Gene–environment interaction and risk of breast cancer

Anja Rudolph[*,1], Jenny Chang-Claude[1,2] and Marjanka K Schmidt[3,4]

[1]Unit of Genetic Epidemiology, Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany; [2]University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Martinistraße 52, 20246 Hamburg, Germany; [3]Division of Molecular Pathology, Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital; Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands and [4]Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute - Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX Amsterdam, The Netherlands

Hereditary, genetic factors as well as lifestyle and environmental factors, for example, parity and body mass index, predict breast cancer development. Gene–environment interaction studies may help to identify subgroups of women at high-risk of breast cancer and can be leveraged to discover new genetic risk factors. A few interesting results in studies including over 30 000 breast cancer cases and healthy controls indicate that such interactions exist. Explorative gene–environment interaction studies aiming to identify new genetic or environmental factors are scarce and still underpowered. Gene–environment interactions might be stronger for rare genetic variants, but data are lacking. Ongoing initiatives to genotype larger sample sets in combination with comprehensive epidemiologic databases will provide further opportunities to study gene–environment interactions in breast cancer. However, based on the available evidence, we conclude that associations between the common genetic variants known today and breast cancer risk are only weakly modified by environmental factors, if at all.

Worldwide, breast cancer is the most common cancer in women with an estimated 1.7 million cases and 521 900 deaths in 2012 (Torre *et al*, 2015). The heritability of breast cancer, that is, the proportion of the phenotypic variance that can be attributed to the germline genotype, has been estimated to be about 30% (Lichtenstein *et al*, 2000; Mucci *et al*, 2013). Furthermore, women with a first-degree relative affected by breast cancer have a twofold higher risk of developing the disease (Pharoah *et al*, 1997). Known genetic factors contributing to a higher lifetime risk of breast cancer comprise rare variants with moderate to high penetrance in *BRCA1*, *BRCA2*, *PALB2*, *ATM* and *CHEK2*, as well as over 90 common variants with low penetrance (Michailidou *et al*, 2015). Taken together, these variants explain about 37% of the excess familial risk (Michailidou *et al*, 2015) (Figure 1).

Apart from genetic factors, there are various well-established environmental factors (including endogenous and exogenous exposure to hormones and other substances, and lifestyle factors) associated with breast cancer risk. These include reproductive factors such as age at menarche, age at first birth, parity, breastfeeding and age at menopause. Reproductive factors are generally considered to be non-modifiable risk factors, that is, they

cannot be controlled or changed as public health measures. Other risk factors are considered modifiable such as extended use of menopausal hormone therapy, high alcohol consumption, lack of physical activity and high body mass index (BMI).

Women carrying a certain genetic variant may be at greater risk for breast cancer, if a related environmental factor is present or absent. Such gene–environment interactions have been identified for bladder cancer, where smokers carrying variants in two carcinogen-metabolising genes (*NAT2* and *GSTM1*) are at higher risk of disease compared with non-smokers (Garcia-Closas *et al*, 2005), as well as for oesophageal squamous cell carcinoma, where alcohol users carrying variants in the alcohol-metabolising pathway (*ADH1B* and *ALDH2*) are at higher risk compared with non-users (Wu *et al*, 2012).

GENERAL CONCEPT OF GENE–ENVIRONMENT INTERACTION STUDIES

Knowledge of gene–environment interaction is important for risk prediction and the identification of certain high-risk populations to

---

*Correspondence: Dr A Rudolph; E-mail: a.rudolph@dkfz-Heidelberg.de
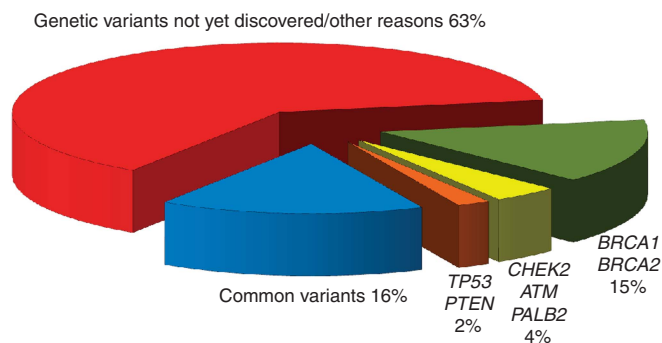
BJC OPEN

Figure 1. Proportions of familial risk of breast cancer explained by hereditary variants.

inform public health strategies for targeted prevention. Gene–environment interaction studies may also help to discover novel genetic risk factors; for example, variants in *NAT2* and *GSTM1* may not have been identified as relevant genetic risk factors, if their association with bladder cancer risk would have been evaluated in a population of non-smokers. In the same way, novel environmental risk factors might be identified when the environmental exposure is a mixture of several composites (Hunter, 2005). For example, when investigating whole-body fat distribution, with new techniques such as MRI, as a breast cancer risk factor, investigating gene–environment interactions may help to determine probable risk increasing or protective relationships between breast cancer risk and different adipose tissue compartments, such as subcutaneous or visceral fat. Gene–environment interactions studies may also be used to gain insight into the biological mechanisms underlying an association between a risk factor and the disease of interest. In the bladder cancer example, gene–environment interaction studies pointed to aromatic amines as the most important bladder carcinogen in tobacco smoke.

To investigate gene–environment interaction, information on both components of the relationship has to be available. Most often, gene–environment interaction is evaluated on the multiplicative scale, that is, investigators assess whether the joint effect measured as relative risks or odds ratios of the genetic and the environmental factor is significantly greater or smaller than would be expected by multiplying the relative risks for the genetic or environmental factor alone. However, gene–environment interaction can also be assessed as departure from the additive model, when looking at risk differences rather than ratio measures.

One of the biggest problems to overcome is to gather the sample size needed to investigate gene–environment interaction with sufficient power. As a general rule of thumb, studies aiming at evaluating departure from multiplicativity have to be at least four times the size needed to investigate effects of the genetic or environmental factor alone (Smith and Day, 1984). This review focuses therefore on studies analysing a large sample coming from a single study, and on studies in which data from multiple studies were pooled to achieve greater power.

STUDIES ASSESSING THE POTENTIAL MODIFICATION OF KNOWN BREAST CANCER RISK VARIANTS BY ENVIRONMENTAL FACTORS

Several studies have investigated whether breast cancer risk associated with the common susceptibility variants might be modified by environmental factors. Here the investigators formally tested for gene–environment interaction on the multiplicative scale. Details on the studies can be found in Table 1. Prior to the large studies, a number of smaller studies investigated possible interactions between the known susceptibility variants in *FGFR2*

and menopausal hormone therapy and yielded inconclusive results (Travis *et al*, 2010). In the beginning, studies included a limited set of single-nucleotide polymorphism (SNPs), which were identified by the first genome-wide association studies (GWAS) to be associated with breast cancer risk.

Travis *et al* (2010), investigated gene–environment interactions between 12 SNPs and several risk factors, analysing a sample of 7610 breast cancer cases and 10 196 controls from the Million Women Study. After accounting for the number of tests performed, none of the gene–environment interactions investigated reached statistical significance (Travis *et al*, 2010). The study of Milne *et al* (2010), which was based on data from up to 26 349 breast cancer cases and 32 208 controls from 21 case–control studies participating in the Breast Cancer Association Consortium (BCAC), confirmed the absence of significant gene–environment interactions after accounting for multiple testing. This study investigated 12 SNPs, 9 of them overlapping with those included in the study by Travis *et al* (2010) and (Milne *et al*, 2010). An investigation in 8576 breast cancer cases and 11 892 controls nested within the Breast and Prostate Cancer Cohort Consortium (BPC3) also found no evidence for gene–environment interactions between 17 common susceptibility loci and 9 environmental factors investigated (Campa *et al*, 2011).

