Short Communication

DNA Repair Gene Variants Associated with Benign Breast Disease in High Cancer Risk Women

Timothy J. Jorgensen,1,2 Kathy J. Helzlsouer,2,4 Sandra C. Clipp,2 Judy Hoffman Bolton,2 Rosa M. Crum,2 and Kala Visvanathan2,3

1Department of Radiation Medicine, Lombardi Comprehensive Cancer Center, Georgetown University, Washington, District of Columbia and 2Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health; 3Department of Oncology, Johns Hopkins University School of Medicine; 4Prevention and Research Center, Mercy Medical Center, Baltimore, Maryland

Abstract

Benign breast disease (BBD) is a risk factor for breast cancer and may have a heritable component. Deficient DNA repair has been implicated in breast cancer etiology and may exert its effect before BBD, a known precursor. The association between allelic variants in DNA repair genes and BBD was examined in a cohort of women in Washington County, Maryland. BBD was defined by two criteria: (a) a physician diagnosis of BBD or fibrocystic disease and/or (b) a benign breast biopsy. 3,212 women without BBD at baseline were genotyped for 12 candidate single nucleotide polymorphisms in seven DNA repair genes. Of these women, 482 subsequently reported a diagnosis of BBD. The Cox model was used to calculate hazard ratios (HR). Variant alleles of XRCC1 Arg194Trp (rs1799782) and ERCC4 Arg139Gln (rs1800067) were significantly associated with BBD [HR, 1.36; 95% confidence interval (95% CI), 1.09-1.76]. Similar estimates were also observed for each of the BBD criterion used. The BBD association for ERCC4 was even stronger among women with a family history of breast cancer (HR, 2.68; 95% CI, 1.52-4.66; \( P_{\text{interaction}} = 0.02 \)). This study suggests that variant alleles in DNA repair genes may modify BBD risk, a potential intermediate marker of breast cancer risk, particularly among high-risk subgroups. (Cancer Epidemiol Biomarkers Prev 2009;18(1):346–50)

Introduction

Benign breast disease (BBD) is a known precursor of breast cancer. However, unlike breast cancer, the importance of heritability in its etiology has been hard to discern. Positive associations between women with a family history of breast cancer and biopsy-proven BBD have been observed in both prospective and case-control studies (1-3). To what extent these positive findings are due to increased breast surveillance and a greater tendency for health providers to suggest a biopsy in women with a family history compared with those without has been unclear (4). Evidence against a detection bias includes a large population-based study where BBD risk was observed to be significantly increased among first-degree relatives of probands with breast cancer diagnosed at age <50 years, even before the diagnosis of breast cancer (2). This finding strengthens the evidence of a role for genetic factors in the etiology of BBD. Women with BBD are ~2-fold as likely to develop breast cancer, with subtypes of BBD showing different levels of risk. For example, the relative risk of non-proliferative BBD has been reported to range between 1.27 and 1.95, whereas proliferative BBD with atypia ranges from 2.58 to 4.00 (3, 5-8). However, little is known about the genetic determinants of BBD and its subtypes.

DNA repair genes are known to be important in hereditary breast cancer. The two known familial breast cancer genes (BRCA1 and BRCA2) are both DNA repair genes. Polymorphic variants in DNA repair genes have also been associated with many cancers (9), including breast, and with some precursor lesions such as colon adenomas (10) and oral premalignant lesions (11). Because DNA repair occurs early in the breast carcinogenesis pathway, it is logical that polymorphic forms of DNA repair genes may alter BBD risk. No prior studies have reported on the association between DNA repair variants and BBD.

We describe here a large population-based cohort of women who were followed for the development of BBD over a 13-year period. Twelve potentially functional single nucleotide polymorphisms (SNP) in seven candidate genes, within DNA repair pathways implicated previously in cancer etiology, were assessed for their association with BBD incidence.
Materials and Methods

Clue II (n = 30,726) is a population-based cohort of men and women established in 1989 among individuals residing in the Washington County, Maryland area (12). Blood and brief medical histories were collected at baseline. From Clue II, a subset of women with no prior history of BBD or cancer (other than nonmelanoma skin cancer) at baseline was identified (n = 7,657). From this larger sample of 7,657, our study population was composed of 3,212 women who had been genotyped for 12 SNPs in the coding regions of seven DNA repair genes suspected to be involved in cancer, as part of an earlier study (13, 14).

