A potential role for the XRCC2 R188H polymorphic site in DNA-damage repair and breast cancer

Saeed Rafii1, Paul O'Regan3, George Xinarianos1, Iman Azmy2, Tim Stephenson4, Malcolm Reed2, Mark Meuth1, John Thacker3 and Angela Cox1,*

1Institute for Cancer Studies, Division of Genomic Medicine and 2Section of Surgical and Anaesthetic Sciences, Division of Clinical Sciences, University of Sheffield Medical School, Sheffield, S10 2RX, UK, 3Medical Research Council, Radiation and Genome Stability Unit, Harwell, OX11 0RD, UK, and 4Department of Histopathology, Royal Hallamshire Hospital, Sheffield, S10 2JF, UK

Received February 6, 2002; Revised and Accepted March 27, 2002

An acquired genetic instability, resulting from the loss of some types of DNA repair, is an early event in the development of a subset of human cancers. The involvement of BRCA1 and BRCA2 in the homologous recombination repair (HRR) of double-strand breaks in DNA implicates this pathway in the suppression of breast cancer. A family of proteins related to human RAD51, including XRCC2, are essential components of this repair pathway. Using site-directed mutagenesis of XRCC2, we show that non-conservative substitution or deletion of amino acid 188 of XRCC2 can significantly affect cellular sensitivity to DNA damage, and that a polymorphic variant at this site (R188H), present on 6% of chromosomes in the population, has a weak effect on damage sensitivity. We tested the hypothesis that the R188H polymorphism could be a low-penetrance susceptibility factor for breast cancer, by genotyping 521 women with breast cancer and a total of 895 control women. Carriage of the rare allele of XRCC2 R188H was associated with breast cancer overall [odds ratio 1.3; 95% confidence interval (CI) = (1.0, 1.8)] and when younger-onset cases with a positive family history were compared with older controls with no family history [odds ratio 1.9; 95% CI = (1.0, 3.8)]. These results support the hypothesis that subtle variation in DNA repair capacity may influence cancer susceptibility in the population.

INTRODUCTION

There is a large body of evidence supporting the premise that an acquired genetic instability is fundamental to tumour development. Defects in DNA-repair mechanisms are known to be responsible for inherited cancer-prone syndromes such as xeroderma pigmentosum and hereditary non-polyposis colon cancer (HNPPC) (1,2). Since the positional cloning of the familial breast cancer genes BRCA1 and BRCA2, it has become clear that the proteins encoded by these genes have a role to play in the homologous recombination DNA-repair pathway, which repairs double-strand breaks in DNA (3). Murine Brca1 or Brca2 null cells show defects in homologous recombination repair and radiation hypersensitivity (4,5). In addition, ionizing radiation induces an S-phase co-localization to nuclear foci of BRCA1 and BRCA2 with RAD51, the human homologue of the Escherichia coli RecA protein, which is considered a central protein in homologous recombination (6,7). Further evidence for the role of defects in DNA double-strand break repair in breast cancer is also provided by the observation that lymphocytes and fibroblasts from breast cancer patients and their relatives show increased sensitivity to ionizing radiation compared with controls (8,9).

Studies of the RecA and Rad51 proteins of E. coli and Saccharomyces cerevisiae respectively have yielded much of what is currently known about the mechanism of homologous recombination repair. In yeast, the Rad55 and Rad57 proteins (which have sequence similarity with Rad51) form a heterodimer that stimulates the efficiency of Rad51-catalysed DNA strand exchange (10). In humans, there are five genes identified so far with some sequence similarity to RAD51; these are RAD51L1 (RAD51B), RAD51L2 (RAD51C), RAD51L3 (RAD51D), XRCC2 and XRCC3 (11). The XRCC2 and XRCC3 genes were cloned by complementation of mutant hamster cell lines hypersensitive to a number of DNA-damaging agents, including ionizing radiation (12,13). Subsequently, both XRCC2 and XRCC3 proteins have been shown to be necessary for homologous recombination repair (14-16) and are required for RAD51 focus formation (17,18).