The second gene–environment interaction study within BCAC (Nickels *et al*, 2013) assessed 11 additional newly identified susceptibility SNPs (23 SNPs in total) using a larger sample size than before, including up to 34 793 breast cancers and 41 099 controls. Additional environmental risk factors were also investigated, such as use of menopausal hormone therapy and alcohol consumption. This study replicated two potential interactions that had previously been reported by Travis *et al* (2010) and Milne *et al* (2010) but deemed statistically non-significant. In the Million Women Study, the per-allele odds ratio (OR) for *CASP8*-rs1045485 was 0.99 (95% confidence interval (CI) 0.92–1.07) in women who reported to consume on average less than one drink per day and 1.23 (95% CI 1.09–1.38) in those who reported to consume one or more drinks per day (*P*-value for interaction = 0.003; Travis *et al*, 2010). In BCAC, *CASP8*-rs1045485 was also associated with an increased breast cancer risk in women who consumed $\geqslant 20$ g per day alcohol (per-allele OR 1.45, 95% CI 1.14–1.85). In women who consumed less, an inverse association was present (per-allele OR 0.91, 95% CI 0.84–0.98, *P*-value for interaction = 0.0003). The second confirmed interaction was between *LSP1*-rs3817198 and number of live births. In Milne *et al* (2010), the association between *LSP1*-rs3817198 and breast cancer was strongest in women having had at least four live births (per-allele OR 1.24, 95% CI 1.11–1.38) and weaker for those having less births, with the weakest association in women having only one live birth (per-allele OR 1.04, 95% CI 0.97–1.11, *P*-value for interaction = 0.002). The per-allele ORs for *LSP1*-rs3817198 was very similar in the larger data set analysed by Nickels *et al* (2013) with per-allele ORs 1.26 (95% CI 1.16–1.37) and 1.08 (95% CI 1.01–1.16), respectively (*P*-value for interaction = 2.4×10^{-6}). When restricting the data set to 6266 cases and 3899 controls not included in the previous report, the interaction remained significant (*P*-value for interaction = 0.002). Furthermore, Nickels *et al* (2013) reported a novel potential interaction between 1p11.2-rs11249433 and ever being parous (*P*-value for interaction = 5.3×10^{-5}), with a per-allele OR of 1.14 (95% CI 1.11–1.17) in parous women and 0.98 (95% CI 0.92–1.05) in nulliparous women (Nickels *et al*, 2013).

The reported interactions were followed up in another study conducted within BPC3 (Barrdahl *et al*, 2014). The study comprised up to 16 285 breast cancer cases and 19 376 controls and investigated 39 SNPs. A meta-analysis with results for 19 SNPs reported by BCAC was also included. The results for the two interactions confirmed by Nickels *et al* (2013) between *CASP8*-

Table 1. Overview of studies investigating gene–environment interaction for breast cancer risk in the general population

Study	N cases	N controls	Study design	Genetic markers	Environmental risk factors	Summary of findings	Assumptions power G × Parity[a]	ORs for interaction of G × Parity, detectable with 80% power, at a nominal $\alpha = 0.05$ and at 'genome-wide significant' $\alpha = 5 \times 10^{-8}$
Studies investigating gene–environment interaction with breast cancer susceptibility loci								
Milne et al, 2010	26349	32208	Pooled analysis of 21 case–control studies	12 SNPs	Age at menarche, ever having had a live birth, number of live births, age at first birth, BMI	No significant evidence for G × E after accounting for multiple testing	N cases = 20000; 1.30 controls per case; prevalence nulliparity = 0.13; OR nulliparity = 1.19	1.12; 1.29
Travis et al, 2010	7610	10196	Case-cohort (cases + randomly selected women without BC)	12 SNPs	Age at menarche, parity, age at first birth, breastfeeding, menopausal status, age at menopause, use of menopausal hormone therapy, BMI, height, alcohol consumption	No significant evidence for G × E after accounting for multiple testing	N cases = 6537; 1.37 controls per case; prevalence nulliparity = 0.13; OR nulliparity = 1.27	1.21; 1.53
Campa et al, 2011	8576	11892	Pooled analysis of 6 cohort studies	17 SNPs	Age at menarche, parity, age at menopause, use of menopausal hormone therapy, family history, height, BMI, smoking status, alcohol consumption	No significant evidence for G × E after accounting for multiple testing	N cases = 8389; 1.39 controls per case; prevalence nulliparity = 0.13; OR nulliparity = 1.23	1.19; 1.46
Nickels et al, 2013	34793	41099	Pooled analysis of 22 case–control and 2 nested case–control studies	23 SNPs	Age at menarche, parity, breastfeeding, BMI, height, oral contraceptive use, use of menopausal hormone therapy, alcohol consumption, cigarette smoking, physical activity	Replication: LSP1-rs3817198 and number of full-term pregnancies (Milne et al, 2010), CASP8-rs17468277 and alcohol consumption (Travis et al, 2010), novel: 1p11.2-rs11249433 and ever being parous	N cases = 28469; 1.02 controls per case; prevalence nulliparity = 0.14; OR nulliparity = 1.25	1.11; 1.24
Barrdahl et al, 2014	16285	19376	Pooled analysis of 8 nested case–control studies, meta-analysis with BCAC data (Nickels et al, 2013)	39 SNPs	Age at menarche, age at menopause, parity, BMI, height, smoking, alcohol consumption, family history	No significant evidence for G × E after accounting for multiple testing	N cases = 13812; 1.21 controls per case; prevalence nulliparity = 0.10; OR nulliparity = 1.22	1.17; 1.41
Rudolph et al, 2015	26633	30119	Pooled analysis of 20 case–control and 2 nested case–control studies	47 SNPs (41 for overall breast cancer risk, 6 for oestrogen receptor negative breast cancer)	Age at menarche, ever having had a full-term pregnancy, number of full-term pregnancies, breastfeeding, age at first full-term pregnancy, BMI in premenopausal women, height, current use of combined oestrogen-progesterone therapy, current use of oestrogen-only therapy, alcohol consumption, current smoking, smoking amount	No significant evidence for G × E after accounting for multiple testing	N cases = 26633; 1.13 controls per case; prevalence nulliparity = 0.15; OR nulliparity = 1.31	1.10; 1.24

Table 1. (Continued)

Exploratory studies of gene–environment interaction for breast cancer risk

Study	N cases	N controls	Study design	Genetic markers	Environmental risk factors	Summary of findings	Assumptions power G × Parity[a]	ORs for interaction of G × Parity, detectable with 80% power, at a nominal $\alpha = 0.05$ and at a 'genome-wide significant' $\alpha = 5 \times 10^{-8}$
Hein et al, 2013	Stage I: 731; Stage II: 2170; Replication: 5795	Stage I: 0; Stage II: 2738; Replication: 5390	Stage I: Case-only analysis; Stage II: Pooled analysis of two case–control studies; meta-analysis of stage I+II, Replication stage: Pooled analysis of 9 case–control studies	Genome-wide (300K SNPs)	Use of menopausal hormone therapy	No significant evidence for G × E after accounting for multiple testing.		
Rudolph et al, 2013	Discovery: 2920; Replication: 7689	Discovery: 0; eplication Replication: 9266	Discovery: 4 case-only GEWIS, meta-analysis; Replication: Pooled analysis of 11 case–control studies	Genome-wide (2.5M SNPs)	Use of menopausal hormone therapy	No significant evidence for G × E after accounting for multiple testing.		
Schoeps et al, 2014	34 475	34 786	Pooled analysis of 21 case–control and 2 nested case–control studies	71527 SNPs enriched for association with breast cancer	Number of full-term pregnancies, age at menarche, height, BMI in postmenopausal women, BMI in premenopausal women, duration of oral contraceptive use, duration of menopausal oestrogen–progesterone therapy in current users, duration of menopausal oestrogen-only therapy in current users, alcohol consumption, family history	rs10483028, rs2242714 on 21q22.12 and postmenopausal BMI		

