Genetic Variants in Hormone-Related Genes and Risk of Breast Cancer

Clendenen, Tess; Zeleniuch-Jacquotte, Anne; Virgin, Isaac; Koenig, Karen L.; Afanasyeva, Yelena; Lundin, Eva; Arslan, Alan A.; Axelsson, Tomas; Försti, Asta; Hallmans, Goran; Hemminki, Kari; Lenner, Per; Roy, Nirmal; Shore, Roy E.; Chen, Yu

Published in:
PLoS ONE

DOI:
10.1371/journal.pone.0069367

2013

Link to publication

Citation for published version (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 12. Feb. 2019
Abstract

Sex hormones play a key role in the development of breast cancer. Certain polymorphic variants (SNPs and repeat polymorphisms) in hormone-related genes are associated with sex hormone levels. However, the relationship observed between these genetic variants and breast cancer risk has been inconsistent. We conducted a case-control study nested within two prospective cohorts to assess the relationship between specific genetic variants in hormone-related genes and breast cancer risk. In total, 1164 cases and 2111 individually-matched controls were included in the study. We did not observe an association between potential functional genetic polymorphisms in the estrogen pathway, SHBG rs6259, ESR1 rs2234693, CYP19 rs10046 and rs4775936, and UGT1A1 rs8175347, or the progesterone pathway, PGR rs1042838, with the risk of breast cancer. Our results suggest that these genetic variants do not have a strong effect on breast cancer risk.

Methods

Ethics Statement

The Institutional Review Board of New York University School of Medicine and the Regional Ethical Committee of the University...
The concordance between samples from the same participant was duplicate samples from NSMSC participants (n = 164 duplicates). Interspersed throughout the plate with the case-control samples.

**Study Subjects**

We conducted a case-control study nested within two prospective cohorts: the Northern Sweden Mammary Screening Cohort (NSMSC) [40] and the New York University Women’s Health Study (NYUWHS). Details about the parent cohorts and breast cancer case ascertainment have been reported previously [3].

Briefly, the NYUWHS cohort includes 14,274 healthy women (ages 34–65) enrolled between 1985–1991 at a mammography screening clinic in New York City and the NSMSC cohort includes over 28,000 healthy women (ages 40–69) enrolled between 1995–2006 during a population-based breast cancer screening program in Vasterbotten County, Sweden. For the present study, only women self-described as Caucasian, African American, or Hispanic were included. All incident cases of invasive breast cancer, a total of 1164 cases, were included in our study (630 cases from NYUWHS and 506 cases from NSMSC). Two controls were individually matched to each case. Controls were selected at random from members of the same cohort who were alive and free of cancer at the time of diagnosis of the case, and who matched the case on age at enrollment (± 6 months) and date of blood donation (± 3 months). NYUWHS cases and controls were also matched on menopausal status. Most of the cases from the NSMSC had at least one control matched on menopausal status (92%). In total, 2111 controls were included in the study (1099 from NYUWHS and 1012 from NSMSC).

**Laboratory Methods**

For the NYUWHS participants, DNA was extracted from blood clots or cell precipitates (prepared by centrifugation of whole blood collected at blood donation) for 42% of participants. For the remainder of NYUWHS participants, DNA was extracted from serum. Samples were genotyped using the TaqMan® approach [41,42] with an ABI 7900 Real-Time PCR instrument (Applied Biosystems, Foster City, CA). The percent of successful genotyping calls was ≥98% for all genetic variants. Prior to the case-control study, a pilot study was conducted to examine genotype concordance across sample types (serum, clots, cell precipitates) for the NYUWHS study. For samples from the same participant (n = 50 subjects with all three sample types plus n = 68 subjects with two sample types), genotype concordance was ≥99%.

For the NSMSC participants, DNA was isolated from buffy coats. Genotyping was performed at the SNP Technology Platform at Uppsala University Hospital (www.genotyping.se) for five SNPs (rs4775936, rs10046, rs6259, rs2234693, rs1042838). Four of these SNPs (rs4775936, rs10046, rs6259, rs2234693) were assayed using the GenomeLab SNPStream 12plex-system (Beckman Coulter) and one (rs1042838) using the FP-TDI system. UGT1A1 rs8175347 was assayed at the German Cancer Research Center in Heidelberg using fluorescent fragment analysis on an ABI PRISM 3100 Genetic analyzer with the GeneMapper software version 3.0 (Applied Biosystems). The percentage of samples with successful calls was ≥98% for all polymorphisms. We conducted a pilot study to assess genotype concordance across duplicate samples from NSMSC participants (n = 164 duplicates). The concordance between samples from the same participant was ≥99% for all loci.