Whilst severe homozygous mutations of these repair genes give embryonic lethality (19), mutations or polymorphisms having more subtle effects on DNA repair capacity may be...
low-penetration risk factors for breast cancer, as has been illustrated by the association of a common variant of BRCA2 with sporadic breast cancer (20). We tested this hypothesis in relation to a recently identified coding variant, XRCC2 R188H, by examining (i) the effect of altering or deleting this amino acid on cell survival after DNA damage, and (ii) the association of the polymorphism with breast cancer in the population.

RESULTS

The effect of alterations at the XRCC2 R188H polymorphic site on cell survival after mitomycin C-induced DNA damage

Sequencing of the XRCC2 gene in DNA from a number of individuals identified a G-to-A transition resulting in the alteration of amino acid arginine 188 to histidine. We carried out cellular complementation tests to investigate the functional effects of substitution at this amino acid site in the XRCC2-deficient cell line, irs1. Using site-directed mutagenesis, we created different alterations at amino acid position 188 and cloned these into a bicistronic IRES (internal ribosome entry site) mammalian expression vector, to ensure expression of the cloned gene along with the selectable marker. Mutant XRCC2 genes, generating R188H, or R188A, or a deletion of this residue (ΔR188), as well as the wild-type gene, were then separately transfected into irs1 cells. Pools of selected transfectants were treated with mitomycin C, and the survival of cells transfected with the mutant genes was compared with that of cells transfected with the wild-type gene. The data in Figure 1 show that mutation of amino acid 188 to a non-conservative amino acid, alanine, or its deletion, has a substantial effect on cell survival (ANOVA P = 0.0058 and 0.0026 respectively when data were fitted to a linear-quadratic curve). Mutation to R188H shows a much smaller difference in survival from the wild type, even at the highest dose of mitomycin C (ANOVA P = 0.054; overall linear-quadratic fit P = 0.071), indicating that the arginine-to-histidine change has a much more subtle effect on the cells’ ability to repair mitomycin C-induced damage.

XRCC2 R188H genotypes in breast cancer cases and blood donor controls

The allele frequency of XRCC2 R188H in 398 anonymous healthy blood donors was 0.06 (95% confidence interval (CI) = (0.04, 0.08)). Genotype frequencies were very consistent with those expected under Hardy–Weinberg equilibrium ($\chi^2 = 0.182, P = 0.67$).

There was significant evidence of an association between carriage of one or two copies of the rare (A) allele of XRCC2 R188H and breast cancer when breast cancer cases were compared with blood donor controls [odds ratio 1.52; 95% CI = (1.04, 2.22); Table 1]. In addition, there was evidence of a dose effect for the number of copies of the A allele [odds ratios 1.48 and 2.44 for one and two copies respectively (Table 1), score test for trend $P = 0.02$]. Since the risk of breast cancer varies considerably with age and menopausal status, and the age range of recruitment of the blood donor controls was significantly lower than the age of diagnosis for the breast cancer cases (median 39 and 58 years respectively, $P < 10^{-12}$), we employed a second control group from women attending mammography screening clinics. This cohort provided a resource that could be used for age-matched comparisons.

Characterization of breast cancer cases and mammography screening controls

Table 2 shows the summary statistics for comparison of the breast cancer cases and mammography screening controls. There were no significant differences between the breast cancer cases and mammography screening controls for age at diagnosis/sampling ($P = 0.31$), body mass index ($P = 0.14$), or age at menarche ($P = 0.33$). However, the breast cancer cases had fewer children ($P = 0.0009$), a higher age at first pregnancy ($P = 0.0021$) and a higher age at menopause.

