

***FGFR2* and other loci identified in genome-wide association studies are associated with breast cancer in African-American and younger women**

Jill S. Barnholtz-Sloan^{1,2,*}, Priya B. Shetty^{1,2},
Xiaowei Guan^{1,2}, Sarah J. Nyante³, Jingchun Luo⁴,
Donal J. Brennan⁵ and Robert C. Millikan^{4,6}

¹Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, Cleveland, OH 44106-5065, USA, ²Department of Epidemiology and Biostatistics, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA, ³Hormonal and Reproductive Epidemiology Branch, National Cancer Institute, National Institutes of Health, Rockville, MD 20852, USA, ⁴Lineberger Comprehensive Cancer Center, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA, ⁵UCD School of Biomolecular and Biomedical Science, UCD Conway Institute, University College Dublin, Ireland and ⁶Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

*To whom correspondence should be addressed. Case Comprehensive Cancer Center, Case Western Reserve University School of Medicine, 11100 Euclid Ave, Wearn 152, Cleveland, OH 44106-5065, USA. Tel: +1 216 368 1506; Fax: +1 216 844 7832; Email: jsb42@case.edu

Twenty-nine single-nucleotide polymorphisms (SNPs) from previously published genome-wide association studies (GWAS) and multiple ancestry informative markers were genotyped in the Carolina Breast Cancer Study (CBCS) (742 African-American (AA) cases, 1230 White cases; 658 AA controls, 1118 White controls). In the entire study population, 9/10 SNPs in fibroblast growth factor receptor 2 (*FGFR2*) were significantly associated with breast cancer after adjusting for age, race and European ancestry [odds ratios (OR) range 1.17–1.81]. Associations were observed for SNPs in *FGFR2*, *LSP1*, *H19*, *TLR1/TLR6* and *RELN* for AA; *FGFR2*, *TNRC9*, *H19* and *MAP3K1* for Whites; *FGFR2*, *TNRC9*, *Msc5A1* and chromosome 8q for women ≥ 50 years old and *FGFR2* and *TNRC9* for women < 50 years old. *FGFR2* haplotypes based upon rs11200014, rs2981579, rs1219648 and rs2420946 were associated with increased risk of breast cancer, including the GTGT haplotype in AAs [OR = 1.27, 95% confidence interval (CI) 1.04–1.56] and younger women of either race [OR = 1.35, 95% CI 1.02–1.78] and the ATGT haplotype in Whites (OR = 1.30, 95% CI 1.15–1.46). Recent GWAS hits for breast cancer in Europeans and Whites (i.e. women of European descent) thus showed evidence of replication among AAs and Whites in the CBCS. Several new haplotypes were associated with breast cancer in AA and younger women, particularly the *FGFR2* GTGT haplotype. These results highlight the need to conduct GWAS among younger women and in a variety of racial–ethnic populations.

Introduction

Until recently, the search for reproducible common, low-penetrance susceptibility genes for breast cancer yielded few positive findings (1). A turning point was reached with the advent of genome-wide association studies (GWAS) (2). Two GWAS of breast cancer were published in 2007 using data collected from European and White (i.e. of European descent) women (3,4). Easton *et al.* (3) discovered five breast cancer susceptibility loci, including fibroblast growth factor

receptor 2 (*FGFR2*) on chromosome 10q26, *TNRC9/TOX3* at 16q12, *MAP3K1* at 5q11, *LSP1* at 11p15 and a locus on 8q. Hunter *et al.* (4) confirmed an association between *FGFR2* and sporadic postmenopausal breast cancer and also identified additional susceptibility loci at *RELN* on 7q and *TLR1/TLR6* on 4p. More recent GWAS conducted in European or Whites, and a few studies among Asians, have discovered loci on chromosomes 2q25 (5,6), 6q22 (7), 6q25 (8), 3p24 and 17q23 (9), as well as 1p11 and 14q24 (10). Within these regions of interest, relative risks ranging from 1.1 to 1.5 have been estimated for single-nucleotide polymorphisms (SNPs) located in high linkage disequilibrium (LD) blocks ranging in size from 25 to 600 kb. Minor allele frequencies for SNPs showing the strongest signals range from 0.13 to 0.50, indicating that the alleles may contribute substantially to breast cancer susceptibility on a population level (6).

Most previous GWAS of breast cancer focused on European or White women and included primarily postmenopausal women. In women under age 45, the incidence of breast cancer is higher in African-American (AA) women compared with White women. Among older women, breast cancer incidence is higher in Whites. Breast cancer mortality is higher among AA women compared with White women across all age groups (11). Identifying variants, particularly in key genes like *FGFR2*, that are specific to AAs or younger women may aid in improving knowledge about breast cancer development and clinically relevant pathways for targeted prevention (12). However, to date, only three studies have addressed the role of *FGFR2* in AAs (13–15) and few studies included younger women. AAs have shorter LD blocks on average and exhibit greater haplotype diversity compared with Europeans and Whites (6), which may facilitate detection of additional risk haplotypes, mapping of GWAS loci and location of potential causal alleles. Using the Carolina Breast Cancer Study (CBCS), a population-based case–control of AA and White women, we evaluated SNPs and haplotypes in *FGFR2* and other previous GWAS-identified loci for their association with breast cancer. We aimed to evaluate GWAS risk genotypes and/or haplotypes in AA women and women diagnosed at age < 50 .

Materials and methods

Study population

The CBCS is a population-based case–control study of breast cancer conducted in North Carolina (16,17). Briefly, eligible cases included women ages 20–74 who were diagnosed with primary invasive breast cancer from 1993 to 2001 and lived within a 24 county study area. Cases were identified using rapid case ascertainment in cooperation with the North Carolina Central Cancer Registry. Randomized recruitment was used to oversample AAs and women < 50 years of age (18). Women diagnosed with breast carcinoma *in situ* were also enrolled in the study from 1996 to 2001. Eligible controls were women aged 20–74 years, residing within the study area, with no history of breast cancer and were identified using Division of Motor Vehicles lists (for women < 65 years) and Medicare records (for women aged 65–74 years). Controls were frequency matched to cases according to race within 5 years age categories.

Women who agreed to participate in the study provided informed consent and completed an in-home interview regarding known and suspected breast cancer risk factors. Women were also asked to provide a 30 ml blood sample. DNA was extracted from the blood samples and stored at -80°C . The interview participation rates for invasive cases and controls were 76 and 55%, respectively, and for carcinoma *in situ*, cases and controls were 83 and 65%, respectively.

Age was defined as age in years at breast cancer diagnosis for cases or at the time of sampling for controls and was dichotomized as < 50 (younger) and ≥ 50 (older) for analysis purposes. Self-identified race was reported by each study participant during the study interview and was used to classify study participants for statistical analyses. Of 3748 CBCS participants with genotyping data, 2293 identified themselves as Caucasian (61%) and 1400 self-identified as AA (37%). Less than 2% of participants ($N = 53$) reported that

Abbreviations: AIM, ancestry informative marker; AA, African-American; CBCS, Carolina Breast Cancer Study; CI, confidence interval; *FGFR2*, fibroblast growth factor receptor 2; GWAS, genome-wide association studies; HWE, Hardy–Weinberg equilibrium; LD, linkage disequilibrium; OR, odds ratio; SNP, single-nucleotide polymorphism.

they were Hispanic, mixed race or other race/ethnicity. For regression analyses, these women and the Caucasian women were grouped together as White.

