Estrogen Receptor Polymorphism at Codon 325 and Risk of Breast Cancer in Women Before Age Forty

Melissa C. Southey, Leigh E. Batten, Margaret R. E. McCredie, Graham G. Giles, Gillian Dite, John L. Hopper, Deon J. Venter*  

Background: The estrogen receptor (ER) protein is believed to play a central role in the development and progression of breast cancer. In a previously published U.S. clinic-based study, a polymorphism in the ER gene (codon 325, CCG → CCG) was found to be more common in 34 case subjects with a family history of breast cancer than in 154 case subjects without such a history (mean allele frequencies ± standard error = 0.28 ± 0.05 versus 0.11 ± 0.02; P < .001). To determine whether this polymorphism is a risk factor for early-onset breast cancer, we conducted a population-based, case–control-family study in Australia. Methods: Case subjects under the age of 40 years with a first primary breast cancer and control subjects, frequency-matched to the case subjects on the basis of age, and their relatives were interviewed to assess the family history of breast cancer. Polymorphism status of the ER gene was determined for 388 case subjects and 294 control subjects. All statistical tests were two-tailed. Results: There was no association between ER gene polymorphism status and breast cancer, before or after adjustment for risk factors. There was no difference in allele frequencies between case subjects and control subjects (0.232 ± 0.015 versus 0.209 ± 0.017; P = .4) or between women with and without a family history of breast cancer (P = .3), irrespective of case–control status. The findings were not altered when different definitions of family history of breast cancer were used and when allele frequencies were adjusted for residence and country of birth. Conclusion: We found no evidence that the ER codon 325 polymorphism is associated with breast cancer before the age of 40 years or with a family history of breast cancer, despite ample power to detect effects half the magnitude of those previously reported. [J Natl Cancer Inst 1998;90:532–6]

The estrogen receptor (ER) protein is believed to play a central role in the development and progression of most breast cancers (1). In addition, the presence of detectable ER protein in tumors is associated with responsiveness to adjuvant hormone treatment and a favorable prognosis. Thirty percent of primary breast cancers in postmenopausal women and 70% in premenopausal women are ER negative (1). Nearly all ER-negative tumors and approximately 40% of ER-positive tumors are resistant to hormone therapy, a factor that confers an adverse prognosis (2). The molecular mechanisms determining ER negativity or the absence of hormone responsiveness in ER-positive tumors are still poorly understood (3–6). In a small percentage of cases, however, mutations in the coding region and splice variants have been described. The pivotal activity of the ER protein makes the ER gene a candidate locus for studies of breast cancer genetic susceptibility.

In a U.S. clinic-based series of women with primary invasive breast cancer, Roodi et al. (3) claimed that a polymorphism in exon 4 (codon 325, CCG → CCG) of the ER gene showed “a strong association” with a family history of breast cancer. Their study was based on the analysis of DNA extracted from tumor material. Roodi et al. found that the allele frequency (± standard error [SE]) in 34 case subjects with a family history of breast cancer (defined as a report of at least one first- or second-degree relative with breast cancer) was 0.28 (±0.05), which was greater than the 0.11 (±0.02) observed in the remaining 154 case subjects without a family history (P < .001). In addition, they obtained evidence of deviation from Hardy–Weinberg equilibrium (P = .003); they observed no case subjects homozygous for the polymorphism (versus an expected 3.8 under Hardy–Weinberg equilibrium). This finding, as well as the possible linkage of the ER locus to breast cancer observed in a family with several cases of this disease (7), raised the possibility that the polymorphism is linked to a mutation involved in the development of breast cancer (3).

Our goal was to determine if the finding by Roodi et al. (3) could be confirmed in a population-based sample of incident case subjects diagnosed before the age of 40 years. We obtained information on family history of breast cancer from interviews of both case subjects and living relatives. In addition, we sought to compare the allele frequency in these case subjects with that in a population-based sample of control subjects, assessed by the same procedure, to determine if the polymorphism was a risk factor for early-onset breast cancer.

