Short Communication

Common Polymorphisms in Checkpoint Kinase 2 Are not Associated with Breast Cancer Risk


Cancer Research UK Human Cancer Genetics Research Group, Department of Oncology [B. K., A. A., C. S. H., B. A. J. P., A. M. D., P. D. P. F.], Cancer Research UK Genetic Epidemiology Group [D. F. E.], and Department of Public Health and Primary Care [N. E. D.], Strangeways Research Laboratories, Cambridge CB1 8RN, United Kingdom; Department of Obstetrics and Gynecology, Technical University Munich, 81675 Munich, Germany [B. K.]; and Department of Obstetrics and Gynecology, Turku University, Finland [A. A.]

Abstract

A substantial proportion of the familial risk of breast cancer may be attributable to genetic variants each contributing a small effect. Polymorphisms in DNA repair genes are good candidates for such low penetrance breast cancer susceptibility alleles. Checkpoint kinase 2 (CHEK2) is a kinase in which the yeast counterpart regulates a cell cycle checkpoint and causes cells to arrest proliferation after DNA damage. A rare, protein truncating mutation in the CHEK2 gene has recently been shown to confer a modest risk of breast cancer. The aim of this study was to determine whether common polymorphic variants in CHEK2 are associated with an increase in breast cancer risk. We assessed two variants in CHEK2 using a case control study design (n = 1786 cases and 1828 controls). No differences in genotype frequencies were found between cases and controls for either the IVS1 + 38insa or the a1013g polymorphisms (P = 0.3 and 0.2 respectively), and no genotype-specific risk was significantly different from unity. The haplotype frequency distribution in cases and controls were also similar (P = 0.3). We conclude that the CHEK2 polymorphisms IVS + 1a and a1013g do not confer an increased risk of breast cancer. It is also unlikely that other, as yet unidentified, common polymorphisms that affect risk are present in the gene in the British population.

Introduction

Mutations in the known breast cancer predisposition genes, BRCA1 and BRCA2, confer a high risk of breast and ovarian cancer but account for <25% of the excess familial risk of breast cancer (1, 2). These and other data suggest that there are other breast cancer susceptibility genes. A recent segregation analysis found that the familial aggregation of breast cancer in the population that is not because of BRCA1 or BRCA2 was best explained by a polygenic model (3). According to this model, susceptibility to breast cancer is conferred by a large number of alleles each contributing a small effect.

Polymorphisms in genes involved in the DNA double-strand break repair pathways are good candidates for such low penetrance breast cancer susceptibility alleles. A functional link between DNA damage and BRCA1 function has been observed when cells are treated with ionizing radiation, an effective method of inducing double-strand breaks in DNA. This leads to phosphorylation of BRCA1, a response characteristic of functional activation (4). CHEK2 is a kinase in which the yeast counterpart regulates a cell cycle checkpoint and causes cells to arrest proliferation after DNA damage. A protein truncating mutation, responsible for the activating phosphorylation of BRCA1 and also TP53, a protein with a central role in DNA damage checkpoint pathways. In mammalian cells, CHEK2 is regulated by phosphorylation by the product of the ATM gene, which also activates BRCA1 and TP53. A protein-truncating mutation, 1100delc, which abolishes the kinase function of CHEK2, has been found to be associated with breast cancer risk in families negative for BRCA1/2 mutations (6).

The aim of this study was to determine whether common polymorphic variants in CHEK2 are associated with an increase in breast cancer risk. We examined two variants: an insertion of a single nucleotide in intron 1 [dbSNP:3841692 RefSeq NT011520 8058insa or IVS1+38insa (5a/6a)] and a single nucleotide a to g substitution (dbSNP:1805129 RefSeq NM007914 a1013g), which is a silent alteration in codon 84 of the coding sequence.

Materials and Methods

Patients and Controls. Cases and controls used in this study have been previously described in detail (7, 8). In brief, cases were drawn from the Anglian Breast Cancer Study, an ongoing population based study, with cases ascertained through the East Anglian Cancer Registry. All patients diagnosed below age 55 years since 1991 and still alive in 1996 (prevalent cases, median age 48 years), together with all those <65 years diagnosed between 1996 and present (incident cases, median age 52 years), were eligible to take part. All study participants completed an epidemiological questionnaire and provided a blood sample for DNA analysis. Female controls were randomly selected from EPIC-Norfolk, a component of EPIC. EPIC is a prospective study of diet and cancer being carried out in nine European countries. The EPIC-Norfolk cohort comprises

1 To whom requests for reprints should be addressed, at Cancer Research UK Human Cancer Genetics Research Group, Department of Oncology, Strange-ways Research Laboratories, Words Causeway, Cambridge CB1 8RN, United Kingdom. Phone: 38-01223-740166; Fax: 38-01223-411609; E-mail: pharoah@rl.cam.ac.uk. Received 11/25/02; revised 5/3/27/03; accepted 5/8/03.