Overall, 2311 cases (894 AA, 1417 White) and 2022 controls (788 AA, 1234 White) enrolled in the study and 2045 (89%) cases and 1818 (90%) controls provided DNA for genotyping. For the present analysis, the dataset included 742 AA cases (387 AA cases ≥ 50 years of age and 355 AA cases < 50 years), 658 AA controls (344 AA controls ≥ 50 years of age and 314 AA controls < 50 years), 1230 White cases (619 White cases ≥ 50 years of age and 611 White cases < 50 years) and 1118 White controls (607 White controls ≥ 50 years of age and 511 White controls < 50 years). All study procedures involving human subjects were approved by the University of North Carolina at Chapel Hill Institutional Review Board and in accordance with an assurance filed with and approved by the Department of Health and Human Services.

SNP selection

Previously reported SNPs showing associations with breast cancer in one or more GWAS were selected for genotyping, including SNPs in *CASP8*, *FGFR2*, *TNRC9*, *LSP1*, *H19*, *Msc5A1*, *TLR1/TLR6*, *MAP3K1*, *RELN* and markers on chromosomes 2p, 2q35, 5p, 5q and 8q (Tables I and II) (3–10). A panel of 158 ancestry informative markers (AIMs) was chosen to be maximally informative for distinguishing between African and European ancestries, which have been shown to be the two relevant ancestral populations for AAs and Whites (i.e. of European descent) (19–21) (supplementary Table 1 is available at *Carcinogenesis* Online). AIMs were selected to maximize the difference in allele frequencies between ancestral populations and the Fisher's information criterion (22) for distinguishing between African and European ancestry, based upon ancestral allele frequencies from Yoruba in Ibadan, Nigeria and CEPH (Utah residents with ancestry from northern and western Europe) populations in HapMap (www.hapmap.org). AIMs were prioritized based on having the highest Fisher's information criterion values in the following order: 90% European/10% African, 10% European/90% African and 50% European/50% African (supplementary Table 1 is available at *Carcinogenesis* Online). This prioritization allowed the AIMs to be chosen to represent the whole expected ancestral distribution of this population.

Genotyping results and quality control

The SNPs in this study were genotyped as part of a larger panel of 1536 SNPs. SNPs were chosen for replication of previous GWAS hits based upon the most significant published findings in Europeans and Whites. Genotyping was performed by the University of North Carolina Mammalian Genotyping Core using the Illumina GoldenGate assay (Illumina, San Diego, CA). Assay intensity data and genotype cluster images for all SNPs were reviewed individually. One hundred and sixty-three SNPs were excluded from the dataset (11%) due to low signal intensity or inability to distinguish between genotype clusters. Blind duplicates of 169 samples were genotyped in order to verify the reproducibility of genotype calls. No SNPs were excluded at this step (i.e. no SNP had > 2 genotype call errors between blind duplicates). Exact tests of Hardy–Weinberg equilibrium (HWE) were conducted in controls stratified by race to check for potential genotyping errors, and HWE test statistics and *P*-values were calculated in Plink v1.05 (23). In order to confirm that HWE deviations were not due to erroneous genotype calls, genotyping cluster images were rereviewed for all SNPs with HWE *P*-values < 0.01 . All but four SNPs rereviewed during this process had acceptable signal intensity and genotype cluster definitions.

Overall, 1373 of 1536 (89%) SNPs passed quality control, including 2/2 (100%) for *CASP8*, 10/11 (91%) for *FGFR2* and 17/17 (100%) for *TNRC9*, *LSP1*, *H19*, *Msc5A1*, *MAP3K1*, *RELN*, *TLR1/TLR6* and chromosomes 2p, 2q35, 5p, 5q and 8q. One hundred and forty-four of 158 (91%) AIMs passed quality control. Participants were excluded because of genotype calls for $< 95\%$ of SNPs ($N = 103$), gender mismatch ($N = 5$) and suspected contamination ($N = 1$). Following quality control, 1776 of 2022 (88%) controls and 1972 of 2311 (85%) cases were successfully genotyped. Participants without genotype data were more probably to be cases, recruited during later years of the study and AA. Among cases, the presence of genotype data was not associated with stage at diagnosis, lymph node status or other clinical variables.

Statistical analysis

Individual estimates of European ancestry were calculated from 144 AIMs using maximum likelihood methods [as described in (19,20)] (supplementary Table 1 is available at *Carcinogenesis* Online). Average and median proportions of European ancestry were 0.93 and 0.94 for Whites and 0.22 and 0.19 for AA, respectively (Figure 1). AIMs were selected to differentiate between African and European ancestry only, so ancestry proportions for the two groups add up to 1.0. Allele frequencies, genotype frequencies and *D'* (as a measure of LD) were calculated using SAS Genetics (24).

All single SNP and haplotype analyses were performed for the total study population and stratified by either self-reported race (AA or White) or age at

diagnosis (< 50 or ≥ 50). All models were adjusted for age and/or race as appropriate as well as individual European ancestry proportion. Individual European ancestry proportion was included in all models in order to control for potential residual confounding due to population stratification. An offset term was also included in all models to account for the randomized recruitment probabilities used to sample eligible study participants (15). Single SNP genotypic associations with breast cancer were modeled using an unconditional logistic regression model with two degrees of freedom (i.e. codominant model). Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated from these models for homozygote and heterozygote genotypes, as well as for additive (per-allele) models. While the additive model is a special case nested within the general codominant model, we present results using both models because the codominant model makes no assumption of a dose–response relationship across genotype categories and allows each genotype category to give a different and non-additive risk. The additive model assumes homozygotes have double the risk compared with heterozygotes on the log scale and can be used as a test for linear trend. Results did not differ when we excluded cases with *in situ* breast cancer or participants with mixed, Hispanic or other race/ethnicities.

Haplotypes were estimated and tested for association with breast cancer risk using haplo.stats (25). SNPs were selected for haplotype analysis if *D'* values were ≥ 0.7 within any gene; hence haplotypes were estimated for *FGFR2* using SNPs rs11200014, rs2981579, rs1219648 and rs2420946 (*D'* ranging from 0.822 to 1.000), *TNRC9* using SNPs rs12443621 and rs3803662 (*D'* = 0.714), *CASP8* using SNPs rs1045485 and rs17468277 (*D'* = 0.999). For *FGFR2*, haplotypes were estimated using the four SNPs listed above in order to directly compare our results with previous studies (*D'* values between 0.82 and 1.0) (4). Haplotypes were also estimated using other genotyped SNPs in *FGFR2* and results were similar since these SNPs were in strong LD with the four SNP haplotype (data not shown). In addition, these four SNPs in *FGFR2* are all in strong LD with each other, including rs2981582, which is the original SNP found to be associated with increased risk of breast cancer in women of European ancestry (3,4). The most common haplotype (GCAC) was associated with lower breast cancer risk compared with the other haplotypes and designated the reference group.