Table 1. XRCC2 genotype frequencies in breast cancer cases and blood donor controls

<table>
<thead>
<tr>
<th>Genotype frequencies (%)</th>
<th>Odd ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G/G vs A/G</td>
</tr>
<tr>
<td>Breast cancer cases</td>
<td>G/G</td>
</tr>
<tr>
<td></td>
<td>431 (83.0)</td>
</tr>
<tr>
<td>Blood donor controls</td>
<td>351 (88.2)</td>
</tr>
</tbody>
</table>
Table 2. Characteristics of breast cancer patients and mammography screening controls

<table>
<thead>
<tr>
<th>Continuous variables</th>
<th>Breast cancer cases</th>
<th>Mammography screening controls</th>
<th>p&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Median</td>
<td>Range</td>
</tr>
<tr>
<td>Age at diagnosis (yr)</td>
<td>521</td>
<td>58</td>
<td>28-89</td>
</tr>
<tr>
<td>BMI</td>
<td>498</td>
<td>25.9</td>
<td>9.2-44.2</td>
</tr>
<tr>
<td>Age at first pregnancy (yr)</td>
<td>419</td>
<td>24</td>
<td>16-41</td>
</tr>
<tr>
<td>Number of children</td>
<td>486</td>
<td>2</td>
<td>0-8</td>
</tr>
<tr>
<td>Age at menarche (yr)</td>
<td>508</td>
<td>13</td>
<td>9-18</td>
</tr>
<tr>
<td>Age at menopause (yr)</td>
<td>420</td>
<td>50</td>
<td>29-61</td>
</tr>
<tr>
<td>Binary variables</td>
<td>n</td>
<td>Frequency %</td>
<td>n</td>
</tr>
<tr>
<td>Children</td>
<td>486</td>
<td>89.3</td>
<td>460</td>
</tr>
<tr>
<td>HRT</td>
<td>521</td>
<td>19.8</td>
<td>496</td>
</tr>
<tr>
<td>OCP</td>
<td>521</td>
<td>40.1</td>
<td>496</td>
</tr>
<tr>
<td>BRCAFH</td>
<td>521</td>
<td>32.3</td>
<td>496</td>
</tr>
</tbody>
</table>

<sup>a</sup>Age at diagnosis indicates age at diagnosis for cases and age at blood sampling for controls; BMI, body mass index; HRT, hormone replacement therapy; OCP, oral contraceptive pill; BRCAFH, breast cancer family history (at least one first- or second-degree relative with breast cancer).

<sup>b</sup>n = number of observations available for that variable.

<sup>c</sup>Wilcoxon rank-sum test.

(\(P = 0.0069\)), and a higher proportion of them had no children (\(P = 3.9 \times 10^{-6}\)). In addition, a higher proportion of breast cancer cases had at least one first- or second-degree relative with breast cancer (\(P = 0.00074\)). Although there was no significant difference in median age between the breast cancer cases and mammography screening controls, the age range of the controls was narrower (45–77 years for controls versus 28–89 years for cases; Table 2), since the majority of the controls were within the age range of women routinely invited for screening (see Materials and Methods). The difference in age ranges probably accounts for the much higher proportion of controls who had received hormone-replacement therapy and oral contraception (\(P < 10^{-12}\) and \(P = 2.3 \times 10^{-11}\) respectively). For example, amongst those individuals born after 1940, for whom oral contraception would have been available during their reproductive years, there was no significant difference in oral contraception use between cases and controls (69.7% versus 71.9%; \(P = 0.58\)).