Abbreviations: BMI = body mass index; G × E = gene–environment interaction; GEWIS = Genome-wide gene–environment interaction study; OR = odds ratio; SNP = single-nucleotide polymorphism.
[a]Assuming a minor-allele frequency of 0.35, log-additive inheritance mode, an OR for the marginal association between the genetic variant and breast cancer risk of 1.10 and population prevalence of disease of 1%.

rs1045485 and alcohol consumption, and between *LSP1*-rs3817198 and number of live births were not replicated due to inconsistent results between BCAC and BPC3. The OR for interaction between *LSP1*-rs3817198 and number of live births in BPC3 was 0.97 (95% CI 0.94–1.00), whereas in BCAC a positive departure from a multiplicative model was observed (OR for interaction = 1.06 (95% CI 1.04–1.09). Similarly, no interaction between *CASP8*-rs1045485 and alcohol consumption was found in BPC3 (OR for interaction = 0.96 (95% CI 0.81–1.15)), whereas in BCAC the joint risks of alcohol intake and the *CASP8*-variant were super-multiplicative (OR for interaction = 1.59 (95% CI 1.24–2.05).

For the interaction between 1p11.2-rs11249433 and ever being parous, the OR for interaction in BPC3 was in the same direction as in BCAC (OR for interaction BCAC = 1.16 (95% CI 1.08–1.24); OR for interaction BPC3 = 1.07 (95% CI 0.95–1.20); OR for interaction META = 1.13 (95% CI 1.07–1.20)). Here the lack of replication could be attributed to the smaller sample size of the study by Barrdahl *et al* (2014) in BPC3 compared with the study by Nickels *et al* (2013) in BCAC.

The heterogeneity in results of *CASP8*-rs1045485 and *LSP1*-rs3817198 between BCAC and BPC3 are unlikely to be explained by differences in the genetic make-up of the populations, since both were restricted to Caucasians and the minor-allele frequencies were comparable between BCAC and BPC3 samples. Although self-reports of the number of live births and other parity-related factors are generally accurate (Olson *et al*, 1997), this is often less the case for alcohol intake (Thiebaut *et al*, 2007). This may be one reason for the lack of replication of the interaction between *CASP8*-rs1045485 and alcohol consumption since misclassification biases the multiplicative interaction parameter towards the null (Garcia-Closas *et al*, 1999). Apart from that, many other sources of heterogeneity between the included studies in BCAC and BPC3 could potentially lead to inconsistent results, for example, differences in ranges of exposure levels, time period in which the recruitment took place (e.g. 80 s or early 2000 s), and study design. Also, the interaction between *LSP1*-rs3817198 and number of pregnancies was modelled similarly in BCAC and BPC3, but this was not the case for the interaction between *CASP8*-rs1045485 and alcohol consumption. In BCAC, alcohol consumption was modelled as a binary variable (<20 g per day versus $\geqslant 20$ g per day; Nickels *et al*, 2013), whereas in BPC3 alcohol consumption was modelled as a linear continuous variable (in g per day) (Barrdahl *et al*, 2014).

Further efforts in BCAC led to the identification of 41 novel susceptibility variants for breast cancer (Michailidou *et al*, 2013). Potential modifications by environmental factors of the associations of these 41 variants as well as 6 additional variants associated with oestrogen receptor negative (ER−) breast cancer were evaluated in another study within BCAC (Rudolph *et al*, 2015). The data set was fairly large for some of the investigated gene–environment interactions (26 633 breast cancer cases and 30 119 controls) and, as can be seen from *post hoc* power calculations presented in Table 1, the power was sufficient to detect moderately sized gene–environment interactions. However, the interactions identified were small, so that none of them reached statistical significance after accounting for the number of tests performed (Rudolph *et al*, 2015). The power presented in Table 1 for the different studies are calculated for an assumed minor-allele frequency of 0.35. Figure 2 illustrates that power varies tremendously in gene–environment interaction studies depending on the frequency of the minor allele of the SNP investigated.

Despite power issues, studies of gene–environmental interaction could contribute towards gaining insight into the biological mechanisms underlying modifying effects of susceptibility alleles with respect to breast cancer risk. In Nickels *et al* (2013), all findings with a *P*-value for interaction below 10^{-4} were reported, including that between 2q35-rs13387042 and current combined
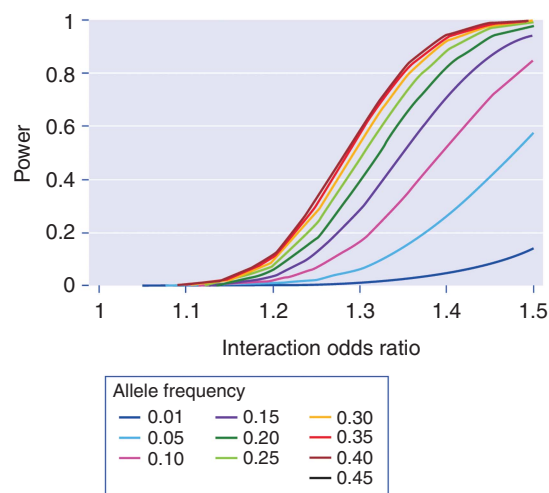

Figure 2. Power for detecting gene–environment interaction given different allele frequencies. Power was calculated with Quanto 1.2.4, assuming a log-additive inheritance mode, a population prevalence of disease of 1%, an OR of 1.10 for the marginal association between SNP and disease, an OR of 1.20 for the marginal association between environmental factor and disease, a prevalence of the environmental factor of 0.15, a sample of 10 000 unmatched case–control pairs and an two-sided alpha of 5×10^{-6}.

oestrogen/progestagen menopausal hormone therapy use, which was restricted to oestrogen-receptor-positive disease. At that time, the gene involved at this locus was still unknown. The closest known genes were *TNP1* (transition protein 1), *IGFBP5* (insulin-like growth-factor-binding protein 5), *IGFBP2* (insulin-like growth-factor-binding protein 2) and *TNS1* (tensin 1/matrix-remodelling-associated protein 6). The authors wrote that 'observed effect modification would suggest that the gene involved may be responsive to steroid hormones' (Nickels *et al*, 2013), which would have implicated the involvement of *IGFBP5* or *IGFBP2*. Indeed, fine mapping of the 2q35 locus and functional studies led to the recognition that the G-allele downregulates *IGFBP5* (Ghoussaini *et al*, 2014).

Interaction effects may be different for rarer variants with higher penetrance. A number of studies have thus assessed the effects of environmental factors on cancer occurrence among *BRCA1/2* mutation carriers (e.g. Brohet *et al*, 2007). No large studies, that is, with at least 10 000 cases, investigated the interactions between mutations in *BRCA1/2* or other high-risk genes and environmental factors. Therefore, gene–environment interaction was not formally investigated as departure from a multiplicative or additive model. One meta-analysis of relatively small studies was published, in which modifiers of cancer risk in *BRCA1* or *BRCA2* mutation carriers were reviewed (Friebel *et al*, 2014). Late age at first birth was identified as a probable factor modifying the risk of *BRCA1* mutation carriers. A first birth at age 30 years or older was associated with a 35% (pooled estimate 0.65, 95% CI 0.42–0.99) reduction in risk of breast cancer compared with a first birth at an age younger than 30 years (Friebel *et al*, 2014). Friebel *et al* (2014), mention two possible explanations for this observation: being older when giving birth for the first time may have differential effects in women carrying an *BRCA1* mutation compared with women without *BRCA1* mutation, or the observed associations may be due to risk-reducing oophor-ectomy or bias in the ascertainment. In the general population, an older age at first birth is associated with an increased breast cancer risk (Reeves *et al*, 2009). Thus, if the associations between age at first birth and *BRCA1* mutation are unbiased and differ depending on the presence of a *BRCA1* mutation, then gene–environment

interaction is present. However, data to assess the modification of breast cancer risk associated with rare variants like *BRCA1* and *BRCA2* are sparse and there is even less data for other rare variants like *CHEK2*. Whether gene–environment interactions are more common or stronger with rarer genetic variants has to be shown in future investigations.