2 The abbreviations used are: CHEK2, checkpoint kinase 2; EPIC, European Prospective Investigation of Cancer; CI, confidence interval; OR, odds ratio.
25,000 individuals resident in Norfolk (East Anglia), ages 45–74 years. The ethnic background of both cases and controls is similar, with >95% being white.

**Genotyping.** The genomic sequence of chromosome 22 was used for primer generation (RefSeq NT011520). We genotyped all patient and control samples for the CHEK2 polymorphisms using the ABI Prism 7700 sequence detection system (TaqMan; Applied Biosystems). TaqMan primers (IVS1 forward, 5'-CTCCCTGGGCTCGATTAT-3'; and reverse, 5'-TATCACGTTTTGTTCAG-3') and probes (IVS1 + 38insVic, 5'-TGTCCTAACCTAATCATC-3') and reverse (IVS1 + 38ins) were designed using Primer Express Oligo Design Software v1.0 (Applied Biosystems). Fifteen-μl assays were carried out on 20 ng of genomic DNA according to manufacturer’s instructions. Primers were used at 900 nm concentrations, IVS1 probes at 100 nm, and for a1013g, probes at 200 nm. Amplifications were carried out on MJ Tetrad thermal cyclers (NJ Research) at annealing temperatures of 64°C for IVS1 and 62°C for a1013g. Plates were read on the ABI Prism 7700 Sequence detector in end point mode using the Allelic Discrimination Sequence Detection software (Applied Biosystems). For the software to recognize the genotype, we included nontemplate controls and positive controls for each polymorphism.

**Statistical Methods.** For each polymorphism, deviation of the genotype frequencies from those expected under Hardy-Weinberg equilibrium was assessed in the controls by χ² tests. Genotype frequencies in cases and controls were compared by χ² tests or Fisher’s exact tests where expected frequencies were small. The genotypic specific risks were estimated as ORs with associated 95% confidence limits by unconditional logistic regression. Floated CIs were also calculated by treating ORs as floating absolute risks (9). This method does not alter the relative risk estimates but reduces the variances attributed to the ORs that are not defined as 1.0 and reduces unwanted covariance between them. This allows a valid comparison between the nonbaseline groups. Haplotype frequencies were estimated with the Estimating Haplotype program (10).

**Results**

Neither genotype distribution in the controls differed significantly from that expected under Hardy-Weinberg equilibrium. The genotypic frequencies were similar in the prevalent (ABC R) and incident (ABC P) cases for both polymorphisms (Table 1). There was no significant different in genotype frequency between cases and controls for either polymorphism (Table 1). All but one genotypic specific OR was close to 1.0, with upper 95% confidence limits of ≤1.9. The only exception was the rare homozygote risk for a1013g, but this did not differ significantly from unity (OR = 3.1, 95% CI 0.3–30).

The two biallelic polymorphisms generate four possible haplotypes (IVS1 + 38 5a/H11001a, IVS1 + 38 5a/H11013a, IVS1 + 38 5a/H11349a, and IVS1 + 38 6a/H11013a). The most common and likely ancestral haplotype is IVS1 + 38 5a/H11013a. However, there is strong linkage disequilibrium between the two loci (D' = −0.94) with one haplotype being extremely rare (IVS1 + 38 6a/H11013a). There was no significant difference in the haplotype frequency distribution between cases and controls (Table 2, χ² = 3.92, degrees of freedom = 3, P = 0.27). Haplotype-specific risks are also given in Table 2.

**Discussion**

Polymorphic variants in CHEK2 are good candidates for breast cancer susceptibility because of its role in the double-strand break DNA repair pathways. Support for this comes from the observation that a rare variant in the gene, 1100delc, is associated with a ~2-fold increased breast cancer risk. We have estimated the breast cancer risks associated with two polymorphisms in CHEK2 using a case-control study design and found no statistically significant differences in frequency between
cases and controls. If one of the assayed polymorphism had been a true functional variant, our study had >90% power at a significance of $10^{-4}$ to detect an allele with a frequency of 15%, similar to the IVS1+38insa polymorphism, that confers a dominant relative risk of 1.5 or a recessive relative risk of 2.8. For a rarer allele such as a1013g with frequency 0.018, the study would have the same power to detect a dominant allele with a relative risk of 2.3.