Subjects and Methods

Subjects

The protocol outlined by Hopper et al. (8) was followed to conduct a population-based case–control-family study of early-onset breast cancer in Melbourne and Sydney, Australia, during the period 1992 to 1995 (9). Women under the age of 40 years at diagnosis of a first primary breast cancer (case subjects) were identified through the Victorian and New South Wales cancer registries. Women without breast cancer (control subjects) were selected from the electoral roll (registration for voting is compulsory in Australia) by use of stratified random sampling and were frequency matched to the case subjects for age. Case subjects, control subjects, and relatives completed the same risk factor questionnaire.

For each case subject and control subject, a detailed family history of cancer as described by McCredie et al. (9) was systematically recorded for all first- and second-degree relatives and subsequently checked with their living relatives at the time of their

*Affiliations of authors: M. C. Southey, L. E. Batten, Department of Pathology and Research, Peter MacCallum Cancer Institute, Melbourne, Australia; M. R. E. McCredie, Cancer and Epidemiology Research Unit, NSW Cancer Council, Kings Cross, Australia, and Department of Preventative and Social Medicine, University of Otago, New Zealand; G. G. Giles, Cancer Epidemiology Centre, Anti-Cancer Council of Victoria, Australia; G. Dite, J. L. Hopper, Departments of Public Health and Community Medicine, The University of Melbourne, Parkville, Australia; D. J. Venter, Department of Pathology and Research, Peter MacCallum Cancer Institute, and Department of Pathology, The University of Melbourne.

Correspondence to: Melissa C. Southey, Ph.D., Grad. Dip.Law., Pathology, Peter MacCallum Cancer Institute, Locked Bag No. 1, A’Beckett St., Melbourne, 3000, Australia. E-mail: m.southey@pmci.unimelb.edu.au

See “Notes” following “References.”

© Oxford University Press
results, we analyzed several samples, including all those homozygous for the polymorphism and 100 randomly selected heterozygous and wild-type homozygous individuals, by cycle sequencing using an Amplicycle sequencing kit (The Perkin-Elmer Corp.).

Fig. 1 shows an example of allele-specific oligonucleotide hybridization used to determine the ER codon 325 polymorphism carrier status of individuals.

Statistical Methods

Under Hardy–Weinberg equilibrium, the maximum likelihood estimator of the allele frequency (i.e., the proportion of ‘‘1’’ alleles) is \( f = (2n_{11} + n_{10})/2n \), where \( n = n_{11} + n_{10} + n_{00} \), \( n_{ij} \) is the observed number of subjects with the ‘‘ij’’ genotype (\( i_j = 0.1 \), respectively) and has asymptotic SE \( \sqrt{(f(1−f)/2n)^{1/2}} \). The Hardy–Weinberg equilibrium assumption was assessed by a comparison of the observed numbers of individuals with different genotypes with those expected for the estimated allele frequency and a comparison of the Pearson goodness-of-fit statistic with a chi-squared distribution with 1 degree of freedom.

Given no evidence of departure from Hardy–Weinberg equilibrium, the allele frequency was analyzed and modeled as a function of potential covariates by use of linear logistic regression, assuming that the number of ‘‘1’’ alleles was a binomial variable with \( n = 2 \).

The influence of polymorphism status on risk of breast cancer was assessed by standard case–control analyses with the use of unconditional multiple linear logistic regression, with and without adjustment for the risk factors identified in this study (9). Polymorphism status was modeled three ways: 1) by genotype (two parameters), 2) by a linear effect per ‘‘1’’ allele (one parameter), and 3) by the presence or absence of the ‘‘1’’ allele (one parameter).

Logistic regression analyses were performed by maximum likelihood methods with the use of the statistical package GLIM (11). Following conventional, all statistical tests and \( P \) values were two-tailed, and statistical significance was taken as a nominal \( P \) value of less than .05.

Results

Table 1 shows that there was no difference in allele frequency between case subjects and control subjects overall (\( P = .4 \)), in those with a reported family history of breast cancer (\( P = .6 \)), or in those without such a history (\( P = .1 \)). In a comparison of women with and without a reported family history of breast cancer, there was no difference overall (\( P = .3 \)), in case subjects (\( P = .09 \)), or in control subjects (\( P = .8 \)).

