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

Rebecca Hein[#],
Unit of Genetic Epidemiology, Division of Cancer Epidemiology (C020), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany; PMV Research Group at the Department of Child and Adolescent Psychiatry and Psychotherapy, University of Cologne, Cologne, Germany

Dieter Flesch-Janys[#],
Department of Cancer Epidemiology, Clinical Cancer Registry, University Cancer Center and Institute for Medical Biometrics and Epidemiology, University Clinic Hamburg-Eppendorf, Hamburg, Germany

Norbert Dahmen,
Congenics AG, Hamburg, Germany

Lars Beckmann,
Unit of Genetic Epidemiology, Division of Cancer Epidemiology (C020), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany; Foundation for Quality and Efficiency in Health Care (IQWIG), Cologne, Germany

Sara Lindström,
Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, MA, USA; Department of Epidemiology, Harvard, School Of Public Health, Boston, MA 02138, USA; Department of Biostatistics, Harvard School Of Public Health, Boston, MA 02138, USA

Nils Schoof,
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Kamila Czene,
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Kirstin Mittelstraß,
Research Unit of Molecular Epidemiology, Helmholtz Center Munich, Munich, Germany

Thomas Illig,
Research Unit of Molecular Epidemiology, Helmholtz Center Munich, Munich, Germany; Hannover Unified Biobank, Hannover Medical School, Hannover, Germany

Petra Seibold,

Unit of Genetic Epidemiology, Division of Cancer Epidemiology (C020), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany

Sabine Behrens,
Unit of Genetic Epidemiology, Division of Cancer Epidemiology (C020), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany

Keith Humphreys,
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Jingmei Li,
Human,Genetics, Genome Institute of Singapore, 60 Biopolis St, Singapore 138672, Singapore

Jianjun Liu,
Human,Genetics, Genome Institute of Singapore, 60 Biopolis St, Singapore 138672, Singapore

Janet E. Olson,
Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

Xianshu Wang,
Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

Susan E. Hankinson,
Department of Epidemiology, Harvard, School Of Public Health, Boston, MA 02138, USA

Thérèse Truong,
Inserm (National Institute of Health and Medical Research), CESP (Center for Research in Epidemiology and Population Health), U1018, Environmental Epidemiology of Cancer, Villejuif, France; University Paris-Sud, UMRS 1018, Villejuif, France

Florence Menegaux,
Inserm (National Institute of Health and Medical Research), CESP (Center for Research in Epidemiology and Population Health), U1018, Environmental Epidemiology of Cancer, Villejuif, France; University Paris-Sud, UMRS 1018, Villejuif, France

Isabel dos Santos Silva, and
Department of Non-Communicable, Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Nichola Johnson
Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, 237 Fulham Road, London SW3 6JB, UK

The GENICA Network
Dr. Margarete Fischer-Bosch-Institute, of Clinical Pharmacology, Stuttgart, University of Tübingen, Tübingen, Germany; Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany; Institute for Prevention and Occupational Medicine of the German Social Accident Insurance (IPA), Bochum, Germany; Institute and Outpatient Clinic of Occupational Medicine, Saarland University Medical Center and Saarland University Faculty of Medicine, Homburg, Germany; Institute of Pathology, Medical Faculty of the University of Bonn, Bonn, Germany; Department of Internal Medicine, Evangelische Kliniken Bonn gGmbH, Johanniter Krankenhaus, Bonn, Germany

Shou-Tung Chen,
Department of Surgery, Changhua Christian Hospital, Changhua, Taiwan

Jyh-Cherng Yu,
Department of Surgery, Tri-Service General Hospital, Taipei, Taiwan

Argyrios Ziogas,

Department of Epidemiology, School of Medicine, University of California, Irvine, CA, USA

Vesa Kataja,
School of Medicine, Institute of Clinical Medicine, Oncology, Biocenter Kuopio, Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland; Cancer Center, Kuopio University Hospital, Kuopio, Finland

Veli-Matti Kosma,
School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland; Biocenter Kuopio, Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland; Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland

Arto Mannermaa,
School of Medicine, Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland; Biocenter Kuopio, Cancer Center of Eastern Finland, University of Eastern Finland, Kuopio, Finland; Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland

Hoda Anton-Culver,
Department of Epidemiology, University of California Irvine, Irvine, CA, USA

Chen-Yang Shen,
Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan; Taiwan Biobank, Taipei, Taiwan

Hiltrud Brauch,
Dr. Margarete Fischer-Bosch-Institute, of Clinical Pharmacology, Stuttgart, University of Tübingen, Tübingen, Germany

Julian Peto,
Department of Non-Communicable, Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Pascal Guénel,
Inserm (National Institute of Health and Medical Research), CESP (Center for Research in Epidemiology and Population Health), U1018, Environmental Epidemiology of Cancer, Villejuif, France; University Paris-Sud, UMRS 1018, Villejuif, France

Peter Kraft,
Program in Molecular and Genetic Epidemiology, Harvard School of Public Health, Boston, MA, USA; Department of Epidemiology, Harvard School Of Public Health, Boston, MA 02138, USA; Department of Biostatistics, Harvard School Of Public Health, Boston, MA 02138, USA

Fergus J. Couch,
Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN, USA

Douglas F. Easton,
Centre for Cancer Genetic Epidemiology, Departments of Public Health and Primary Care and Oncology, University of Cambridge, Cambridge, UK

Per Hall, and
Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Jenny Chang-Claude
Unit of Genetic Epidemiology, Division of Cancer Epidemiology (C020), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Germany

[#] These authors contributed equally to this work.

Abstract

Menopausal hormone therapy (MHT) is associated with an elevated risk of breast cancer in postmenopausal women. To identify genetic loci that modify breast cancer risk related to MHT use in postmenopausal women, we conducted a two-stage genome-wide association study (GWAS) with replication. In stage I, we performed a case-only GWAS in 731 invasive breast cancer cases from the German case-control study Mammary Carcinoma Risk Factor Investigation (MARIE). The 1,200 single nucleotide polymorphisms (SNPs) showing the lowest P values for interaction with current MHT use (within 6 months prior to breast cancer diagnosis), were carried forward to stage II, involving pooled case-control analyses including additional MARIE subjects (1,375 cases, 1,974 controls) as well as 795 cases and 764 controls of a Swedish case-control study. A joint P value was calculated for a combined analysis of stages I and II. Replication of the most significant interaction of the combined stage I and II was performed using 5,795 cases and 5,390 controls from nine studies of the Breast Cancer Association Consortium (BCAC). The combined stage I and II yielded five SNPs on chromosomes 2, 7, and 18 with joint P values $<6 \times 10^{-6}$ for effect modification of current MHT use. The most significant interaction was observed for rs6707272 ($P = 3 \times 10^{-7}$) on chromosome 2 but was not replicated in the BCAC studies ($P = 0.21$). The potentially modifying SNPs are in strong linkage disequilibrium with SNPs in *TRIP12* and *DNER* on chromosome 2 and *SETBP1* on chromosome 18, previously linked to carcinogenesis. However, none of the interaction effects reached genome-wide significance. The inability to replicate the top SNP $\times$ MHT interaction may be due to limited power of the replication phase. Our study, however, suggests that there are unlikely to be SNPs that interact strongly enough with MHT use to be clinically significant in European women.

Keywords

Introduction

Menopausal hormone therapy (MHT) is associated with an elevated risk of breast cancer in postmenopausal women [1,2]. Estrogen and its metabolic compounds are involved in breast carcinogenesis through their influence on cell growth and proliferation [3, 4]. Progesterone increases breast cancer risk by provoking an elevated formation of estradiol caused by induction of 17 betahydroxysteroid dehydrogenase [5] and regulating estrogen receptor levels in breast cancer cells by progesterone metabolites [6]. Furthermore, progesterone influences breast carcinogenesis directly by regulating proliferation, apoptosis, and cell adhesion in breast cells [7, 8]. Current use of any MHT compared to never use is associated with about 1.5 times higher breast cancer risk, whereas past use does not show any association. [1, 9, 10]. The magnitude of risk varies with histological type of breast cancer, being more than twofold higher for lobular/tubular than for ductal invasive cancer. In a population-based case-control study, we found the odds ratios (ORs) of association between current MHT use and risk of invasive breast cancer to be 2.95 for lobular, 3.75 for tubular, and 1.39 for ductal tumors [10]. Risk also differs by type of therapy and is higher for combined estrogen-progestagen therapy (EPT) than estrogen-only therapy (ET) [9-17]. For example, a meta-analysis reported ORs of 1.39 and 1.16 for the association between EPT, respectively, ET and breast cancer [17].

