Published in final edited form as:
Int J Cancer. 2015 March 15; 136(6): E685–E696. doi:10.1002/ijc.29188.

An investigation of gene-environment interactions between 47 newly identified breast cancer susceptibility loci and environmental risk factors

A full list of authors and affiliations appears at the end of the article.

Abstract

A large genotyping project within the Breast Cancer Association Consortium (BCAC) recently identified 41 associations between single nucleotide polymorphisms (SNPs) and overall breast cancer (BC) risk. We investigated whether the effects of these 41 SNPs, as well as six SNPs associated with estrogen receptor (ER) negative BC risk are modified by 13 environmental risk factors for BC.

Data from 22 studies participating in BCAC were pooled, comprising up to 26,633 cases and 30,119 controls. Interactions between SNPs and environmental factors were evaluated using an empirical Bayes-type shrinkage estimator.

Six SNPs showed interactions with associated *p*-values (p_{int}) <1.1×10^{-3}. None of the observed interactions was significant after accounting for multiple testing. The Bayesian False Discovery Probability was used to rank the findings, which indicated three interactions as being noteworthy at 1% prior probability of interaction. SNP rs6828523 was associated with increased ER-negative BC risk in women ≥170cm (OR=1.22, p=0.017), but inversely associated with ER-negative BC risk in women <160cm (OR=0.83, p=0.039, p_{int}=1.9×10^{-4}). The inverse association between rs4808801 and overall BC risk was stronger for women who had had four or more pregnancies (OR=0.85, p=2.0×10^{-4}), and absent in women who had had just one (OR=0.96, p=0.19, p_{int} = 6.1×10^{-4}). SNP rs11242675 was inversely associated with overall BC risk in never/former smokers (OR=0.93, p=2.8×10^{-5}), but no association was observed in current smokers (OR=1.07, p=0.14, p_{int} = 3.4×10^{-4}).

In conclusion, recently identified breast cancer susceptibility loci are not strongly modified by established risk factors and the observed potential interactions require confirmation in independent studies.

Keywords

gene-environment interaction; breast cancer; risk factor; genetic susceptibility

Corresponding author: Anja Rudolph, PhD, Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany, Tel: +49-6221 42 3181, Fax: +49-6221 42 2203, a.rudolph@dkfz.de.

Conflict of Interest Statement
The authors declare no conflict of interest.

Introduction

Genetic and environmental factors are known to contribute to the risk of breast cancer. The biological interplay between them may lead to varying associations of the genetic factors with breast cancer risk depending on the exposure to an environmental factor. This can be assessed as departure from multiplicativity of the risk ratios of the genetic variant and the environmental factor (gene-environment (G×E) interaction). Several studies have investigated whether the relative risks associated with common genetic breast cancer susceptibility loci are modified by environmental risk factors[1–5]. In the most recent investigation of 23 single nucleotide polymorphisms (SNPs) using data from the Breast Cancer Association Consortium (BCAC), we were able to replicate a previously reported interaction between rs3817198 in *LSP1* and number of full-term pregnancies in parous women, and identify an interaction between rs17468277 in *CASP8* and varying levels of mean lifetime alcohol consumption (>20 g/day vs. ≥20 g/day)[1]. The identification of G×E interactions may improve our understanding of breast cancer aetiology by suggesting potential biological pathways involved.

A recently conducted large genotyping project (Collaborative Oncological Gene-environment Study (COGS)) identified 41 novel genetic susceptibility loci for breast cancer, explaining an additional 5% of the familial breast cancer risk[6]. The project also led to the identification of four loci associated with risk of estrogen receptor (ER) negative breast cancer[7] additional to the three previously established ER-negative breast cancer susceptibility loci[8–10]. G×E interactions with these newly identified variants have not been investigated so far.

Here, we evaluated G×E interactions on overall breast cancer risk between 47 single nucleotide polymorphisms (SNPs) and the following environmental factors: age at menarche, parity, age at first birth, breastfeeding, use of menopausal hormone therapy (MHT), body-mass index (BMI), adult height, smoking and alcohol consumption. The 47 SNPs represent 41 newly identified genetic susceptibility loci for overall breast cancer as well as 6 loci associated with risk for ER negative breast cancer (genotype data for the seventh ER-negative breast cancer SNP (rs2284378) was not available). We also assessed G×E interactions regarding risk for ER-positive and ER-negative breast cancer separately, as different pathways may be involved in the development of these subtypes. This investigation uses the largest dataset available at present, including genotype data on the newly identified breast cancer susceptibility loci and comprehensive data on environmental risk factors.

Materials and Methods

Study samples

We pooled data from 22 studies participating in BCAC (20 case-control studies, 2 cohort studies), which mainly recruited participants of European descent (Supplementary Table 1). Selected studies comprised at least 200 cases and 200 controls with genotype data and information on at least one of the environmental risk factors of interest.

We excluded participants from this analysis if they were male, were prevalent cases at recruitment (in MCCS and pKARMA), were not of European descent, or had a missing value for reference age (age at diagnosis/interview), the specific environmental variable of interest, or the related adjustment variables. Therefore, the number of participants available for analysis varied depending on the investigated environmental factor. The dataset with subjects included in at least one of the analyses comprised 31,850 cases and 34,816 controls. The largest sample was available for the G×E interaction analysis between SNPs and ever being parous, which included 26,633 cases and 30,119 controls and the smallest sample was available for the analysis involving lifetime average intake of alcohol, which included 3,811 cases and 4,053 controls.

All studies were approved by the relevant ethics committees and informed consent was obtained from all participants.

Data harmonization and variable definitions

Data from the different studies were harmonized in a multi-step process according to a common data dictionary. In both case-control and cohort studies, time-dependent variables were assessed at reference date, which was defined as the date of diagnosis for cases and the date of interview for controls: for controls and cases from the two included cohort studies, data from the baseline interview were considered, or if available, follow-up information[1]. The median time between last interview/questionnaire and diagnosis was 7.6 years in the MCCS cohort and 2.0 years in the UKBGS cohort.

Current use of any MHT was defined as use within 6 months prior to the reference date and current smoking as smoking within one year prior to the reference date. An age surrogate was used to define menopausal status. Women aged ≤ 54 years at reference date were considered to be premenopausal and women aged > 54 years postmenopausal[1]. BMI was calculated based on usual adult weight or weight one year prior to the reference date (studies ABCFS, BREOGAN, CECILE, GENICA, kConFab/AOCS, KBCP, MARIE, MCBCS, OFBCR, PBCS, SASBAC) or weight in early adulthood (age around 20 years, studies ESTHER, pKARMA, SEARCH). For the two cohort studies (MCCS, UKBGS), we used the weight reported at baseline interview.

Genetic information

The genotyping data used in this study for all studies except BREOGAN were generated as part of the COGS project (www.nature.com/icogs). Participants from studies in BCAC were genotyped using an Illumina iSelect array (iCOGS)[6]. Approximately 61,000 of the 211,155 SNPs included on the iCOGS array were selected to follow-up on a meta-analysis of nine breast cancer genome-wide association studies (GWAS). A subsequent association study in 45,290 cases and 41,880 controls led to the identification of 41 SNPs associated with overall breast cancer risk[6]. Similarly, three GWAS studies were meta-analysed to identify loci associated with ER-negative breast cancer risk, and 13,276 SNPs were selected to be genotyped on the iCOGS array for the replication stage, comprising 6,514 ER-negative breast cancer cases and 41,455 controls. Four new loci showing a specific association with ER-negative breast cancer were detected[7].

In the current study, we use the original quality-controlled genotype data that was used for identification of the 41 SNPs and the four SNPs associated with ER-negative breast cancer risk, for all studies except for subjects in BREOGAN. Two of the three previously identified SNPs specifically associated with ER-negative breast cancer risk were also genotyped with the iCOGS array (rs10069690 in *TERT* on chromosome 5, rs8170 in *BABAM1* on chromosome 19); genotype data for rs2284378 on chromosome 20q11 were not available.

Study participants were excluded from analyses if the overall genotyping call rate was below 95% over the whole iCOGS array or if heterozygosity deviated significantly from that expected in the general population (either lower or higher, $p < 10^{-6}$).