There was no evidence for a deviation from Hardy–Weinberg equilibrium in case subjects (\( \chi^2 = 0.7; P > .4 \)), in control subjects (\( \chi^2 = 1.9; P = .16 \)), and overall (\( \chi^2 = 2.5; P = .12 \)) or in any group or subgroup defined by case–control status and family history of breast cancer.

Table 2 shows that the allele frequency did not differ by family history status overall or in control subjects, although in case subjects there was a marginal indi-
Table 1. Distribution of estrogen receptor codon 325 alleles and estimated allele frequency in breast cancer case subjects and control subjects, by family history of breast cancer

<table>
<thead>
<tr>
<th>No. of alleles</th>
<th>Case subjects</th>
<th>Control subjects</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Family history</td>
<td>Family history</td>
<td></td>
</tr>
<tr>
<td></td>
<td>No (%)</td>
<td>Yes (%)</td>
<td>Total (%)</td>
</tr>
<tr>
<td>0</td>
<td>148 (55)</td>
<td>78 (64)</td>
<td>226 (58)</td>
</tr>
<tr>
<td>1</td>
<td>105 (39)</td>
<td>39 (32)</td>
<td>144 (37)</td>
</tr>
<tr>
<td>2</td>
<td>14 (5)</td>
<td>4 (3)</td>
<td>18 (5)</td>
</tr>
</tbody>
</table>

Allele frequency [standard error]

- Case subjects
  - Reported: 0.212 [0.040]
  - First-degree relatives: 0.235 [0.016]
  - First-degree relatives verified: 0.195 [0.044]
  - First- or second-degree relatives: 0.230 [0.016]
  - First- or second-degree relatives verified: 0.194 [0.027]

- Control subjects
  - Reported: 0.333 [0.079]
  - First-degree relatives: 0.201 [0.017]
  - First-degree relatives verified: 0.308 [0.091]
  - First- or second-degree relatives: 0.205 [0.017]
  - First- or second-degree relatives verified: 0.216 [0.032]

- Total
  - Reported: 0.243 [0.036]
  - First-degree relatives: 0.220 [0.012]
  - First-degree relatives verified: 0.222 [0.040]
  - First- or second-degree relatives: 0.214 [0.012]
  - First- or second-degree relatives verified: 0.203 [0.020]

*P values refer to comparison of subjects with a family history of breast cancer with those without.
†P values refer to comparisons between case subjects and control subjects with regard to absence of family history (P = .1) and presence of family history (P = .6) and all case subjects with or without a family history and all control subjects with or without a family history (P = .4).

Table 2. Estimated allele frequency in breast cancer case subjects and control subjects for different definitions and measures of family history of breast cancer

<table>
<thead>
<tr>
<th>Definition of family history of breast cancer</th>
<th>Allele frequency [standard error]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case subjects</td>
<td>Control subjects</td>
</tr>
<tr>
<td>Reported</td>
<td>0.212 [0.040]</td>
</tr>
<tr>
<td>First-degree relatives</td>
<td>0.235 [0.016]</td>
</tr>
<tr>
<td>Verified</td>
<td>0.195 [0.044]</td>
</tr>
<tr>
<td>First-degree relatives verified</td>
<td>0.230 [0.016]</td>
</tr>
<tr>
<td>First- or second-degree relatives</td>
<td>0.194 [0.027]</td>
</tr>
<tr>
<td>First- or second-degree relatives verified</td>
<td>0.250 [0.019]</td>
</tr>
<tr>
<td>Reported</td>
<td>0.175 [0.031]</td>
</tr>
<tr>
<td>First-degree relatives</td>
<td>0.246 [0.017]</td>
</tr>
<tr>
<td>First- or second-degree relatives</td>
<td>0.246 [0.017]</td>
</tr>
</tbody>
</table>

*P value refers to comparison of subjects with a family history of breast cancer with those without.