**XRCC2 R188H genotypes in breast cancer cases and mammography screening controls**

To control as much as possible for age and birth cohort effects, we compared XRCC2 R188H genotype frequencies in cases and mammography screening controls diagnosed/recruited between the ages of 50 and 65 years, the age range for which women routinely attend mammography screening. The numbers of A homozygotes in this age stratum were small, thus yielding very wide confidence intervals, but Table 3 shows that the trend in odds ratios for carriage of one or two copies of the A allele (1.24 and 1.78 respectively) was consistent with what was observed for the blood donor controls, although the results did not reach statistical significance (score test for trend \(P = 0.28\)). Comparable results were obtained when all 519 cases and 493 mammography screening controls were analysed, or when the analysis was restricted to only postmenopausal women (data not shown). We carried out one subgroup analysis using an extreme sampling approach, in which we compared cases of younger onset and positive family history with older controls with no family history. The odds ratio for carriage of the A allele was 1.93 [95% CI = (0.99, 3.76)] for this comparison. Combining all available data, regardless of age, the odds ratios obtained by comparing all breast cancer cases with both entire blood donor and mammography screening control groups combined were 1.26 [95% CI = (0.93, 1.72)] and 2.14 [95% CI = (0.65, 7.07)], for one and two copies of the A allele respectively (score test for trend \(P = 0.06\)), with an odds ratio of 1.30 [95% CI = (0.96, 1.75)] for carriage of at least one copy of the A allele (Table 3).

Table 3. XRCC2 genotype frequencies in breast cancer cases and mammography screening controls

<table>
<thead>
<tr>
<th>Genotype frequencies (%)</th>
<th>Odds ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G/G vs G/G</td>
</tr>
<tr>
<td>BCC 50-65 years&lt;sup&gt;a&lt;/sup&gt;</td>
<td>207 (82.1)</td>
</tr>
<tr>
<td>MSC 50-65 years&lt;sup&gt;a&lt;/sup&gt;</td>
<td>368 (85.2)</td>
</tr>
<tr>
<td>All data: BCC</td>
<td>431 (83.0)</td>
</tr>
<tr>
<td>BCC + MSC combined</td>
<td>770 (86.4)</td>
</tr>
</tbody>
</table>

<sup>a</sup>Age at recruitment for controls, age at diagnosis for cases.
DISCUSSION
Defects in DNA repair, resulting in genetic instability, have been shown to be important in the development of both inherited and sporadic tumours (2,21). We have tested the hypothesis that population variation in genes involved in DNA-repair pathways could have subtle effects on DNA-repair capacity and thereby affect an individual's susceptibility to cancer. Our results show that non-conservative (R188A) mutation or deletion of amino acid position 188 of XRCC2 has a significant effect on cell survival following DNA damage. Consistent with our evidence of the importance of this amino acid site, a naturally occurring polymorphism at this position, R188H, also has a weak effect on XRCC2 function. We present data suggesting that the same polymorphism is associated with breast cancer.

The role of position 188 in the functioning of the XRCC2 protein is unknown. Indeed, very few sites of functional significance within the protein have been identified. We have shown, for example, that the integrity of a well-characterized domain normally associated with ATP binding (22) does not seem to be required for XRCC2 function (18). We have recently identified regions of XRCC2 that are involved in the interaction with another RAD51-like protein, RAD51L3, but these do not overlap with position 188 (K. Spink and J. Thacker, unpublished data). Further, it has been shown that XRCC2 is part of a heterotetrameric complex with RAD51L1, RAD51L2 and RAD51L3, which binds to single-stranded and nicked DNA, perhaps facilitating the activity of RAD51 (23). Therefore other functional sites remain to be identified, and it is noteworthy that comparisons of human, mouse and hamster XRCC2 proteins show that the arginine at position 188 is invariant (12, and J. Thacker, unpublished data), suggesting a conserved function.

In our case-control study, the overall results were of borderline statistical significance (Table 3), and thus require confirmation in an independently ascertained dataset. The odds ratios were broadly less striking for the age-matched analysis using the mammography screening controls than for the preliminary analysis with the blood donor controls (Table 3; cf. Table 1). These differences, whilst within the limits of sampling variation, could also be due to confounding by age in the blood donor controls, or bias in one or both groups. The former is unlikely, since the use of extreme sampling (i.e. younger cases and older mammography controls (24)) provided support for an association [odds ratio 1.9, 95% CI = (1.0, 3.8)]. We cannot rule out the latter, but the mammography screening controls should provide reasonable representation of females in Sheffield aged 50-65 years, since the average attendance rate for screening in Sheffield is 81.2%. Population stratification is another well-documented potential confounder in case-control studies, but is unlikely to be a factor in the present study, since both control groups were carefully ethnically matched to cases. Our case series is based on hospital patients having surgical treatment for breast cancer. The recruitment rate and age distribution in our study were broadly consistent with population-based Trent Cancer Registry data for 1999 (data not shown), with the exception that cases diagnosed over 80 years old were less well represented in our cohort, since women over this age are less likely to have surgery.