Maximum likelihood estimates of haplotype frequencies and the posterior probabilities of the pairs of haplotypes for each individual were calculated using the expectation–maximization algorithm as implemented in haplo.stats (26). The posterior probabilities were then incorporated into generalized linear models to compute the score statistics for the association between a haplotype and breast cancer risk. The most common haplotype for each locus formed the reference group. ORs and 95% CIs were calculated from these models, adjusting for age and/or race as appropriate, individual European ancestry proportion and offset term. Haplotype *P*-values were two sided and were derived from the *t*-statistic using the beta coefficient in logistic regression models with a significance level of 0.05. Rare haplotypes with a frequency $< 0.5\%$ were binned as a group of rare haplotypes, but associations for this group are not reported due to the difficulty in deriving meaningful interpretations.

Results

In the entire study population, combining AA and White study participants, the minor alleles in nine of 10 SNPs in *FGFR2* were significantly associated with increased odds of breast cancer using codominant and additive models after adjusting for age, race, European ancestry and study offset term with adjusted ORs ranging from 1.17 to 1.81 (supplementary Table 2 is available at *Carcinogenesis* Online). Variants in SNPs in *TNRC9* (rs8051542 and rs3803662) and *TLR1/TLR6* (rs7696175) were also associated with increased risk of breast cancer in the entire study population. No associations were observed for SNPs in *CASP8*, *LSP1*, *H19*, *Msc5A1*, *MAP3K1*, *RELN* or loci on chromosomes 2p, 2q35, 5p, 5q or 8q. In the haplotype analysis for the entire study population, only haplotypes in *FGFR2* showed positive associations with breast cancer. Three *FGFR2* haplotypes (ATGT, GTGT and GTAT) were associated with increased odds of breast cancer after controlling for race, age, European ancestry and study offset term (supplementary Table 3 is available at *Carcinogenesis* Online).

Single SNP results adjusted for age, European ancestry and study offset term stratified by self-reported race are presented in Table I (see supplementary Table 4, available at *Carcinogenesis* Online, for genotype frequencies stratified by self-reported race). Per-allele ORs for *FGFR2* using an additive model were slightly stronger in White

