19). Among NHW women, there was an inverse association with the TT genotype of rs11770531 (OR, 0.53; 95% CI 0.31-0.90, Wald p=0.02). Overall, none of the associations between the *EGFR* SNPs and breast cancer risk remained statistically significant after adjustment for multiple comparisons (**Table 3**). In analyses stratified by percentage of Native American ancestry (data not shown), decreased breast cancer risk was associated with the TT genotype vs. the CC/TT genotypes of rs11770531 among women with 0-28% Native American ancestry (OR, 0.54; 95% CI 0.32-0.89 Wald p=0.02) and with the AG/GG genotypes of rs6944906 among women with 71-100% ancestry (OR, 0.74; 95% CI 0.57-0.97, Wald p=0.03), after adjusting for age and study site. These results were no longer significant after adjustment for multiple comparisons.

Table 4 shows associations of *EGFR* polymorphisms associated with breast cancer risk (p <0.05), prior to adjustment for multiple comparisons, by ER/PR tumor phenotype. We found inverse associations for the GG genotype of rs12671550 and ER+/PR- tumors (OR, 0.55; 95% CI 0.32-0.94, Wald p=0.10) and for the TT genotype versus the CC/TT genotypes of rs11770531 and ER+/PR+ tumors (OR, 0.44; 95% CI 0.23-0.83; Wald p=0.01). The following SNPs were associated with ER-/PR+ tumors: rs2075112$_{GA/AA}$ (OR, 0.34; 95% CI 0.18-0.63, Wald p<0.001, p adj=0.01); rs12671550$_{CG}$ (OR, 0.42; 95% CI 0.21-0.84, Wald p=0.18); and rs2472520$_{GG}$ (OR, 2.55; 95% CI 1.01-6.43, Wald p=0.05). The GT/TT genotypes of rs12718945 were associated with increased risk for ER-/PR- tumors (OR, 1.28; 95% CI 1.02-1.61 Wald p=0.04). After adjusting for multiple comparisons, only the association with rs2075112 remained statistically significant.

Table 3. Associations with *EGFR* polymorphisms and breast cancer risk overall and interaction by ethnicity, the Breast Cancer Health Disparities Study

		Cases		Controls		All Women Combined		Non-Hispanic Whites		Hispanics			
EGFR SNPs	Genotype	N	%	N	%	OR	95% CI	OR	95% CI	OR	95% CI	p-int	p_{adj}
rs6944906													
	AA	1211	(46.83)	1375	(53.17)	1.00		1.00		1.00		0.10	1.00
	AG/GG	2311	(44.93)	2833	(55.07)	0.93	(0.85-1.03)	1.03	(0.88-1.19)	0.87	(0.77-0.99)		
P-value: Wald; p adj						0.15;	1.00	0.72;	1.00	0.03;	0.61		
rs6954351													
	GG/GA	3460	(45.44)	4154	(54.56)	1.00		1.00		1.00		0.01	0.17
	AA	59	(53.15)	52	(46.85)	1.30	(0.89-1.90)	0.86	(0.52-1.40)	2.49	(1.31-4.72)		
P-value: Wald; p adj						0.17;	1.00	0.53;	1.00	0.01;	0.12		
rs11770531													
	CC/CT	3489	(45.73)	4140	(54.27)	1.00		1.00		1.00		0.70	1.00
	TT	34	(33.01)	69	(66.99)	0.56	(0.37-0.84)	0.53	(0.31-0.90)	0.61	(0.32-1.17)		
P-value: Wald; p adj						0.01;	0.13	0.02;	0.45	0.13;	1.00		
rs3752651													
	TT/TC	3406	(45.44)	4089	(54.56)	1.00		1.00		1.00		0.05	0.79
	CC	117	(49.58)	119	(50.42)	1.15	(0.88-1.49)	0.90	(0.63-1.29)	1.51	(1.03-2.20)		
P-value: Wald; p adj						0.30;	1.00	0.57;	1.00	0.03;	0.65		
rs845558													
	GG	1092	(44.14)	1382	(55.86)	1.00		1.00		1.00		0.21	1.00
	GA/AA	2386	(46.24)	2774	(53.76)	1.09	(0.99-1.20)	1.02	(0.87-1.18)	1.15	(1.01-1.30)		
P-value: Wald; p adj						0.07;	1.00	0.84;	1.00	0.03;	0.61		
rs2293348													
	GG	1494	(45.73)	1773	(54.27)	1.00		1.00		1.00		0.17	1.00
	GA	1535	(44.38)	1924	(55.62)	0.96	(0.87-1.06)	1.06	(0.91-1.23)	0.91	(0.80-1.03)		
	AA	482	(49.23)	497	(50.77)	1.20	(1.04-1.38)	1.10	(0.86-1.41)	1.24	(1.03-1.48)		
P-value: trend; p adj						0.12;	1.00	0.36;	1.00	0.19;	1.00		