For both cohorts, each case and her individually-matched controls were analyzed as a set on the same 96-well plate. Quality control samples (10%) were included on each plate and were interspersed throughout the plate with the case-control samples.

Laboratory personnel were blinded as to case-control status and the identity of the quality control samples.

**Statistical Methods**

We assessed deviation from Hardy-Weinberg equilibrium (HWE) in each cohort for each genetic variant with a chi-square test. Odds ratios and 95% confidence intervals for breast cancer risk were estimated using the conditional logistic regression model, as appropriate for the matched study design. All models were adjusted for race/ethnicity (Caucasian, African American, or Hispanic) and through matching, were also adjusted for age at blood donation, duration of sample storage, and menopausal status. Multivariate-adjusted models included other known risk factors for breast cancer: family history of breast cancer (in a first degree relative), age at menarche, age at first birth/parity (≥20 years, 21–25 years, 26–30 years, >30 years, nulliparous), ever use of hormone replacement therapy, and body mass index (BMI). For covariates with missing data (age at menarche, age at first birth, use of hormone replacement therapy, and BMI), we performed multiple imputation of missing data for each cohort separately using a fully conditional specification model [43] including family history of breast cancer and case-control status along with the imputed variables. Each of the imputed variables had <4% missing data. We also conducted a logistic regression including all of the estrogen-related variants simultaneously (all variants except PGR-12), where homozygous genotypes for the variant associated with higher estrogen was coded as one and the other genotypes were coded as zero and the sum of the scores was modeled as the independent variable (women with four or five variants were grouped because there were only 3 cases and 4 controls who had five high estrogen variants). Heterogeneity between cohorts was assessed by comparing models with cross-product terms (cohort × genotype) to models excluding them using the likelihood ratio test. Analyses were also conducted separately for Caucasians (n = 1067 cases, 1931 controls) and for estrogen receptor positive (ER+) breast cancer (n = 625 cases, 1091 controls).

**Results**

Descriptive statistics for the cases and controls are shown in Table 1. The expected relationships between breast cancer and the traditional risk factors were observed. Among postmenopausal women, cases had a higher mean BMI than controls. Cases were more likely than controls to be nulliparous, to have ever used HRT, and to have a family history of breast cancer. Among parous women, average age at first birth was greater for cases than controls. The frequency of genotypes within controls did not deviate from Hardy-Weinberg Equilibrium by cohort (all p-values >0.05).

Table 2 shows that there were no statistically significant associations between the selected genetic variants in hormone-related genes and risk of breast cancer in age- or multivariate-adjusted models. For each genotype, variants are listed in order of expected increasing estrogen (or progesterone for the PGR SNP) exposure. There was no association with risk for individuals with multiple genotypes associated with high estrogen levels. Tests for heterogeneity by cohort were not significant. The odds ratios were not appreciably different in analyses restricted to Caucasians (data not shown). Tests for interaction between each genetic polymorphism and age at diagnosis were not significant. ORs were not significant and were generally similar in magnitude and direction for analyses restricted to ER+ breast cancer (data not shown), except that the OR estimates were no longer greater than one for the CYP19 rs10046 TT and rs4775936 AA genotypes. For SHBG...
rs6259, the OR was somewhat greater for ER+ tumors (OR for GG vs. GA/AA: 1.22, 95% CI: 0.94–1.60) than for all tumors combined (for GG vs. GA/AA: 1.02, 95% CI: 0.84–1.23), though the association was not significant.

**Discussion**

Our results are in agreement with those of large meta-analyses of hormone-related genetic variants and breast cancer risk (which included over 10,000 cases and 10,000 controls) that did not observe an association for **SHBG** rs6259, **PGR** rs1042838, **CYP19** rs10046 and rs4775936, or **UGT1A1** rs8175347 variants [32,44]. Two meta-analyses (both with over 10,000 cases and 10,000 controls), observed a borderline inverse association for **ESR1** rs2234693 (OR for C vs. T: 0.97, 95% CI: 0.93–1.00 [32], p = 0.055 and OR for CC vs. TT: 0.92, 95% CI: 0.86–0.99 [28]), which we did not observe in our study.