BBD incident cases, family history, and breast cancer risk factors were identified from follow-up questionnaires completed in 1996, 1998, 2000, and 2003. Two criteria were used to define BBD cases: (a) the respondent reported having a benign breast biopsy on a date after baseline (n = 500) or (b) the respondent reported being told by a physician that they had BBD or fibrocystic disease of the breast on a date after baseline (n = 924). More than 85% accuracy of the self-reported BBD was able to be confirmed by pathologic report from a sampling of the respondents to the 1996 questionnaire (n = 110). Sixty-eight percent of these pathology reports indicated a diagnosis of nonproliferative BBD, 28% reported a diagnosis of proliferative disease without atypia, and 4% reported a diagnosis of proliferative disease with atypia (15). For women who reported a benign biopsy, their noncancer status was confirmed through linkage to both the Washington County and the Maryland State cancer registries and by examination of available pathologic records (15). A total of 1,163 women were defined as BBD cases by one or both criteria. Women who gave no date of diagnosis for either criterion were excluded from the analysis (n = 202). Of the 961 remaining BBD cases, 482 were among the genotyped group. Note: The Institutional Review Board at Johns Hopkins Bloomberg School of Public Health approved this study.

All genotypes were assessed using the patented fluorogenic method for nucleic acid analysis commonly known as the TaqMan 5′-nucleotide assay (Applied Biosystems Division, Perkin-Elmer) as described previously (13). To assess the reliability of the genotyping methodology, the ERCC2 codon 751 and XRCC1 codons 399 and 194 genotyping was also conducted using a PCR-restriction fragment length polymorphism technique with primers as described previously (16). A total of 230 samples were genotyped using both platforms. Concordance was >94%, suggesting that our findings were not genotyping platform specific.

BBD free survival time was measured similar to earlier BBD cohort studies (17-20). Individuals were censored either at the date of their last questionnaire response or at the date of a diagnosis of breast cancer. Time-to-event data were fitted to the Cox proportional hazards model, with BBD diagnosis being the failure event. Hazard ratios (HR) were calculated for carriers of the variant allele, with the common allele (wild-type) as the reference. The risk associated with the variant allele was fitted to dominant, codominant, or recessive models of inheritance based on principles of Mendelian inheritance at a single locus (21). In the dominant model, carriers with either one or two variant alleles are considered at equal risk, and the HR reflects the BBD risk associated with being either heterozygous or homozygous for the variant allele. In the recessive model, only carriers with two variant alleles are considered to be at increased risk, and the HR reflects the risk associated with being homozygous for the variant allele. In the codominant (linear) model, carriers with two variant alleles are considered to have more risk than carriers with only one variant allele, and the HR is the measure of association per variant allele carried. Associations with P ≤ 0.05 were considered to be statistically significant. All models were adjusted for age at baseline. Known breast cancer risk factors were evaluated using forward and backward stepwise selection. None of the covariates significantly altered the point estimates presented.

Table 1. Association between the variant (minor) allele of each DNA repair gene SNP and BBD, 1989-2003