Even though the prevalence of MHT use decreased substantially after publication of the results of the Women's Health Initiative trial in 2002, postmenopausal MHT can still be considered as a relevant—in terms of prevalence of use (e.g., ET/EPT prevalence reported

for Germany: 3/5 %)—and effective therapy for the treatment of menopausal symptoms [18-21]. Thus, the identification of potential genetic modifiers of the MHT associated breast cancer risk can be of great importance from a clinical (keyword: personalized medicine) as well as a public health point on view. Furthermore, from a scientific perspective, the detection of such genetic modifiers may help to elucidate the mechanisms underlying the involvement of sexual hormones in breast carcinogenesis and additionally have the potential to improve risk prediction models. Previous studies showed that functionally relevant polymorphisms in genes involved in steroid metabolism may influence hormone levels [22, 23] which are associated with breast cancer risk. Thus, risk after exposure to MHT may vary according to individual susceptibility. However, only a limited number of studies have examined this and no established interactions between single nucleotide polymorphisms (SNPs) and MHT use exist to date [24-35]. For example, initial findings of interaction between variants in the established breast cancer susceptibility gene *FGFR2* and MHT could not be confirmed in subsequent studies [24,25, 27-29, 35].

Case-only studies are more efficient than case-control studies for detecting gene-environment interactions if the genetic and environmental factors are not associated in the population from which the cases were drawn [36]. To efficiently identify novel genetic loci that modify breast cancer risk associated with current MHT use in postmenopausal women, we conducted a two-stage genome-wide association study (GWAS), consisting of a case-only GWAS and a case-control analysis. We attempted to replicate the most significant interactions from the combined stage I and II in nine studies within the Breast Cancer Association Consortium (BCAC; http://www.srl.cam.ac.uk/consortia/bcac/index.html). Additionally, fine mapping of the three most interesting loci from the two-stage GWAS was conducted.

Subjects and methods

The studies, number of individuals and SNPs involved in each analysis stage of our study are shown in Fig. 1.

Study population and data collection

The GWAS was carried out within the Mammary Carcinoma Risk Factor Investigation (MARIE) study. The MARIE study is a population-based case-control study of primarily European postmenopausal women carried out in two regions in Germany with incident breast cancer cases and controls matched by birth year and study region [10].

For the GWAS first stage, 500 ductal and 300 lobular tumor cases were randomly selected from the breast cancer patients with known age at menopause from the MARIE study, whereby lobular cases were over-sampled.

Stage II comprised 1,375 additional cases and 1,974 matched controls from the MARIE study—excluding those cases already included in stage I—as well as a random subset of 795 cases and 764 controls of European origin from the Singapore and Sweden Breast Cancer Study (SASBAC) [37]. SASBAC is a subset of the Swedish nationwide population-based case-control study, including 1,801 incident cases and 1,712 controls, aged 50–74 years [38].

The replication study was conducted using nine case-control studies with information on current MHT use and genotype data within BCAC. Thus, a total of 5,795 cases and 5,390 controls of the BCAC studies BBCS, CECILE, GENICA, KBCP, MCBCS, NHS, SASBAC, TWBCS, and UCIBCS (see Supplementary Table 1 for an explanation of acronyms and

study details) were included in the replication stage. For the SASBAC study, only subjects not included in stage II were incorporated in the replication stage.

Fine mapping was performed in all MARIE participants with available blood sample (2,516 cases, 4,854 controls) including those involved in stages I and II.

Supplementary Table 1 contains further information on these studies.

Menopausal hormone therapy exposure definition

MHT use was defined as use of any type of menopausal hormone replacement therapy, including EPT and ET. Only women using MHT for more than 3 months were considered to be ever users of MHT. The available data on MHT use from MARIE and SASBAC were used to determine the categories "current use of any MHT" and "past/never use of any MHT", where current use was defined as use within the last six months before reference date (i.e., date of diagnosis for cases, date of interview for controls). Current use of EPT was accordingly defined as "current use of any combination of estrogens and progestagens" versus "past/never use of EPT". From the other BCAC studies, data on current versus past/never MHT and EPT use was centrally collected according to the BCAC data dictionary and underwent quality assurance.

Laboratory methods

In stage I, 318,237 SNPs were genotyped with Illumina Humancnv370-duo chip (370 K) on quality-checked DNA samples from 799 MARIE cases (one sample was excluded due to low DNA concentration). Scanning of the chips was done by BeadArrayReaders, the calling by Beadstudio 2.0 (Illumina). Samples with completion rate <90 % were assayed a second time. Twenty-three samples were subsequently excluded due to failure in the second run. Fifteen pairs of individuals with identical by state (IBS) score >0.8 were excluded from the analysis. There was no significant evidence of population stratification, i.e., the difference between IBS scores of groups defined by MHT exposure was not significant ($P = 0.33$). Further 15 individuals with >10 % missing genotypes were removed and genotyping rate in the remaining individuals was 99 %. Additionally, SNPs were excluded if minor allele frequencies (MAFs) were <0.01 (224 SNPs excluded) or if the rate of missing genotypes exceeded 10 % (1,043 SNPs). Thus, 731 cases and 316,974 SNPs remained for statistical analyses.

Genotyping capacities allowed for a total of 1,200 SNPs to be genotyped in stage II. Therefore, the 1,200 SNPs with the lowest P values for statistical interaction with current MHT (SNP × MHT interaction) in stage I were genotyped in the 3,349 MARIE participants using Illumina Golden Gate AssayStage II (Illumina, San Diego, CA). After excluding 101 SNPs that failed, genotyping rate in remaining individuals was 99.8 %. Two SNPs were excluded due to genotyping rate <90 %. We did not exclude SNPs based on MAFs (MAF<0.05 for 10 SNPs) or deviation from Hardy–Weinberg equilibrium (HWE) ($P < 0.01$ for 23 SNPs). No individuals were excluded due to missing genotype data since all individuals were geno-typed for >90 % of the SNPs. The SASBAC participants were already genotyped with Illumina HumanHap550. Quality control (QC) checked genotype data was provided for 1,154 of the 1,200 SNPs. Of the initial 1,200 SNPs, only SNPs that were successfully genotyped and also passed the QC in MARIE (1,097 SNPs) as well as SAS-BAC (1,154 SNPs) were eligible for inclusion in our analyses, i.e., 99 SNPs were discordant. Thus, a total of 1,055 concordant SNPs were included in the stage II analysis for a total of 2,170 cases and 2,738 controls.

For a case–control study nested within the NHS study (1,090 cases, 1,078 controls), which was involved in the replication phase, genotypes of the most significant SNP of the combined stage I and II were obtained from a GWAS conducted with the Illumina HumanHap500 as part of the Cancer Genetic Markers of Susceptibility Project (CGEMS) (for further information on genotyping and QC see [39]).