Genotyping of the 47 SNPs for BREOGAN was performed at the CeGen-ISCIII (Spanish National Genotyping Center), using Sequenom MassARRAY Genotyping system (technology iPLEX GOLD) and following the manufacturer's instructions. DNA was dispensed in 384 well plates by a Tecan Freedom Evo robot, each plate included case and control samples, a trio of Coriell samples: Na10830, Na10831 and Na12147, and negative controls (minimum 6 per plate). We included >5% concordant duplicates. The laboratory was equipped with Life Technologies GeneAmp 9700 dual cyclers, a RS1000 Nanodispenser and a MA4 mass spectrometer. Data analysis was done using the software Typer analyzer v4.0.20. The SNPs were analyzed in 4 assays (Assay Design v4 software). The genotyping data were quality checked using the same criteria as for iCOGS[6].

To evaluate potential functional implications of selected SNPs and SNPs in high linkage disequilibrium (LD) with selected SNPs we used HaploReg v2[11] and the UCSC genome browser[12].

Statistical analysis

We employed an efficient empirical Bayes procedure to calculate the interaction log odds ratio that corresponds to a weighted average of the case-only and case-control estimators. In this way, the method makes use of the greater precision of the case-only estimator by simultaneously reducing the chance of generating biased estimates due to violations of the assumption of gene-environment independence in controls[13]. The method is implemented in the R package "CGEN", version 2.2, which was used within R 2.15.2.

In total, 13 variables representing the environmental risk factors of interest were used in G×E analyses (Supplementary Figure 1). The variables were: age at menarche (per 2 years), ever parous (no vs. yes), number of full-term pregnancies (among parous, 1, 2, 3, ≥4 pregnancies), ever breastfed (yes vs. no), age at first full-term pregnancy (per 5 years), adult BMI in premenopausal women (per 5 kg/m^2), adult BMI in postmenopausal women (per 5 kg/m^2), adult height (per 5 cm), current use of combined estrogen-progesterone therapy (no vs. yes), current use of estrogen-only therapy (no vs. yes), lifetime average intake of alcohol (per 10 g/day), current smoking (no vs. yes), smoking amount (per 10 pack-years). The variables age at menarche, number of full-term pregnancies, age at first full-term pregnancy, adult BMI, adult height, lifetime intake of alcohol, and smoking amount were entered into models as linear continuous variables. For SNPs, we included the number of minor alleles

(0-1-2) as a continuous variable. Subjects with missing data for a particular SNP or an environmental factor were excluded from the respective analysis.

The models were adjusted for study, reference age and seven principal components to account for population substructure. Additionally, to account for potential differential main effects of environmental variables by study design, we included in all models an interaction term between the environmental variable of interest and an indicator variable for study design (non-population-based vs. population-based).

MHT was classified into estrogen-only therapy and estrogen-progesterone therapy. Models used to assess associations with current use of the type of MHT of interest were further adjusted for former use of any MHT and use of MHT preparations other than the one of interest, and the analysis was restricted to postmenopausal women. Also, when assessing G×E interactions with adult BMI in postmenopausal women, the study sample was restricted to never or former users of any MHT.

All analyses were conducted with overall breast cancer risk as the outcome, as well as with ER-negative and ER-positive breast cancer risk. Heterogeneity between risk associations for ER-negative and ER-positive breast cancer was evaluated using case-case analysis with ER-status as the dependent variable and the SNP, the environmental variable, the multiplicative interaction term, ancestry informative principal components and study as independent variables. The association between SNP and breast cancer risk in strata defined by categories of the environmental risk factor was evaluated using logistic regression. Stratified analyses were conducted using SAS 9.2.

To evaluate between-study heterogeneity in G×E interaction OR estimates, we calculated these by study and performed Cochrane's *Q*-test and calculated the I^2 index, using the R package "meta" (version 2.2).

We selected G×E interactions showing *p*-values for interaction $<1.1\times10^{-3}$ for overall breast cancer (all subtypes combined) or a subtype of breast cancer, but in the latter case, only if significant subtype heterogeneity (*p*-values for heterogeneity between ER-positive and ER-negative disease <0.05) was also observed. The *p*-value threshold for selection was derived by dividing the conventional p-value threshold of 0.05 by the number of SNPs investigated. To account for chance findings due to multiple hypothesis testing, we applied the Bayesian False Discovery Probability (BFDP)[14] to assess noteworthiness of selected G×E interactions in terms of generating new hypotheses. We assumed a four-fold cost of a false non-discovery compared to the cost of a false discovery, considering interactions with a BFDP of less than 80% as being noteworthy, as suggested by Wakefield et al. [14]. The OR corresponding to the 97.5% point of the prior was 1.50 for positive G×E interactions and 0.66 for negative G×E interactions, i.e. we assumed that the prior probability of observing an OR for interaction larger than 1.5 or smaller than 0.66 was 5%. We calculated the BFDP for each selected interaction assuming six different prior probabilities for true interaction (20%, 10%, 5%, 1%, 0.1% and 0.01%).

Results

A brief description of the BCAC studies included in this analysis of G×E interactions is provided in Supplementary Table 1. The number of included cases and controls as well as the mean reference age for each study is shown in Table 1. Overall, the mean age was 56.7 years for cases and 55.6 years for controls. Further descriptions of the environmental variables are displayed in Supplementary Table 2.

The associations between SNPs and breast cancer risk were very similar in the study sample used for G×E interaction analysis (N = 66,666) to those reported by Michailidou et al. (sample size N = 87,170)[6] (Supplementary Table 3). The largest difference was observed for rs11814448, which was previously reported to be associated with overall breast cancer with an OR of 1.26 (95% CI 1.18 – 1.35) and showed a slightly attenuated effect size in the G×E dataset (odds ratio (OR) = 1.21, 95% confidence interval (CI) 1.13 – 1.31). Supplementary Table 3 shows further information for each SNP such as the minor allele frequency and SNP location.

Forest plots for meta-analyses of the associations between the 13 environmental risk factors and breast cancer risk by study can be found in Supplementary Figure 1. The risk factor associations based on the population-based studies were consistent with previous reports. Age at menarche, ever been parous, and number of full-term pregnancies among parous women were significantly associated with a decreased breast cancer risk. Significant associations with increased breast cancer risk were observed for breastfeeding (among parous women, no vs. yes), age at first full-term pregnancy, BMI in postmenopausal women not currently using MHT, body height, current use of postmenopausal combined estrogen-progesterone therapy, and average lifetime intake of alcohol. No significant associations were found between breast cancer risk and BMI in premenopausal women, current use of estrogen-only therapy, current smoking and smoking amount (pack-years).

We identified six G×E interactions with *p*-values for interaction (p_{int}) $<1.1\times10^{-3}$ (Figure 1). Estimates for each investigated G×E interaction and ORs for association between SNP and breast cancer stratified by categories of the environmental factors are presented in Supplementary Table 4. Estimates from empirical Bayes and case-control analysis (data not shown) were very similar, but *p*-values for interaction from empirical Bayes analysis were usually more extreme, possibly reflecting a small gain in power. Three of the six interactions were considered noteworthy according to a BFDP <80% at a 1% prior probability of interaction (Table 2). However, none of the observed interactions was noteworthy by this criterion assuming more conservative prior probabilities for interaction <1%.

The interaction with the lowest BFDP (BFDP = 36.0% at 1% prior probability of interaction) was observed regarding ER-negative breast cancer risk, between the SNP rs6828523 located in an intron of *ADAM29* and adult height (ER-negative OR for interaction (OR_{int}) = 1.14, 95% CI 1.06 – 1.22, $p_{int} = 1.9\times10^{-4}$). The interaction was not observed for ER-positive breast cancer risk (OR_{int} = 1.00, 95% CI 0.96 – 1.03, $p_{int} = 9.0\times10^{-1}$, *p*-value for heterogeneity by ER status (p_{het}) = 0.003). SNP rs6828523 was associated with increased risk for ER-negative breast cancer in women of 170cm height or taller (ER-

negative OR = 1.22, 95% CI 1.04 – 1.44, $p = 0.017$), but showed an inverse association in women shorter than 160cm (ER-negative OR = 0.83, 95% CI 0.70 – 0.99, $p = 0.039$) (Figure 1A). The five additional G×E interactions are reported below, ordered by their corresponding BFDP, as reported in Table 2.

Regarding overall breast cancer risk, an interaction between the number of full-term pregnancies and rs4808801 was observed ($OR_{int} = 0.96$, 95% CI 0.94 – 0.98, $p_{int} = 6.1\times10^{-4}$, BFDP = 51.6% at 1% prior probability of interaction). The interaction did not differ by ER status ($p_{het} = 0.40$) (Supplementary Table 5). The SNP is located in an intron of *ELL* on chromosome 19. The association between breast cancer risk and rs4808801 was stronger in women with four or more full-term pregnancies (OR = 0.85, 95% CI 0.77 – 0.93, $p = 2.0\times10^{-4}$), and weaker in women with one full-term pregnancy (OR = 0.96, 95% CI 0.90 – 1.02, $p = 0.19$) (Figure 1B).