Models adjusted for age, study site, and genetic ancestry.

Table 4. Associations with *EGFR* polymorphisms and risk of breast cancer by tumor phenotype, the Breast Cancer Health Disparities Study

	ER+/PR+			ER+/PR-			ER-/PR+			ER-/PR-			
EGFR SNPs	N	OR	(95% CI)	N	OR	(95% CI)	N	OR	(95% CI)	N	OR	(95% CI)	*P*-value (P adj)
rs12718945													
GG	449	1.00		71	1.00		16	1.00		121	1.00		0.15 (1.00)
GT/TT	818	0.97	(0.84, 1.11)	158	1.21	(0.90, 1.63)	27	0.97	(0.51, 1.82)	281	1.28	(1.02, 1.61)	
Wald P; p adj		0.62;	1.00		0.20;	1.00		0.92;	1.00		0.04;	0.75	
rs2075112													
GG	423	1.00		77	1.00		26	1.00		144	1.00		0.01 (0.24)
GA/AA	845	1.05	(0.91, 1.20)	153	1.05	(0.79, 1.39)	17	0.34	(0.18, 0.63)	258	0.93	(0.75, 1.16)	
P-trend; p adj		0.50;	1.00		0.76;	1.00		0.00;	0.01		0.83;	1.00	
rs12671550													
CC	564	1.00		105	1.00		26	1.00		178	1.00		0.18 (1.00)
CG	554	0.95	(0.83, 1.10)	109	1.00	(0.75, 1.32)	12	0.42	(0.21, 0.84)	181	0.96	(0.77, 1.20)	
GG	150	0.99	(0.80, 1.23)	16	0.55	(0.32, 0.94)	5	0.64	(0.24, 1.69)	43	0.86	(0.61, 1.23)	
P-trend; p adj		0.72;	1.00		0.10;	1.00		0.06;	1.00		0.44;	1.00	
rs763317													
GG	460	1.00		80	1.00		15	1.00		169	1.00		0.30 (1.00)
GA/AA	808	0.94	(0.82, 1.08)	150	1.03	(0.78, 1.38)	28	1.11	(0.58, 2.11)	233	0.76	(0.61, 0.95)	
Wald P; p adj		0.36;	1.00		0.82;	1.00		0.76;	1.00		0.01;	0.32	
rs6954351													
GG	953	1.00		173	1.00		38	1.00		308	1.00		0.09 (1.00)
GA	280	1.12	(0.95, 1.31)	55	1.22	(0.89, 1.68)	5	0.51	(0.20, 1.31)	90	1.12	(0.87, 1.45)	
AA	32	1.67	(1.06, 2.63)	1	0.29	(0.04, 2.15)	0	0.00	(0.00, 0.00)	4	0.68	(0.24, 1.91)	
P-trend; p adj		0.02;	0.52		0.63;	1.00		0.11;	1.00		0.66;	1.00	
rs11770531													
CC/CT	1256	1.00		228	1.00		41	1.00		400	1.00		0.01 (0.27)
TT	11	0.44	(0.23, 0.83)	2	0.45	(0.11, 1.84)	2	2.91	(0.68, 12.47)	2	0.27	(0.06, 1.10)	
Wald P; p adj		0.01;	0.26		0.26;	1.00		0.15;	1.00		0.07;	1.00	
rs2472520													
CC	386	1.00		77	1.00		7	1.00		114	1.00		0.32 (1.00)
CG	620	0.99	(0.86, 1.16)	104	0.84	(0.62, 1.14)	23	1.98	(0.84, 4.63)	199	1.07	(0.84, 1.36)	
GG	262	1.01	(0.84, 1.22)	49	0.95	(0.65, 1.38)	13	2.55	(1.01, 6.43)	89	1.12	(0.83, 1.50)	
P-trend; p adj		0.93;	1.00		0.65;	1.00		0.05;	0.99		0.46;	1.00	

Models adjusted for age, study site, and genetic ancestry.