For the remaining eight BCAC studies that were incorporated in the replication phase (i.e., all nine BCAC studies except for the NHS), the most significant SNP of the combined stage I and II was genotyped along with other SNPs in BCAC by matrix-assisted laser desorption/ ionization time-of-flight mass spectrometry (MALDI-TOF MS) using iPLEX technology (www.sequenom.com). Standard QC guidelines were applied. Data were excluded for: (1) Any sample that consistently failed genotyping for >20 % of the SNPs; (2) All study data for any SNP with overall call rate <95 % or duplicate concordance <98 % (based on at least 2 % of samples in each study being genotyped in duplicate); (3) For any study with evidence for departure of SNP genotype distribution from HWE in controls ($P < 0.005$). Clustering of the intensity plots was reviewed manually and the data excluded if the clusters were poorly separated. In addition, all genotyping centres assayed an identical plate of 80 control DNA samples from the Coriell Institute for Medical Research, Camden, NJ, USA (referred to as the Coriell plate; which also included 14 internal duplicates) and had to achieve call rates and duplicate concordance >98 % in order for their data to be included. For these eight BCAC studies (4,705 cases, 4,312 controls) genotyping quality complied with the QC criteria.

For fine mapping of the three loci containing the five most significant SNPs of the combined stage I and II, a total of 33 tagging SNPs with high linkage disequilibrium (LD) ($r^2 > 0.8$) based on Paul de Bakker's Tagger algorithm [40] and a MAF > 0.05 was selected, using HapMap CEU (http://hapmap.ncbi.nlm.nih.gov). These 38 SNPs were genotyped in the MARIE subjects by MALDI-TOF MS using the Sequenom MassARRAY platform and iPLEX GOLD methodology according to manufacturer's instructions (http:// www.sequenom.com). Three tagging SNPs on chromosome 2 could not be genotyped and two SNPs on chromosome 2 with call rate <90 % were excluded. 274 individuals (92 cases, 182 controls) were excluded because of more than 3 genotypes missing, MAFs of all SNPs were >0.01. One SNP, rs6460021, showed deviation from HWE at a $P < 0.01$. 33 SNPs in 2,424 cases and 4,672 controls remained for the statistical analysis. The 22 SNPs for the locus on chromosome 2 (including the three original top SNPs) tagged a haplotype block of about 500 kb (Fig. 2). The locus on chromosome 7 was covered by the top SNP of this locus and five additional SNPs spanning a region of 20 kb on both sides of the top SNP (Supplementary Fig. 1). The locus on chromosome 18 was captured by five SNPs, including the top SNP of this locus, spanning a region of 10 kb on each side of the top SNP (Supplementary Fig. 2). Coverage of chromosomes 2, 7, and 18 was 80, 87, and 86 %, respectively (HapMap CEU, http://hapmap.ncbi.nlm.nih.gov, r^2 0.8).

Statistical analysis

In stage I analyses, we tested for SNP × MHT/SNP × EPT interactions on the genome-wide level in cases only using logistic regression with current hormone use as the outcome variable and the SNP as the explanatory variable. Each SNP was assessed for interaction in four different models. Interactions between SNP and current MHT use were investigated using models 1–3, whereby the SNP was assessed according to a log-additive model, i.e., a test for trend by number of minor alleles for all cases (model 1) or lobular cases only (model 2) was conducted. In addition, a genotypic model, i.e., two dummy variables were used to code for the heterozygous and homozygous mutant genotype (2 df test), was investigated for all cases (model 3). In model 4, interactions between SNP and current EPT use assuming a

log-additive model were assessed in all invasive cases. All models were adjusted for the covariates age at reference date (date of diagnosis; continuously) and study centre.

The 1,200 SNPs carried forward to stage II were selected initially based on the 1,200 SNP × MHT interactions with the lowest P values obtained from model 1. Then the SNPs with the highest P values were replaced by non-overlapping SNPs showing SNP × MHT (model 2 and 3)/SNP × EPT (model 4) interactions with $P < 0.001$ for the models 2, 3, and 4.

In stage II, a case-control logistic regression analysis was performed. The model included SNP and current hormone use (vs. past/never) as well as the multiplicative SNP × MHT/ SNP × EPT interaction term. Additionally, potential breast cancer risk factors or confounders entered the model as continuous [i.e., year of birth (in 5-year categories), BMI (22.4, 22.5–24.9, 25–29.9, 30 kg/m^2), and number of full-term pregnancies [0, 1, 2, 3; coded continuously]] or categorical covariates [i.e., study, study centre (for MARIE), age at menarche (< 12, 12–14, 15 years), type of menopause (natural, induced, hysterectomy, other), ever benign breast disease (yes/no), first degree family history of breast cancer (yes/ no), and number of mammograms (0, 1–4, 5–9, 10, unknown)]. Four different models corresponding to those used in stage I were employed to test for SNP × MHT (model 1–3)/ SNP × EPT (model 4) interaction in the case-control scenario. Analyses in stages I and II were conducted with plink version 1.0.

For a combined analysis of stages I and II, joint test statistics z_{joint} and P values P_{joint} were calculated, separately for the different models and only if interaction ORs of the two stages were in the same direction, according to Skol et al. [41]. The critical value C_{joint} for z_{joint} which corresponds to an overall significance level of 10^{-7}, was calculated using the program CaTS (http://www.sph.umich.edu/csg/abecasis/cats/).

For the replication study, a case-control logistic regression analysis was performed according to model 1 (since lowest P values for the interaction of the top SNP with current hormone use were obtained with model 1) and adjusted for age at reference date (date of diagnosis for cases/date of interview for controls; continuously) to analyze the NHS data. Similarly, a pooled case-control logistic regression analysis of the remaining eight BCAC studies according to model 1, adjusted for study and age at reference date was conducted. A meta-analysis (fixed effects model) was performed to combine these results. The study specific analysis of NHS data as well as the meta-analysis was performed using the R software version 2.9.2 [42] and the R-package meta. For the pooled analysis the SAS software version 9.2 of the SAS System for Windows was employed.

The fine mapping SNPs were assessed for SNP × MHT interaction according to model 1 (as lowest P values for interaction were obtained with model 1) in a case-control logistic regression framework as described for stage II, stratified by year of birth (in 5-year categories) and study centre and adjusted for the covariates used in stage II analyses except for study, year of birth, and study centre. Analyses were performed using the SAS software version 9.2. of the SAS System for Windows.

To test for independent SNP × MHT interaction effects for highly correlated SNPs (as measured in terms of LD) logistic regression analyses were performed using models that contained two correlated SNPs at a time along with the respective SNP × MHT interaction terms and the covariates that were also used in the single SNP models that led to the detection of the respective interaction effects. SNP × MHT interaction effects were considered to be independent if both interaction terms appeared to be significant in the respective regression model that included both interaction effects at a time. These analyses were performed using the SAS software version 9.2. of the SAS System for Windows.

Results

The prevalence of MHT and EPT use in stages I and II individuals is presented in Supplementary Tables 2 and 3. The estimate of the main effect of MHT and EPT, respectively, obtained from the pooled stage II analysis was OR = 1.49 (95 % confidence interval (CI) = 1.31–1.70, $P = 2.89 \times 10^{-9}$) and 1.54 (95 % CI = 1.32–1.79, $P = 1.86 \times 10^{-8}$) for breast cancer overall, 2.37 (95 % CI = 1.80–3.12, $P = 8.20 \times 10^{-10}$) and 2.08 (95 % CI = 1.54–2.81, $P = 2.05 \times 10^{-6}$) for lobular carcinoma and 1.24 (95 % CI = 1.06–1.43, $P = 0.005$), and 1.39 (95 % CI = 1.18–1.64, $P = 1.08 \times 10^{-4}$) for ductal carcinoma.

The set of 1,200 SNPs selected from stage I and carried forward to stage II was composed of the 767 SNPs with lowest P values for SNP × MHT interaction obtained from models 1 and 433 additional SNPs with $P < 0.001$ for SNP × MHT interactions selected from models 2–4 (105, 206, 168 SNPs selected from models 2, 3, 4, respectively; 44 out of 433 SNPs were selected by more than one of the three models).