Another interaction on overall breast cancer risk was found between current smoking and rs11242675, located on chromosome 6 near *FOXQ1* ($OR_{int} = 1.13$, 95% CI 1.06 – 1.21, $p_{int} = 3.4\times10^{-4}$, BFDP = 60.5% at 1% prior probability of interaction). Again, the interaction was not substantially different for ER-negative and ER-positive breast cancer ($p_{het} = 0.82$) (Supplementary Table 5). As shown in Figure 1C, rs11242675 was associated with a decreased breast cancer risk in women who did not smoke at reference time (OR = 0.93, 95% CI 0.89 – 0.96, $p = 2.8\times10^{-5}$), but this association was not observed in women who smoked at reference time (OR = 1.07, 95% CI 0.98 – 1.16, $p = 0.14$).

The three remaining interactions of the six G×E interactions in total with $p_{int} < 1.1\times10^{-3}$ could not be considered noteworthy according to their BFDP estimated using a prior probability of interaction of 1% or lower (Table 2). One of the interactions was observed for ER-positive breast cancer risk, between rs16857609 and adult height (ER-positive $OR_{int} = 0.95$, 95% CI 0.93 – 0.98, $p_{int} = 1.7\times10^{-4}$, $p_{het} = 0.018$; Supplementary Table 5). The variant rs16857609 is located in an intron of *DIRC3* on chromosome 2. In the stratified analysis, rs16857609 was associated with an increased risk of estrogen receptor positive breast cancer in women shorter than 160cm (ER-positive OR = 1.15, 95% CI 1.07 – 1.23, $p = 2.0\times10^{-4}$), whereas it was not associated with breast cancer risk in women of 170cm height or taller (ER-positive OR = 0.97, 95% CI 0.90 – 1.04, $p = 0.40$) (Figure 1D).

Two further G×E interactions were observed specifically for ER-negative breast cancer risk, one between rs12422552 located on chromosome 12 and adult height (ER-negative $OR_{int} = 1.09$, 95% CI 1.04 – 1.15, $p_{int} = 7.4\times10^{-4}$, $p_{het} = 0.006$, Supplementary Table 5). The minor allele of rs12422552 was associated with risk for ER-negative breast cancer in women of 170cm height or taller (ER-negative OR = 1.18, 95% CI 1.04 – 1.34, $p = 0.011$), but not in women shorter than 160cm (ER-negative OR = 0.92, 95% CI 0.81 – 1.04, $p = 0.16$) (Figure 4E). The other interaction specific for ER-negative breast cancer risk was between rs941764 located in an intron of *CCDC88C* on chromosome 14 and alcohol consumption (ER-negative $OR_{int} = 0.53$, 95% CI 0.36 – 0.76, $p_{int} = 6.8\times10^{-4}$, $p_{het} = 0.042$, Figure 4F, Supplementary Table 5). As shown in Figure 4F, rs941764 was inversely associated with risk of ER-negative breast cancer risk in women having an lifetime average consumption of at least 20 g alcohol per day (ER-negative OR = 0.61, 95% CI 0.38 – 0.97, $p = 0.037$), while

this association was not present in women with a lower lifetime average consumption of alcohol (ER-negative OR = 0.96, 95% CI 0.84 – 1.10, $p = 0.59$).

There was no significant heterogeneity between study-wise estimates for G×E interactions: p-values from Q-test ranged from 0.36 to 0.78 (Supplementary Figure 2).

Discussion

The present study identified six G×E interactions with $p_{int} < 1.1 \times 10^{-3}$, two regarding risk for overall breast cancer, one regarding risk for ER-positive breast cancer and three regarding risk for ER-negative breast cancer. After calculating the BFDP, none of the six interactions could be considered as being noteworthy at prior probabilities for interaction smaller than one percent although three G×E interactions were considered noteworthy at 1% prior probability of interaction. Our results do not suggest that the relative risks associated with 47 recently identified breast cancer susceptibility loci are strongly modified by environmental risk factors for breast cancer.

For some the effect modifications assessed, our findings are based on the largest available dataset at present. The number of studies with available data was relatively small for other environmental risk factors such as alcohol consumption and use of MHT. Power was also likely diminished due to the fact that we studied mostly tag-SNPs, rather the true genetic variants affecting breast cancer risk. The power was even further reduced when looking at subtype specific associations, especially for ER-negative breast cancer risk. Although the environmental data of the contributing studies were harmonized in a standardized fashion, we still observed heterogeneity in marginal effect associations with breast cancer risk (Supplementary Figure 1). Associations were less heterogeneous between population-based studies, and we included an interaction term between study design and the environmental variable in the models to account for potentially biased estimates from non-population-based studies. The assessment of associations between environmental factors and breast cancer risk was restricted to population-based studies and the estimates were comparable to those reported in the literature[15–22]. The association was not significant for BMI in premenopausal women, which may in part be attributed to the small sample available when using only population-based studies, however the direction of association was as expected. Another limitation of our study was that the sample consisted primarily of case-control studies and comprised only two cohort studies. While case-control studies have the advantage of being able to assess exposure close to the reference date, for example, for current MHT use, the retrospective assessment of exposure is prone to recall bias. However, we did not observe any heterogeneity between study-wise estimates for G×E interactions. Also, G×E interaction estimates derived from the whole study sample and from a sensitivity analysis restricted to population-based studies were similar (Supplementary Table 6). The robustness of our findings is also supported by the fact that, given reasonable assumptions, selection bias is unlikely to influence the assessment of multiplicative G×E interactions[23]. Also, both non-differential and differential misclassifications of environmental risk factors would lead to a reduction in power rather than increasing the probability of a spurious finding of an interaction[24]. The magnitude of the interactions for which strongest evidence was observed was comparable to those previously reported between breast cancer risk SNPs and

environmental factors[1]. When taking into account the number of tests performed, the identified G×E interactions were not statistically significant and further evidence is needed for confirmation. However, not all of the performed tests can be considered independent as we looked at different variables that are highly correlated (e.g. parity and number of full term pregnancies) and also tests for interaction concerning all cases and subgroups of cases defined by ER status are related. We therefore calculated the BFDP to be able to rate the noteworthiness of the observed G×E interactions.

It should be noted, that the investigated susceptibility loci have been identified in a sample of European descent, and that this investigation of gene-environment interaction was also restricted to subjects with European ancestry. The potential gene-environment interactions detected here do not necessarily have to be present in study populations of different ancestry due to the varying genetic structure and possible different prevalence of risk factors.

The interaction with the lowest BFDP was found between rs6828523 and adult height on ER-negative breast cancer risk. The SNP rs6828523 itself was associated with a decreased risk of ER-positive breast cancer, but not for ER-negative breast cancer, showing significant heterogeneity by ER status in the analysis identifying the variant ($p_{het} = 1.2\times10^{-7}$)[6], and also in sample analysed here ($p_{het} = 9.5\times10^{-6}$). SNP rs6828523 showed a positive association with ER-negative breast cancer risk in women taller than 164 cm (the median height in the study sample) (ER-negative OR = 1.13, 95% CI 1.01 – 1.26, $p = 0.036$). Current evidence suggests that adult body height is a risk factor for both ER-positive and ER-negative breast cancer, although the estimates for ER-negative breast cancer are not entirely consistent across studies[25–29]. The variant rs6828523 is located in an intron of *ADAM29*. The potentially functional implications of rs6828523 or SNPs highly correlated with rs6828523 ($r^2>0.6$) are unclear as they are not located within any strong regulatory elements (Supplementary Figure 3). Also, a more comprehensive investigation of the functional effects of the 41 SNPs associated with overall breast cancer risk did not identify a SNP in LD with rs6828523 coinciding with a regulatory genomic feature[30]. *ADAM29* encodes a disintegrin-metalloproteinase. Metalloproteinases are involved in the modification of the extracellular matrix and growth factor bioavailability, and changes in expression of metalloproteinases have been linked to breast cancer progression[31]. It is unclear however, how factors involved in growth and adult height might interplay with variants in *ADAM29* to influence risk of ER-negative breast cancer.