Table 5. Associations between *EGFR* polymorphisms and breast cancer-specific mortality and interaction by self-reported ethnicity, the Breast Cancer Health Disparities Study

						Ethnicity									
	All women combined					Non-Hispanic Whites					Hispanics				
EGFR	Death	Person Years	HR	95% CI		Death	Person Years	HR	95% CI		Death	Person Years	HR	95% CI		P (P adj)
rs6978771																
CC	93	9976	1.00			47	5539	1.00			46	4438	1.00			0.64 (1.00)
CT	93	8516	1.11	0.83	1.49	45	3861	1.26	0.83	1.91	48	4656	0.99	0.66	1.50	
TT	31	1960	1.68	1.11	2.56	10	705	1.94	0.97	3.88	21	1254	1.50	0.88	2.54	
p-trend; p adj			0.03;	0.55				0.06;	1.00				0.24;	1.00		
rs6944906																
AA	67	6890	1.00			39	3390	1.00			28	3500	1.00			0.02 (0.50)
AG/GG	150	13554	1.15	0.86	1.54	63	6715	0.83	0.55	1.25	87	6839	1.62	1.05	2.48	
Wald p; p adj			0.34;	1.00				0.37;	1.00				0.03;	0.55		
rs1344307																
AA	167	14545	1.00			73	6643	1.00			94	7902	1.00			0.75 (1.00)
AG	48	5253	0.78	0.56	1.08	28	3046	0.90	0.58	1.41	20	2207	0.72	0.43	1.18	
GG	2	668	0.24	0.06	0.97	1	416	0.22	0.03	1.60	1	252	0.29	0.04	2.07	
p-trend; p adj			0.01;	0.31				0.18;	1.00				0.08;	1.00		
rs9642391																
GG	90	9674	1.00			38	5015	1.00			52	4659	1.00			0.28 (1.00)
GC	102	8983	1.25	0.94	1.66	54	4354	1.53	1.00	2.34	48	4629	1.01	0.68	1.51	
CC	25	1808	1.64	1.04	2.58	10	735	1.93	0.93	4.01	15	1073	1.37	0.77	2.47	
p-trend; p adj			0.02;	0.45				0.02;	0.46				0.41;	1.00		
rs4947971																0.02 (0.48)
CC	125	10886	1.00			64	4986	1.00			61	5900	1.00			
CT/TT	92	9580	0.81	0.62	1.07	38	5119	0.59	0.40	0.89	54	4461	1.13	0.78	1.65	
Wald p; p adj			0.14;	1.00				0.01;	0.27				0.52;	1.00		
rs6593205																
GG	85	8122	1.00			42	3729	1.00			43	4394	1.00			0.06 (1.00)
GA	94	9519	0.94	0.70	1.27	45	4712	0.85	0.55	1.30	49	4807	1.01	0.66	1.53	
AA	38	2815	1.37	0.93	2.01	15	1664	0.77	0.42	1.40	23	1152	2.08	1.24	3.48	
p-trend; p adj			0.25;	1.00				0.33;	1.00				0.03;	0.55		
rs17151957																

GG	132	11014	1.00			65	5877	1.00			67	5137	1.00			0.30 (1.00)
GA/AA	85	9451	0.69	0.52	0.91	37	4228	0.79	0.53	1.20	48	5224	0.61	0.42	0.88	
Wald p; p adj			0.01;	0.18				0.27;	1.00				0.01	0.20		
rs917880																
CC	69	6279	1.00			20	2754	1.00			49	3525	1.00			0.01 (0.13)
CT/TT	148	14187	0.90	0.67	1.20	82	7351	1.48	0.90	2.42	66	6836	0.62	0.42	0.90	
Wald p; p adj			0.47;	1.00				0.12;	1.00				0.01	0.28		
rs4947979																
AA	131	13572	1.00			59	6509	1.00			72	7063	1.00			0.50 (1.00)
AG/GG	85	6889	1.36	1.03	1.79	42	3591	1.23	0.82	1.84	43	3298	1.47	1.00	2.16	
Wald p; p adj			0.03	0.51				0.31	1.00				0.05	0.93		
rs845552																
AA	38	4799	1.00			19	2509	1.00			19	2289				
AG	99	9086	1.43	0.98	2.09	52	5146	1.28	0.75	2.19	47	3940				
GG	47	3857	1.62	1.05	2.49	28	2172	1.86	1.02	3.37	19	1685				
p-trend; p adj			0.03	0.51				0.04	0.78							

Models adjusted for age, study, admixture, SEER summary stage.