In the combined stage I and II, we observed no SNP × MHT interaction with a z_{joint} value greater than the critical value of $|C_{\text{joint}}| \sim 4.9$ which corresponds to an overall significance level of 1×10^{-7} (calculated for the following parameters: 2,901 cases (25 % referred to stage I), 2,738 controls, 0.03 % SNPs referred to second stage). The greatest absolute z_{joint} value 4.12 was obtained from model 1 for the modification of the current MHT related breast cancer risk by rs6707272 on chromosome 2 ($P_{\text{combined}} = 2.98 \times 10^{-7}$; see Table 1). Furthermore, model 1 yielded additional four SNP × MHT interactions with $P < 6 \times 10^{-6}$ involving rs560304, rs13014061, rs2909969 as well as rs1942574 on chromosomes 2, 7, and 18, respectively. These two SNPs on chromosome 2 were, however, highly correlated with the most significant SNP rs6707272 ($r^2 = 0.73$ for rs560304, $r^2 = 0.74$ for rs13014061). When testing for the modification of the current MHT use related breast cancer risk by SNPs rs560304 and rs13014061, respectively, on chromosome 2 in a model additionally adjusted for rs6707272 and the interaction between current MHT use and rs6707272, P values for interaction were >0.05 (data not shown).

The genotype distribution for rs6707272 in the BCAC studies that were used in the replication analysis is provided in Supplementary Table 4. Prevalence of MHT use in the respective study populations is shown in Supplementary Table 5. The OR for breast cancer associated with current use of MHT obtained from the pooled BCAC studies and the NHS were 0.88 (95 % CI = 0.79–0.97), $P = 0.01$ and OR = 1.07 (95 % CI = 1.02–1.11), $P = 4.30 \times 10^{-3}$, respectively.

The SNP rs6707272 was not associated with breast cancer risk in the replication sample, with OR = 0.96 (95 % CI = 0.90–1.04), $P = 0.33$, in the pooled studies, and OR = 0.99 (95 % CI = 0.96–1.02), $P = 0.37$, in NHS. It did not modify the effect of current MHT in the pooled studies [OR = 1.06 (95 % CI = 0.90–1.24), $P_{\text{interaction}} = 0.51$] or in NHS [OR = 1.19 (95 % CI = 0.92–1.55), $P_{\text{interaction}} = 0.19$]. Thus, the meta-analysis of the nine studies did not yield a significant SNP × MHT interaction result [OR = 1.09 (95 % CI = 0.95–1.25), $P_{\text{interaction}} = 0.21$]. Restriction to subjects of European ancestry did not change the results for interaction substantially [OR = 1.09 (95 % CI = 0.94–1.25), $P_{\text{interaction}} = 0.25$].

Further analyses were carried out restricted to the four *population-based* BCAC studies, CECILE, GENICA, NHS, and SASBAC (excluding cases and controls involved in stage II) for which recruitment times for cases and controls were overlapping and estimates of MHT main effects were risk-enhancing. The corresponding SNP × MHT interaction effect was non-significant [OR = 1.03 (95 % CI = 0.87–1.23), $P_{\text{interaction}} = 0.73$; 2,886 cases, 3,169 controls].

The genotype distributions of the 33 fine mapping SNPs along with main effect ORs and corresponding P values are given in Supplementary Table 6. Key characteristics of the study population are shown in Supplementary Table 7. The OR for breast cancer associated with current use of MHT was 1.60 (95 % CI = 1.44–1.79; $P = 5.14 \times 10^{-17}$).

SNP × MHT interactions involving rs6707272 [chromosome 2; OR = 0.77 (95 % CI = 0.66–0.90), $P_{interaction} = 1.01 \times 10^{-3}$] and rs1942574 [chromosome 18; OR = 0.71 (95 % CI = 0.59–0.87), $P_{interaction} = 7.08 \times 10^{-4}$] already investigated in the combined stage I and II, were significant when using a Bonferroni-corrected significance level of   = 0.002 in the fine mapping analysis (see Table 1). In the combined stages as well as in the fine mapping, rs6707272 and rs1942574 were observed to reduce the risk enhancing effect of MHT in relation to breast cancer. The locus on chromosome 7 could not be replicated in the fine mapping analysis. SNP rs11886634 on chromosome 2 which was not included in previous stages I and II was observed to be significant after Bonferroni-correction [OR = 0.75 (95 % CI = 0.64–0.68), $P_{interaction} = 4.57 \times 10^{-4}$]. However, the modification of the breast cancer risk associated with current MHT use by SNPs rs11886634 and rs6707272, respectively, did not remain significant in a model including both SNPs (data not shown). Table 2 presents the effect of current MHT use in strata defined by genotypes of the five top SNPs of the combined stage I and II. In correspondence with the observed protective SNP × MHT interaction effects, the main effect of MHT use decreases with the number of minor alleles of SNPs rs6707272 and rs1942574 and is non-significant for individuals homozygous for the respective minor allele.

Discussion

Our results suggest possible modifications of breast cancer risk associated with current MHT use by SNPs on chromosomes 2 and 18, though none of the observed interaction effects reached genome-wide significance in the detection stage, i.e., the combined stage I and II. Furthermore, the most significant interaction of combined stage I and II, the modification of the current MHT use related breast cancer risk by rs6707272, was not significant in the replication studies. Analyses of chromosome 2 including two SNPs at a time suggest that—if at all—a single causative variant may modify breast cancer risk associated with current MHT use on chromosome 2.

SNP rs6707272 lies in a ~ 100 kb LD block on 2q36.3 that contains one known gene, delta/notch-like epidermal growth factor-like repeat containing (*DNER*). Recent experimental studies suggest that the *DNER* signalling pathway has differentiating and tumor-suppressing actions on glioma-derived stem-l cells [43] and may also be involved in modulating adipogenetic differentiation [44]. Although *DNER* is an unorthodox ligand of the single-pass transmembrane receptors *NOTCH*, it is interesting to note that *NOTCH2* has been associated with breast cancer susceptibility [45]. The other two SNPs on chromosome 2, rs13014061 and rs560304, lie in intronic regions of the thyroid hormone receptor interactor 12 (*TRIP12*) and of the F-box protein 36 (*FBXO36*) gene, respectively (see Fig. 2). No association with cancer has been reported for *FBXO36*, while *TRIP12* may increase cancer risk by causing a ubiquitin-mediated degradation of ARF (alternate reading frame of the INK4a/CDKN2A locus), the key activator of the tumor suppressor p53 [46].

SNP rs1942574 on chromosome 18, which showed the strongest interaction in the fine mapping analysis, is located approximately 80 kb downstream of the suppressor of variegation, enhancer of zeste, and Trithorax binding protein (*SETBP1*) (see Supplementary Fig. 1), which was proposed to be inversely associated with cancer through its involvement in the mechanism of SET-related leukogenesis and tumorigenesis [47].

To date, GWAS are still rarely being conducted to agnostically search for gene-environment interactions [48]. These studies confront the researchers with issues that similarly apply to GWAS screening for genetic marginal effects, such as the misspecification of the true genetic model. In our study, we analysed $SNP \times MHT$ interactions under the assumption of a log-additive (models 1, 2, and 4) as well as a genotypic (model 3) genetic model. On the one hand, we thereby allowed for different genetic models while restricting the number of tests performed by focussing on specific genetic models and, thus, reducing the problem of multiple testing. On the other hand, our study might have lost power due to a potential misspecification of the true genetic model.

In addition to the issues that arise when searching for genetic marginal effects in GWAS, further challenges come up when aiming to identify gene-environment interactions on a genome-wide scale. These include misspecification of the environmental exposure as well as magnified problems due to the indirect measurement of causal markers through LD and multiple testing which, in turn, imply the requirement of larger samples [49, 50].