The association of rs4808801 located on chromosome 19 in an intron of *ELL* with overall breast cancer risk appeared to vary according to the number of full-term pregnancies in parous women. Risk of breast cancer associated with the SNP decreased with an increasing number of pregnancies. Several SNPs in LD with rs4808801 ($r^2 > 0.6$) are located in regulatory regions (enhancer elements, DNAse hypersensitive sites, transcription factor binding sites) in the proximity of *ELL* and two closely-located genes, *SSBP4* and *ISYNA1* (Supplementary Figure 3). Three SNPs in LD with rs4808801 ($r^2 \geq 0.9$) are located in exons of *SSBP4* (rs10405636) and *ISYNA1* (rs2303697, rs4595905), and all result in synonymous codon changes. *ELL* encodes the eleven-nineteen lysine-rich leukaemia protein, which was first identified as part of a fusion gene *MLL-ELL* in acute myeloid leukaemia cells, caused by a t(11;19)(q23;p13.1) translocation[32]. ELL is part of the super elongation complex, an

important regulator of transcriptional elongation[33]. Furthermore, ELL has been found to be essential for the transcription of rapidly induced genes, and therefore plays a key role in quick responses to environmental changes[34]. Rhie et al. identified another SNP (rs2303696) in LD with rs4808801 ($r^2 = 0.79$) located in the promoter region of *ISYNA1*, and likely to affect *ISYNA1* expression[30]. *ISYNA1* encodes an inositol-3-phosphate synthase enzyme that catalyses the synthesis of inositol 1-phosphate from glucose 6-phosphate. *ISYNA1* expression has been found to be reduced in breast cancer[30]. Inositol containing compounds are involved in many biological processes and act as essential second messenger molecules in signalling pathways[35], as components of cellular membranes[35] and regulators of chromatin remodelling[36, 37]. Less is known about the role of *SSBP4. SSBP4* is a putative tumour suppressor, as chromosomal regions containing members of the *SSBP* gene family are often found to be deleted in solid tumours[38]. How biological changes associated with multiple pregnancies in women potentially interplay with rs4808801 to influence its association with breast cancer risk is unknown.

We also observed that current smoking may modify the risk associated with rs11242675, located in close proximity to *FOXQ1* on chromosome 6. Two SNPs in LD with rs11242675 are located in enhancer regions[30]. *FOXQ1* is a transcription factor, which has been found to be involved in the epithelial-mesenchymal transition of tumour cells, a process initiating metastasis[39]. Overexpression of *FOXQ1* was observed in colorectal cancer[40] and metastatic breast cancer cell lines[41] and a subsequent study suggested that *FOXQ1* overexpression is caused by aberrant Wnt signalling[42]. A potential biological implication of the interaction between current smoking and rs11242675 is suggested by the observation that cigarette smoking deregulates nitric oxide synthesis[43], which in turn may decrease the expression of the Wnt/β-catenin regulator Dickkopf-1 (DKK1) and release Wnt signalling[44].

This is the first evaluation of multiplicative G×E interactions between these 47 newly identified breast cancer susceptibility loci and environmental risk factors. For most of the investigated pairs of SNPs and environmental factors, there was no indication of multiplicative G×E interaction. However, despite the overall very large study sample, we cannot exclude the existence of real G×E interactions of smaller magnitude with some environmental risk factors, for which power in this study was still limited. The six potential interactions identified are largely hypothesis generating and have to be confirmed in independent studies of sufficient size. Overall, our study does not suggest that the associations between recently identified breast cancer susceptibility loci and breast cancer risk are strongly modified by environmental risk factors.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Anja Rudolph[1], Roger L. Milne[2,3], Thérèse Truong[4,5], Julia A. Knight[6,7], Petra Seibold[1], Dieter Flesch-Janys[8,9], Sabine Behrens[1], Ursula Eilber[1], Manjeet K. Bolla[10], Qin Wang[10], Joe Dennis[10], Alison M. Dunning[11], Mitul Shah[11], Hannah R.

Munday[11], Hatef Darabi[12], Mikael Eriksson[12], Judith S. Brand[12], Janet Olson[13], Celine M. Vachon[13], Emily Hallberg[13], J. Esteban Castelao[14], Angel Carracedo[15,16,17], Maria Torres[16], Jingmei Li[18], Keith Humphreys[12], Emilie Cordina-Duverger[4,5], Florence Menegaux[4,5], Henrik Flyger[19], Børge G. Nordestgaard[20,21], Sune F. Nielsen[20,21], Betul T. Yesilyurt[22], Giuseppe Floris[23], Karin Leunen[23], Ellen G. Engelhardt[24], Annegien Broeks[25], Emiel J. Rutgers[26], Gord Glendon[27], Anna Marie Mulligan[28,29], Simon Cross[30], Malcolm Reed[31], Anna Gonzalez-Neira[32], José Ignacio Arias Perez[33], Elena Provenzano[34,35], Carmel Apicella[36], Melissa C. Southey[37], Amanda Spurdle[38], kConFab Investigators[39], AOCS Group[40], Lothar Häberle[41], Matthias W. Beckmann[41], Arif B. Ekici[42], Aida Karina Dieffenbach[43,44], Volker Arndt[43], Christa Stegmaier[45], Catriona McLean[46], Laura Baglietto[2,3], Stephen J. Chanock[47], Jolanta Lissowska[48], Mark E. Sherman[47,49], Thomas Brüning[50], Ute Hamann[51], Yon-Dschun Ko[52], Nick Orr[53], Minouk Schoemaker[54], Alan Ashworth[53], Veli-Matti Kosma[55,56], Vesa Kataja[57,58,59], Jaana M. Hartikainen[55,56], Arto Mannermaa[55,56], Anthony Swerdlow[54,60], GENICA-Network[61], Graham G. Giles[2,3], Hermann Brenner[43,44], Peter A. Fasching[41,62], Georgia Chenevix-Trench[38], John Hopper[36], Javier Benítez[63], Angela Cox[31], Irene L. Andrulis[27,64], Diether Lambrechts[22], Manuela Gago-Dominguez[15], Fergus Couch[65], Kamila Czene[12], Stig E. Bojesen[20,21], Doug F. Easton[10,11], Marjanka K. Schmidt[66], Pascal Guénel[4,5], Per Hall[12], Paul D. P. Pharoah[11,10], Montserrat Garcia-Closas[54,67], and Jenny Chang-Claude[1]

Affiliations

[1]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany [2]Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia [3]Centre for Epidemiology & Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia [4]Inserm (National Institute of Health and Medical Research), CESP (Center for Research in Epidemiology and Population Health), U1018, Environmental Epidemiology of Cancer, Villejuif, France [5]Unité Mixte de Recherche Scientifique (UMRS) 1018, University Paris-Sud, Villejuif, France [6]Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada [7]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada [8]Department of Cancer Epidemiology/Clinical Cancer Registry, University Clinic Hamburg-Eppendorf, Hamburg, Germany [9]Institute for Medical Biometrics and Epidemiology, University Clinic Hamburg-Eppendorf, Hamburg, Germany [10]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK [11]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK [12]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden [13]Department of Health Sciences Research, Mayo Clinic, Rochester, MN, USA [14]Oncology and Genetics Unit, Biomedical Research Institute of Vigo (IBIV), Complejo Hospitalario Universitario de Vigo, Servicio Galego de Saude (SERGAS), Vigo, Spain [15]Genomic Medicine Group, Galician Foundation of Genomic Medicine,