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNP codon (rs no.)*</th>
<th>MAF †</th>
<th>HWE ‡</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Dominant† inheritance model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Codominant‡ inheritance model</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Recessive‡ inheritance model</td>
</tr>
<tr>
<td>ERCC2</td>
<td>Asp312Asn (rs1799793)</td>
<td>0.34</td>
<td>0.91</td>
<td>1.13 (0.94-1.37)</td>
</tr>
<tr>
<td>ERCC2</td>
<td>Lys791Gln (rs13181)</td>
<td>0.37</td>
<td>0.32</td>
<td>1.08 (0.90-1.31)</td>
</tr>
<tr>
<td>ERCC4</td>
<td>Arg415Gln (rs1800067)</td>
<td>0.07</td>
<td>0.60</td>
<td>1.39 (1.09-1.76)</td>
</tr>
<tr>
<td>ERCC5</td>
<td>Asp1108His (rs17655)</td>
<td>0.21</td>
<td>0.71</td>
<td>0.98 (0.81-1.17)</td>
</tr>
<tr>
<td>RAD23B</td>
<td>Ala799Val (rs1805329)</td>
<td>0.18</td>
<td>0.36</td>
<td>0.89 (0.73-1.09)</td>
</tr>
<tr>
<td>XPC</td>
<td>Val446Asp (rs2227999)</td>
<td>0.05</td>
<td>0.003</td>
<td>0.85 (0.62-1.16)</td>
</tr>
<tr>
<td>XPC</td>
<td>Ala868Val (rs2228000)</td>
<td>0.24</td>
<td>0.23</td>
<td>0.98 (0.81-1.18)</td>
</tr>
<tr>
<td>XPC</td>
<td>Arg697Arg (rs3731151)</td>
<td>0.23</td>
<td>0.15</td>
<td>1.05 (0.87-1.26)</td>
</tr>
<tr>
<td>XPC</td>
<td>Lys899Gln (rs2228001)</td>
<td>0.42</td>
<td>0.34</td>
<td>1.02 (0.85-1.24)</td>
</tr>
<tr>
<td>XRCC1</td>
<td>Lys149Trp (rs1799782)</td>
<td>0.07</td>
<td>0.03</td>
<td>1.36 (1.06-1.74)</td>
</tr>
<tr>
<td>XRCC1</td>
<td>Gin399Arg (rs25487)</td>
<td>0.35</td>
<td>0.24</td>
<td>0.91 (0.76-1.09)</td>
</tr>
<tr>
<td>XRCC2</td>
<td>Arg386His (rs3218536)</td>
<td>0.08</td>
<td>0.22</td>
<td>0.87 (0.66-1.12)</td>
</tr>
</tbody>
</table>

*SNP codon and reference SNP number (rs no.) from National Center for Biotechnology Information of NIH.
†Minor allele frequencies.
‡Hardy-Weinberg equilibrium test P values are shown.
§In the dominant model, carriers with one or two variant alleles are at equal risk.
¶In the codominant model, carriers with two variant alleles have more risk than carriers with just one variant allele. The HR for the codominant model is a measure of the associated risk per variant allele carried.
∥In the recessive model, carriers with two variant alleles are at increased risk.


Downloaded from cebp.aacrjournals.org on June 20, 2017. © 2009 American Association for Cancer Research.
estimates (all had \( P \) values > 0.05). Therefore, only age
was included in the final Cox models.

**Results**

BBD was associated with a significantly increased risk
of developing breast cancer [odds ratio, 2.13; 95%
confidence interval (95% CI), 1.60-2.82] in the overall
group (\( n = 7,657 \)), which is consistent with previous
studies in other populations (3, 5-7, 22). Age-adjusted
comparisons of breast cancer risk factors showed no
significant difference in risk factors between the geno-
typed (\( n = 3,212 \)) and the nongenotyped (\( n = 4,445 \)),
suggesting that the genotyped sample was representative
of the overall group. All SNP genotypes were tested for
Hardy-Weinberg equilibrium. Only one, XPC Val492Ile
(rs2227999), had a significant departure from equilibrium
(\( P = 0.0003 \); Table 1). However, three other SNPs within
the same gene had nonsignificant \( P \) values for the Hardy-
Weinberg equilibrium test, suggesting that the alleles of
this gene actually are in population equilibrium and that
some type of genotyping call bias is likely responsible for
the apparent disequilibrium of this SNP.

None of the variant alleles were associated with
significantly elevated risk under a recessive inheritance
model (two variant alleles required to confer risk).
However, for dominant and codominant inheritance
models (one variant allele sufficient to confer risk), two
SNPs had significant associations with BBD (Table 1).
These SNPs were XRCC1 Arg194Trp (rs1799782) and
ERCC4 Arg415Gln (rs1800067), and their HRs (95% CIs)
for dominant inheritance were 1.36 (1.06-1.74) and 1.39
(1.09-1.76), respectively. XRCC1 codes for a protein
that interacts with DNA ligase III and is a central player in
the base excision repair pathway. ERCC4 (also known as XPF)
is an essential gene in the nucleotide excision pathway.
Both are important to the repair of base damage in DNA.