Whilst the possible effects of XRCC2 R188H in terms of both in vitro function and genetic association are small, they are consistent with what might be expected for polymorphisms in genes involved in the homologous recombination repair pathway. For example, Healey et al. (20) identified an odds ratio of 1.3 overall for the N372H BRCA2 polymorphism in breast cancer, using five independent datasets, and this effect was not statistically significant in all datasets analysed. The present study had approximately 80% power (α = 0.05) to detect an odds ratio of 1.6 for carriage of the rare allele (A) of XRCC2, but only 33% power to detect an odds ratio of 1.3. Thus, if the true odds ratio is close to 1.3, it is not surprising that our results are of only borderline statistical significance. Similarly, we are not able to distinguish whether the phenotype is dominant or recessive, since our estimates of the odds ratios for homozygosity lack precision due to the relatively low frequency of this allele. Further large datasets will be required to resolve this issue.

Recent studies have shown that a polymorphism of a related gene, XRCC3 T241M, may be associated with melanoma and bladder cancer, although both of these studies were carried out with relatively small sample sizes and require replication (25,26). The same polymorphism was not associated with lung cancer (27). As alluded to above, very large sample sizes, of the order of several thousand, will be required to quantify the role of this pathway in breast and other cancers. In particular it will be important to investigate the interaction between different components of the pathway at both the functional and genetic level, since it seems that particular combinations of variants in different genes of the pathway may affect overall risk.

It has been argued that low-penetration genes could play a role in genetic susceptibility in many forms of cancer. The results of this study support the hypothesis that naturally occurring polymorphisms in genes involved in DNA repair can subtly affect repair capacity, and have implications for cancer susceptibility in the population. It is not unlikely that these types of effects may extend to individual tumour response to various chemotherapeutic agents and radiotherapy.

MATERIALS AND METHODS
Site-directed mutagenesis and gene transfection
Human XRCC2 cDNA (12) was cloned into pBS as a template for site-directed mutagenesis. Mutagenic PCR was carried out using the Quikchange site-directed mutagenesis protocol with Pfu Turbo polymerase (Stratagene). Mutagenic primers were designed to create the following changes to position 188 of the amino acid sequence of XRCC2: arginine to histidine (R188H), arginine to alanine (R188A) and deletion of R188 (∆R188). Inserts carrying mutations were subcloned into the NheI and BamHI sites of the pIRESneo2 mammalian expression vector (Clontech) and sequenced to ensure that the required changes were present. Mutant and wild-type XRCC2 genes in pIRESneo2 were transfected into XRCC2-deficient irs1 cells, and cells were selected for 14 days in 500 μg/ml G418. The IRES (internal ribosome entry site) vector was used to select against integration events that disrupted the XRCC2 gene; in control experiments, 93% (28/30) of G418-resistant clones also expressed the XRCC2 construct (data not shown).
Clonogenic survival following mitomycin C treatment

Pools of approximately 100 G418-resistant clones carrying mutant or wild-type XRCC2 were grown in medium containing 100 μg/ml G418 and mitomycin C (0–100 nM) for 10–12 days. At this time, surviving colonies were stained with methylene blue and counted.