Table I. Single SNP results stratified by self-reported race

Dataset	Gene or locus	rs no.	Alleles (test/reference)	Codominant model (heterozygous), OR (95% CI) ^a	Codominant model (homozygous), OR (95% CI) ^a	Additive model (per-allele), OR (95% CI) ^a	
AA	<i>CASP8</i>	rs1045485	C/G	0.96 (0.68–1.36)	0.45 (0.05–3.74)	0.93 (0.67–1.30)	
		rs17468277	T/C	0.96 (0.67–1.37)	0.24 (0.02–3.28)	0.92 (0.65–1.29)	
	<i>FGFR2</i>	rs1896395	A/C	0.87 (0.69–1.10)	1.69 (0.94–3.07)	1.01 (0.83–1.23)	
		rs11200014	A/G	1.00 (0.79–1.27)	1.23 (0.71–2.11)	1.04 (0.86–1.26)	
		rs2981579	T/C	1.39 (1.02–1.90)	1.54 (1.12–2.13)	1.22 (1.04–1.42)	
		rs1219648	G/A	1.13 (0.89–1.45)	1.45 (1.05–1.99)	1.19 (1.02–1.39)	
		rs2912774	A/C	1.05 (0.78–1.41)	1.55 (1.13–2.13)	1.27 (1.09–1.49)	
		rs2936870	T/C	1.05 (0.78–1.41)	1.54 (1.12–2.11)	1.27 (1.09–1.48)	
		rs2981582	T/C	1.30 (1.00–1.68)	1.40 (1.02–1.91)	1.19 (1.02–1.39)	
		rs2420946	T/C	1.02 (0.77–1.35)	1.36 (0.99–1.86)	1.17 (1.00–1.37)	
		rs2162540	G/A	1.19 (0.91–1.56)	1.51 (1.11–2.06)	1.23 (1.05–1.44)	
		rs3135718	G/A	1.11 (0.83–1.48)	1.54 (1.14–2.10)	1.26 (1.08–1.46)	
		rs8049149	C/T	0.93 (0.55–1.59)	0.69 (0.04–11.25)	0.92 (0.56–1.52)	
		<i>TNRC9 3' UTR</i>	rs16951186	C/T	0.81 (0.63–1.03)	1.10 (0.61–1.97)	0.90 (0.74–1.09)
		<i>TNRC9 5' UTR</i>	rs8051542	C/T	1.04 (0.83–1.31)	1.45 (0.99–2.11)	1.14 (0.97–1.35)
		<i>TNRC9/TOX3</i>	rs12443621	G/A	1.21 (0.92–1.58)	1.35 (0.99–1.83)	1.16 (0.99–1.35)
		<i>TNRC9 5' UTR</i>	rs3803662	C/T	0.97 (0.74–1.29)	0.90 (0.66–1.23)	0.95 (0.81–1.11)
	<i>TNRC9</i>	rs9940048	G/A	1.20 (0.96–1.51)	1.07 (0.71–1.61)	1.10 (0.93–1.31)	
	<i>LSP1</i>	rs3817198	C/T	1.19 (0.93–1.52)	0.47 (0.23–0.99)	1.01 (0.82–1.24)	
	<i>H19</i>	rs2107425	T/C	1.31 (1.00–1.73)	1.43 (1.04–1.97)	1.20 (1.02–1.40)	
	<i>Msc5A1</i>	rs6476643	T/G	1.13 (0.89–1.42)	1.21 (0.68–2.15)	1.12 (0.92–1.36)	
	<i>TLR1/TLR6</i>	rs7696175	T/C	1.22 (0.88–1.71)	4.11 (1.28–13.24)	1.39 (1.04–1.86)	
	<i>MAP3K1</i>	rs889312	C/A	1.00 (0.79–1.25)	0.84 (0.57–1.25)	0.95 (0.80–1.13)	
	<i>RELN</i>	rs17157903	T/C	0.87 (0.66–1.16)	5.21 (1.48–18.29)	1.07 (0.83–1.37)	
	Chromosome 2p	rs4666451	A/G	1.03 (0.82–1.30)	0.48 (0.28–0.81)	0.87 (0.72–1.05)	
	Chromosome 2q35	rs13387042	G/A	1.00 (0.80–1.26)	0.91 (0.58–1.43)	0.98 (0.82–1.17)	
	Chromosome 5p	rs981782	G/T	0.86 (0.63–1.17)	1.10 (0.39–3.16)	0.90 (0.68–1.19)	
	Chromosome 5q	rs30099	T/C	1.26 (0.98–1.62)	1.28 (0.59–2.74)	1.22 (0.98–1.52)	
	Chromosome 8q	rs13281615	G/A	1.14 (0.89–1.46)	0.97 (0.70–1.34)	1.00 (0.86–1.18)	
	White	<i>CASP8</i>	rs1045485	C/G	0.92 (0.75–1.13)	0.75 (0.4–1.42)	0.90 (0.76–1.08)
			rs17468277	T/C	0.92 (0.75–1.13)	0.75 (0.4–1.42)	0.91 (0.76–1.08)
		<i>FGFR2</i>	rs1896395	A/C	1.51 (0.45–5.03)		2.14 (0.78–5.92)
			rs11200014	A/G	1.30 (1.07–1.58)	1.63 (1.27–2.10)	1.28 (1.13–1.45)
			rs2981579	T/C	1.36 (1.12–1.66)	1.71 (1.33–2.19)	1.31 (1.16–1.48)
			rs1219648	G/A	1.35 (1.11–1.63)	1.65 (1.28–2.12)	1.30 (1.15–1.46)
			rs2912774	A/C	1.39 (1.14–1.68)	1.63 (1.27–2.09)	1.29 (1.14–1.46)
			rs2936870	T/C	1.38 (1.14–1.67)	1.64 (1.27–2.11)	1.30 (1.15–1.46)
			rs2981582	T/C	1.37 (1.13–1.66)	1.62 (1.26–2.09)	1.29 (1.14–1.46)
			rs2420946	T/C	1.38 (1.14–1.67)	1.62 (1.25–2.08)	1.29 (1.14–1.46)
			rs2162540	G/A	1.35 (1.12–1.64)	1.63 (1.26–2.10)	1.29 (1.14–1.46)
			rs3135718	G/A	1.40 (1.15–1.69)	1.64 (1.28–2.12)	1.30 (1.15–1.47)
rs8049149			C/T	0.84 (0.10–6.81)		0.84 (0.10–6.81)	
<i>TNRC9 3' UTR</i>			rs16951186	C/T	1.14 (0.63–2.04)		1.14 (0.63–2.04)
<i>TNRC9 5' UTR</i>			rs8051542	C/T	0.92 (0.76–1.12)	0.26 (0.98–1.61)	1.10 (0.97–1.24)
<i>TNRC9/TOX3</i>			rs12443621	G/A	0.75 (0.61–0.93)	0.74 (0.58–0.95)	0.87 (0.77–0.98)
<i>TNRC9 5' UTR</i>			rs3803662	C/T	1.20 (1.00–1.44)	1.64 (1.20–2.23)	1.25 (1.09–1.42)
<i>TNRC9</i>		rs9940048	G/A	1.07 (0.89–1.28)	1.01 (0.70–1.48)	1.04 (0.90–1.20)	
<i>LSP1</i>		rs3817198	C/T	1.04 (0.87–1.25)	1.27 (0.94–1.71)	1.07 (0.96–1.25)	
<i>H19</i>		rs2107425	T/C	0.90 (0.75–1.07)	0.70 (0.52–0.96)	0.86 (0.75–0.98)	
<i>Msc5A1</i>		rs6476643	T/G	1.07 (0.89–1.28)	0.88 (0.61–1.25)	1.00 (0.87–1.15)	
<i>TLR1/TLR6</i>		rs7696175	T/C	0.96 (0.78–1.16)	1.20 (0.93–1.53)	1.08 (0.95–1.22)	
<i>MAP3K1</i>		rs889312	C/A	1.06 (0.88–1.27)	1.57 (1.16–2.13)	1.18 (1.03–1.34)	
<i>RELN</i>		rs17157903	T/C	0.86 (0.70–1.06)	0.79 (0.46–1.38)	0.87 (0.73–1.03)	
Chromosome 2p		rs4666451	A/G	1.05 (0.87–1.27)	0.96 (0.74–1.24)	0.99 (0.88–1.13)	
Chromosome 2q35		rs13387042	G/A	0.92 (0.75–1.12)	0.88 (0.68–1.12)	0.93 (0.83–1.06)	
Chromosome 5p		rs981782	G/T	0.96 (0.78–1.18)	1.05 (0.82–1.35)	1.02 (0.90–1.15)	
Chromosome 5q		rs30099	T/C	1.07 (0.85–1.34)	0.79 (0.33–1.89)	1.03 (0.84–1.26)	
Chromosome 8q		rs13281615	G/A	1.17 (0.97–1.43)	1.23 (0.95–1.58)	1.12 (0.99–1.26)	

^aAdjusted for age, European ancestry and offset term.

compared with AA women, and 8/10 loci were statistically significant in AA and 9/10 among White women. Among AA women, positive per-allele associations were observed for *H19* (rs2107425) and *TLR1/TLR6* (rs7696175), and significant but imprecise associations were observed for homozygotes carrying the less common allele for *LSP1* (rs3817198) and *RELN* (rs17157903). For White women, positive per-allele associations were observed for *TNRC9* (rs3803662)

and *MAP3K1* (rs889312), whereas inverse per-allele association was found for *TNRC9/TOX3* (rs12443621) and *H19* (rs2107425). Single SNP results adjusted for age, race, European ancestry and study offset term and stratified by age at diagnosis (<50, ≥50 years) are presented in Table II (see supplementary Table 4, available at *Carcinogenesis* Online, for genotype frequencies stratified by age at diagnosis). In general, per-allele ORs for *FGFR2* SNPs were stronger for