Table 5 shows associations of *EGFR* polymorphisms with risk of breast cancer-specific mortality for all invasive breast cancer cases and by ethnicity. The results displayed in this table were all significant (p<0.05) prior to adjustment for multiple comparisons. After adjusting for age, study site, SEER summary stage, and genetic ancestry, the TT genotype of rs6978771 (HR, 1.68; 95% CI 1.11-2.56, p trend=0.03), the CC genotype of rs9642391 (HR, 1.64; 95% CI 1.04-2.58, p trend=0.02), the AG/GG genotypes of rs4947979 (HR, 1.36; 95% CI 1.03-1.79, Wald p=0.03), and the GG genotype of rs845552 (HR, 1.62; 95% 1.05-2.49, p trend=0.03) were associated with increased risk of breast cancer death, whereas the GG genotype of rs1344307 (HR, 0.24; 95% CI 0.06-0.97, p trend=0.01) and the GA/AA genotypes of rs17151957 (HR, 0.69; 95% CI 0.52-0.91, Wald p=0.01) were associated with decreased risk of breast cancer death. Among Hispanics, two SNPs were associated with increased risk of breast cancer death ($rs6593205_{AA}$: HR, 2.08; 95% CI 1.24-3.48, p trend=0.03; $rs6944906_{AA/GG}$: HR, 1.62; 95% CI 1.05-2.48, p trend=0.03), and two SNPs were associated with decreased risk of mortality ($rs17151957_{GA/AA}$: HR, 0.61; 95% CI 0.42-0.88, Wald p=0.01; and $rs917880_{CT/TT}$: HR, 0.62; 95% CI 0.42-0.90, Wald p=0.01). Among NHW women, the CT/TT genotypes of rs4947971 decreased risk of breast cancer death (HR, 0.59; 95% CI 0.40-0.89, Wald p=0.01) and the GG genotype of rs845552 was associated with increased risk (HR, 1.86; 95% CI 1.02-3.37, p trend=0.04). Overall, none of the associations between the *EGFR* SNPs and breast cancer-specific mortality remained statistically significant after adjusting for multiple comparisons (**Table 5**). Interaction between *EGFR* polymorphisms and breast cancer death by genetic ancestry were similar to those reported by ethnicity (data not shown).

We also examined the associations between *EGFR* polymorphisms and risk of breast cancer death by menopausal status (data not shown). There were no significant interactions for risk of breast cancer death by menopausal status within our study population. Our data did suggest an increase in risk of breast cancer mortality among premenopausal women for the following polymorphisms: $rs9642391_{CC}$ HR, 2.08, 95% CI 1.08-4.02, p trend=0.06 and $rs4947979_{AG/GG}$ HR, 1.56, 95% CI 1.03-2.38, Wald p=0.04. The TT genotype of rs6978771 (HR, 1.74, 95% CI 1.02-2.97, p trend=0.17) and the CC genotype of rs3752651 (HR, 2.12, 95% CI 1.01-4.45, Wald p=0.05) were associated with increased risk of breast cancer mortality among postmenopausal women. The GA/AA genotypes of rs17151957 was inversely associated with breast cancer mortality among postmenopausal cases (HR, 0.58, 95% CI 0.39-0.85, Wald p=0.01). None of the results by menopausal status remained statistically significant after multiple comparisons.