Servicio Galego de Saude (SERGAS), Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain [16]National Genotyping Center - Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), University of Santiago de Compostela, Spain [17]Center of Excellence in Genomic Medicine Research, King Abdulaziz University, Jeddah, KSA [18]Human Genetics Division, Genome Institute of Singapore, Singapore, Singapore [19]Department of Breast Surgery, Herlev Hospital, Copenhagen University Hospital, Copenhagen, Denmark [20]Copenhagen General Population Study, Herlev Hospital, Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark [21]Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, University of Copenhagen, Copenhagen, Denmark [22]Vesalius Research Center (VRC), VIB, Leuven, Belgium [23]Multidisciplinary Breast Center, University Hospital Gasthuisberg, Leuven, Belgium [24]Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands [25]Division of Molecular Pathology, Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands [26]Department of Surgery, Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands [27]Ontario Cancer Genetics Network, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada [28]Laboratory Medicine Program, University Health Network, Toronto, Ontario, Canada [29]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada [30]Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, UK [31]Sheffield Cancer Research Centre, Department of Oncology, University of Sheffield, Sheffield, UK [32]Human Genotyping Unit-CEGEN, Human Cancer Genetics Program, Spanish National Cancer Research Centre (CNIO), Madrid, Spain [33]Servicio de Cirugía General y Especialidades, Hospital Monte Naranco, Oviedo, Spain [34]Cancer Research UK Cambridge Institute, Cambridge, UK [35]Cambridge Breast Unit, Addenbrooke's Hospital, Cambridge University Hospital NHS Foundation Trust and NIHR Cambridge Biomedical Research Centre, Cambridge, UK [36]Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of Melbourne, Melbourne, Australia [37]Department of Pathology, University of Melbourne, Melbourne, Australia [38]Department of Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia [39]Peter MacCallum Cancer Center, Melbourne, Australia [40]QIMR Berghofer Medical Research Institute, Brisbane, Australia [41]Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany [42]Institute of Human Genetics, University Hospital Erlangen, Friedrich-Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany [43]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany [44]German Cancer Consortium (DKTK), Heidelberg, Germany [45]Saarland Cancer Registry, Saarbrücken, Germany [46]Anatomical Pathology, The Alfred Hospital, Melbourne, Australia [47]Division of Cancer

Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA [48]Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center & Institute of Oncology, Warsaw, Poland [49]Division of Cancer Prevention, National Cancer Institute, Bethesda, Maryland, USA [50]Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany [51]Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany [52]Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany [53]Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London, UK [54]Division of Genetics and Epidemiology, Institute of Cancer Research, Sutton, Surrey, UK [55]School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine and Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland [56]Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland [57]School of Medicine, Institute of Clinical Medicine, Oncology and Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland [58]Cancer Center, Kuopio University Hospital, Kuopio, Finland [59]Jyväskylä Central Hospital, Jyväskylä, Finland [60]Division of Breast Cancer Research, Institute of Cancer Research, Sutton, Surrey, UK [61]The GENICA Network: Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, and University of Tübingen, Germany; Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany, Institute of Pathology, University of Bonn, Germany, Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany, and Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany; Institute of Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany [62]David Geffen School of Medicine, Department of Medicine Division of Hematology and Oncology, University of California at Los Angeles, CA, USA [63]Human Genetics Group, Human Cancer Genetics Program, Spanish National Cancer Research Centre (CNIO), Madrid, Spain [64]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada [65]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA [66]Netherlands Cancer Institute, Antoni van Leeuwenhoek hospital, Amsterdam, The Netherlands [67]Breakthrough Breast Cancer Research Centre, Division of Breast Cancer Research, The Institute of Cancer Research, London, UK

Acknowledgments

This study would not have been possible without the contributions of the following: Kyriaki Michailidou, Andrew Berchuck (OCAC), Rosalind A. Eeles, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), Antonis Antoniou, Lesley McGuffog, Ken Offit (CIMBA), Alison M. Dunning, Andrew Lee, and Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory, the staff of the CNIO genotyping unit, Jacques Simard and Daniel C. Tessier, Francois Bacot, Daniel Vincent, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, Sune F. Nielsen, Borge G.

Nordestgaard, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility, Maggie Angelakos, Judi Maskiell, Gillian Dite, Sanquin Research, the Netherlands, Sonja Oeser, Silke Landrith, Matthias Rübner, Alexander Hein, staff and participants of the Copenhagen General Population Study. For the excellent technical assistance: Dorthe Uldall Andersen, Maria Birna Arnadottir, Anne Bank, Dorthe Kjeldgård Hansen. CNIO-BCS would like to thank Guillermo Pita, Charo Alonso, Daniel Herrero, Nuria Álvarez, Pilar Zamora, Primitiva Menendez, the Human Genotyping-CEGEN Unit (CNIO). The GENICA-Network: Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart and University of Tuebingen, Germany [Hiltrud Brauch, Wing-Yee Lo, Christina Justenhoven, Molecular Genetics of Breast Cancer, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Germany [UH], Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany [TB, Beate Pesch, Sylvia Rabstein, Anne Lotz], Institute for Occupational Medicine and Maritime Medicine, University Medical Center Hamburg-Eppendorf, Germany [Volker Harth]Institute of Pathology, Medical Faculty of the University of Bonn, Bonn, Germany [Hans-Peter Fischer]. Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany [YDK, Christian Baisch]. KBCP thanks Eija Myöhänen and Helena Kemiläinen for technical assistance. kConFab/AOCS wish to thank Heather Thorne, Eveline Niedermayr, all the kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the Clinical Follow Up Study (funded 2001–2009 by NHMRC and currently by the National Breast Cancer Foundation and Cancer Australia #628333) for their contributions to this resource, and the many families who contribute to kConFab. The Australian group gratefully acknowledges the members of the Australian Ovarian Cancer Study Group. LMBC thanks Gilian Peuteman, Dominiek Smeets, Thomas Van Brussel and Kathleen Corthouts. MCBCS thanks the Mayo Clinic Breast Cancer Patient Registry, David F. and Margaret T. Grohne Family Foundation and Ting Tsung and Wei Fong Chao Foundation. MARIE would like to thank Alina Vrieling, Katharina Buck, Muhabbet Celik, Ursula Eilber and Sabine Behrens. OFBCR would like to thank Teresa Selander, Nayana Weerasooriya and the OFBCR staff and participants. PBCS thanks Louise Brinton, Neonila Szeszenia-Dabrowska, Beata Peplonska, Witold Zatonski, Pei Chao, Michael Stagner. The SBCS thanks Sue Higham, Helen Cramp, Dan Connley, Ian Brock, Sabapathy Balasubramanian. SEARCH thanks Craig Luccarini, Caroline Baynes, Don Conroy, Anne Stafford, Sue Irvine, Barbara Perkins, Val Rhenius. The UKBGS would like to thanks Breakthrough Breast Cancer and the Institute of Cancer Research for support and funding of the Breakthrough Generations Study, and the study participants, study staff, and the doctors, nurses and other health care providers and health information sources who have contributed to the study.

Research Support

This work was supported by multiple funding agencies: Funding for the **iCOGS** infrastructure came from: the European Community's Seventh Framework Programme under grant agreement n° 223175 [HEALTH-F2-2009-223175] (COGS), Cancer Research UK [C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692], the National Institutes of Health [CA128978] and Post-Cancer GWAS initiative [1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative], the Department of Defence [W81XWH-10-1-0341], the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund.

The **ABCFS** work was supported by the United States National Cancer Institute, National Institutes of Health (NIH) under RFA-CA-06-503 and through cooperative agreements with members of the Breast Cancer Family Registry (BCFR) and Principal Investigators, including Cancer Care Ontario [U01 CA69467], Cancer Prevention Institute of California [U01 CA69417], University of Melbourne [U01 CA69638]. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the BCFR, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the BCFR. The ABCFS was also supported by the National Health and Medical Research Council of Australia, the New South Wales Cancer Council, the Victorian Health Promotion Foundation (Australia) and the Victorian Breast Cancer Research Consortium. J.L.H. is a National Health and Medical Research Council (NHMRC) Australia Fellow and a Victorian Breast Cancer Research Consortium Group Leader. M.C.S. is a NHMRC Senior Research Fellow and a Victorian Breast Cancer Research Consortium Group Leader.

The **ABCS** was funded by the Dutch Cancer Society [grant no. NKI2007-3839]; and BBMRI-NL complementation project no.11; BBMRI-NL is a Research Infrastructure financed by the Dutch government [NWO 184.021.007]; M.K.S. was funded by Dutch Cancer Society [grant no. NKI2009-4363].

The work of the **BBCC** was partly funded by ELAN-Program of the University Hospital of Erlangen.

BREOGAN is funded by FIS PI12/02125 and PI13/01136 Acción Estratégica de Salud del Instituto de Salud Carlos III; KAU grant No. (1-117-1434-HiCi); the Botin Foundation's Fund; Programa Grupos Emergentes, Cancer Genetics Unit, CHUVI Vigo Hospital, Instituto de Salud Carlos III, Spain; Consellería de Industria Programa Sectorial de Investigación Aplicada, PEME I+D e I+D Suma del Plan Gallego de Investigación, Desarrollo e Innovación Tecnológica de la Consellería de Industria de la Xunta de Galicia, Spain [grant 10CSA012E]; Fomento

de la Investigación Clínica Independiente, Ministerio de Sanidad, Servicios Sociales e Igualdad, Spain [grant EC11-192]; and Grant FEDER-Innterconecta. Ministerio de Economia y Competitividad, Xunta de Galicia, Spain.

