Short Communication

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

B. Kuschel, A. Auranen, C. S. Gregory, N. E. Day, D. F. Easton, B. A. J. Ponder, A. M. Dunning, and Paul D. P. Pharoah[1]

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. P.], 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 control for either the IVS1 + 38ins*a* or the *a*1013*g* 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 + 1*a* and *a*1013*g* 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[2] is a kinase in which the yeast counterpart regulates a cell cycle checkpoint and causes cells to arrest proliferation after DNA damage to allow repair to the injury before proceeding with mitosis (5). CHEK2 is 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, 1100del*c*, 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 8058ins*a* or IVS1+38ins*a* (5*a*/6*a*)] and a single nucleotide *a* to *g* substitution (dbSNP:1805129 RefSeq NM007914 *a*1013*g*), 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, Strangeways Research Laboratories, Worts Causeway, Cambridge CB1 8RN, United Kingdom. Phone: 38-01223-740166; Fax: 38-01223-411609; E-mail: paul.pharoah@srl.cam.ac.uk.

Received 11/25/02; revised 53/27/03; accepted 5/8/03.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[2] The abbreviations used are: CHEK2, checkpoint kinase 2; EPIC, European Prospective Investigation of Cancer; CI, confidence interval; OR, odds ratio.

Table 1 *CHEK2* polymorphism genotype frequencies and genotypic specific risks

Polymorphism	Series	Rare allele frequency	Common homozygote (*n*)	Heterozygote (*n*)	Rare homozygote (*n*)	χ^2 (*P*) Incident *versus* prevalent cases	χ^2 (*P*) cases *versus* controls
IVS1 + 38ins*a*	Incident cases	0.15	713	282	15		
(5*a*/6*a*)	Prevalent cases	0.16	431	161	15	2.3 (0.3)	2.6 (0.3)
	Controls	0.17	1066	431	41		
	OR		1.0	0.96	0.68		
	95% CI			0.82–1.1	0.42–1.1		
	95% FCI		0.92–1.1	0.84–1.1	0.43–1.1		
*a*1013*g*	Incident cases	0.024	1001	48	1		
	Prevalent cases	0.026	698	34	2	0.8 (0.7)	3.9 (0.2)[a]
	Controls	0.018	1763	64	1		
	OR		1.0	1.3[b]	3.1		
	95% CI			0.96–1.9	0.28–30		
	95% FCI[c]		0.93–1.1	0.96–1.9	0.28–30		

[a] Fisher's exact test, $P = 0.17$.
[b] $P = 0.09$.
[a] FCI = floated confidence interval (see text for explanation).

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+38ins*a* forward, 5′-CCCCCTGGGCTCGATTAT-3′ and reverse 5′-AATCAGAACCTTCCACCTGGTAATA-3′; and *a*1013*g* forward, 5′-CCTGAAGGGCCCATAATCG-3′ and reverse, 5′-GTGTCCACTCAGGAACTCTATTCTAATC-3′) and probes (IVS1 + 38ins*a* Vic, 5′-TATCACGTTTTGTTCAG*AAAAA*CACTTACAGAAAGT-3′; Fam, 5′-TATCACGTTTTGTTCAG*AAAAAA*CACTTACAGAAAGTT-3′; *a*1013*g* Vic, 5′-TAGGCTCCTCAGG*T*TCTTGGTCCTC-3′; and Fam, 5′-TAGGCTCCTCAGG*C*TCTTGGTCCTC-3′) 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 *a*1013*g*, 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 *a*1013*g*. 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 allele of the SNP (eight of each) in a 96-well plate.

Statistical Methods. For each polymorphism, deviation of the genotype frequencies from those expected under Hardy-Weinberg equilibrium was assessed in the controls by χ^2 tests. Genotype frequencies in cases and controls were compared by χ^2 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 Haplotypes program (10).

Table 2 Estimated *CHEK2* haplotype frequencies and risks in cases and controls

Haplotype	Controls	Cases	OR[a] (95% CI)
IVS1+38 5*a*/1013*a*	0.818	0.820	0.99 (0.87–1.1)
IVS1+38 5*a*/1013*g*	0.017	0.024	0.70 (0.49–1.1)
IVS1+38 6*a*/1013*a*	0.164	0.155	1.1 (0.93–1.2)
IVS1+38 6*a*/1013*g*	0.001	0.001	1.0 (0.21–4.8)

[a] OR for each haplotype compared with the other haplotypes combined.

Results

Neither genotype distribution in the controls differed significantly from that expected under Hardy-Weinberg equilibrium. The genotype 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 biallelelic polymorphisms generate four possible haplotypes (IVS1+38 5*a*/1013*a*, IVS1+38 5*a*/1013*g*, IVS1+38 6*a*/1013*a*, and IVS1+38 6*a*/1013*g*). The most common and likely ancestral haplotype is IVS1+38 5*a*/1013a. However, there is strong linkage disequilibrium between the two loci ($D' = -0.94$) with one haplotype being extremely rare (IVS1+38 6*a*/1013*g*). There was no significant difference in the haplotype frequency distribution between cases and controls (Table 2, $\chi^2 = 3.92$, 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, 1100del*c*, 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