In our well-powered combined stage I and II analysis with very good genome coverage, we agnostically searched for potential genetic modifiers of the MHT related breast cancer risk. Thereby, our study design exploits the advantages of the case-only approach and aims to avoid false positive results due to gene-environment dependence when combining the results of the case-only approach in stage I with the results of the case-control approach in stage II [36]. Amongst others, the power of our study depends on the definition of the MHT exposure variable that should reflect the true association of MHT use with breast cancer risk. We restricted our regression analyses to *current* MHT/EPT use since risk among past users was shown to be comparable to risk among never users of any MHT [1, 9,10]. We also exploited that MHT-related breast cancer risk is more pronounced for lobular tumors [10] and for EPT use [9-17]. In our combined stage I and II, we investigated separate regression models for lobular breast cancer cases only (model 2) as well as current EPT use (model 4) with the aim to increase the power of our study. By applying the four different regression models on a genome-wide scale, we magnified the issues of multiple testing. We were more concerned with Type II error, thus, we did not account for the different models (1–4) tested in our evaluation of the findings of the combined stage I and II analysis. However, since none of our findings reached genome-wide significance, correcting for the dependency between tests conducted for models 1–4 would not have changed our conclusions.

We employed a single definition of the MHT variable, i.e. we restricted our investigations to a *binary* MHT variable. Thereby, we did not take into account differences between MHT users with respect to duration of use and medication dosage. Moreover, we defined current MHT use as use within the 6 months prior to reference date. In the logistic regression model, the reference category, thus, included never MHT users as well as past users of MHT for whom last MHT use was more than 6 months prior to reference date. Given that breast cancer may be diagnosed months after disease onset, a limitation of our study is that it most likely includes patients misclassified with respect to MHT status, i.e., according to our definition of current use, some patients truly were current users of MHT at disease onset but were classified as past users since this was the correct classification at the reference date, i.e. the date of diagnosis. For these reasons, it is likely that the estimates of MHT/EPT main effects derived from our study are downwardly biased and that the power of our study to detect MHT main and $SNP \times MHT$ interaction effects is, thus, decreased [51].

In the replication stage, we used data of nine BCAC studies with a harmonized definition of the MHT exposure variable, which corresponds to the one used in the combined stage I and II. The composition of MHT products differs, however, between countries and, thus, between studies involved in our replication analysis. By adjusting for study we were in part

able to take these differences into account. While our stage II as well as fine mapping analyses demonstrate a positive association between MHT use and breast cancer risk, which is in line with results reported in the literature [1, 2], current MHT use was inversely associated with breast cancer risk in the replication stage. Since this may result in biased interaction effects [52, 53], we carried out an additional replication analysis including only *population-based* BCAC studies, all of which showed the established risk enhancing effect of MHT. However, restriction to these studies resulted in a reduction of power (80 % power to detect an interaction effect of 1.41 (or 0.71, respectively) [54]) and a weaker, non-significant interaction effect.

Only a sub-sample of the MARIE cases was genotyped on a genome-wide scale and thereby power to detect interactions was reduced compared to a one-stage design with the whole sample genotyped. For GWAS analyses, Skol et al. proposed the combined analysis of the most significant results of a stage I GWAS analysis and the reanalysis of these findings in an independent sample [41]. They showed that this design provides a trade-off between genotyping costs and power that can preserve much of the power of the corresponding one-stage GWAS design and is almost always more powerful than the corresponding two-stage design. We performed power calculations according to Skol et al. to determine the number of SNPs to be transferred from stage I to stage II. For the design of our study, we originally chose a scenario that was calculated to yield 80 % power to detect at least one true interaction with an OR of 1.7 or 0.59, respectively when transferring 1,500 SNPs to stage II including additional 200 cases and 1,000 controls. Due to decreases in genotyping costs during the study conduct we were able to increase the number of cases (1,375) and controls (1,974) for stage II. Whereas Skol et al. analyzed genetic main effects, we applied their design to the analysis of gene-environment interactions using a case-only design in the first and a case-control design in the second stage. Their results with respect to relative power of the different designs can be assumed to similarly apply to our analysis as estimates of the interaction effects derived from case-only versus case-control analyses are identical in presence of gene-environment independence. The violation of this assumption, however, can result in biased effect estimates derived from case-only analysis.

Our inability to replicate the top SNP × MHT interaction involving rs6707272 in BCAC may indicate that this finding is indeed a null result. However, our replication study based on all nine BCAC studies had 80 % power to detect an interaction effect of 1.20 (or 0.83, respectively) [54]. The corresponding interaction effect estimates obtained in stages I and II were OR = 0.60 and 0.76, i.e. it is likely that we over-estimated the interaction effect in the detection phase (the combined stage I and II)—a phenomenon known as 'the winners curse' — and, thus, our replication phase may have lacked power to detect a true interaction effect [55].

Bonferroni correction was applied to correct the fine mapping results for multiple testing. Thereby, we did not take into account the dependence between the samples used in the combined stage I and II and in the fine mapping, i.e., the case-control sample included in the fine mapping analysis comprised all MARIE subjects involved in stage I and II. Therefore, the significance of the fine mapping results is likely to be limited.

Conclusions

Our study provides a hint for a possible modification of the current MHT use related breast cancer risk by SNPs on chromosomes 2 and 18. We were not successful in replicating our top SNP × MHT interaction on chromosome 2 in an independent sample, which may be attributable to a lack of power. However, since none of the investigated interaction effects reached genome-wide significance in our well-powered combined stage I and II analyses,

we conclude that the presence of clinically relevant SNP × MHT interactions in European women is questionable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Abbreviations

ARF	Alternate reading frame of the INK4a/CDKN2A locus
BCAC	Breast Cancer Association Consortium
CGEMS	Cancer Genetic Markers of Susceptibility Project
CI	Confidence interval
DNE	Delta/notch-like epidermal growth factor-like repeat containing
EPT	Estrogen-progestagen combined therapy
ET	Estrogen-only therapy
FBXO36	F-box protein 36
HWE	Hardy-Weinberg equilibrium
IBS	Identical by state
LD	Linkage disequilibrium
MAF	Minor allele frequency
MALDI-TOF MS	Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MARIE	Mammary Carcinoma Risk Factor Investigation
MHT	Menopausal hormone therapy
OR	Odds ratio
QC	Quality control
SASBAC	Singapore and Sweden Breast Cancer Study
SET	Suppressor of variegation, enhancer of zeste, and Trithorax
SETBP1	SET binding protein
SNP	Single nucleotide polymorphism
TRIP12	Thyroid hormone receptor interactor 12

References

1. Beral V. Breast cancer and hormone-replacement therapy in the Million Women Study. Lancet. 2003; 362(9382):419–427. [PubMed: 12927427]

2. Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, Jackson RD, Beresford SA, Howard BV, Johnson KC, Kotchen JM, Ockene J. Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. JAMA. 2002; 288(3):321–333. [PubMed: 12117397]

3. Horwitz K, Clarke C. Estrogens and progestins in mammary development and neoplasia. Introduction. J Mammary Gland Biol Neoplasia. 1998; 3(1):1–2. [PubMed: 10819499]

4. Zumoff B. Does postmenopausal estrogen administration increase the risk of breast cancer? Contributions of animal, biochemical, and clinical investigative studies to a resolution of the controversy. Proc Soc Exp Biol Med. 1998; 217(1):30–37. [PubMed: 9421204]

5. Poutanen M, Isomaa V, Peltoketo H, Vihko R. Role of 17 beta-hydroxysteroid dehydrogenase type 1 in endocrine and intracrine estradiol biosynthesis. J Steroid Biochem Mol Biol. 1995; 55(5-6):525–532. [PubMed: 8547177]

6. Pawlak KJ, Wiebe JP. Regulation of estrogen receptor (ER) levels in MCF-7 cells by progesterone metabolites. J Steroid Biochem Mol Biol. 2007; 107(3-5):172–179. [PubMed: 17683929]

7. Wiebe JP. Progesterone metabolites in breast cancer. Endocr Relat Cancer. 2006; 13(3):717–738. [PubMed: 16954427]

8. Wiebe JP, Souter L, Zhang G. Dutasteride affects progesterone metabolizing enzyme activity/ expression in human breast cell lines resulting in suppression of cell proliferation and detachment. J Steroid Biochem Mol Biol. 2006; 100(4-5):129–140. [PubMed: 16806904]

9. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormone replacement therapy: collaborative reanalysis of data from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer. Lancet. 1997; 350(9084): 1047–1059. [PubMed: 10213546]

10. Flesch-Janys D, Slanger T, Mutschelknauss E, Kropp S, Obi N, Vettorazzi E, Braendle W, Bastert G, Hentschel S, Berger J, Chang-Claude J. Risk of different histological types of postmenopausal breast cancer by type and regimen of menopausal hormone therapy. Int J Cancer. 2008; 123(4): 933–941. [PubMed: 18506692]

11. Chlebowski RT, Anderson GL, Gass M, Lane DS, Aragaki AK, Kuller LH, Manson JE, Stefanick ML, Ockene J, Sarto GE, Johnson KC, Wactawski-Wende J, Ravdin PM, Schenken R, Hendrix SL, Rajkovic A, Rohan TE, Yasmeen S, Prentice RL. Estrogen plus progestin and breast cancer incidence and mortality in postmenopausal women. JAMA. 2010; 304(15):1684–1692. [PubMed: 20959578]

12. Li CI. Postmenopausal hormone therapy and the risk of breast cancer: the view of an epidemiologist. Maturitas. 2004; 49(1):44–50. [PubMed: 15351095]

13. Newcomb PA, Titus-Ernstoff L, Egan KM, Trentham-Dietz A, Baron JA, Storer BE, Willett WC, Stampfer MJ. Postmenopausal estrogen and progestin use in relation to breast cancer risk. Cancer Epidemiol Biomarkers Prev. 2002; 11(7):593–600. [PubMed: 12101105]

14. Ross RK, Paganini-Hill A, Wan PC, Pike MC. Effect of hormone replacement therapy on breast cancer risk: estrogen versus estrogen plus progestin. J Natl Cancer Inst. 2000; 92(4):328–332. [PubMed: 10675382]

15. Schairer C, Lubin J, Troisi R, Sturgeon S, Brinton L, Hoover R. Menopausal estrogen and estrogen-progestin replacement therapy and breast cancer risk. JAMA. 2000; 283(4):485–491. [PubMed: 10659874]

16. Seeger H, Mueck AO. Are the progestins responsible for breast cancer risk during hormone therapy in the postmenopause? Experimental vs. clinical data. J Steroid Biochem Mol Biol. 2008; 109(1-2):11–15. [PubMed: 18206365]

17. Shah NR, Borenstein J, Dubois RW. Postmenopausal hormone therapy and breast cancer: a systematic review and meta-analysis. Menopause. 2005; 12(6):668–678. [PubMed: 16278609]

18. Rusner C, Bandemer-Greulich U, Engel J, Stegmaier C, Zawinell A, Holleczek B, Katalinic A, Kuss O, Schmidt-Pokrzywniak A, Schubert-Fritschle G, Tillack A, Stang A. Population-based hormone receptor-specific incidence trends of breast cancer in Germany. Maturitas. 2012; 73(2): 152–157. [PubMed: 22785007]

19. Sprague BL, Trentham-Dietz A, Cronin KA. A sustained decline in postmenopausal hormone use: results from the National Health and Nutrition Examination Survey, 1999-2010. Obstet Gynecol. 2012; 120(3):595–603. [PubMed: 22914469]

20. Canfell K, Banks E, Clements M, Kang YJ, Moa A, Armstrong B, Beral V. Sustained lower rates of HRT prescribing and breast cancer incidence in Australia since. Breast Cancer Res Treat. 2009; 117(3):671–673. 2003. [PubMed: 19219631]

21. Hemminki E, Kyyronen P, Pukkala E. Postmenopausal hormone drugs and breast and colon cancer: nordic countries 1995-2005. Maturitas. 2008; 61(4):299–304. [PubMed: 18996655]

22. Barnes KM, Dickstein B, Cutler GB Jr, Fojo T, Bates SE. Steroid treatment, accumulation, and antagonism of P-glycoprotein in multidrug-resistant cells. Biochemistry. 1996; 35(15):4820–4827. [PubMed: 8664272]

23. Kim WY, Benet LZ. P-glycoprotein (P-gp/MDR1)-mediated efflux of sex-steroid hormones and modulation of P-gp expression in vitro. Pharm Res. 2004; 21(7):1284–1293. [PubMed: 15290871]

24. Campa D, Kaaks R, Le ML, 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. Interactions between genetic variants and breast cancer risk factors in the breast and prostate cancer cohort consortium. J Natl Cancer Inst. 2011; 103(16):1252–1263. [PubMed: 21791674]

25. Kawase T, Matsuo K, Suzuki T, Hiraki A, Watanabe M, Iwata H, Tanaka H, Tajima K. FGFR2 intronic polymorphisms interact with reproductive risk factors of breast cancer: results of a case control study in Japan. Int J Cancer. 2009; 125(8):1946–1952. [PubMed: 19582883]

26. Milne RL, Gaudet MM, Spurdle AB, Fasching PA, Couch FJ, Benitez J, rias Perez JI, Zamora MP, Malats N, Dos SS,I, Gibson LJ, Fletcher O, Johnson N, nton-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, Lambrechts D, Reumers J, Smeets A, Wang-Gohrke S, Hall P, Czene K, Liu J, Irwanto AK, Chenevix-Trench G, Holland H, 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. 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(6):R110. [PubMed: 21194473]

27. 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(11):3079–3085. [PubMed: 19861516]

28. Rebbeck TR, DeMichele A, Tran TV, Panossian S, Bunin GR, Troxel AB, Strom BL. Hormone-dependent effects of FGFR2 and MAP3K1 in breast cancer susceptibility in a population-based sample of post-menopausal African-American and European-American women. Carcinogenesis. 2009; 30(2):269–274. [PubMed: 19028704]

29. Travis RC, Reeves GK, Green J, Bull D, Tipper SJ, Baker K, Beral V, Peto R, Bell J, Zelenika D, Lathrop M. Gene–environment interactions in 7610 women with breast cancer: prospective evidence from the Million Women Study. Lancet. 2010; 375(9732):2143–2151. [PubMed: 20605201]

30. Hein R, Abbas S, Seibold P, Salazar R, Flesch-Janys D, Chang-Claude J. Polymorphism Thr160Thr in SRD5A1, involved in the progesterone metabolism, modifies postmenopausal breast cancer risk associated with menopausal hormone therapy. Breast Cancer Res Treat. 2012; 131(2):653–661. [PubMed: 21947678]

31. The MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk. Postmenopausal estrogen monotherapy-associated breast cancer risk is modified by CYP17A1_-34_T > C polymorphism. Breast Cancer Res Treat. 2010; 120(3):737–744. [PubMed: 19672705]

32. The MARIE-GENICA Consortium6 on Genetic Susceptibility for Menopausal Hormone Therapy7 Related Breast Cancer Risk. Genetic polymorphisms in phase I and phase II enzymes and breast cancer risk associated with menopausal hormone therapy in postmenopausal women. Breast Cancer Res Treat. 2010; 119(2):463–474. [PubMed: 19424794]

33. The MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk. Polymorphisms in the BRCA1 and ABCB1 genes modulate menopausal hormone therapy associated breast cancer risk in postmenopausal women. Breast Cancer Res Treat. 2010; 120(3):727–736. [PubMed: 19672706]

34. The MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk. Polymorphisms in genes of the steroid receptor super-family modify postmenopausal breast cancer risk associated with menopausal hormone therapy. Int J Cancer. 2010; 126(12):2935–2946. [PubMed: 19739075]

35. Nickels S, Truong T, Hein R, et al. Evidence of gene–environment interactions between common breast cancer susceptibility Loci and Established Environmental Risk Factors. PLoS Genet. 2013 accepted for publication.