Patients and controls

521 histologically confirmed breast cancer patients were recruited from among those attending surgical outpatient clinics at the Royal Hallamshire Hospital, Sheffield and Rotherham District General Hospital between November 1998 and June 2000. Two groups of healthy controls were included in this study. The first group were 399 anonymous female blood donors recruited during 1996, and the second group were 496 women recruited from among those attending the Sheffield Breast Screening Service from October 2000 to February 2001, and women in this group were included only if their mammogram showed no evidence of a breast lesion. All women on the community health index between the ages of 50 and 65 years are invited for screening every 3 years, and the uptake rate in Sheffield is on average 81.2%. Women older than 65 and younger than 50 also attend on request, or if they are at high risk, respectively. Age and sex were recorded for the blood donors. Demographic, environmental risk factor and family history data were recorded for all breast cancer cases and mammography screening controls, and were obtained by a clinician (I.A.) or research nurse using a standard questionnaire. Histopathological data was reviewed by a histopathologist (T.S.), and clinical data were obtained by a clinician (I.A.) or research nurse from hospital records. All individuals included in the study were white Caucasians of North European descent; only 11 individuals who had been recruited were subsequently excluded from the analysis as being of other ethnicity. This study was approved by the South Sheffield Research Ethics Committee and informed written consent was obtained from all subjects.

DNA extraction and genotyping

Ten millilitres of blood were obtained from each individual by venepuncture into EDTA-vacutainers and stored at −20°C until use. DNA was extracted as described previously (28).

The polymorphism in exon 3 of the XRCC2 gene at nucleotide position 31479 (GenBank accession no. AC003109) is a previously (28). The PCR amplification cycles were 95°C for 10 minutes, followed by 40 cycles of 95°C for 15 seconds and 63°C for 1 minute. Levels of Fam and Tet fluorescence were determined and allelic discrimination was carried out using the ABI 7700 Sequence Detector.

In addition, 96 XRCC2 R188H genotypes were determined independently by SSCP, and in all cases, the results were 100% concordant. Five XRCC2 R188H genotypes were also confirmed by automated sequencing of gel-purified PCR products.

Statistical analysis

All data were entered into a Microsoft Access database and exported to Stata version 6.0 for the statistical analyses. Genotype data were available for 519 breast cancer cases, 493 mammography screening controls and 398 blood donor controls. Odds ratios and 95% confidence intervals were determined for the genotype comparisons. Confidence intervals for allele frequencies were based on the binomial distribution.

The variables age at diagnosis, body mass index, age at first pregnancy, number of children, and ages at menarche and menopause were compared between breast cancer cases and mammography screening controls using the two-sample Wilcoxon rank-sum (Mann-Whitney) test. The proportions of women with children, breast cancer family history (defined as presence of at least one first- or second-degree relative with breast cancer), and hormone replacement therapy or oral contraceptive use were compared between cases and controls by performing χ² tests on 2 × 2 contingency tables. All tests were two-sided.

Cases and mammography screening controls were matched for age by restricting the analysis to those cases and controls diagnosed/recruited between ages 50 and 65 years. This is the age range of women routinely invited for mammography screening. Only 61 out of 493 controls in this study were outside this age range, insufficient for analysis of other age strata. A second analysis was restricted to postmenopausal women only (418 cases and 413 mammography screening controls). In addition, we used extreme sampling, which has been shown to increase efficiency for genes involved in complex inheritance (24). The comparison groups for this analysis were defined as cases with age at diagnosis below the median of 57.5 years, with a positive family history (as defined above) (n = 88), versus controls aged over 57.5 years with no family history (n = 176).

ACKNOWLEDGEMENTS

We should like to thank Dan Connelly for data management, Helen Cramp and Jane McDaid for their invaluable assistance in recruitment and sample collection, Dawn Teare and Nicky Brown for helpful comments on the manuscript, and all the people who took part in this study. This work was supported by Yorkshire Cancer Research, the Breast Cancer Campaign, the Medical Research Council, the European Commission (co-
tract FIGH-CT 1999-00010), and a scholarship from the Iranian government (S.R.).

REFERENCES