Three of the genetic variants we selected directly influence estrogen levels: **CYP19** rs10046 and rs4775936 and **UGT1A1** rs8175347. **CYP19** encodes for the enzyme aromatase, which converts androgens to estrogens. The rs10046 T allele and the rs4775936 A allele have been shown to be associated with higher levels of circulating estrogen or estrogen to androgen ratios in several studies [5,7,29,45], including our own [46]. The 7 base repeat allele of **UGT1A1** rs8175347 is associated with lower transcriptional activity and may result in reduced glucuronidation of estrogens [34]. Circulating estrogen levels were higher among women with 7 base repeat alleles in our study [46] as in most other studies [34,47]. However, despite the relationship between these SNPs and estrogen levels, most studies, including our own, have not observed any association between these SNPs in **CYP19** and **UGT1A1** and breast cancer risk [32,48].

**SHBG** rs6259 influences estrogen bioavailability and the A allele is associated with higher SHBG levels (suggesting lower estrogen bioavailability) in most [5,8,18,19,49], but not all studies [6,50]. Consistent with most previous studies [32], we did not observe an association between rs6259 and risk of breast cancer.

**PGR** rs1042838 influences progesterone signaling and the T allele is in high linkage disequilibrium with the PROGINS allele, which reduces the **PGR** transcript stability and may decrease the response to progesterone [23]. We did not observe an association between **PGR** rs1042838 and breast cancer risk, in agreement with nearly all previous studies [32].
Table 2. Associations between genetic variants in hormone-related genes and breast cancer risk in pre- and post-menopausal women.

<table>
<thead>
<tr>
<th>n (Cases/Controls)</th>
<th>Ethnicity-adjusted ORs (95% CI)</th>
<th>Multivariate-adjusted ORs (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SHBG (rs6259)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>15/28</td>
<td>1.00</td>
</tr>
<tr>
<td>GA</td>
<td>197/351</td>
<td>1.00 (0.53–1.92)</td>
</tr>
<tr>
<td>GG</td>
<td>933/1655</td>
<td>1.02 (0.54–1.91)</td>
</tr>
<tr>
<td>p-trend</td>
<td>0.90</td>
<td>0.89</td>
</tr>
<tr>
<td>GG vs. GA/AA</td>
<td>1.01 (0.84–1.22)</td>
<td>1.02 (0.84–1.23)</td>
</tr>
<tr>
<td><strong>PGR-12 (rs1042838)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>26/54</td>
<td>1.00</td>
</tr>
<tr>
<td>GT</td>
<td>288/523</td>
<td>1.14 (0.70–1.86)</td>
</tr>
<tr>
<td>GG</td>
<td>846/1516</td>
<td>1.20 (0.74–1.93)</td>
</tr>
<tr>
<td>p-trend</td>
<td>0.42</td>
<td>0.44</td>
</tr>
<tr>
<td>GG vs. GT/TT</td>
<td>1.06 (0.90–1.25)</td>
<td>1.05 (0.89–1.24)</td>
</tr>
<tr>
<td><strong>ESR1 (rs2234693)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>334/660</td>
<td>1.00</td>
</tr>
<tr>
<td>CT</td>
<td>585/1010</td>
<td>1.14 (0.96–1.35)</td>
</tr>
<tr>
<td>CC</td>
<td>244/436</td>
<td>1.10 (0.90–1.35)</td>
</tr>
<tr>
<td>p-trend</td>
<td>0.28</td>
<td>0.28</td>
</tr>
<tr>
<td><strong>CYP19 3’UTR (rs10046)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>306/549</td>
<td>1.00</td>
</tr>
<tr>
<td>CT</td>
<td>548/1032</td>
<td>0.95 (0.80–1.13)</td>
</tr>
<tr>
<td>TT</td>
<td>308/523</td>
<td>1.07 (0.87–1.31)</td>
</tr>
<tr>
<td>p-trend</td>
<td>0.53</td>
<td>0.71</td>
</tr>
<tr>
<td><strong>CYP19 5’Flank (rs4775936)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>361/654</td>
<td>1.00</td>
</tr>
<tr>
<td>GA</td>
<td>531/1011</td>
<td>0.94 (0.79–1.12)</td>
</tr>
<tr>
<td>AA</td>
<td>271/438</td>
<td>1.13 (0.92–1.39)</td>
</tr>
<tr>
<td>p-trend</td>
<td>0.33</td>
<td>0.46</td>
</tr>
<tr>
<td><strong>UGT1A1 (rs8175347)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6/6</td>
<td>510/938</td>
<td>1.00</td>
</tr>
<tr>
<td>6/7</td>
<td>478/846</td>
<td>1.03 (0.88–1.21)</td>
</tr>
<tr>
<td>7/7</td>
<td>151/257</td>
<td>1.06 (0.84–1.34)</td>
</tr>
<tr>
<td>p-trend</td>
<td>0.60</td>
<td>0.51</td>
</tr>
<tr>
<td>Combined model for estrogen pathway variants²</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0 high E genotypes</td>
<td>108/190</td>
<td>1.00</td>
</tr>
<tr>
<td>1 high E genotypes</td>
<td>485/911</td>
<td>0.93 (0.72, 1.21)</td>
</tr>
<tr>
<td>2 high E genotypes</td>
<td>267/441</td>
<td>1.07 (0.81, 1.41)</td>
</tr>
<tr>
<td>3 high E genotypes</td>
<td>194/310</td>
<td>1.12 (0.83, 1.50)</td>
</tr>
<tr>
<td>4–5 high E genotypes</td>
<td>64/106</td>
<td>1.06 (0.71, 1.57)</td>
</tr>
<tr>
<td>p-trend</td>
<td>0.15</td>
<td>0.21</td>
</tr>
</tbody>
</table>