The **CECILE** study was funded by the Fondation de France, the French National Institute of Cancer (INCa), The National League against Cancer, the National Agency for Environmental and Occupational Health and Food Safety (ANSES), the National Agency for Research (ANR), and the Association for Research against Cancer (ARC)

The **CGPS** was supported by the Chief Physician Johan Boserup and Lise Boserup Fund, the Danish Medical Research Council and Herlev Hospital

The **CNIO-BCS** was supported by the Genome Spain Foundation, the Red Temática de Investigación Cooperativa en Cáncer and grants from the Asociación Española Contra el Cáncer and the Fondo de Investigación Sanitario [PI11/00923 and PI081120]. The Human Genotyping-CEGEN Unit, CNIO is supported by the Instituto de Salud Carlos III.

ESTHER was supported in part by the Baden-Württemberg State Ministry of Science, Research and Arts; and by the German Federal Ministry of Education and Research. Additional cases were recruited in the context of the VERDI study, which was supported by a grant from the German Cancer Aid (Deutsche Krebshilfe).

The **GENICA** was funded by the Federal Ministry of Education and Research (BMBF) Germany [grants 01KW9975/5, 01KW9976/8, 01KW9977/0, 01KW0114, 01KH0401, 01KH0410, and 01KH0411], the Robert Bosch Foundation, Stuttgart, Deutsches Krebsforschungszentrum (DKFZ), Heidelberg, Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University (IPA), Bochum, as well as the Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany

KBCP was supported by special Government Funding (EVO) of Kuopio University Hospital grants, Cancer Fund of North Savo, the Finnish Cancer Organizations, the Academy of Finland and by the strategic funding of the University of Eastern Finland

kConFab is supported by grants from the National Breast Cancer Foundation, the National Health and Medical Research Council (NHMRC) and by the Queensland Cancer Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia. G.C.T. is an NHMRC Senior Principal Research Fellow. A.B.S. is an NHMRC Senior Research Fellow. **AOCS** is supported by the US Army Medical Research and Materiel Command [DAMD 170110729 and W81XWH0610220 [AUS]

LMBC is supported by the ‘Stichting tegen Kanker’ [232-2008 and 196-2010].

MCBCS is supported by the NCI specialized program of research excellence (SPORE) in breast cancer [P50 CA116201, NIH R01 CA128978], and the Breast Cancer Research Foundation

The **MARIE** study was supported by the Deutsche Krebshilfe e.V. [70-2892-BR I], the Hamburg Cancer Society, the German Cancer Research Center and the Federal Ministry of Education and Research (BMBF) Germany [01KH0402].

MCCS cohort recruitment was funded by VicHealth and Cancer Council Victoria. The MCCS was further supported by Australian NHMRC [grants 209057, 251553 and 504711] and by infrastructure provided by Cancer Council Victoria.

The work of the **OFBCR** was supported by the National Cancer Institute [grant UM1 CA164920]. The content of this manuscript does not necessarily reflect the views or policies of the National Cancer Institute or any of the collaborating centers in the Breast Cancer Family Registry (BCFR), nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government or the BCFR.

The **PBCS** was funded by Intramural Research Funds of the National Cancer Institute, Department of Health and Human Services, USA.

pKARMA is a combination of the KARMA and LIBRO-1 studies. KARMA was supported by Märit and Hans Rausings Initiative Against Breast Cancer. KARMA and LIBRO-1 were supported the Cancer Risk Prediction Center (CRisP; www.crispcenter.org), a Linnaeus Centre [Contract ID 70867902] financed by the Swedish Research Council.

SASBAC was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the US National Institute of Health (NIH) and the Susan G. Komen Breast Cancer Foundation. KC was financed by the Swedish Cancer Society [5128-B07-01PAF].

The **SBCS** was supported by Yorkshire Cancer Research [S305PA].

SEARCH is funded by a programme grant from Cancer Research UK [C490/A10124] and supported by the UK National Institute for Health Research Biomedical Research Centre at the University of Cambridge.

The **UKBGS** is funded by Breakthrough Breast Cancer and the Institute of Cancer Research (ICR). ICR acknowledges NHS funding to the NIHR Biomedical Research Centre.

Abbreviations

BCAC	Breast Cancer Association Consortium
BFDP	Bayesian False Discovery Probability
BMI	body-mass index
CI	confidence interval
ER	estrogen receptor
G×E	gene-environment
GWAS	genome-wide association studies
MHT	menopausal hormone therapy
OR	odds ratio
pint	p-value for interaction
phet	p-value for heterogeneity
SNP	single nucleotide polymorphism

References

1. Nickels S, Truong T, Hein R, Stevens K, Buck K, Behrens S, Eilber U, Schmidt M, Haberle L, Vrieling A, Gaudet M, Figueroa J, et al. Evidence of Gene-Environment Interactions between Common Breast Cancer Susceptibility Loci and Established Environmental Risk Factors. PLoS Genet. 2013; 9:e1003284. [PubMed: 23544014]
2. Campa D, Kaaks R, Le Marchand L, Haiman CA, Travis RC, Berg CD, Buring JE, Chanock SJ, Diver WR, Dostal L, Fournier A, Hankinson SE, et al. Interactions between genetic variants and breast cancer risk factors in the breast and prostate cancer cohort consortium. J Natl Cancer Inst. 2011; 103:1252–63. [PubMed: 21791674]
3. Travis RC, Reeves GK, Green J, Bull D, Tipper SJ, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M. Million Women Study C. Gene-environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study. Lancet. 2010; 375:2143–51. [PubMed: 20605201]
4. Prentice RL, Huang Y, Hinds DA, Peters U, Pettinger M, Cox DR, Beilharz E, Chlebowski RT, Rossouw JE, Caan B, Ballinger DG. Variation in the FGFR2 gene and the effects of postmenopausal hormone therapy on invasive breast cancer. Cancer Epidemiol Biomarkers Prev. 2009; 18:3079–85. [PubMed: 19861516]
5. Milne RL, Gaudet MM, Spurdle AB, Fasching PA, Couch FJ, Benitez J, Arias Perez JI, Zamora MP, Malats N, Dos Santos Silva I, Gibson LJ, Fletcher O, et al. Assessing interactions between the associations of common genetic susceptibility variants, reproductive history and body mass index with breast cancer risk in the breast cancer association consortium: a combined case-control study. Breast Cancer Res. 2010; 12:R110. [PubMed: 21194473]