Table II. Single SNP analysis stratified by age at diagnosis

Dataset	Gene or locus	rs no.	Alleles (test/reference)	Codominant model (heterozygous), OR (95% CI) ^a	Codominant model (homozygous), OR (95% CI) ^a	Additive model (per-allele), OR (95% CI) ^a	
Age ≥ 50	<i>CASP8</i>	rs1045485	C/G	0.80 (0.62–1.03)	0.86 (0.35–2.13)	0.83 (0.66–1.03)	
		rs17468277	T/C	0.79 (0.62–1.02)	0.82 (0.32–2.06)	0.82 (0.65–1.02)	
	<i>FGFR2</i>	rs1896395	A/C	0.97 (0.71–1.34)	1.76 (0.81–3.83)	1.10 (0.85–1.43)	
		rs11200014	A/G	1.43 (1.16–1.77)	1.65 (1.20–2.26)	1.32 (1.14–1.53)	
		rs2981579	T/C	1.66 (1.31–2.10)	1.90 (1.44–2.49)	1.38 (1.20–1.58)	
		rs1219648	G/A	1.46 (1.18–1.80)	1.78 (1.35–2.35)	1.35 (1.18–1.55)	
		rs2912774	A/C	1.43 (1.14–1.79)	1.95 (1.49–2.56)	1.40 (1.22–1.60)	
		rs2936870	T/C	1.45 (1.15–1.82)	1.94 (1.48–2.55)	1.40 (1.22–1.60)	
		rs2981582	T/C	1.52 (1.23–1.89)	1.77 (1.35–2.34)	1.36 (1.18–1.55)	
		rs2420946	T/C	1.47 (1.18–1.83)	1.81 (1.38–2.37)	1.35 (1.18–1.55)	
		rs2162540	G/A	1.49 (1.20–1.86)	1.89 (1.44–2.48)	1.39 (1.21–1.59)	
		rs3135718	G/A	1.53 (1.22–1.92)	1.96 (1.50–2.56)	1.41 (1.23–1.61)	
		rs8049149	C/T	0.81 (0.37–1.79)		0.98 (0.47–2.06)	
		<i>TNRC9 3' UTR</i>	rs16951186	C/T	0.90 (0.66–1.23)	1.11 (0.50–2.47)	0.93 (0.72–1.21)
		<i>TNRC9 5' UTR</i>	rs8051542	C/T	0.80 (0.65–0.99)	1.58 (1.18–2.11)	1.14 (0.99–1.31)
	<i>TNRC9/TOX3</i>	rs12443621	G/A	0.98 (0.77–1.23)	1.00 (0.77–1.30)	1.02 (0.89–1.16)	
	<i>TNRC9/TOX3</i>	rs3803662	C/T	1.05 (0.85–1.30)	1.06 (0.78–1.42)	1.03 (0.89–1.19)	
	<i>TNRC9 5' UTR</i>	rs9940048	G/A	1.18 (0.97–1.43)	1.08 (0.73–1.60)	1.10 (0.94–1.28)	
	<i>LSP1</i>	rs3817198	C/T	0.98 (0.80–1.21)	1.00 (0.68–1.46)	1.00 (0.85–1.16)	
	<i>H19</i>	rs2107425	T/C	1.05 (0.85–1.29)	0.82 (0.61–1.11)	0.95 (0.82–1.09)	
	<i>Msc5A1</i>	rs6476643	T/G	1.23 (1.01–1.50)	0.93 (0.61–1.40)	1.07 (0.91–1.25)	
	<i>TLR1/TLR6</i>	rs7696175	T/C	0.95 (0.75–1.20)	0.88 (0.63–1.23)	0.96 (0.82–1.12)	
	<i>MAP3K1</i>	rs889312	C/A	1.13 (0.92–1.38)	1.05 (0.75–1.47)	1.06 (0.92–1.22)	
	<i>RELN</i>	rs17157903	T/C	0.87 (0.69–1.10)	1.40 (0.70–2.78)	0.94 (0.77–1.15)	
	Chromosome 2p	rs4666451	A/G	0.98 (0.80–1.21)	0.81 (0.59–1.12)	0.93 (0.80–1.07)	
	Chromosome 2q35	rs13387042	G/A	0.97 (0.79–1.02)	0.85 (0.63–1.14)	0.93 (0.81–1.07)	
	Chromosome 5p	rs981782	G/T	0.96 (0.76–1.22)	1.20 (0.86–1.67)	1.07 (0.91–1.25)	
	Chromosome 5q	rs30099	T/C	1.27 (1.00–1.62)	0.99 (0.48–2.05)	1.18 (0.96–1.46)	
	Chromosome 8q	rs13281615	G/A	1.29 (1.04–1.59)	1.28 (0.97–1.70)	1.15 (1.01–1.32)	
	Age < 50	<i>CASP8</i>	rs1045485	C/G	1.10 (0.85–1.42)	0.64 (0.27–1.49)	1.01 (0.81–1.27)
			rs17468277	T/C	1.10 (0.85–1.43)	0.64 (0.27–1.49)	1.01 (0.81–1.27)
		<i>FGFR2</i>	rs1896395	A/C	0.81 (0.58–1.13)	2.07 (0.84–5.11)	0.99 (0.74–1.31)
			rs11200014	A/G	0.96 (0.77–1.19)	1.35 (0.98–1.86)	1.10 (0.95–1.28)
			rs2981579	T/C	1.16 (0.91–1.47)	1.39 (1.05–1.84)	1.18 (1.03–1.36)
			rs1219648	G/A	1.10 (0.88–1.38)	1.37 (1.03–1.84)	1.16 (1.01–1.34)
			rs2912774	A/C	1.16 (0.92–1.46)	1.40 (1.06–1.86)	1.18 (1.03–1.36)
			rs2936870	T/C	1.13 (0.89–1.42)	1.42 (1.07–1.89)	1.19 (1.03–1.37)
			rs2981582	T/C	1.21 (0.97–1.51)	1.30 (0.98–1.74)	1.15 (1.00–1.33)
			rs2420946	T/C	1.08 (0.86–1.36)	1.32 (0.99–1.75)	1.14 (0.99–1.32)
			rs2162540	G/A	1.14 (0.91–1.43)	1.34 (1.01–1.78)	1.16 (1.00–1.33)
rs3135718			G/A	1.12 (0.89–1.41)	1.39 (1.05–1.84)	1.18 (1.02–1.35)	
rs8049149			C/T	1.02 (0.51–2.04)		0.87 (0.45–1.65)	
<i>TNRC9 3' UTR</i>			rs16951186	C/T	0.81 (0.58–1.13)	1.21 (0.51–2.90)	0.94 (0.72–1.24)
<i>TNRC9 5' UTR</i>			rs8051542	C/T	1.16 (0.93–1.44)	1.14 (0.84–1.53)	1.08 (0.93–1.24)
<i>TNRC9/TOX3</i>		rs12443621	G/A	0.83 (0.65–1.05)	0.85 (0.64–1.12)	0.92 (0.80–1.06)	
<i>TNRC9 5' UTR</i>		rs3803662	C/T	1.26 (1.01–1.57)	1.41 (1.05–1.91)	1.20 (1.04–1.39)	
<i>TNRC9</i>		rs9940048	G/A	1.07 (0.87–1.32)	1.04 (0.70–1.54)	1.02 (0.87–1.19)	
<i>LSP1</i>		rs3817198	C/T	1.17 (0.95–1.45)	1.27 (0.86–1.88)	1.14 (0.98–1.33)	
<i>H19</i>		rs2107425	T/C	0.94 (0.75–1.17)	1.10 (0.81–1.51)	1.02 (0.88–1.18)	
<i>Msc5A1</i>		rs6476643	T/G	0.99 (0.80–1.22)	1.06 (0.67–1.66)	1.02 (0.86–1.20)	
<i>TLR1/TLR6</i>		rs7696175	T/C	1.12 (0.88–1.44)	1.84 (1.32–2.58)	1.28 (1.09–1.49)	
<i>MAP3K1</i>		rs889312	C/A	0.94 (0.76–1.15)	1.50 (1.05–2.15)	1.12 (0.94–1.30)	
<i>RELN</i>		rs17157903	T/C	0.84 (0.66–1.08)	1.15 (0.58–2.30)	0.91 (0.74–1.11)	
Chromosome 2p		rs4666451	A/G	1.08 (0.87–1.34)	0.89 (0.64–1.24)	0.99 (0.85–1.14)	
Chromosome 2q35		rs13387042	G/A	0.94 (0.75–1.17)	0.91 (0.67–1.23)	0.95 (0.82–1.10)	
Chromosome 5p		rs981782	G/T	0.88 (0.68–1.12)	0.91 (0.65–1.28)	0.94 (0.80–1.11)	
Chromosome 5q		rs30099	T/C	1.06 (0.83–1.35)	1.05 (0.40–2.75)	1.05 (0.84–1.31)	
Chromosome 8q		rs13281615	G/A	1.06 (0.84–1.32)	0.94 (0.74–1.25)	1.00 (0.87–1.15)	

^aAdjusted for age, race, European ancestry and offset term.

older compared with younger women, with very few other single SNPs showing a positive association with breast cancer risk in either age group.