36. Piegorsch WW, Weinberg CR, Taylor JA. Non-hierarchical logistic models and case-only designs for assessing susceptibility in population-based case-control studies. Stat Med. 1994; 13(2):153–162. [PubMed: 8122051]

37. Wedren S, Lovmar L, Humphreys K, Magnusson C, Melhus H, Syvanen AC, Kindmark A, Landegren U, Fermer ML, Stiger F, Persson I, Baron J, Weiderpass E. Oestrogen receptor alpha gene haplotype and postmenopausal breast cancer risk: a case control study. Breast Cancer Res. 2004; 6(4):R437–R449. [PubMed: 15217512]

38. Magnusson C, Baron JA, Correia N, Bergstrom R, Adami HO, Persson I. Breast-cancer risk following long-term oestrogen- and oestrogen–progestin-replacement therapy. Int J Cancer. 1999; 81(3):339–344. [PubMed: 10209946]

39. Hunter DJ, Kraft P, Jacobs KB, Cox DG, Yeager M, Hankinson SE, Wacholder S, Wang Z, Welch R, Hutchinson A, Wang J, Yu K, Chatterjee N, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF Jr, Hoover RN, Thomas G, Chanock SJ. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet. 2007; 39(7):870–874. [PubMed: 17529973]

40. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet. 2005; 37(11):1217–1223. [PubMed: 16244653]

41. Skol AD, Scott LJ, Abecasis GR, Boehnke M. Joint analysis is more efficient than replication-based analysis for two-stage genome-wide association studies. Nat Genet. 2006; 38(2):209–213. [PubMed: 16415888]

42. R Development Core Team R. A Language and Environment for Statistical Computing [Version 2.12.2.]. R Foundation for Statistical Computing; 2011.

43. Sun P, Xia S, Lal B, Eberhart CG, Quinones-Hinojosa A, Maciaczyk J, Matsui W, Dimeco F, Piccirillo SM, Vescovi AL, Laterra J. DNER, an epigenetically modulated gene, regulates glioblastoma-derived neurosphere cell differentiation and tumor propagation. Stem Cells. 2009; 27(7):1473–1486. [PubMed: 19544453]

44. Park JR, Jung JW, Seo MS, Kang SK, Lee YS, Kang KS. DNER modulates adipogenesis of human adipose tissue-derived mesenchymal stem cells via regulation of cell proliferation. Cell Prolif. 2010; 43(1):19–28. [PubMed: 20070733]

45. Thomas G, Jacobs KB, Kraft P, Yeager M, Wacholder S, Cox DG, Hankinson SE, Hutchinson A, Wang Z, Yu K, Chatterjee N, Garcia-Closas M, Gonzalez-Bosquet J, Prokunina-Olsson L, Orr N, Willett WC, Colditz GA, Ziegler RG, Berg CD, Buys SS, McCarty CA, Feigelson HS, Calle EE, Thun MJ, Diver R, Prentice R, Jackson R, Kooperberg C, Chlebowski R, Lissowska J, Peplonska B, Brinton LA, Sigurdson A, Doody M, Bhatti P, Alexander BH, Buring J, Lee IM, Vatten LJ, Hveem K, Kumle M, Hayes RB, Tucker M, Gerhard DS, Fraumeni JF Jr. Hoover RN, Chanock SJ, Hunter DJ. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). Nat Genet. 2009; 41(5):579–584. [PubMed: 19330030]

46. Collado M, Serrano M. The TRIP from ULF to ARF. Cancer Cell. 2010; 17(4):317–318. [PubMed: 20385357]

47. Minakuchi M, Kakazu N, Gorrin-Rivas MJ, Abe T, Copeland TD, Ueda K, Adachi Y. Identification and characterization of SEB, a novel protein that binds to the acute undifferentiated leukemia-associated protein SET. Eur J Biochem. 2001; 268(5):1340–1351. [PubMed: 11231286]

48. Aschard H, Lutz S, Maus B, Duell EJ, Fingerlin TE, Chatterjee N, Kraft P, Van SK. Challenges and opportunities in genome-wide environmental interaction (GWEI) studies. Hum Genet. 2012; 131(10):1591–1613. [PubMed: 22760307]

49. Dempfle A, Scherag A, Hein R, Beckmann L, Chang-Claude J, Schafer H. Gene–environment interactions for complex traits: definitions, methodological requirements and challenges. Eur J Hum Genet. 2008; 16(10):1164–1172. [PubMed: 18523454]

50. Hein R, Beckmann L, Chang-Claude J. Sample size requirements for indirect association studies of gene–environment interactions (G × E). Genet Epidemiol. 2008; 32(3):235–245. [PubMed: 18163529]

51. Lindstrom S, Yen YC, Spiegelman D, Kraft P. The impact of gene–environment dependence and misclassification in genetic association studies incorporating gene–environment interactions. Hum Hered. 2009; 68(3):171–181. [PubMed: 19521099]

52. Morimoto LM, White E, Newcomb PA. Selection bias in the assessment of gene–environment interaction in case-control studies. Am J Epidemiol. 2003; 158(3):259–263. [PubMed: 12882948]

53. Wacholder S, Chatterjee N, Hartge P. Joint effect of genes and environment distorted by selection biases: implications for hospital-based case-control studies. Cancer Epidemiol Biomarkers Prev. 2002; 11(9):885–889. [PubMed: 12223433]

54. Gauderman WJ. Sample size requirements for matched case-control studies of gene–environment interaction. Stat Med. 2002; 21(1):35–50. [PubMed: 11782049]

55. Kraft P. Curses—winner's and otherwise—in genetic epidemiology. Epidemiology. 2008; 19(5): 649–651. [PubMed: 18703928]
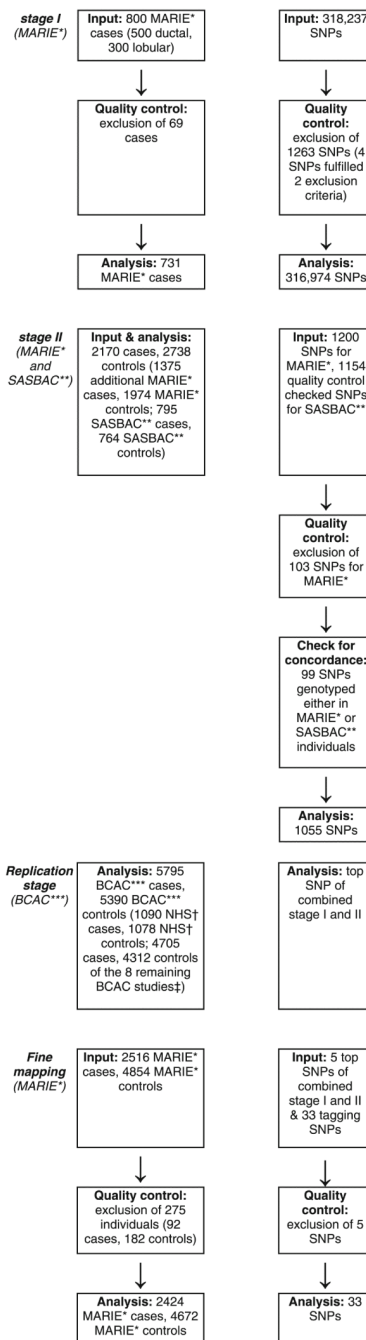
Fig. 1.
Flow chart outlining studies, number of individuals, and number of SNPs involved in each analysis stage. *_MARIE_ Mammary Carcinoma Risk Factor Investigation, **_SASBAC_ Case–Control Study Sample of the Singapore and Sweden Breast Cancer Study, ***_BCAC_ Breast Cancer Association Consortium, _NHS_ Nurses Health Study, _BBCS_ British Breast Cancer Study, _CECILE_ CECILE Breast Cancer Study, _GENICA_ Gene Environment Interaction and Breast Cancer in Germany, _KBCP_ Kuopio Breast Cancer Project, _MCBCS_ Mayo Clinic Breast Cancer Study, _SASBAC_ Case–Control Study Sample of the Singapore and Sweden Breast Cancer Study, _TWBCS_ Taiwanese Breast Cancer Study, _UCIBCS_ UCI Breast Cancer Study. Remark: MARIE cases included in stage II are distinct from those