Similar point estimates and 95% CIs were also
obtained for codominant inheritance (Table 1), suggest-
ing that the findings of associations are robust and not
dependent on the assumption of a dominant mode of
inheritance. Also, Weibull and logarthmic survival
models gave similar results to the Cox model, suggesting
that the findings were not survival model dependent.
Point estimates for these variant alleles were positively
associated with BBD irrespective of the criterion used to
define the case (Table 2). Furthermore, point estimates
did not change when both variant alleles were entered
simultaneously into the model compared with the
univariate analysis, suggesting that they were independ-
ent risk factors.

Because having a family history of breast cancer,
which is thought to have a genetic basis, is a strong risk
factor for BBD (6), we stratified our analysis by family
history (Table 3). For the XRCC1 gene, the HRs in the
two strata was similar. In contrast, for ERCC4, we
found that, for those women with a family history of
breast cancer, the HR (95% CI) for the variant allele was
even greater \([2.68 (1.52-4.66)]\) when compared with
those without a positive family history \([1.27 (0.98-1.66); \]
\( P \) interaction = 0.02).

**Discussion**

BBD is a risk factor for breast cancer that persists for at
least 25 years (5-7) and is likely to have an inherited
component (1-3). The cumulative risk of breast cancer
after a diagnosis of atypical hyperplasia may be as high as
18% at 25 years for a single focus and 40% for multiple foci
(23). Furthermore, BBD can be prevented by tamoxifen,
a proven chemoprevention agent for breast cancer (24).
Therefore, identification of genetic risk factors associated
with a diagnosis of BBD and its high risk subtypes may
allow for the early introduction of prevention strategies
that target BBD and, in turn, breast cancer.

Cellular DNA damage is thought to initiate carcino-
genesis. Because BBD is an intermediate phenotype
within the breast carcinogenesis pathway, deficiencies
in DNA repair associated with breast cancer are likely to
be associated with BBD as well. Our results support the
contention that allelic variants in the nucleotide excision
and base excision repair pathways (ERCC4 and XRCC1,
respectively) may be linked to BBD. ERCC4 codes for a
5’-endonuclease in the nucleotide excision repair path-
way, whereas XRCC1 codes for an important component
in the base excision repair pathway (25), suggesting that
base damage repair pathways may be important in
protecting against breast carcinogenesis. The 194 codon

**Table 2. Association between the XRCC1 and ERCC4 variant SNP alleles and BBD by different criteria, 1989-2003**

<table>
<thead>
<tr>
<th>Case criteria*</th>
<th>No. BBD cases</th>
<th>XRCC1 codon 194 HR (95% CI)</th>
<th>ERCC4 codon 415 HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criterion 1: negative breast biopsy</td>
<td>212</td>
<td>1.52 (1.06-2.16)</td>
<td>1.12 (0.77-1.65)</td>
</tr>
<tr>
<td>Criterion 2: informed by physician</td>
<td>380</td>
<td>1.38 (1.05-1.82)</td>
<td>1.35 (1.03-1.77)</td>
</tr>
<tr>
<td>Criterion 1 or 2</td>
<td>479</td>
<td>1.36 (1.06-1.74)</td>
<td>1.39 (1.09-1.76)</td>
</tr>
</tbody>
</table>

*HRs were calculated for the risk allele of each SNP using the prone model.

**Table 3. Association between the XRCC1 and ERCC4 variant SNP allele and BBD stratified by family history, 1989-2003**

<table>
<thead>
<tr>
<th>Family history*</th>
<th>n</th>
<th>HR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>XRCC1 codon 194</td>
</tr>
<tr>
<td>Negative</td>
<td>2,699</td>
<td>1.35 (1.03-1.77)</td>
</tr>
<tr>
<td>Positive</td>
<td>423</td>
<td>1.47 (1.01-2.08)</td>
</tr>
</tbody>
</table>

*HRs were calculated for each stratum using the Cox proportional model and dominant inheritance.