6. Michailidou K, Hall P, Gonzalez-Neira A, Ghoussaini M, Dennis J, Milne RL, Schmidt MK, Chang-Claude J, Bojesen SE, Bolla MK, Wang Q, Dicks E, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. Nat Genet. 2013; 45:353–61. [PubMed: 23535729]
7. Garcia-Closas M, Couch FJ, Lindstrom S, Michailidou K, Schmidt MK, Brook MN, Orr N, Rhie SK, Riboli E, Feigelson HS, Le Marchand L, Buring JE, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. Nat Genet. 2013; 45:392–8. [PubMed: 23535733]
8. Haiman CA, Chen GK, Vachon CM, Canzian F, Dunning A, Millikan RC, Wang X, Ademuyiwa F, Ahmed S, Ambrosone CB, Baglietto L, Balleine R, et al. A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. Nat Genet. 2011; 43:1210–4. [PubMed: 22037553]
9. Siddiq A, Couch FJ, Chen GK, Lindstrom S, Eccles D, Millikan RC, Michailidou K, Stram DO, Beckmann L, Rhie SK, Ambrosone CB, Aittomaki K, et al. A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. Hum Mol Genet. 2012; 21:5373–84. [PubMed: 22976474]
10. Stevens KN, Fredericksen Z, Vachon CM, Wang X, Margolin S, Lindblom A, Nevanlinna H, Greco D, Aittomaki K, Blomqvist C, Chang-Claude J, Vrieling A, et al. 19p13. 1 is a triple-negative-specific breast cancer susceptibility locus. Cancer Res. 2012; 72:1795–803. [PubMed: 22331459]
11. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 2012; 40:D930–4. [PubMed: 22064851]
12. Meyer LR, Zweig AS, Hinrichs AS, Karolchik D, Kuhn RM, Wong M, Sloan CA, Rosenbloom KR, Roe G, Rhead B, Raney BJ, Pohl A, et al. The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res. 2013; 41:D64–9. [PubMed: 23155063]
13. Mukherjee B, Chatterjee N. Exploiting gene-environment independence for analysis of case-control studies: an empirical Bayes-type shrinkage estimator to trade-off between bias and efficiency. Biometrics. 2008; 64:685–94. [PubMed: 18162111]
14. Wakefield J. A Bayesian measure of the probability of false discovery in genetic epidemiology studies. Am J Hum Genet. 2007; 81:208–27. [PubMed: 17668372]
15. Collaborative Group on Hormonal Factors in Breast C. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. Lancet Oncol. 2012; 13:1141–51. [PubMed: 23084519]
16. Reeves GK, Pirie K, Green J, Bull D, Beral V. Million Women Study C. Reproductive factors and specific histological types of breast cancer: prospective study and meta-analysis. Br J Cancer. 2009; 100:538–44. [PubMed: 19190634]
17. Bernier MO, Plu-Bureau G, Bossard N, Ayzac L, Thalabard JC. Breastfeeding and risk of breast cancer: a metaanalysis of published studies. Hum Reprod Update. 2000; 6:374–86. [PubMed: 10972524]
18. Suzuki R, Orsini N, Saji S, Key TJ, Wolk A. Body weight and incidence of breast cancer defined by estrogen and progesterone receptor status--a meta-analysis. Int J Cancer. 2009; 124:698–712. [PubMed: 18988226]
19. Green J, Cairns BJ, Casabonne D, Wright FL, Reeves G, Beral V. Million Women Study c. Height and cancer incidence in the Million Women Study: prospective cohort, and meta-analysis of prospective studies of height and total cancer risk. Lancet Oncol. 2011; 12:785–94. [PubMed: 21782509]
20. Marjoribanks J, Farquhar C, Roberts H, Lethaby A. Long term hormone therapy for perimenopausal and postmenopausal women. Cochrane Database Syst Rev. 2012; 7:CD004143. [PubMed: 22786488]
21. Seitz HK, Pelucchi C, Bagnardi V, La Vecchia C. Epidemiology and pathophysiology of alcohol and breast cancer: Update 2012. Alcohol Alcohol. 2012; 47:204–12. [PubMed: 22459019]
22. Gaudet MM, Gapstur SM, Sun J, Diver WR, Hannan LM, Thun MJ. Active smoking and breast cancer risk: original cohort data and meta-analysis. J Natl Cancer Inst. 2013; 105:515–25. [PubMed: 23449445]

23. Morimoto LM, White E, Newcomb PA. Selection bias in the assessment of gene-environment interaction in case-control studies. Am J Epidemiol. 2003; 158:259–63. [PubMed: 12882948]
24. Garcia-Closas M, Rothman N, Lubin J. Misclassification in case-control studies of gene-environment interactions: assessment of bias and sample size. Cancer Epidemiol Biomarkers Prev. 1999; 8:1043–50. [PubMed: 10613335]
25. Sellers TA, Davis J, Cerhan JR, Vierkant RA, Olson JE, Pankratz VS, Potter JD, Folsom AR. Interaction of waist/hip ratio and family history on the risk of hormone receptor-defined breast cancer in a prospective study of postmenopausal women. Am J Epidemiol. 2002; 155:225–33. [PubMed: 11821247]
26. John EM, Phipps AI, Sangaramoorthy M. Body size, modifying factors, and postmenopausal breast cancer risk in a multiethnic population: the San Francisco Bay Area Breast Cancer Study. Springerplus. 2013; 2:239. [PubMed: 23762816]
27. Fagherazzi G, Vilier A, Boutron-Ruault MC, Clavel-Chapelon F, Mesrine S. Height, sitting height, and leg length in relation with breast cancer risk in the E3N cohort. Cancer Epidemiol Biomarkers Prev. 2012; 21:1171–5. [PubMed: 22623708]
28. Ritte R, Lukanova A, Tjonneland A, Olsen A, Overvad K, Mesrine S, Fagherazzi G, Dossus L, Teucher B, Steindorf K, Boeing H, Aleksandrova K, et al. Height, age at menarche and risk of hormone receptor-positive and -negative breast cancer: a cohort study. Int J Cancer. 2013; 132:2619–29. [PubMed: 23090881]
29. Colditz GA, Rosner BA, Chen WY, Holmes MD, Hankinson SE. Risk factors for breast cancer according to estrogen and progesterone receptor status. J Natl Cancer Inst. 2004; 96:218–28. [PubMed: 14759989]
30. Rhie SK, Coetzee SG, Noushmehr H, Yan C, Kim JM, Haiman CA, Coetzee GA. Comprehensive functional annotation of seventy-one breast cancer risk Loci. PLoS One. 2013; 8:e63925. [PubMed: 23717510]
31. Hojilla CV, Wood GA, Khokha R. Inflammation and breast cancer: metalloproteinases as common effectors of inflammation and extracellular matrix breakdown in breast cancer. Breast Cancer Res. 2008; 10:205. [PubMed: 18394187]
32. Thirman MJ, Levitan DA, Kobayashi H, Simon MC, Rowley JD. Cloning of ELL, a gene that fuses to MLL in a t(11;19)(q23;p13. 1) in acute myeloid leukemia. Proc Natl Acad Sci U S A. 1994; 91:12110–4. [PubMed: 7991593]
33. Smith E, Lin C, Shilatifard A. The super elongation complex (SEC) and MLL in development and disease. Genes Dev. 2011; 25:661–72. [PubMed: 21460034]
34. Byun JS, Fufa TD, Wakano C, Fernandez A, Haggerty CM, Sung MH, Gardner K. ELL facilitates RNA polymerase II pause site entry and release. Nat Commun. 2012; 3:633. [PubMed: 22252557]
35. Berridge MJ, Irvine RF. Inositol phosphates and cell signalling. Nature. 1989; 341:197–205. [PubMed: 2550825]
36. Steger DJ, Haswell ES, Miller AL, Wente SR, O'Shea EK. Regulation of chromatin remodeling by inositol polyphosphates. Science. 2003; 299:114–6. [PubMed: 12434012]
37. Shen X, Xiao H, Ranallo R, Wu WH, Wu C. Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. Science. 2003; 299:112–4. [PubMed: 12434013]
38. Castro P, Liang H, Liang JC, Nagarajan L. A novel, evolutionarily conserved gene family with putative sequence-specific single-stranded DNA-binding activity. Genomics. 2002; 80:78–85. [PubMed: 12079286]
39. Qiao Y, Jiang X, Lee ST, Karuturi RK, Hooi SC, Yu Q. FOXQ1 regulates epithelial-mesenchymal transition in human cancers. Cancer Res. 2011; 71:3076–86. [PubMed: 21346143]
40. Kaneda H, Arao T, Tanaka K, Tamura D, Aomatsu K, Kudo K, Sakai K, De Velasco MA, Matsumoto K, Fujita Y, Yamada Y, Tsurutani J, et al. FOXQ1 is overexpressed in colorectal cancer and enhances tumorigenicity and tumor growth. Cancer Res. 2010; 70:2053–63. [PubMed: 20145154]
41. Zhang H, Meng F, Liu G, Zhang B, Zhu J, Wu F, Ethier SP, Miller F, Wu G. Forkhead transcription factor foxq1 promotes epithelial-mesenchymal transition and breast cancer metastasis. Cancer Res. 2011; 71:1292–301. [PubMed: 21285253]

42. Christensen J, Bentz S, Sengstag T, Shastri VP, Anderle P. FOXQ1, a novel target of the Wnt pathway and a new marker for activation of Wnt signaling in solid tumors. PLoS One. 2013; 8:e60051. [PubMed: 23555880]

43. Vleeming W, Rambali B, Opperhuizen A. The role of nitric oxide in cigarette smoking and nicotine addiction. Nicotine Tob Res. 2002; 4:341–8. [PubMed: 12215243]

44. Du Q, Zhang X, Liu Q, Zhang X, Bartels CE, Geller DA. Nitric Oxide Production Upregulates Wnt/beta-Catenin Signaling by Inhibiting Dickkopf-1. Cancer Res. 2013; 73:6526–37. [PubMed: 24008318]

Novelty and Impact Statement

Relative risks associated with 47 recently identified susceptibility loci for overall or estrogen receptor negative breast cancer may vary depending on exposure levels of environmental (non-genetic) risk factors. In this study, gene-environment interactions between these 47 single nucleotide polymorphisms and 13 established environmental risk factors were investigated. Relative risks of breast cancer associated with the susceptibility loci were not strongly modified by environmental risk factors. This finding may have important implications for risk prediction.