comprised in stage I. SASBAC cases and controls included in the replication are distinct from those comprised in stage II. MARIE cases and controls included in the fine mapping comprise all cases and controls also included in stages I and II
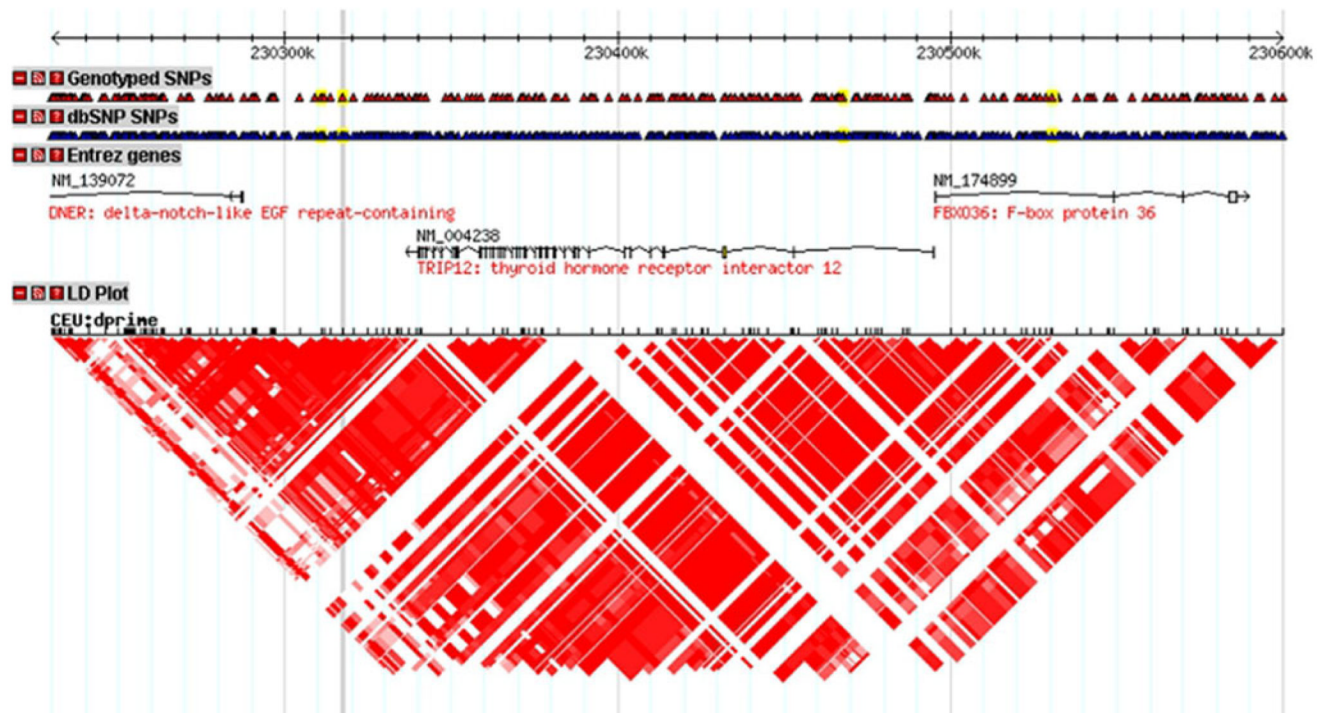
Fig. 2.
Linkage disequilibrium (LD) blocks on chromosome 2, position 230,230,000–230,600,000.
LD in terms of D' between SNPs rs6707272, rs11886634, rs2909969, and rs1942574 with
the genes *DNER*, *TRIP12*, and *FBXO36* on chromosome 2. LD blocks were generated using
data from the HapMap project. The intensity of the colour is proportional to the strength of
the LD for the SNP pair. *Dark red* indicates D =1

Table 1

Effect estimates of the combined stage I and II with P values $<10^{-5}$ for the interaction of current use of MHT and SNP on postmenopausal breast cancer risk as well as corresponding results of the fine mapping

Chromosome	Genes[a]	SNP	Position[b]	Allele change	MAF[c]	P value (HWE[d])	Stage I OR (95 % CI)[e]	Stage II OR (95 % CI)[e]	Combined stage I and II P value	Fine mapping OR (95 % CI)[e]	P value
2	DNER	rs6707272	230317393	C > T	0.31	0.64	0.60 (0.47–0.76)	0.76 (0.62–0.92)	2.98×10^{-7}	0.77 (0.66–0.90)	1.01×10^{-3}
2	TRIP 12	rsl3014061	230467943	C > A	0.33	0.82	0.59 (0.47–0.75)	0.82 (0.67–1.00)	6.61×10^{-7}	0.79 (0.68–0.93)	3.56×10^{-3}
2	FBX036	rs560304	230530540	A > G	0.33	0.95	0.58 (0.46–0.73)	0.83 (0.68–1.00)	5.27×10^{-7}	0.80 (0.68–0.93)	4.39×10^{-3}
7	CALN1/ TYW1B	rs2909969	71620860	G > A	0.49	0.11	0.60 (0.48–0.74)	0.86 (0.72–1.03)	8.14×10^{-7}	1.04 (0.90–1.20)	5.73×10^{-1}
18	SETBP1	rs 1942574	40979527	G > A	0.18	0.38	0.52 (0.40–0.68)	0.93 (0.74–1.18)	5.32×10^{-6}	0.71 (0.59–0.87)	7.08×10^{-4}

[a] DNER delta/notch-like EGF repeat containing. TRIP12: thyroid hormone receptor interactor 12. FBX036: F-box protein 36. CALN1: calneuron 1; TYW1B: tRNA-yW synthesizing protein 1 homolog B (*S. cerevisiae*). SETBP1: SET binding protein 1

[b] Position based on HapMap data Release 27, Genome Build 36

[c] *MAF* minor allele frequency. The MAF presented is based on genotype data of stage II

[d] *HWE* Hardy–Weinberg equilibrium, *P*-values for HWE are presented for the genotype data of stage II

[e] *OR* odds ratio, *CI* confidence interval, for the interaction between SNPs and current MHT, both derived using model 1

Table 2

Estimates of the main effect of current use of **MHT** on postmenopausal breast cancer risk in the fine mapping analysis in strata defined by genotypes of **SNP**s with **P**-values $< 10^{-5}$ for the interaction with current **MHT** use in the combined stage **I** and **II**

SNP	Allele change	Genotype					
		Homozygous major		**Heterozygous**		**Homozygous minor**	
		OR (95 % CI)[a]	**P value**	**OR (95 % CI)**[a]	**P value**	**OR (95 % CI)**[a]	**P value**
rs6707272	C > T	1.92 (1.66–2.23)	4.95×10^{-18}	1.44 (1.23–1.69)	4.28×10^{-6}	1.29 (0.91–1.81)	1.50×10^{-1}
rs13014061	C > A	1.89 (1.63–2.20)	1.70×10^{-16}	1.51 (1.29–1.76)	2.18×10^{-7}	1.25 (0.91–1.73)	1.64×10^{-1}
rs560304	A > G	1.87 (1.61–2.18)	5.66×10^{-16}	1.52 (1.30–1.78)	1.23×10^{-7}	1.25 (0.91–1.72)	1.74×10^{-1}
rs2909969	G > A	1.74 (1.41–2.13)	1.63×10^{-7}	1.44 (1.24–1.66)	1.09×10^{-6}	1.94 (1.58–2.38)	1.60×10^{-10}
rs1942574	G> A	1.83 (1.62–2.07)	7.94×10^{-22}	1.28 (1.06–1.55)	1.20×10^{-2}	1.09 (0.59–2.02)	7.88×10^{-1}

[a]**OR** odds ratio, **CI** confidence interval for the main effect of current **MHT** use