Table 3 Relative risks of susceptibility alleles excluded by marker haplotypes

True allele frequency	LD coefficient					
	1.0	0.9	0.8	0.7	0.6	0.5
Haplotype IVS1+38 5*a*/1013*a* (frequency 82%)						
0.05	3.0	3.2	3.6	4.0	4.6	5.4
0.50	1.2	1.2	1.3	1.3	1.4	1.4
Haplotype IVS1+38 6*a*/1013*a* (frequency 16%)						
0.05	1.7	1.8	1.9	2.0	2.2	2.4
0.10	1.3	1.4	1.4	1.5	1.6	1.7

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+38ins*a* polymorphism, that confers a dominant relative risk of 1.5 or a recessive relative risk of 2.8. For a rarer allele such as *a*1013*g* 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 1100del*c* occurs only on the IVS1+38 6*a*/1013*a* 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 1100del*c* through LD with either IVS1+38ins*a* or *a*1013*g*. However, the power to detect an association by LD is weakened by the fact that 1100del*c* represents only a small fraction of the observed haplotype. The power is dependent on Δ (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/Δ^2. Δ between 1100del*c* and the haplotypes defined by IVS1+38ins*a* and *a*1013*g* is −0.045. Thus, the sample size to detect the effect of the rare variant 1100del*c* by assaying IVS1+38ins*a* would need to be increased >400-fold compared with that needed if 1100del*c* were assayed directly. The known association of breast cancer with 1100del*c* 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 risks 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 5*a*/1013*a* (82%) and IVS1+38 6*a*/1013*a* (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 5*a*/1013*a* 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 6*a*/1013*a* 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 +1*a* and *a*1013*g* 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

$$s = p(D' - p'D' + p')$$

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

References

1. Anglian Breast Cancer Study Group. Prevalence and penetrance of *BRCA1* and *BRCA2* in a population based series of breast cancer cases. Br. J. Cancer, *83:* 1301–1308, 2000.

2. Peto, J., Collins, N., Barfoot, R., Seal, S., Warren, W., Rahman, N., Easton, D. F., Evans, C., Deacon, J., and Stratton, M. R. Prevalence of *BRCA1* and *BRCA2* gene mutations in patients with early-onset breast cancer. J. Natl. Cancer Inst. (Bethesda), *91:* 943–949, 1999.

3. Antoniou, A. C., Pharoah, P. D. P., McMullen, G., Day, N. E., Stratton, M. R., Peto, J., Ponder, B. A. J., and Easton, D. F. A comprehensive model for familial breast cancer incorporating *BRCA1*, *BRCA2* and other genes. Br. J. Cancer, *86:* 76–83, 2002.

4. Scully, R. Association of BRCA1 with RAD51 in mitotic and meiotic cells. Cell, *8:* 265–275, 1997.

5. Matsuoka, S., Huang, M., and Elledge, S. J. Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. Science (Wash. DC), *282:* 1893–1897, 1998.

6. Meijers-Heijboer, H., van den Ouweland, A., Klijn, J., Wasielewski, M., de Snoo, A., Oldenburg, R., Hollestelle, A., Houben, M., Crepin, E., van Veghel-Plandsoen, M., Elstrodt, F., van Duijn, C., Bartels, C., Meijers, C., Schutte, M., McGuffog, L., Thompson, D., Easton, D., Sodha, N., Seal, S., Barfoot, R.,

Mangion, J., Chang-Claude, J., Eccles, D., Eeles, R., Evans, D. G., Houlston, R., Murday, V., Narod, S., Peretz, T., Peto, J., Phelan, C., Zhang, H. X., Szabo, C., Devilee, P., Goldgar, D., Futreal, P. A., Nathanson, K. L., Weber, B., Rahman, N., and Stratton, M. R. Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat. Genet., *31:* 55–59, 2002.

7. Pharoah, P. D. P., Lipscombe, J. M., Redman, K., Easton, D. F., Ponder, B. A. J., and Day, N. E. Familial predisposition to breast cancer in a British population: implications for prevention. Eur. J. Cancer, *36:* 773–779, 2000.

8. Day, N., Oakes, S., Luben, R., Khaw, K. T., Bingham, S., Welch, A., and Wareham, N. EPIC-Norfolk: study design and characteristics of the cohort. European Prospective Investigation of Cancer. Br. J. Cancer, *80* (Suppl. 1)*:* 95–103, 1999.

9. Easton, D. F., Peto, J., and Babiker, A. G. A. G. Floating absolute risk: an alternative to relative risk in survival and case-control analysis avoiding an arbitrary reference group. Stat. Med., *10:* 1025–1035, 1991.

10. Xie, X., and Ott, J. Testing linkage disequilibrium between a disease gene and marker loci. Am. J. Hum. Genet., *53* (Suppl)*:* 1107, 1993.

11. Goode, E. L., Dunning, A. M., Healey, C. S., Kuschel, B., Durocher, F., Teare, M. D., Pharoah, P. D. P., Ponder, B. A. J., Day, N. E., and Easton, D. F. Assessment of population stratification in a large population-based cohort. 10th Annual Meeting of the International Genetic Epidemiology Society. Garmisch Patenkirchen, Germany, 2001.

12. Goode, E. L., Dunning, A. M., Kuschel, B., Healey, C. S., Day, N. E., Ponder, B. A. J., Easton, D. F., and Pharoah, P. D. P. Effect of germline genetic variation on breast cancer survival in a population-based study. Cancer Res., *62:* 3052–3057, 2002.

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

B. Kuschel, A. Auranen, C. S. Gregory, et al.

Cancer Epidemiol Biomarkers Prev 2003;12:809-812.