ORs for haplotypes adjusted for age, European ancestry and study offset term stratified by self-reported race are presented in Table III. Among AA women, only the *FGFR2* GTGT haplotype (based upon rs11200014, rs2981579, rs1219648 and rs2420946) was positively

associated with breast cancer (OR = 1.27, 95% CI 1.04–1.56). The *FGFR2* GTGT haplotype was present in 23.1% of AAs (24.6% of AA cases and 21.4% of AA controls) but only 0.6% of Whites (0.8% of White cases and 0.6% of White controls). Among White women, only the *FGFR2* ATGT haplotype showed a positive association with breast cancer (OR = 1.30, 95% CI 1.15–1.46) and was present in 39.4% of participants (42.1% of White cases and 39.2% of White

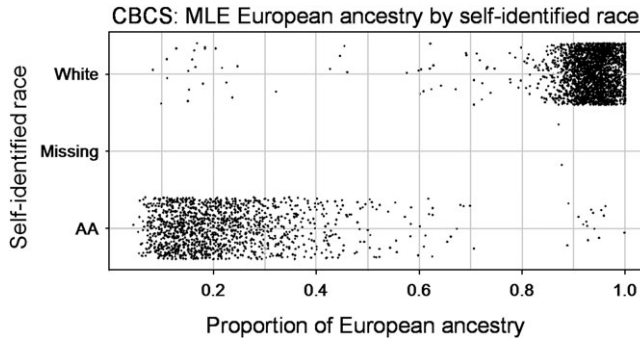


Fig. 1. Distribution of individual maximum likelihood estimates (MLEs) of European ancestry proportions by self-reported race; AA.

Table III. Haplotype results stratified by self-reported race

Dataset	Gene	Haplotype	Haplotype frequency (%)	OR (95% CI) ^a	P-value (t-statistic) ^a
AA	CASP8 ^b	GC	94.1	Reference	—
		CT	5.3	0.91 (0.65–1.27)	0.58
	FGFR2 ^c	GCAC	35.4	Reference	—
		ATGT	15.8	1.11 (0.88–1.39)	0.38
		GTGT	23.1	1.27 (1.04–1.56)	2.05E-02
		GTAC	7.4	1.08 (0.78–1.49)	0.66
		GTAT	8.9	1.25 (0.93–1.67)	0.14
		ATAC	3.2	1.11 (0.68–1.80)	0.68
	TNRC9 ^d	GCAT	2.2	0.77 (0.43–1.38)	0.38
		GCGT	2.4	1.10 (0.65–1.89)	0.72
		AC	39.5	Reference	—
		AT	12.0	1.01 (0.77–1.31)	0.96
		GC	7.9	0.83 (0.61–1.13)	0.24
		GT	40.4	0.91 (0.77–1.07)	0.25
White	CASP8 ^b	GC	86.9	Reference	—
		CT	13.1	0.90 (0.76–1.07)	0.24
	FGFR2 ^c	GCAC	55.7	Reference	—
		ATGT	39.4	1.30 (1.15–1.46)	1.91E-05
		ATAC	2.0	1.32 (0.87–2.01)	0.19
		ATGC	1.1	1.00 (0.57–1.75)	1.00
	TNRC9 ^d	AC	49.6	Reference	—
		AT	1.2	0.66 (0.35–1.22)	0.18
		GC	20.7	1.01 (0.87–1.18)	0.87
		GT	28.5	1.25 (1.09–1.42)	1.37E-03

^aAdjusted for age, European ancestry and offset term.

^bCASP8 SNPs are rs1045485 and rs17468277 in relative nucleotide location order.

^cFGFR2 SNPs are rs11200014, rs2981579, rs1219648 and rs2420946 in relative nucleotide location order.

^dTNRC9 SNPs are rs12443621 and rs3803662 in relative nucleotide location order.

controls). The *FGFR2* ATGT haplotype was present in 15.8% of AA participants (15.8% of AA cases and 15.9% of AA controls). Additionally, among White women, the *TNRC9* GT haplotype (based upon rs12443621 and rs3803662) was associated with increased odds of breast cancer (OR = 1.25, 95% CI 1.09–1.42) and was common in all study participants (28.5% of Whites and 40.4% of AAs).

Haplotype analyses adjusted for age, race, European ancestry and study offset term and stratified by age at diagnosis are presented in Table IV. Among women aged ≥50, the *CASP8* CT haplotype (based upon rs1045485 and rs17468277) was associated with decreased odds of breast cancer (OR = 0.79, 95% CI 0.64–0.98). The *TNRC9* GT haplotype was associated with increased odds of breast cancer among younger (OR = 1.17, 95% CI 1.01–1.36) but not older women. The *FGFR2* ATGT (OR = 1.37, 95% CI 1.18–1.58) and GTGT (OR = 1.39, 95%

Table IV. Haplotype results stratified by age at diagnosis

Dataset	Gene	Haplotype	Haplotype frequency (%)	OR (95% CI) ^a	P-value (t-statistic) ^a
Age ≥ 50	CASP8 ^b	GC	89.7	Reference	—
		CT	10.0	0.79 (0.64–0.98)	0.03
	FGFR2 ^c	GCAC	48.3	Reference	—
		ATGT	30.2	1.37 (1.18–1.58)	2.51E-05
		GTGT	9.0	1.39 (1.08–1.80)	0.01
		GTAC	2.9	1.25 (0.83–1.87)	0.29
		GTAT	3.9	1.29 (0.90–1.86)	0.16
		ATAC	2.2	1.55 (0.98–2.44)	0.06
	TNRC9 ^d	GCAT	1.0	0.85 (0.43–1.68)	0.63
		ATGC	1.3	1.35 (0.75–2.45)	0.32
		AC	45.9	Reference	—
		AT	5.6	0.96 (0.70–1.33)	0.82
		GC	16.6	0.95 (0.79–1.15)	0.61
		GT	31.8	1.03 (0.89–1.19)	0.69
Age < 50	CASP8 ^b	GC	89.5	Reference	—
		CT	10.3	1.04 (0.84–1.29)	0.74
	FGFR2 ^c	GCAC	48.0	Reference	—
		ATGT	31.2	1.14 (0.98–1.32)	0.10
		GTGT	8.8	1.35 (1.02–1.78)	0.03
		GTAC	3.1	0.99 (0.64–1.53)	0.95
		GTAT	3.0	1.30 (0.85–1.99)	0.22
		ATAC	2.4	1.01 (0.64–1.59)	0.98
	TNRC9 ^d	GCAT	1.1	0.61 (0.31–1.22)	0.16
		GCGT	1.2	0.80 (0.43–1.48)	0.47
		AC	44.9	Reference	—
		AT	5.5	1.10 (0.79–1.54)	0.58
		GC	16.0	0.97 (0.79–1.19)	0.76
		GT	33.5	1.17 (1.01–1.36)	0.04

^aAdjusted for age, race, European ancestry and offset term.