Discussion

Our study is one of the first using a tag-SNP approach to examine the associations of *EGFR* polymorphisms with risk of breast cancer-specific mortality and risk of breast cancer by ER/PR tumor phenotype. Nonetheless, only one association from the present analysis remained statistically significant after adjusting for multiple comparisons; rs2075112 was associated with significantly reduced risk for ER-/PR+ tumor phenotype. Prior to adjustment for multiple comparisons, two *EGFR* SNPs were found to be associated with overall breast cancer risk. With respect to breast cancer-specific mortality, we identified associations with four *EGFR* SNPs (rs6978771, rs9642391, rs4947979, and rs845552); and, after stratifying by ethnicity, we found rs6944906 and rs6593205 to be uniquely associated with increased risk of breast cancer death among Hispanic women. Only rs845552 was associated with increased risk of breast cancer death among NHW women from our sample, prior to adjustment for multiple comparisons.

EGFR is known to transfer extra-cellular mitogenic signals, such as EGF and transforming growth factor-alpha (TGF-α), by activating numerous downstream signaling cascades, which involve phospholipase C-c, Ras, and phosphatidylinositol-3 kinase (PI-3K) [31]. Apoptosis usually occurs after activation of the *EGFR*-mediated downstream pathways [31]. However, within cancer cells, there are altered gene activities leading to uncontrolled tumor proliferation. The mechanisms behind these outcomes are thought to involve Akt, also known as protein kinase B (PKB) [31]. When Akt is activated in breast cells, it phosphorylates cell cycle regulators, such as $p21^{Cip/WAF1}$, and subsequently promotes tumor survival by eradi-

cating the cell cycle checkpoints and apoptosis [31, 32]. Other research has suggested the existence of a more direct mode of the *EGFR* pathway which involves cellular transport of *EGFR* from the cell-surface to the cell nucleus, association of nuclear *EGFR* complex with gene promoters, and transcriptional regulation of the target genes [31]. Furthermore, evidence suggests that the *EGFR* pathway itself is associated with increased tumor cell proliferation and poor survival rate in women with breast cancer [31, 33].

Previous studies that examined the associations between *EGFR* polymorphisms and breast cancer risk have produced mixed findings. Several recent studies found no association between *EGFR* SNPs and breast cancer risk [9, 12, 16]. In a two-stage breast cancer case-control study using data from the Shanghai Breast Cancer Study [15], Hong et al. assessed associations with 51 *EGFR* polymorphisms, using a tagSNP approach. Stage 1 included 1,062 cases and 1,069 controls; and Stage 2 included 1,932 cases and 1,857 controls. Of the 51 *EGFR* SNPs, we examined the following SNPs: rs9642391, rs884419, rs6978771, rs6593205, rs763317, rs917880, rs11977660, rs3752651, rs2472520, and rs2293348. The Shanghai Study found significant associations with ten SNPs in Stage 1 (rs3735064, rs845562, rs845560, rs17172434, rs7780270, rs9642391, rs11976696, rs15543848, rs7808697, and rs884419). However, in a Stage 2 analysis in an independent study population, associations with the 10 SNPs could not be confirmed, suggesting that the results detected in Stage 1 were perhaps chance findings [15]. Similarly, we found no significant associations with rs9642391 and rs884419.

As previously reported, Jami and colleagues reported that the short/short genotype, compared to the long/long genotype, of rs11568315 was associated with an almost two-fold increased risk of breast cancer overall, and a nearly three-fold increased risk among women aged <55 years [2]. Brandt et al. also examined the relationship between *EGFR* and breast cancer risk in young women diagnosed at age <50 years and found no association for the main effects between the polymorphic CA repeat located at the 5-regulatory sequence in the intron 1 of *EGFR* and breast cancer risk; however, having two long alleles (≥19 CA) was associated with a significantly increased risk of breast cancer among women with a first degree family history of breast cancer (OR, 10.4; 95% CI 1.85-58.70, p interaction=0.015) [34]. We also investigated interaction effects by menopausal status; however, we did not find significant differences in results between premenopausal and postmenopausal women.

EGFR overexpression has been found in approximately 50% of triple negative breast cancer cases [35] and Hispanic women with breast cancer compared to NHWs are more likely to have triple negative disease [13, 14]. Our study is the first to investigate the associations between *EGFR* polymorphisms and breast cancer risk among Hispanic women. Although we were unable to assess risk for triple negative breast cancer due to incomplete data on HER2 status, our results suggest, prior to multiple comparisons, an association between one *EGFR* SNP and ER-/PR- tumor phenotype. This analysis, however, was limited to the two U.S. study centers, given the lack of tumor phenotype data for the MBCS.