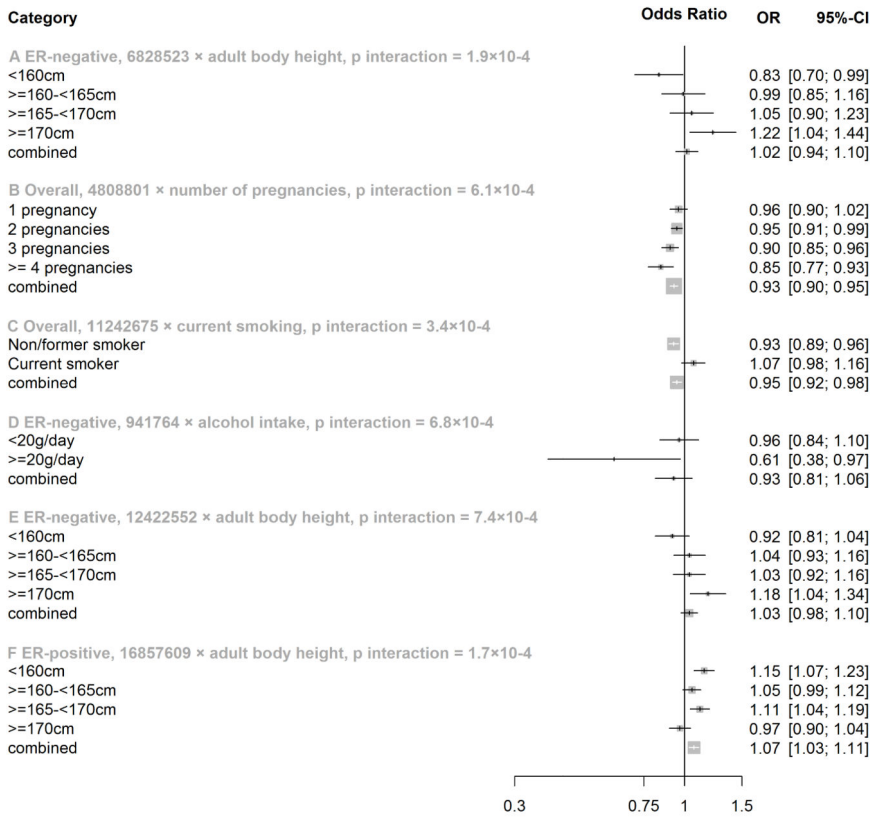

Figure 1.

Odds ratios and 95% confidence intervals for association between SNP and overall breast cancer (B, C), estrogen receptor positive breast cancer (D), and estrogen receptor negative breast cancer (A, E, F) stratified by categories of environmental factors.

Table 1

List of participating studies and number of Caucasian subjects included in at least one G×E analysis

Study acronym	Study Name	Country	Design category[1]	Cases	ER+ cases	ER− cases	Controls	Mean age (sd) cases	Mean age (sd) controls
ABCFS	Australian Breast Cancer Family Study	Australia	Population-based	790	456	261	551	40.0 (6.7)	42.4 (9.3)
ABCS	Amsterdam Breast Cancer Study	Netherlands	Mixed	1143	420	152	1177	43.4 (8.7)	48.1 (12.2)
BBCC	Bavarian Breast Cancer Cases and Controls	Germany	Mixed	554	456	82	458	61.2 (12.1)	57.6 (10.9)
BREOGAN	Breast Oncology Galicia Network	Spain	Mixed	1216	819	194	1806	56.9 (12.4)	43.3 (14.7)
CECILE	CECILE Breast cancer study	France	Population-based	900	743	130	999	54.8 (10.8)	55.3 (11)
CGPS	Copenhagen General Population Study	Denmark	Mixed	2811	1919	357	4086	62.3 (12.4)	58.4 (15.5)
CNIO-BCS	Spanish National Cancer Centre Breast Cancer Study	Spain	Mixed	704	213	76	834	54.7 (11.8)	50.5 (11.1)
ESTHER	ESTHER Breast Cancer Study	Germany	Population-based	471	302	98	502	61.0 (8.9)	62.8 (7.2)
GENICA	Gene Environment Interaction & Breast Cancer in Germany	Germany	Population-based	465	328	119	427	57.5 (10.9)	57.8 (11.8)
KBCP	Kuopio Breast Cancer Project	Finland	Population-based	410	288	89	250	59.4 (14.5)	52.8 (11.6)
kConFab/AOCS	Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer/Australian Ovarian Cancer Study	Australia	Mixed	410	135	50	897	46.2 (9.5)	58.3 (11.2)
LMBC	Leuven Multidisciplinary Breast Centre	Belgium	Mixed	2522	2069	378	1386	57.1 (12.4)	44.4 (9.1)
MARIE	Mammary Carcinoma Risk Factor Investigation	Germany	Population-based	1656	1279	370	1777	62.8 (6.3)	62.3 (6.1)
MCBCS	Mayo Clinic Breast Cancer Study	USA	Mixed	1546	1271	250	1931	57.5 (12.6)	57.1 (14)
MCCS	Melbourne Collaborative Cohort Study	Australia	Population-based	454	330	110	511	64.3 (8.7)	56.4 (8.3)
OFBCR	Ontario Familial Breast Cancer Registry	Canada	Mixed	1157	629	267	511	53.5 (10.2)	52.4 (9.2)
PBCS	NCI Polish Breast Cancer Study	Poland	Population-based	519	519	0	424	56.8 (10)	56.9 (9.8)
pKARMA	Karolinska Mammography Project for Risk Prediction of Breast Cancer - prevalent cases	Sweden	Mixed	2700	2238	387	5529	59.3 (10.5)	53.9 (9.5)
SASBAC	Singapore and Sweden Breast Cancer Study	Sweden	Population-based	1163	663	144	1378	63.6 (6.5)	63.8 (6.4)
SBCS	Sheffield Breast Cancer Study	UK	Mixed	751	358	104	848	60.0 (12.4)	58.0 (5.8)
SEARCH	Study of Epidemiology and Risk factors in Cancer Heredity	UK	Mixed	9095	5130	1170	8064	54.9 (9.2)	58.2 (8.6)
UKBGS	UK Breakthrough Generations Study	UK	Population-based	413	88	18	470	56.8 (10.2)	54.7 (10)
Total				31850	20653	4806	34816	56.7 (11.5)	55.6 (12)

[1]Population-based design was defined as recruiting a random sample of all cases occurring in a geographically defined population during a specified period of time, and recruiting controls that were a random sample of the same source population as cases during the same period of time. Mixed design was defined as not strictly population-based or hospital-based.

Table 2

Bayesian False Discovery Probability of G×E interactions showing p-value for interaction $<1.1\times10^{-3}$

Breast cancer subtype	Environmental Factor × SNP (locus)	ORint (95% CI)[1]	Assumed 95% probability range interaction OR[2]	BFDP Prior of βGE ≠ 0					
				0.2	0.1	0.05	0.01	0.001	0.0001
ER-negative	Adult height × rs6828523 (ADAM29)	1.14 (1.06 – 1.22)	0.66–1.50	0.022	0.049	0.097	0.360	0.850	0.983
Overall	Number of full-term pregnancies × rs4808801 (ELL)	0.96 (0.94 – 0.98)	0.66–1.50	0.041	0.089	0.170	0.516	0.915	0.991
Overall	Current smoking × rs11242675 (4kb 3′ of FOXQ1)	1.13 (1.06 – 1.21)	0.66–1.50	0.058	0.122	0.227	0.605	0.939	0.994
ER-negative	Alcohol intake × rs941764 (CCDC88C)	0.53 (0.36 – 0.76)	0.66–1.50	0.177	0.327	0.506	0.842	0.982	0.998
ER-negative	Adult height × rs12422552 (105kb 5′ of ATF7IP)	1.09 (1.04 – 1.15)	0.66–1.50	0.188	0.342	0.523	0.851	0.983	0.998
ER-positive	Adult height × rs16857609 (DIRC3)	0.95 (0.93 – 0.98)	0.66–1.50	0.224	0.394	0.579	0.877	0.986	0.999

[1] Adjusted for reference age, study, principal components to adjust for population stratification and an interaction term between environmental factor and study design (population-based vs. non-population-based). Model used to assess association with current smoking have been adjusted for former smoking.

[2] To calculate the BFDP it was assumed that with probability 0.95, the interaction OR lies within the given rage.

ER: estrogen receptor; OR_{int}: odds ratio for interaction; CI: confidence interval; BFDP: Bayesian False Discovery Probability