To summarise, the associations between the susceptibility loci known today and breast cancer risk are not likely to be strongly modified by environmental factors. The potential gene–environment interactions that have been identified are of small to moderate magnitude. With interaction effects of this size, differences between additive and multiplicative interactions are subtle, as departures from either the multiplicative or additive model are minimal. A simulation study by Aschard *et al* (2012), showed that the inclusion of gene–gene and gene–environment interactions with small effects (relative risks between 0.5 and 2.0) is unlikely to strongly improve risk prediction for breast cancer. Apart from that, for most breast cancer susceptibility loci the causal variants have still to be to be discovered as well as the underlying biological mechanisms.

It has recently been shown that combining small effects from common genetic susceptibility loci into a genetic risk score helps to discriminate women with low and high breast cancer risk (Mavaddat *et al*, 2015). In a study by Li *et al* (2013), among Chinese women, the association between a genetic risk score constructed with 10 SNPs and breast cancer risk was not modified by any of the investigated environmental risk factors. One example, where several small single-SNP gene–environment interactions translated into a significant interaction between the respective genetic risk score and the environmental risk factor is a recent study by Qi *et al* (2014), with respect to BMI. In this study, consuming fried foods four or more times a week compared with consuming fried foods less than once a week was associated with an increase of $1 \, \mathrm{kg \, m^{-2}}$ (s.e. $0.2 \, \mathrm{kg \, m^{-2}}$) in BMI among women within the upper tertile of the genetic score. Among women in the lower tertile, frequent consumption of fried foods was associated with an increase in BMI of $0.5 \, \mathrm{kg \, m^{-2}}$ (SE $0.2 \, \mathrm{kg \, m^{-2}}$), *P* for interaction = 0.005). Similar results were observed for men. Of the 32 variants included in the genetic score, 4 (*FTO*, *GNPDA2*, *NEGR1*, and *SEC16B*) showed nominally significant interactions with total fried-food consumption on BMI (*P* for interaction < 0.05) and all interactions were in the same direction (that is, stronger association between fried-food consumption and BMI in carriers of one or more BMI increasing alleles). Therefore, if multiple unidirectional gene–environment interactions are present, then considering these may further improve discrimination, although the effect of each interaction may be small.

EXPLORATORY STUDIES AIMING TO FIND NEW GENETIC RISK FACTORS AND GAIN AETIOLOGICAL INSIGHT

Instead of variants known to be associated with breast cancer risk, explorative gene–environment interaction studies investigate large numbers of variants or use a genome-wide approach. This approach offers the possibility to identify new genetic risk factors and to gain knowledge on the biological mechanisms underlying the associations between environmental risk factors and breast cancer risk.

In 2013, Hein *et al* (2013) published a GWAS that aimed to identify potential genetic modifiers of the association between menopausal hormone therapy use and ductal and lobular breast cancer risk. In the first stage, a case-only GWAS was conducted with current use of menopausal hormone therapy as the dependent variable in 731 breast cancer cases of a German study (MARIE). The 1200 SNPs with lowest *P*-values for interaction were selected for stage two. In stage two, gene–environment interactions were tested using case–control analyses in pooled sample of additional MARIE cases ($N = 1375$) and controls ($N = 1974$) as well as 795 cases and 764 controls from a Swedish case–control study (SASBAC). The lowest joint *P*-value for stage one and two was used to identify the interaction between variant rs6707272 on chromosome 2 and current use of MHT (joint *P*-value for interaction = 3.0×10^{-7}). This result was followed up in a replication stage using 5795 cases and 5390 controls from nine studies of the BCAC but the interaction could not be replicated (*P*-value for interaction = 0.21). Most likely this study suffered from an underpowered replication stage, and from an overestimation of the interaction effect in stages one and two, a phenomenon known as 'the winners curse' (Kraft, 2008). Rudolph *et al* (2013), used a larger sample both for detection and replication, conducting a second genome-wide study with a case-only design in the first stage to identify potential interactions between SNPs and menopausal hormone therapy use on breast cancer risk. A meta-analysis of four studies comprising 2920 cases in total yielded 1391 variants showing interactions with use of menopausal hormone therapy with *P*-values for interaction < 0.003, which were followed up. The replication stage based on case–control analysis comprising 7689 cases and 9266 controls from 11 studies in BCAC failed to detect any interactions at genome-wide significance level (Rudolph *et al*, 2013). These results indicate that strong modifiers of the established association between menopausal hormone therapy use and breast cancer risk are unlikely and larger studies will be required to detect more moderate interactions.

An exploratory study with the aim to identify novel genetic susceptibility loci for breast cancer was published in 2014 (Schoeps *et al*, 2014). In total, 71 527 SNPs were tested for interaction with 10 environmental risk factors of breast cancer. The SNPs were enriched for association with breast cancer risk as they had been proposed by BCAC consortium members to be genotyped using a custom Illumina iSelect genotyping array (iCOGS). To gain power, three recently developed two-step methods to test gene–environment interactions as well as a joint test of association and interaction were employed in the study [22]. Two SNPs (rs10483028 and rs2242714) in strong linkage disequilibrium located on chromosome 21 showed statistically significant interactions with BMI in postmenopausal women. The variants were identified by all four methods applied. A third variant, rs12197388 on chromosome 6, was detected through the joint test (*P*-value 2df test $< 7.0 \times 10^{-7}$), based on its association with breast cancer but did not show significant interaction. Overall, this study exemplifies that new risk loci can be discovered by accounting for gene–environment interaction.

For breast cancer, gene–environment interaction studies aiming at identifying new environmental risk factors are lacking. The main reason is the difficulty to obtain comprehensive reliable data on environmental exposure for the large sample sizes that would be needed, and lack of sufficient studies with biological specimens for for example, metabolomics or proteomics. Further technical development may help to overcome this problem similarly to the vast advances that have been made in measuring genomic variability.

CONCLUSION

Studying gene–environment interaction offers the opportunity to obtain new knowledge on several aspects of breast cancer aetiology and it may help to improve prevention. Although several comprehensive large-scale gene–environment interaction studies have been conducted for breast cancer risk, relevant gene–environment interaction has not been established to date. The potential interactions that have been identified with common

genetic variants are of small to moderate size, and they are unlikely to improve risk prediction individually, but may be helpful when combined. Stronger interactions could be present with rarer variants, but this has still to be examined. Presently, we should consider hereditary variants and environmental factors as multiplicative/additive factors in the prediction of breast cancer risk. Hence in risk prediction models both should be included and modelled as independent factors pending further research.

Although the few exploratory studies that attempted to identify new genetic susceptibility loci by accounting for gene–environment interaction showed interesting results, this approach is likely to be more successful with the availability of larger data resources. The large genotyping experiment currently being conducted by the OncoArray consortium (2013) together with the resources already available in BCAC and BPC3 will create opportunities to further study gene–environment interactions.

REFERENCES

(2013) Consortium launches genotyping effort. *Cancer Discov* 3(12): 1321–1322.