It is also possible to detect an association using neutral markers (alleles or haplotypes) that are in linkage disequilibrium with a true functional variant. In particular, the known, rare breast cancer susceptibility allele 1100delc occurs only on the IVS1+38 6a/1013a haplotype, consistent with the hypothesis that 1100delc arose on this haplotype with no subsequent recombination. Thus, given a large enough study, it would have been theoretically possible to identify the effect of the 1100delc through LD with either IVS1+38insa or a1013g. However, the power to detect an association by LD is weakened by the fact that 1100delc represents only a small fraction of the observed haplotype. The power is dependent on $\Delta$ (a measure of linkage disequilibrium that takes into account both recombination and relative allele frequencies). If the sample size needed to detect a susceptibility locus when the locus itself is assayed is N, it can be shown that the sample size needed to detect that locus by assaying a nearby marker in LD is approximated by $N/\Delta^2$. $\Delta$ between 1100delc and the haplotypes defined by IVS1+38insa and a1013g is $\sim 0.045$. Thus, the sample size to detect the effect of the rare variant 1100delc by assaying IVS1+38insa would need to be increased >400-fold compared with that needed if 1100delc were assayed directly. The known association of breast cancer with 1100delc is also consistent with the limits shown in Table 3 because it has a frequency of <1% in the United Kingdom population.

There are no known nonsynonymous SNPs in the coding sequence with a frequency of ≥5%. However, it is possible that variants in the regulatory sequences may affect breast cancer risk by altering gene expression. No such regulatory polymorphisms have been identified, but based on our results, it is possible to quantify the range of effects that might be associated with any true susceptibility allele in LD with the markers typed. Two of the four possible haplotypes occurred with a frequency of ≥5%. IVS1+38 5a/1013a (82%) and IVS1+38 6a/1013a (16%). If we assume that a true dominant susceptibility allele is carried as a subgroup of either of these haplotypes, the risk detectable will depend on the true susceptibility allele frequency and the extent of linkage disequilibrium between it and its haplotype marker (see “Appendix”). The upper 95% confidence limit for the relative risk associated with IVS1+38 5a/1013a haplotype, assuming it acts dominantly was 1.12. This imposes boundaries on the relative risk associated with any functional locus in complete LD with this haplotype (i.e., $D’ = 1$ or $-1$). On this basis, we can exclude the presence of a true susceptibility allele with frequency of 0.05 that confers a relative risk of ≥3.0 (Table 3). Similarly, we can exclude a true susceptibility allele with a smaller risk if it is more common, e.g., an allele with frequency 0.5 that confers a risk of ≥1.2. Although the markers we tested showed little evidence of recombination ($D’ \sim 1$), it is possible that LD across the whole gene is weaker. The range of risks excluded for different degrees of LD are given in Table 3. If LD is complete, a true susceptibility allele of frequency 0.05 and relative risk 1.7 can be excluded from being carried on the IVS1+38 6a/1013a haplotype (frequency 16%).

In interpreting these results, some limitations of the study need to be considered. Our controls were not individually matched for ethnic group, but given the very high proportion of both cases and controls that were white European, it is highly unlikely that population stratification would have any effect on the results (11). A second potential source of bias to be considered is survival bias, which may affect genotype frequencies in the prevalent cases if genotype were associated with different survival. However, we found no evidence for a difference in genotype frequency between the incident and prevalent case series and have previously found no association for either polymorphism with survival (12).

In conclusion, we have found no evidence that the CHEK2 polymorphisms IVS +1a and a1013g confer an increased risk of breast cancer. It is also unlikely that other, as yet unidentified, common polymorphisms that affect risk are present in the gene in the British population.

### Appendix

Suppose the true dominant susceptibility allele has a frequency $p$ and confers a relative risk $r$ (unknown) relative to noncarriers. Suppose this allele is associated with a marker haplotype that has frequency $p’$ and for which the estimated relative risk of disease is $r’$. $D’$ is the linkage disequilibrium coefficient between marker and true susceptibility allele. Then

$$r’ = \frac{(1 - p)(rs + p’ - s)}{(p - s)r + (1 - p - p’ + s)p’}$$

where $s = p(D’ - pD’ + p’).$

Thus, for given values of $p$, $p’$, $D’$, and $r’$, the corresponding value of $r$ can be determined.

### References


---

Cancer Epidemiology, Biomarkers & Prevention 811

Table 3 Relative risks of susceptibility alleles excluded by marker haplotypes

<table>
<thead>
<tr>
<th>True allele frequency</th>
<th>LD coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>Haplotype IVS1+38 5a/1013a (frequency 82%)</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>3.0</td>
</tr>
<tr>
<td>0.50</td>
<td>1.2</td>
</tr>
<tr>
<td>Haplotype IVS1+38 6a/1013a (frequency 16%)</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>1.7</td>
</tr>
<tr>
<td>0.10</td>
<td>1.3</td>
</tr>
</tbody>
</table>


Common Polymorphisms in Checkpoint Kinase 2 Are not Associated with Breast Cancer Risk


Updated version  Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/12/8/809

Cited articles  This article cites 9 articles, 2 of which you can access for free at:
http://cebp.aacrjournals.org/content/12/8/809.full#ref-list-1

Citing articles  This article has been cited by 2 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/12/8/809.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.