^bCASP8 SNPs are rs1045485 and rs17468277 in relative nucleotide location order.

^cFGFR2 SNPs are rs11200014, rs2981579, rs1219648 and rs2420946 in relative nucleotide location order.

^dTNRC9 SNPs are rs12443621 and rs3803662 in relative nucleotide location order.

CI 1.08–1.80) haplotypes were positively associated with breast cancer in older women. Among younger women, only the *FGFR2* GTGT haplotype (OR = 1.35, 95% CI 1.02–1.78) was associated with breast cancer. In older women, 30.2% had the ATGT haplotype (32.5% of cases and 33.8% of controls) and 9.0% had the GTGT haplotype (9.8% of cases and 5.2% of controls). In younger women, 31.2% had the ATGT haplotype (32.3% of cases and 34.8% of controls) and 8.8% had the GTGT haplotype (9.4% of cases and 5.0% of controls). In AA women, the GTGT haplotype was more strongly associated with breast cancer risk in older women (OR = 1.34, 95% CI 1.01–1.76) compared with younger women (OR = 1.20, 95% CI 0.89–1.62). In White women, the ATGT haplotype was more strongly associated with breast cancer risk in older women (OR = 1.42, 95% CI 1.23–1.64) compared with younger women (OR = 1.19, 95% CI 1.02–1.39).

Discussion

We investigated the role of *FGFR2* and other loci identified in breast cancer GWAS using a population-based case–control study that included large proportions of AA and younger study participants. Our results replicate previous findings for *FGFR2* and contribute new information for AA and younger women (diagnosed at age <50).

Our results for the *FGFR2* ATGT haplotype in the overall study population (OR = 1.25, 95% CI 1.15–1.36) and in Whites only (OR = 1.30, 95% CI 1.15–1.46) are similar to the results of Hunter *et al.* (4). Hunter *et al.* (4) used four SNPs (rs11200014, rs2981579, rs1219648 and rs2420946) to identify high- and low-risk *FGFR2* intron 2 haplotypes, given that the strongest signal in a previous

GWAS was in a LD block of 25 kb in intron 2, which is thought to mediate differential expression of *FGFR2* (3). A positive association was observed for the AAGT haplotype pooled across several studies (OR = 1.26, 95% CI 1.17–1.35) (4). Positive signals were not found elsewhere in the gene or in neighboring regions. The authors designated the rs2981579 alleles (the second SNP in the *FGFR2* haplotype) using antisense coding, so that the GGAC haplotype in Hunter *et al.* (4) corresponds to GCAC in our study and AAGT corresponds to ATGT. However, the biological/functional significance of these haplotypes as they relate to breast cancer risk is unknown. Causal alleles in *FGFR2* remain to be identified, but variation in rs7895676 and rs2981578 in *FGFR2* intron 2 was found to be associated with increased *FGFR2* expression (26). rs2981578 is in close proximity to and in strong LD with rs2981579, which is included in our four SNP *FGFR2* intron 2 haplotype, and this was also reported by Hunter *et al.* (4).

In addition, we detected a *FGFR2* intron 2 haplotype in AAs (GTGT) that was associated with increased risk of breast cancer (OR = 1.27, 95% CI 1.04–1.56). Three previous studies have addressed the role of *FGFR2* in AAs (13–15). Udler *et al.* (13) employed dense SNP mapping and targeted resequencing to further narrow and characterize the region of interest in *FGFR2* intron 2 in AAs. The authors employed a regression approach to calculate likelihood ratios for observed as well as imputed SNP genotypes and detected the strongest signal for the SNP rs2981578 among 1253 AA breast cancer cases and 1245 AA controls from four US epidemiologic studies. In the CBCS, the breast cancer-associated haplotypes in *FGFR2* intron 2 (ATGT, GTGT and GTAT) shared the T allele at rs2981579, providing further evidence for a potential causal locus in this region. Rebbeck *et al.* (14) found no association for two SNPs in *FGFR2* intron 2 (rs1219648 and rs2981582) in 157 AA cases and 427 controls. Zheng *et al.* (15) examined the role of 6 SNPs in *FGFR2* in 810 AA cases and 1784 controls and observed a positive association for rs1219648, which is also included in our four SNP *FGFR2* intron 2 haplotype.

In the two other studies of AAs, Rebbeck *et al.* (14) reported a positive association for *MAP3K1* (rs889312), whereas Zheng *et al.* (15) reported a positive association for chromosome 2q35 (rs13387042). While our study evaluated these SNPs in the AA women in CBCS, we did not replicate these previous results at these loci but instead found new associations at *H19* (rs2107425) and *TLR1/TLR6* (rs7696175). Additional studies of AAs are needed in order to increase sample size and achieve sufficient power to examine the role of GWAS-identified loci among AAs. The combined sample size of the four published studies of *FGFR2* and other GWAS loci in AAs [CBCS, Udler *et al.* (13), Rebbeck *et al.* (14), Zheng *et al.* (15)] is 2962 cases and 4114 controls compared with much larger studies of Europeans and Whites (i.e. women of European descent) [e.g. >25 000 cases and >25 000 controls in a validation study by Gaudet *et al.* (27)].

We also evaluated the association between these GWAS SNPs and risk of breast cancer for young women (diagnosed age <50). While both the *FGFR2* ATGT and GTGT haplotypes were associated with breast cancer in older women (ATGT: OR = 1.37, 95% CI 1.18–1.58; GTGT: OR = 1.39, 95% CI 1.08–1.80), only the GTGT haplotype was associated with breast cancer among younger women (OR = 1.35, 95% CI 1.02–1.78). In addition, we found that the *CASP8* CT haplotype was associated with decreased odds of breast cancer (OR = 0.79, 95% CI 0.64–0.98) in older women only, whereas the *TNRC9* GT haplotype was associated with breast cancer among younger women only (OR = 1.17, 95% CI 1.01–1.36). The same haplotypes were associated with disease risk when age-specific results were further stratified by race (data not shown). While our results are suggestive, none of the GWAS studies presented results for younger women to which we can compare our results.