Aschard H, Chen J, Cornelis MC, Chibnik LB, Karlson EW, Kraft P (2012) Inclusion of gene-gene and gene-environment interactions unlikely to dramatically improve risk prediction for complex diseases. *Am J Hum Genet* 90(6): 962–972.

Barrdahl M, Canzian F, Joshi AD, Travis RC, Chang-Claude J, Auer PL, Gapstur SM, Gaudet M, Diver WR, Henderson BE, Haiman CA, Schumacher FR, Le Marchand L, Berg CD, Chanock SJ, Hoover RN, Rudolph A, Ziegler RG, Giles GG, Baglietto L, Severi G, Hankinson SE, Lindstrom S, Willet W, Hunter DJ, Buring JE, Lee IM, Zhang S, Dossus L, Cox DG, Khaw KT, Lund E, Naccarati A, Peeters PH, Quiros JR, Riboli E, Sund M, Trichopoulos D, Prentice RL, Kraft P, Kaaks R, Campa D (2014) Post-GWAS gene-environment interplay in breast cancer: results from the Breast and Prostate Cancer Cohort Consortium and a meta-analysis on 79,000 women. *Hum Mol Genet* 23(19): 5260–5270.

Brohet RM, Goldgar DE, Easton DF, Antoniou AC, Andrieu N, Chang-Claude J, Peock S, Eeles RA, Cook M, Chu C, Nogues C, Lasset C, Berthet P, Meijers-Heijboer H, Gerdes AM, Olsson H, Caldes T, van Leeuwen FE, Rookus MA (2007) Oral contraceptives and breast cancer risk in the international BRCA1/2 carrier cohort study: a report from EMBRACE, GENEPSO, GEO-HEBON, and the IBCCS Collaborating Group. *J Clin Oncol* 25(25): 3831–3836.

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, Henderson BE, Hoover RN, Isaacs C, Johansson M, Kolonel LN, Kraft P, Lee IM, McCarty CA, Overvad K, Panico S, Peeters PH, Riboli E, Sanchez MJ, Schumacher FR, Skeie G, Stram DO, Thun MJ, Trichopoulos D, Zhang S, Ziegler RG, Hunter DJ, Lindstrom S, Canzian F (2011) Interactions between genetic variants and breast cancer risk factors in the breast and prostate cancer cohort consortium. *J Natl Cancer Inst* 103(16): 1252–1263.

Friebel TM, Domchek SM, Rebbeck TR (2014) Modifiers of cancer risk in BRCA1 and BRCA2 mutation carriers: systematic review and meta-analysis. *J Natl Cancer Inst* 106(6): dju091.

Garcia-Closas M, Malats N, Silverman D, Dosemeci M, Kogevinas M, Hein DW, Tardon A, Serra C, Carrato A, Garcia-Closas R, Lloreta J, Castano-Vinyals G, Yeager M, Welch R, Chanock S, Chatterjee N, Wacholder S, Samanic C, Tora M, Fernandez F, Real FX, Rothman N (2005) NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. *Lancet* 366(9486): 649–659.

Garcia-Closas M, Rothman N, Lubin J (1999) Misclassification in case-control studies of gene-environment interactions: assessment of bias and sample size. *Cancer Epidemiol Biomarkers Prev* 8(12): 1043–1050.

Ghoussaini M, Edwards SL, Michailidou K, Nord S, Cowper-Sal Lari R, Desai K, Kar S, Hillman KM, Kaufmann S, Glubb DM, Beesley J, Dennis J, Bolla MK, Wang Q, Dicks E, Guo Q, Schmidt MK, Shah M, Luben R, Brown J, Czene K, Darabi H, Eriksson M, Klevebring D, Bojesen SE, Nordestgaard BG, Nielsen SF, Flyger H, Lambrechts D, Thienpont B, Neven P, Wildiers H, Broeks A, Van't Veer LJ, Th Rutgers EJ, Couch FJ, Olson JE, Hallberg E, Vachon C, Chang-Claude J, Rudolph A, Seibold P, Flesch-Janys D, Peto J, Dos-Santos-Silva I, Gibson L, Nevanlinna H, Muranen TA, Aittomaki K,

Blomqvist C, Hall P, Li J, Liu J, Humphreys K, Kang D, Choi JY, Park SK, Noh DY, Matsuo K, Ito H, Iwata H, Yatabe Y, Guenel P, Truong T, Menegaux F, Sanchez M, Burwinkel B, Marme F, Schneeweiss A, Sohn C, Wu AH, Tseng CC, Van Den Berg D, Stram DO, Benitez J, Zamora MP, Perez JI, Menendez P, Shu XO, Lu W, Gao YT, Cai Q, Cox A, Cross SS, Reed MW, Andrulis IL, Knight JA, Glendon G, Tchatchou S, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N, Haiman CA, Henderson BE, Schumacher F, Le Marchand L, Lindblom A, Margolin S, Teo SH, Yip CH, Lee DS, Wong TY, Hooning MJ, Martens JW, Collee JM, van Deurzen CH, Hopper JL, Southey MC, Tsimiklis H, Kapuscinski MK, Shen CY, Wu PE, Yu JC, Chen ST, Alnaes GG, Borresen-Dale AL, Giles GG, Milne RL, McLean C, Muir K, Lophatananon A, Stewart-Brown S, Siriwanarangsan P, Hartman M, Miao H, Buhari SA, Teo YY, Fasching PA, Haeberle L, Ekici AB, Beckmann MW, Brenner H, Dieffenbach AK, Arndt V, Stegmaier C, Swerdlow A, Ashworth A, Orr N, Schoemaker MJ, Garcia-Closas M, Figueroa J, Chanock SJ, Lissowska J, Simard J, Goldberg MS, Labreche F, Dumont M, Winqvist R, Pylkas K, Jukkola-Vuorinen A, Brauch H, Bruning T, Koto YD, Radice P, Peterlongo P, Bonanni B, Volorio S, Dork T, Bogdanova NV, Helbig S, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Devilee P, Tollenaar RA, Seynaeve C, Van Asperen CJ, Jakubowska A, Lubinski J, Jaworska-Bieniek K, Durda K, Slager S, Toland AE, Ambrosone CB, Yannoukakos D, Sangrajrang S, Gaborieau V, Brennan P, McKay J, Hamann U, Torres D, Zheng W, Long J, Anton-Culver H, Neuhausen SL, Luccarini C, Baynes C, Ahmed S, Maranian M, Healey CS, Gonzalez-Neira A, Pita G, Alonso MR, Alvarez N, Herrero D, Tessier DC, Vincent D, Bacot F, de Santiago I, Carroll J, Caldas C, Brown MA, Lupien M, Kristensen VN, Pharoah PD, Chenevix-Trench G, French JD, Easton DF, Dunning AM. Australian Ovarian Cancer Management G, Australian Ovarian Cancer Management G (2014) Evidence that breast cancer risk at the 2q35 locus is mediated through IGFBP5 regulation. *Nat Commun* 4: 4999.

Hein R, Flesch-Janys D, Dahmen N, Beckmann L, Lindström S, Schoof N, Czene K, Mittelstraβ K, Illig T, Seibold P, Behrens S, Humphreys K, Li J, Liu J, Olson JE, Wang X, Hankinson SE, Truong T, Menegaux F, Dos Santos Silva I, Johnson N; GENICA Network, Chen ST, Yu JC, Ziogas A, Kataja V, Kosma VM, Mannermaa A, Anton-Culver H, Shen CY, Brauch H, Peto J, Guénel P, Kraft P, Couch FJ, Easton DF, Hall P, Chang-Claude J (2013) A genome-wide association study to identify genetic susceptibility loci that modify ductal and lobular postmenopausal breast cancer risk associated with menopausal hormone therapy use: a two-stage design with replication. *Breast Cancer Res Treat* 138(2): 529–542.