There are other strengths and limitations to the present analysis. Our study was able to compare the associations between 34 *EGFR* polymorphisms and breast cancer risk by ethnicity and levels of Native American ancestry. However, given the large number of SNPs analyzed, almost all of the detected associations were no longer significant after adjustment for multiple comparisons. This adjustment may have contributed to false negative associations [36], thus replication of our findings is warranted.

We also were able to examine associations of *EGFR* polymorphisms with breast cancer-specific mortality. There are well documented disparities in breast cancer outcomes between Hispanic and NHW women [13, 37], and we examined whether differences in associations with *EGFR* polymorphisms could possibly explain some of the breast cancer survival disparities. This analysis, however, was limited to the U.S. study centers and thus we were not able to evaluate the full range of Native American ancestry. A strength of the survival analyses is the length of follow-up time, approximately 10 years for the SFBCS and approximately 8 years for the 4-CBCS.

In conclusion, we observed significant associations of specific SNPs in the *EGFR* gene with breast cancer risk and with breast cancer-specific mortality, before adjustment for multiple comparisons. Our results also suggest that these associations may differ according to ER/PR tumor phenotype. Some of our findings also suggest that differences between Hispanic and NHW women for breast cancer risk and mortality might be influenced by specific *EGFR* SNPs. These findings provide additional insight for the role of *EGFR* in breast cancer development and prognosis. Further research is needed to elucidate the contribution of *EGFR* to ethnic disparities in breast cancer.

Acknowledgements

The Breast Cancer Health Disparities Study was funded by grant CA14002 from the National Cancer Institute to Dr. Slattery. The San Francisco Bay Area Breast Cancer Study was supported by grants CA63446 and CA77305 from the National Cancer Institute, grant DAMD17-96-1-6071 from the U.S. Department of Defense and grant 7PB-0068 from the California Breast Cancer Research Program. The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN2612010000-36Cawarded to the Cancer Prevention Institute of California; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement #1U58 DP000807-01 awarded to the Public Health Institute. The 4-Corner's Breast Cancer Study was funded by grants CA078682, CA078762, CA078552, and CA078802 from the National Cancer Institute. The research also was supported by the Utah Cancer Registry, which is funded by contract N01-PC-67000 from the National Cancer Institute, with additional support from the State of Utah Department of Health, the New Mexico Tumor Registry, and the Arizona and Colorado cancer registries, funded by the Centers for Disease Control and Prevention National Program of Cancer Registries and additional state support. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official view of the National Cancer Institute or endorsement by the State of California Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors. The Mexico Breast Cancer Study was funded by Consejo Nacional de Ciencia y Tecnología (CONACyT) (SALUD-2002-C01-7462). We would also like to acknowledge the contributions of the following individuals to the study: Jennifer Herrick and Sandra Edwards for data harmonization oversight; Erica Wolff and Michael Hoffman for laboratory support; Carolina Ortega for her assistance with data management for the Mexico Breast Cancer Study, Jocelyn Koo for data management for the San Francisco Bay Area Breast Cancer Study, Dr. Tim Byers for his contribution to the 4-Corner's Breast Cancer Study, and Dr. Josh Galanter for assistance in selection of AIMs markers for the study.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Avonne E Connor, Department of Epidemiology and Population Health, School of Public Health and Information Sciences, University of Louisville, 485 E. Gray St. Louisville, KY 40202, USA. Tel: 502-852-6441; Fax: 502-852-3294; E-mail: aeconn02@louisville.edu; Dr. Martha L Slattery, Department of Medicine, University of Utah, 383 Colorow, Salt Lake City, Utah 84108, USA. E-mail: marty.slattery@hsc.utah.edu

References

[1] Yano S, Kondo K, Yamaguchi M, Richmond G, Hutchison M, Wakeling A, Averbuch S and Wadsworth P. Distribution and function of EGFR in human tissue and the effect of EGFR tyrosine kinase inhibition. Anticancer Res 2003; 23: 3639-3650.

[2] Jami MS, Hemati S, Salehi Z and Tavassoli M. Association between the length of a CA dinucleotide repeat in the EGFR and risk of breast cancer. Cancer Invest 2008; 26: 434-437.