NOTE: The Institutional Review Board at the Johns Hopkins Bloomberg School of Public Health approved this study.

*Family history of breast cancer was defined as women with a first- or second-degree relative with breast cancer.
of XRCC1 is located in a hydrophobic linker region between its DNA polymerase β domain and poly(ADP-ribose) polymerase-interacting domains, so the change from arginine to tryptophan could alter the interaction of XRCC1 with either or both of these DNA repair proteins within the base excision repair complex. The functional role of codon 415 in ERCC4 is less clear, but some disease-related mutations in ERCC4 map close to Arg^{415}Gln within exon 8 suggesting functional significance. Our findings of a strong interaction between ERCC4 and family history of breast cancer is of particular interest. Having a family history of breast cancer is a major BBD and breast cancer risk factor (1, 2, 26-28), but it is not clear what a family history of breast cancer reflects in terms of carcinogenic mechanisms. It could be related to differential estrogen metabolism. Estrogen metabolites have been shown to form base adducts in DNA (29). These adducts can either be removed by the nucleotide excision repair pathway or spontaneously depurate to produce apurinic sites that are then repaired by base excision repair. Thus, estrogens metabolites may directly produce many of the base lesions that ERCC4 and XRCC1 proteins repair. It has also been shown that estrogens can induce reactive oxygen species that are well-known base-damaging agents (30-33). Alternatively, it could be that family history of breast cancer may partly reflect a heritable propensity to increased intracellular oxidative stress, which is known to produce base damage (34), thus rendering these women vulnerable to deficiencies in base damage repair.

In this genotyped cohort, our ability to also evaluate the association between these two DNA repair SNPs and breast cancer was limited by the number of breast cancer cases accrued thus far (n = 35). Longer follow-up is needed to evaluate the association between the SNPs and BBD progression to breast cancer within this population. However, nonstatistically significant positive associations were observed previously in an earlier case-control study within Clue II between these two DNA repair SNPs and breast cancer (13, 35). The variant alleles for both of these genes have also been reported previously to be breast cancer-associated, with marginal significance, in a hospital-based case-control study (36). Furthermore, the ERCC4 gene has recently shown to be associated with premenopausal breast cancer in a nested case-control study within the Nurses’ Health Study II (37). Taken together, these studies support the notion that DNA repair genes may have a common role in the etiology of both BBD and breast cancer.

Although the sample size of this prospective study is large, one limitation is the lack of information on histologic subtypes of BBD. Nevertheless, if the association reported here is restricted to proliferative BBD, which is a strong risk factor for invasive breast cancer, the magnitude of the association would be even larger than our observed overall point estimates. Subsequent investigations should explore potential differences in association by subtype.

In summary, variant ERCC4 and XRCC1 genotypes are statistically significantly associated with BBD in this population. For ERCC4, the risk is particularly increased in women with both the variant allele and a family history of breast cancer, suggesting that this subgroup of women is at much higher risk for DNA damage-induced BBD than the general population. Studying precursor lesions, such as BBD, which appear earlier than cancerous tumors, may also represent a powerful tool to explore potential associations between genotypes and cancer risk, particularly within high-risk subpopulations.

**Disclosure of Potential Conflicts of Interest**
No potential conflicts of interest were disclosed.

**Acknowledgments**
We acknowledge and thank all the participants of Clue II.

**References**
20. Webb PM, Byrne C, Schnitt SJ, et al. A prospective study of diet and...


DNA Repair Gene Variants Associated with Benign Breast Disease in High Cancer Risk Women

Timothy J. Jorgensen, Kathy J. Helzlsouer, Sandra C. Clipp, et al.


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/18/1/346

Cited articles
This article cites 35 articles, 16 of which you can access for free at:
http://cebp.aacrjournals.org/content/18/1/346.full#ref-list-1

Citing articles
This article has been cited by 3 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/18/1/346.full#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.