Hunter DJ (2005) Gene-environment interactions in human diseases. *Nat Rev Genet* 6(4): 287–298.

Kraft P (2008) Curses—winner's and otherwise–in genetic epidemiology. *Epidemiology* 19(5): 649–651discussion 657–658.

Li HX, Beeghly-Fadiel A, Wen WQ, Lu W, Gao YT, Xiang YB, Cai QY, Long JR, Shi JJ, Chen KX, Zheng Y, Shu XO, Zheng W (2013) Gene-environment interactions for breast cancer risk among Chinese women: a report from the Shanghai Breast Cancer Genetics Study. *Am J Epidemiol* 177(2): 161–170.

Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, Pukkala E, Skytthe A, Hemminki K (2000) Environmental and heritable factors in the causation of cancer–analyses of cohorts of twins from Sweden, Denmark, and Finland. *N Engl J Med* 343(2): 78–85.

Mavaddat N, Pharoah PD, Michailidou K, Tyrer J, Brook MN, Bolla MK, Wang Q, Dennis J, Dunning AM, Shah M, Luben R, Brown J, Bojesen SE, Nordestgaard BG, Nielsen SF, Flyger H, Czene K, Darabi H, Eriksson M, Peto J, Dos-Santos-Silva I, Dudbridge F, Johnson N, Schmidt MK, Broeks A, Verhoef S, Rutgers EJ, Swerdlow A, Ashworth A, Orr N, Schoemaker MJ, Figueroa J, Chanock SJ, Brinton L, Lissowska J, Couch FJ, Olson JE, Vachon C, Pankratz VS, Lambrechts D, Wildiers H, Van Ongeval C, van Limbergen E, Kristensen V, Grenaker Alnaes G, Nord S, Borresen-Dale AL, Nevanlinna H, Muranen TA, Aittomaki K, Blomqvist C, Chang-Claude J, Rudolph A, Seibold P, Flesch-Janys D, Fasching PA, Haeberle L, Ekici AB, Beckmann MW, Burwinkel B, Marme F, Schneeweiss A, Sohn C, Trentham-Dietz A, Newcomb P, Titus L, Egan KM, Hunter DJ, Lindstrom S, Tamimi RM, Kraft P, Rahman N, Turnbull C, Renwick A, Seal S, Li J, Liu J, Humphreys K, Benitez J, Pilar Zamora M, Arias Perez JI, Menendez P, Jakubowska A, Lubinski J, Jaworska-Bieniek K, Durda K, Bogdanova NV, Antonenkova NN, Dork T, Anton-Culver H, Neuhausen SL, Ziogas A, Bernstein L, Devilee P, Tollenaar RA, Seynaeve C, van Asperen CJ, Cox A, Cross SS, Reed MW, Khusnutdinova E, Bermisheva M, Prokofyeva D, Takhirova Z, Meindl A, Schmutzler RK, Sutter C, Yang R, Schurmann P, Bremer M, Christiansen H, Park-Simon TW, Hillemanns P, Guenel P, Truong T, Menegaux F, Sanchez M, Radice P,

Peterlongo P, Manoukian S, Pensotti V, Hopper JL, Tsimiklis H, Apicella C, Southey MC, Brauch H, Bruning T, Ko YD, Sigurdson AJ, Doody MM, Hamann U, Torres D, Ulmer HU, Forsti A, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N, Andrulis IL, Knight JA, Glendon G, Marie Mulligan A, Chenevix-Trench G, Balleine R, Giles GG, Milne RL, McLean C, Lindblom A, Margolin S, Haiman CA, Henderson BE, Schumacher F, Le Marchand L, Eilber U, Wang-Gohrke S, Hooning MJ, Hollestelle A, van den Ouweland AM, Koppert LB, Carpenter J, Clarke C, Scott R, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Brenner H, Arndt V, Stegmaier C, Karina Dieffenbach A, Winqvist R, Pylkas K, Jukkola-Vuorinen A, Grip M, Offit K, Vijai J, Robson M, Rau-Murthy R, Dwek M, Swann R, Annie Perkins K, Goldberg MS, Labreche F, Dumont M, Eccles DM, Tapper WJ, Rafiq S, John EM, Whittemore AS, Slager S, Yannoukakos D, Toland AE, Yao S, Zheng W, Halverson SL, Gonzalez-Neira A, Pita G, Rosario Alonso M, Alvarez N, Herrero D, Tessier DC, Vincent D, Bacot F, Luccarini C, Baynes C, Ahmed S, Maranian M, Healey CS, Simard J, Hall P, Easton DF, Garcia-Closas M (2015) Prediction of breast cancer risk based on profiling with common genetic variants. *J Natl Cancer Inst* **107**(5).

Michailidou K, Beesley J, Lindstrom S, Canisius S, Dennis J, Lush MJ, Maranian MJ, Bolla MK, Wang Q, Shah M, Perkins BJ, Czene K, Eriksson M, Darabi H, Brand JS, Bojesen SE, Nordestgaard BG, Flyger H, Nielsen SF, Rahman N, Turnbull C, Bocs, Fletcher O, Peto J, Gibson L, Dos-Santos-Silva I, Chang-Claude J, Flesch-Janys D, Rudolph A, Eilber U, Behrens S, Nevanlinna H, Muranen TA, Aittomaki K, Blomqvist C, Khan S, Aaltonen K, Ahsan H, Kibriya MG, Whittemore AS, John EM, Malone KE, Gammon MD, Santella RM, Ursin G, Makalic E, Schmidt DF, Casey G, Hunter DJ, Gapstur SM, Gaudet MM, Diver WR, Haiman CA, Schumacher F, Henderson BE, Le Marchand L, Berg CD, Chanock SJ, Figueroa J, Hoover RN, Lambrechts D, Neven P, Wildiers H, van Limbergen E, Schmidt MK, Broeks A, Verhoef S, Cornelissen S, Couch FJ, Olson JE, Hallberg E, Vachon C, Waisfisz Q, Meijers-Heijboer H, Adank MA, van der Luijt RB, Li J, Liu J, Humphreys K, Kang D, Choi JY, Park SK, Yoo KY, Matsuo K, Ito H, Iwata H, Tajima K, Guenel P, Truong T, Mulot C, Sanchez M, Burwinkel B, Marme F, Surowy H, Sohn C, Wu AH, Tseng CC, Van Den Berg D, Stram DO, Gonzalez-Neira A, Benitez J, Zamora MP, Perez JI, Shu XO, Lu W, Gao YT, Cai H, Cox A, Cross SS, Reed MW, Andrulis IL, Knight JA, Glendon G, Mulligan AM, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N. kConFab Investigators, Group A, Lindblom A, Margolin S, Teo SH, Yip CH, Taib NA, Tan GH, Hooning MJ, Hollestelle A, Martens JW, Collee JM, Blot W, Signorello LB, Cai Q, Hopper JL, Southey MC, Tsimiklis H, Apicella C, Shen CY, Hsiung CN, Wu PE, Hou MF, Kristensen VN, Nord S, Alnaes GI. NBCS, Giles GG, Milne RL, McLean C, Canzian F, Trichopoulos D, Peeters P, Lund E, Sund M, Khaw KT, Gunter MJ, Palli D, Mortensen LM, Dossus L, Huerta JM, Meindl A, Schmutzler RK, Sutter C, Yang R, Muir K, Lophatananon A, Stewart-Brown S, Siriwanarangsan P, Hartman M, Miao H, Chia KS, Chan CW, Fasching PA, Hein A, Beckmann MW, Haeberle L, Brenner H, Dieffenbach AK, Arndt V, Stegmaier C, Ashworth A, Orr N, Schoemaker MJ, Swerdlow AJ, Brinton L, Garcia-Closas M, Zheng W, Halverson SL, Shrubsole M, Long J, Goldberg MS, Labreche F, Dumont M, Winqvist R, Pylkas K, Jukkola-Vuorinen A, Grip M, Brauch H, Hamann U, Bruning T. GENICA Network, Radice P, Peterlongo P, Manoukian S, Bernard L, Bogdanova NV, Dork T, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Devilee P, Tollenaar RA, Seynaeve C, Van Asperen CJ, Jakubowska A, Lubinski J, Jaworska K, Huzarski T, Sangrajrang S, Gaborieau V, Brennan P, McKay J, Slager S, Toland AE, Ambrosone CB, Yannoukakos D, Kabisch M, Torres D, Neuhausen SL, Anton-Culver H, Luccarini C, Baynes C, Ahmed S, Healey CS, Tessier DC, Vincent D, Bacot F, Pita G, Alonso MR, Alvarez N, Herrero D, Simard J, Pharoah PP, Kraft P, Dunning AM, Chenevix-Trench G, Hall P, Easton DF (2015) Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat Genet* **47**(4): 373–380.