*Models were adjusted for ethnicity (Caucasian, African American, Hispanic) and through matching, were also adjusted for age at blood donation, duration of sample storage, and menopausal status.

²Models were adjusted for ethnicity (Caucasian, African American, Hispanic), age at first birth/parity (≤20 years, 21–25 years, 26–30 years, >30 years, nulliparous), age at menarche, family history of breast cancer, ever use of HRT, and BMI, and through matching, were also adjusted for age at blood donation, duration of sample storage, and menopausal status.

For each genetic variant (except PGR-12 rs1042838), the genotype associated with higher estrogen exposure (see below) was assigned a value of 1 and other genotypes (homozygous and heterozygous for the lower estrogen exposure allele) were assigned 0. A score was created by adding the values. Women with four or five high estrogen variables were grouped because there were too few women with five high estrogen variables to assess separately (3 cases/4 controls).

**Notes:** For each genotype, variants are listed in order of expected increasing estrogen (or progesterone for PGR-12) exposure:

SHBG (rs6259) A allele is associated with higher SHBG levels. SHBG binds to estrogens and reduces their bioavailability. G allele = higher estrogen exposure.

PGR-12 (rs1042838) T allele has been shown to reduce the PGR transcript stability and the response to progesterone. G allele = higher progesterone exposure.
The function of the \textit{ESR1} rs2234693 SNP has not been clearly demonstrated, though the C allele produces a binding site for the B-myb transcription factor, which may result in an alternative demonstration, though the C allele produces a binding site for the B-myb transcription factor, which may result in an alternative transcriptional activity of aromatase (converts androgens to estrogens). T allele = higher estrogen exposure. \textit{CYP19} 3’ Flank (rs4775936) A allele is associated with increased transcriptional activity of aromatase (converts androgens to estrogens). A allele = higher estrogen exposure.

\textit{UGT1A1} (rs8175347) 7 repeat allele is associated with reduced glucuronidation (and clearance) of estrogens. 7 repeat allele = higher estrogen exposure.

doi:10.1371/journal.pone.0069367.t002

Table 2 Cont.

\begin{table}[h]
\centering
\begin{tabular}{|l|l|}
\hline
\textbf{Gene} & \textbf{Description} \\
\hline
\textit{ESR1} & rs2234693 C allele may be associated with increased E\(_2\)R3 transcription. C allele = possible higher exposure to estrogen signaling. \\
\textit{CYP19} 3’ UTR (rs10046) T allele is associated with increased transcriptional activity of aromatase (converts androgens to estrogens). T allele = higher estrogen exposure. 7 repeat allele = higher estrogen exposure. \\
\hline
\end{tabular}
\end{table}

References