Our analysis was based on previously published GWAS papers that included mainly postmenopausal European or White women. We did not attempt to identify tagSNPs for *FGFR2* or other loci among AAs, as our goal was limited to replication of GWAS hits

previously reported for Whites. There may be other relevant SNPs within these loci that are associated with breast cancer, particularly for AA and younger women, that were not examined. In addition, our sample size was relatively small for AAs. Future research will require pooling of data from AAs breast cancer studies, particularly for fine mapping of GWAS hits and to study gene–gene and gene–environment interaction. Strengths of our analysis include the population-based study design of the CBCS, a study that includes a large group of AA and a large group of women <50, which affords the necessary statistical power to examine genetic associations within these subgroups. We utilized individual European ancestry estimates derived from 144 AIMs, chosen to robustly distinguish between European and African ancestry, in order to control for potential residual population stratification. Failure to adjust for population admixture can result in false-positive and false-negative associations (as reviewed in ref. 20). In our study, the SNP effect estimates changed by up to 43% and haplotype effect estimates changed by up to 7% with the inclusion of ancestry information (data not shown). The strongest effects of ancestry adjustment were seen for the following SNPs: the effect estimates for rs1896395 in *FGFR2* in Whites changed from OR = 1.077 to OR = 1.509 in the codominant heterozygous model and from OR = 1.502 to OR = 2.143 in the additive model after including the ancestry estimates. For *TNRC9* in Whites, the effect estimate for haplotype AT changed from OR = 0.611 to OR = 0.655 after adjustment for ancestry. The majority of effect estimates changed by ~0.5% due to the ancestry adjustment.

In conclusion, we demonstrated that recent GWAS hits for breast cancer in Europeans and Whites showed evidence of replication among AAs and Whites in the CBCS. We identified several new haplotypes that were associated with breast cancer in AA and younger women, particularly the *FGFR2* GTGT haplotype. Our results and those of the three previously published studies of *FGFR2* in AAs (13–15) highlight the need to conduct GWAS and GWAS validation in a variety of racial–ethnic populations as well as among younger women.

Supplementary material

Supplementary Tables 1–4 can be found at <http://carcin.oxfordjournals.org/>

Funding

Specialized Program of Research Excellence in Breast Cancer (NIH/NCI P50-CA58223 to R.C.M.); Center for Environmental Health and Susceptibility (NIEHS P30-ES10126 to R.C.M.); Lineberger Comprehensive Cancer Center Core Grant (P30-CA16086 to R.C.M.); Case Comprehensive Cancer Center Core Grant (P30-CA043703 to J.S.B.-S.).

Acknowledgements

The authors acknowledge VanAhn Tran and Chiu-Kit Tse for assistance with statistical analysis and ancestry estimations and Michael Andre, Amanda Flood, Patricia Basta and Hendrik Dejong for laboratory expertise.

Conflict of Interest Statement: None declared.

References

- Ioannidis, J.P. (2006) Common genetic variants for breast cancer: 32 largely refuted candidates and larger prospects. *J. Natl Cancer Inst.*, **98**, 1350–1353.
- Hardy, J. *et al.* (2009) Genomewide association studies and human disease. *N. Engl. J. Med.*, **360**, 1759–1768.
- Easton, D.F. *et al.* (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature*, **447**, 1087–1093.

4. Hunter,D.J. *et al.* (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat. Genet.*, **39**, 870–874.
5. Stacey,S.N. *et al.* (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat. Genet.*, **39**, 865–869.
6. Turnbull,C. *et al.* (2008) Genetic predisposition to breast cancer: past, present, and future. *Annu. Rev. Genomics Hum. Genet.*, **9**, 321–345.
7. Gold,B. *et al.* (2008) Genome-wide association study provides evidence for a breast cancer risk locus at 6q22.33. *Proc. Natl Acad. Sci. USA*, **105**, 4340–4345.
8. Zheng,W. *et al.* (2009) Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat. Genet.*, **41**, 324–328.
9. Ahmed,S. *et al.* (2009) Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat. Genet.*, **41**, 585–590.
10. Thomas,G. *et al.* (2009) A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat. Genet.*, **41**, 579–584.
11. American Cancer Society. (2009) *Breast Cancer Facts & Figures 2009–2010*. American Cancer Society, Inc, Atlanta, GA.
12. Turner,N. *et al.* (2010) Fibroblast growth factor signalling: from development to cancer. *Nat. Rev. Cancer*, **10**, 116–129.
13. Udler,M.S. *et al.* (2009) FGFR2 variants and breast cancer risk: fine-scale mapping using African American studies and analysis of chromatin conformation. *Hum. Mol. Genet.*, **18**, 1692–1703.
14. Rebbeck,T.R. *et al.* (2009) Hormone-dependent effects of FGFR2 and MAP3K1 in breast cancer susceptibility in a population-based sample of post-menopausal African-American and European-American women. *Carcinogenesis*, **30**, 269–274.
15. Zheng,W. *et al.* (2009) Evaluation of 11 breast cancer susceptibility loci in African-American women. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 2761–2764.
16. Millikan,R.C. *et al.* (1998) Cigarette smoking, N-acetyltransferases 1 and 2, and breast cancer risk. *Cancer Epidemiol. Biomarkers Prev.*, **7**, 371–378.
17. Newman,B. *et al.* (1995) The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. *Breast Cancer Res. Treat.*, **35**, 51–60.
18. Weinberg,C.R. *et al.* (1991) Randomized recruitment in case-control studies. *Am. J. Epidemiol.*, **134**, 421–432.
19. Barnholtz-Sloan,J.S. *et al.* (2005) Examining population stratification via individual ancestry estimates versus self-reported race. *Cancer Epidemiol. Biomarkers Prev.*, **14**, 1545–1551.
20. Barnholtz-Sloan,J.S. *et al.* (2008) Ancestry estimation and correction for population stratification in molecular epidemiologic association studies. *Cancer Epidemiol. Biomarkers Prev.*, **17**, 471–477.
21. Tian,C. *et al.* (2006) A genomewide single-nucleotide-polymorphism panel with high ancestry information for African American admixture mapping. *Am. J. Hum. Genet.*, **79**, 640–649.
22. Pfaff,C.L. *et al.* (2004) Information on ancestry from genetic markers. *Genet. Epidemiol.*, **26**, 305–315.
23. Purcell,S. *et al.* (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.*, **81**, 559–575.
24. SAS Institute Inc. (2007) *Statistical Analysis Software, Version 9.2*. Cary, NC.
25. Schaid,D.J. (2004) Evaluating associations of haplotypes with traits. *Genet. Epidemiol.*, **27**, 348–364.
26. Meyer,K.B. *et al.* (2008) Allele-specific up-regulation of FGFR2 increases susceptibility to breast cancer. *PLoS Biol.*, **6**, e108.
27. Gaudet,M.M. *et al.* (2009) Five polymorphisms and breast cancer risk: results from the Breast Cancer Association Consortium. *Cancer Epidemiol. Biomarkers Prev.*, **18**, 1610–1616.

Received April 16, 2010; revised May 28, 2010; accepted June 11, 2010