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, Lee A, Turnbull C, Rahman N. Breast and Ovarian Cancer Susceptibility Collaboration, Fletcher O, Peto J, Gibson L, Dos Santos Silva I, Nevanlinna H, Muranen TA, Aittomaki K, Blomqvist C, Czene K, Irwanto A, Liu J, Waisfisz Q, Meijers-Heijboer H, Adank M. Hereditary B, Ovarian Cancer Research Group Netherlands, van der Luijt RB, Hein R, Dahmen N, Beckman L, Meindl A, Schmutzler RK, Muller-Myhsok B, Lichtner P, Hopper JL, Southey MC, Makalic E, Schmidt DF, Uitterlinden AG, Hofman A, Hunter DJ, Chanock SJ, Vincent D, Bacot F, Tessier DC, Canisius S, Wessels LF, Haiman CA, Shah M, Luben R, Brown J, Luccarini C, Schoof N, Humphreys K, Li J, Nordestgaard BG, Nielsen SF, Flyger H,

Couch FJ, Wang X, Vachon C, Stevens KN, Lambrechts D, Moisse M, Paridaens R, Christiaens MR, Rudolph A, Nickels S, Flesch-Janys D, Johnson N, Aitken Z, Aaltonen K, Heikkinen T, Broeks A, Veer LJ, van der Schoot CE, Guenel P, Truong T, Laurent-Puig P, Menegaux F, Marme F, Schneeweiss A, Sohn C, Burwinkel B, Zamora MP, Perez JI, Pita G, Alonso MR, Cox A, Brock IW, Cross SS, Reed MW, Sawyer EJ, Tomlinson I, Kerin MJ, Miller N, Henderson BE, Schumacher F, Le Marchand L, Andrulis IL, Knight JA, Glendon G, Mulligan AM. kConFab Investigators, Australian Ovarian Cancer Study Group, Lindblom A, Margolin S, Hooning MJ, Hollestelle A, van den Ouweland AM, Jager A, Bui QM, Stone J, Dite GS. Apicella C, Tsimiklis H, Giles GG, Severi G, Baglietto L, Fasching PA, Haeberle L, Ekici AB, Beckmann MW, Brenner H, Muller H, Arndt V, Stegmaier C, Swerdlow A, Ashworth A, Orr N, Jones M, Figueroa J, Lissowska J, Brinton L, Goldberg MS, Labreche F, Dumont M, Winqvist R, Pylkas K, Jukkola-Vuorinen A, Grip M, Brauch H, Hamann U, Bruning T. GENICA (Gene Environment Interaction and Breast Cancer in Germany) Network, Radice P, Peterlongo P, Manoukian S, Bonanni B, Devilee P, Tollenaar RA, Seynaeve C, van Asperen CJ, Jakubowska A, Lubinski J, Jaworska K, Durda K, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM, Bogdanova NV, Antonenkova NN, Dork T, Kristensen VN, Anton-Culver H, Slager S, Toland AE, Edge S, Fostira F, Kang D, Yoo KY, Noh DY, Matsuo K, Ito H, Iwata H, Sueta A, Wu AH, Tseng CC, Van Den Berg D, Stram DO, Shu XO, Lu W, Gao YT, Cai H, Teo SH, Yip CH, Phuah SY, Cornes BK, Hartman M, Miao H, Lim WY, Sng JH, Muir K, Lophatananon A, Stewart-Brown S, Siriwanarangsan P, Shen CY, Hsiung CN, Wu PE, Ding SL, Sangrajrang S, Gaborieau V, Brennan P, McKay J, Blot WJ, Signorello LB, Cai Q, Zheng W, Deming-Halverson S, Shrubsole M, Long J, Simard J, Garcia-Closas M, Pharoah PD, Chenevix-Trench G, Dunning AM, Benitez J, Easton DF (2013) Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* **45**(4): 353–361.

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, Johnson N, Anton-Culver H, Ziogas A, Figueroa J, Brinton L, Sherman ME, Lissowska J, Hopper JL, Dite GS. Apicella C, Southey MC, Sigurdson AJ, Linet MS, Schonfeld SJ, Freedman DM, Mannermaa A, Kosma VM, Kataja V, Auvinen P, Andrulis IL, Glendon G, Knight JA, Weerasooriya N, Cox A, Reed MW, Cross SS, Dunning AM, Ahmed S, Shah M, Brauch H, Ko YD, Bruning T. GENICA NetworkLambrechts D, Reumers J, Smeets A, Wang-Gohrke S, Hall P, Czene K, Liu J, Irwanto AK, Chenevix-Trench G, Holland H. kConFab, AOCS, Giles GG, Baglietto L, Severi G, Bojensen SE, Nordestgaard BG, Flyger H, John EM, West DW, Whittemore AS, Vachon C, Olson JE, Fredericksen Z, Kosel M, Hein R, Vrieling A, Flesch-Janys D, Heinz J, Beckmann MW, Heusinger K, Ekici AB, Haeberle L, Humphreys MK, Morrison J, Easton DF, Pharoah PD, Garcia-Closas M, Goode EL, Chang-Claude J (2010) 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* **12**(6): R110.

Mucci LA, Kaprio J, Harris J, Czene K, Kraft P, Scheike T, Graff R, Brandt I, Holmes N, Havelick D, Hartman M, Penney K, Pukkala E, Parmigiani G, Skytthe A, Adami HO, Hjelmborg J, (NorTwinCan) NTSoC (2013) Heritability and familial risk of cancer: an update from the Nordic Twin Registry of Cancer (NorTwinCan). In: *The American Society of Human Genetics 63rd Annual Meeting*: Boston, MA, USA.