[3] Brandt B, Meyer-Staeckling S, Schmidt H, Agelopoulos K and Buerger H. Mechanisms of egfr gene transcription modulation: relationship to cancer risk and therapy response. Clin Cancer Res 2006; 12: 7252-7260.

[4] Foley J, Nickerson NK, Nam S, Allen KT, Gilmore JL, Nephew KP and Riese DJ 2nd. EGFR signaling in breast cancer: bad to the bone. Semin Cell Dev Biol 2010; 21: 951-960.

[5] Garcia Castro C, Ravina M, Castro V and Salido EC. Expression of epidermal growth factor receptor (proto-oncogene c-erbB-1) and estrogen receptor in human breast carcinoma. An immunocytochemical study of 70 cases. Arch Gynecol Obstet 1993; 252: 169-77.

[6] Rakha EA, El-Sayed ME, Green AR, Lee AH, Robertson JF and Ellis IO. Prognostic markers in triple-negative breast cancer. Cancer 2007; 109: 25-32.

[7] Teng YH, Tan WJ, Thike AA, Cheok PY, Tse GM, Wong NS, Yip GW, Bay BH and Tan PH. Mutations in the epidermal growth factor receptor (EGFR) gene in triple negative breast cancer: possible implications for targeted therapy. Breast Cancer Res 2011; 13: R35.

[8] Jacot W, Lopez-Crapez E, Thezenas S, Senal R, Fina F, Bibeau F, Romieu G and Lamy PJ. Lack of EGFR-activating mutations in European patients with triple-negative breast cancer could emphasise geographic and ethnic variations in breast cancer mutation profiles. Breast Cancer Res 2011; 13: R133.

[9] Choura M, Frikha F, Kharrat N, Aifa S and Rebai A. Investigating the function of three non-synonymous SNPs in EGFR gene: structural modelling and association with breast cancer. Protein J 2010; 29: 50-54.

[10] Nomura M, Shigematsu H, Li L, Suzuki M, Takahashi T, Estess P, Siegelman M, Feng Z, Kato H, Marchetti A, Shay JW, Spitz MR, Wistuba II, Minna JD and Gazdar AF. Polymorphisms, mutations, and amplification of the EGFR gene in non-small cell lung cancers. PLoS Med 2007; 4: e125.

[11] Merlino GT, Ishii S, Whang-Peng J, Knutsen T, Xu YH, Clark AJ, Stratton RH, Wilson RK, Ma DP, Roe BA, et al. Structure and localization of genes encoding aberrant and normal epidermal growth factor receptor RNAs from A431 human carcinoma cells. Mol Cell Biol 1985; 5: 1722-34.

[12] Sobti RC, Askari M, Nikbakht M, Singh N, Sharma SC and Abitew AM. Genetic variants of EGFR (142285G>A) and ESR1 (2014G>A) gene polymorphisms and risk of breast cancer. Mol Cell Biochem 2012; 369: 217-25.

[13] Ooi SL, Martinez ME and Li CI. Disparities in breast cancer characteristics and outcomes by race/ethnicity. Breast Cancer Res Treat 2011; 127: 729-738.

[14] Lara-Medina F, Perez-Sanchez V, Saavedra-Perez D, Blake-Cerda M, Arce C, Motola-Kuba D, Villarreal-Garza C, Gonzalez-Angulo AM, Bargallo E, Aguilar JL, Mohar A and Arrieta O. Triple-negative breast cancer in Hispanic patients: high prevalence, poor prognosis, and association with menopausal status, body mass index, and parity. Cancer 2011; 117: 3658-3669.

[15] Hong YS, Deming SL, Gao YT, Long JR, Shu XO, Cai Q, Lu W and Zheng W. A two-stage case-control study of EGFR polymorphisms and breast cancer risk. Cancer Epidemiol Biomarkers Prev 2009; 18: 680-683.

[16] Kallel I, Rebai M, Khabir A, Farid NR and Rebai A. Genetic polymorphisms in the EGFR (R521K) and estrogen receptor (T594T) genes, EGFR and ErbB-2 protein expression, and breast cancer risk in Tunisia. J Biomed Biotechnol 2009; 2009: 753683.