Nickels S, Truong T, Hein R, Stevens K, Buck K, Behrens S, Eilber U, Schmidt M, Haberle L, Vrieling A, Gaudet M, Figueroa J, Schoof N, Spurdle AB, Rudolph A, Fasching PA, Hopper JL, Makalic E, Schmidt DF, Southey MC, Beckmann MW, Ekici AB, Fletcher O, Gibson L, Dos Santos Silva I, Peto J, Humphreys MK, Wang J, Cordina-Duverger E, Menegaux F, Nordestgaard BG, Bojesen SE, Lanng C, Anton-Culver H, Ziogas A, Bernstein L, Clarke CA, Brenner H, Muller H, Arndt V, Stegmaier C, Brauch H, Bruning T, Harth V. The Genica Network, Mannermaa A, Kataja V, Kosma VM, Hartikainen JM. Kconfab, AOCS Management Group, Lambrechts D, Smeets D, Neven P, Paridaens R, Flesch-Janys D, Obi N, Wang-Gohrke S, Couch FJ, Olson JE, Vachon CM, Giles GG, Severi G, Baglietto L, Offit K, John EM, Miron A, Andrulis IL, Knight JA, Glendon G, Mulligan AM, Chanock SJ, Lissowska J, Liu J, Cox A, Cramp H, Connley D, Balasubramanian S, Dunning AM, Shah M, Trentham-Dietz A, Newcomb P, Titus L, Egan K, Cahoon EK, Rajaraman P, Sigurdson AJ, Doody MM, Guenel P, Pharoah PD, Schmidt MK, Hall P, Easton DF, Garcia-Closas M, Milne RL, Chang-Claude J (2013) Evidence of gene-

environment interactions between common breast cancer susceptibility loci and established environmental risk factors. *PLoS Genet* **9**(3): e1003284.

Olson JE, Shu XO, Ross JA, Pendergrass T, Robison LL (1997) Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the Children's Cancer Group. *Am J Epidemiol* **145**(1): 58–67.

Pharoah PD, Day NE, Duffy S, Easton DF, Ponder BA (1997) Family history and the risk of breast cancer: a systematic review and meta-analysis. *Int J Cancer* **71**(5): 800–809.

Qi Q, Chu AY, Kang JH, Huang J, Rose LM, Jensen MK, Liang L, Curhan GC, Pasquale LR, Wiggs JL, De Vivo I, Chan AT, Choi HK, Tamimi RM, Ridker PM, Hunter DJ, Willett WC, Rimm EB, Chasman DI, Hu FB, Qi L (2014) Fried food consumption, genetic risk, and body mass index: gene-diet interaction analysis in three US cohort studies. *BMJ* **348**: g1610.

Reeves GK, Pirie K, Green J, Bull D, Beral V. Million Women Study Collaborators (2009) Reproductive factors and specific histological types of breast cancer: prospective study and meta-analysis. *Br J Cancer* **100**(3): 538–544.

Rudolph A, Hein R, Lindstrom S, Beckmann L, Behrens S, Liu J, Aschard H, Bolla MK, Wang J, Truong T, Cordina-Duverger E, Menegaux F, Bruning T, Harth V. GENICA Network, Severi G, Baglietto L, Southey M, Chanock SJ, Lissowska J, Figueroa JD, Eriksson M, Humpreys K, Darabi H, Olson JE, Stevens KN, Vachon CM, Knight JA, Glendon G, Mulligan AM, Ashworth A, Orr N, Schoemaker M, Webb PM. kConFab Investigators, AOCS Management Group, Guenel P, Brauch H, Giles G, Garcia-Closas M, Czene K, Chenevix-Trench G, Couch FJ, Andrulis IL, Swerdlow A, Hunter DJ, Flesch-Janys D, Easton DF, Hall P, Nevanlinna H, Kraft P, Chang-Claude J. Breast Cancer Association Consortium (2013) Genetic modifiers of menopausal hormone replacement therapy and breast cancer risk: a genome-wide interaction study. *Endocr Relat Cancer* **20**(6): 875–887.

Rudolph A, Milne RL, Truong T, Knight JA, Seibold P, Flesch-Janys D, Behrens S, Eilber U, Bolla MK, Wang Q, Dennis J, Dunning AM, Shah M, Munday HR, Darabi H, Eriksson M, Brand JS, Olson J, Vachon CM, Hallberg E, Castelao JE, Carracedo A, Torres M, Li J, Humphreys K, Cordina-Duverger E, Menegaux F, Flyger H, Nordestgaard BG, Nielsen SF, Yesilyurt BT, Floris G, Leunen K, Engelhardt EG, Broeks A, Rutgers EJ, Glendon G, Mulligan AM, Cross S, Reed M, Gonzalez-Neira A, Arias Perez JI, Provenzano E, Apicella C, Southey MC, Spurdle A. kConFab Investigators, AOCS Group, Haberle L, Beckmann MW, Ekici AB, Dieffenbach AK, Arndt V, Stegmaier C, McLean C, Baglietto L, Chanock SJ, Lissowska J, Sherman ME, Bruning T, Hamann U, Ko YD, Orr N, Schoemaker M, Ashworth A, Kosma VM, Kataja V, Hartikainen JM, Mannermaa A, Swerdlow A, Network G, Giles GG, Brenner H, Fasching PA, Chenevix-Trench G, Hopper J, Benitez J, Cox A, Andrulis IL, Lambrechts D, Gago-Dominguez M, Couch F, Czene K, Bojesen SE, Easton DF, Schmidt MK, Guenel P, Hall P, Pharoah PD, Garcia-Closas M, Chang-Claude J (2015) Investigation of gene-environment interactions between 47 newly identified breast cancer susceptibility loci and environmental risk factors. *Int J Cancer* **136**(6): E685–E696.

Schoeps A, Rudolph A, Seibold P, Dunning AM, Milne RL, Bojesen SE, Swerdlow A, Andrulis I, Brenner H, Behrens S, Orr N, Jones M, Ashworth A, Li J, Cramp H, Connley D, Czene K, Darabi H, Chanock SJ, Lissowska J, Figueroa JD, Knight J, Glendon G, Mulligan AM, Dumont M, Severi G, Baglietto L, Olson J, Vachon C, Purrington K, Moisse M, Neven P, Wildiers H, Spurdle A, Kosma VM, Kataja V, Hartikainen JM, Hamann U, Ko YD, Dieffenbach AK, Arndt V, Stegmaier C, Malats N, Arias Perez JI, Benitez J, Flyger H, Nordestgaard BG, Truong T, Cordina-Duverger E, Menegaux F, dos Santos Silva I, Fletcher O, Johnson N, Haberle L, Beckmann MW, Ekici AB, Braaf L, Atsma F, van den Broek AJ, Makalic E, Schmidt DF, Southey MC, Cox A, Simard J, Giles GG, Lambrechts D, Mannermaa A, Brauch H, Guenel P, Peto J, Fasching PA, Hopper J, Flesch-Janys D, Couch F, Chenevix-Trench G, Pharoah PD, Garcia-Closas M, Schmidt MK, Hall P, Easton DF, Chang-Claude J (2014) Identification of new genetic susceptibility loci for breast cancer through consideration of gene-environment interactions. *Genet Epidemiol* **38**(1): 84–93.

Smith PG, Day NE (1984) The design of case-control studies: the influence of confounding and interaction effects. *Int J Epidemiol* **13**(3): 356–365.

Thiebaut AC, Freedman LS, Carroll RJ, Kipnis V (2007) Is it necessary to correct for measurement error in nutritional epidemiology? *Ann Intern Med* **146**(1): 65–67.

Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A (2015) Global cancer statistics, 2012. *CA Cancer J Clin* **65**(2): 87–108.

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 Collaborators (2010) Gene-environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study. *Lancet* **375**(9732): 2143–2151.

Wu C, Kraft P, Zhai K, Chang J, Wang ZM, Li Y, Hu ZB, He ZH, Jia WH, Abnet CC, Liang LM, Hu N, Miao XP, Zhou YF, Liu ZH, Zhan QM, Liu Y, Qiao Y, Zhou YL, Jin GF, Guo CH, Lu CD, Yang HJ, Fu JH, Yu DK, Freedman ND, Ding T, Tan W, Goldstein AM, Wu TC, Shen HB, Ke Y, Zeng YX, Chanock SJ, Taylor PR, Lin DX (2012) Genome-wide association analyses of esophageal squamous cell carcinoma in Chinese identify multiple susceptibility loci and gene-environment interactions. *Nat Genet* **44**(10): 1090–1097.