[17] Slattery ML, John EM, Torres-Mejia G, Lundgreen A, Herrick JS, Baumgartner KB, Hines LM, Stern MC and Wolff RK. Genetic variation in genes involved in hormones, inflammation, and energetic factors and breast cancer risk in an admixed population. Carcinogenesis 2012; 33: 1512-21.

[18] Slattery ML, Sweeney C, Edwards S, Herrick J, Baumgartner K, Wolff R, Murtaugh M, Baumgartner R, Giuliano A and Byers T. Body size, weight change, fat distribution and breast cancer risk in Hispanic and non-Hispanic white women. Breast Cancer Res Treat 2007; 102: 85-101.

[19] Angeles-Llerenas A, Ortega-Olvera C, Perez-Rodriguez E, Esparza-Cano JP, Lazcano-Ponce E, Romieu I and Torres-Mejía G. Moderate physical activity and breast cancer risk: the effect of menopausal status. Cancer Causes Control 2010; 21: 577-86.

[20] John EM, Horn-Ross PL and Koo J. Lifetime physical activity and breast cancer risk in a multiethnic population: the San Francisco Bay area breast cancer study. Cancer Epidemiol Biomarkers Prev 2003; 12: 1143-1152.

[21] John EM, Phipps AI, Davis A and Koo J. Migration history, acculturation, and breast cancer risk in Hispanic women. Cancer Epidemiol Biomarkers Prev 2005; 14: 2905-2913.

[22] Falush D, Stephens M and Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 2003; 164: 1567-1587.

[23] Pritchard JK, Stephens M and Donnelly P. Inference of population structure using multilocus genotype data. Genetics 2000; 155: 945-959.

[24] Hosmer D, Lemeshow S. Applied Logistic Regression. New York: Wiley, 1989.

[25] Flom PL. Multinomial and ordinal logistic regression using PROC LOGISTIC.9. New York: National Development and Research Institutes, Inc., 2010.

[26] Hosmer DW, Lemeshow S. Applied Logistic Regression. 2nd Edition. New York: John Wiley and Sons, Inc., 2000.

[27] Nyholt DR. A simple correction for multiple testing for single-nucleotide polymorphisms in

linkage disequilibrium with each other. Am J Hum Genet 2004; 74: 765-769.
[28] Li J and Ji L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. Heredity (Edinb) 2005; 95: 221-227.
[29] Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scand J Statist 1979; 6: 65-70.
[30] Connor AE, Baumgartner RN, Baumgartner KB, Kerber RA, Pinkston C, John EM, Torres-Mejia G, Hines L, Giuliano A, Wolff RK and Slattery ML. Associations between TCF7L2 polymorphisms and risk of breast cancer among Hispanic and non-Hispanic White women: the Breast Cancer Health Disparities Study. Breast Cancer Res Treat 2012; 136: 593-602.
[31] Lo HW, Hsu SC and Hung MC. EGFR signaling pathway in breast cancers: from traditional signal transduction to direct nuclear translocalization. Breast Cancer Res Treat 2006; 95: 211-218.
[32] Li Y, Dowbenko D and Lasky LA. AKT/PKB phosphorylation of p21Cip/WAF1 enhances protein stability of p21Cip/WAF1 and promotes cell survival. J Biol Chem 2002; 277: 11352-11361.
[33] Lo HW, Xia W, Wei Y, Ali-Seyed M, Huang SF and Hung MC. Novel prognostic value of nuclear epidermal growth factor receptor in breast cancer. Cancer Res 2005; 65: 338-348.
[34] Brandt B, Hermann S, Straif K, Tidow N, Buerger H and Chang-Claude J. Modification of breast cancer risk in young women by a polymorphic sequence in the egfr gene. Cancer Res 2004; 64: 7-12.
[35] Dent R, Trudeau M, Pritchard KI, Hanna WM, Kahn HK, Sawka CA, Lickley LA, Rawlinson E, Sun P and Narod SA. Triple-negative breast cancer: clinical features and patterns of recurrence. Clin Cancer Res 2007; 13: 4429-4434.
[36] Rothman KJ. No adjustments are needed for multiple comparisons. Epidemiology 1990; 1: 43-46.
[37] Li CI, Malone KE and Daling JR. Differences in breast cancer stage, treatment, and survival by race and ethnicity. Arch Intern Med 2003; 163: 49-56.