Table 3. Estimated odds ratio (OR) for breast cancer according to codon 325 polymorphism status before and after adjusting for risk factors

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Unadjusted OR (95% CI)</th>
<th>Adjusted OR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>01</td>
<td>1.09 (0.79–1.50)</td>
<td>1.03 (0.73–1.46)</td>
</tr>
<tr>
<td>11</td>
<td>1.59 (0.70–3.63)</td>
<td>1.58 (0.66–3.77)</td>
</tr>
<tr>
<td>Presence of the ‘‘1’’ allele</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Yes</td>
<td>1.13 (0.83–1.54)</td>
<td>1.08 (0.77–1.50)</td>
</tr>
</tbody>
</table>

*Risk factors adjusted for were city of residence, age, country of birth, marital status, highest level of education, parity, height, weight, and breast cancer in a first-degree relative. 95% CI = 95% confidence interval.

(OR = 1.39; 95% CI = 1.00–1.69; P = .05). After adjustment for the independent predictors of allele frequency, there was no difference in allele frequency between case subjects and control subjects (OR = 1.09; 95% CI = 0.83–1.41; P = .5) or between women with a family history of breast cancer and those without such a history, whether in case subjects and control subjects combined (OR = 0.88; 95% CI = 0.66–1.18; P = .4), in case subjects only (OR = 0.73; 95% CI = 0.50–1.06; P = .09), or in control subjects only (OR = 1.06; 95% CI = 0.68–1.65; P = .8). When analyses were restricted to the 489 Australian-born women or to the 419 of these for whom both parents were also born in Australia, the allele frequency was still dependent on city of residence, but it remained independent of case–control status, of family history, and of family history in case subjects alone and in control subjects alone, before and after adjustment for city of residence.

Table 3 shows that, irrespective of how ER polymorphism status was modeled, there was no association of the ER polymorphism status with breast cancer, either before or after adjustment for the risk factors identified in the full dataset (9). When modeled as a linear effect on the log OR scale, the crude effect per ‘‘1’’ allele was 0.141 (0.136 [SE]; P = .3), equivalent to predicted ORs of 1.15 and 1.32 for one and two alleles, respectively. After adjustment, the effect on log OR was 0.105 (0.146 [SE]; P = .4), with pre-
dicted ORs of 1.11 and 1.23, respectively. The SEs of the log OR estimates were less than 0.2, so that effects equivalent to a log OR of 0.5 (i.e., OR of 1.65) or more would have been detectable with more than 80% power.

For women with a reported family history of breast cancer, the crude OR for presence of the ER polymorphism on breast cancer risk was 0.88 (95% CI = 0.56–1.37); after adjustment for the covariates as in Table 3, it became 0.90 (95% CI = 0.56–1.43). For women without a reported family history, the crude OR was 1.18 (95% CI = 0.89–1.57); after adjustment, it was 1.13 (95% CI = 0.84–1.51). Because all of these 95% CIs included unity, there was no evidence of an effect of the ER polymorphism on risk of breast cancer in women with a family history of breast cancer or in women without such a history. Furthermore, the estimated effect of the ER polymorphism was no different between women with and women without a family history, in terms of either crude or adjusted OR (P = .2 and P = .4, respectively). Finally, after adjustment for both family history and ER polymorphism status, there was no evidence of an additional statistical interaction in their effects on risk of breast cancer, irrespective of whether family history of breast cancer was defined in terms of at least one reported or verified affected first-degree or first- or second-degree relative or whether analysis was in terms of crude or adjusted OR. There was little power, however, to detect nonadditivity or the log odds scale.

Discussion

In this population-based study of Australian women with breast cancer diagnosed before the age of 40 years and control subjects, there was no evidence of an association of the ER polymorphism with cancer or with a family history of breast cancer in case subjects. A previous U.S. study (3) found that the allele frequency in case subjects was greater in those with a family history of breast cancer than in those without such a history (0.28 versus 0.11, respectively). In our study, the allele frequency was 0.194 in case subjects with the same definition of a family history of breast cancer and 0.250 in case subjects without a family history. When we took into account the imprecision of these estimates, the allele frequency was no different between the U.S. and Australian case subjects with a family history (P = .1), but it was lower in the U.S. case subjects without a family history (P < .001). Importantly, we used data from population-based control subjects to show that the allele frequency in our case subjects, whether or not they had a family history of breast cancer by whatever definition or measure, did not differ from that in our control subjects.

The U.S. study found an OR for presence of the “1” allele by family history of 4.3 (95% CI = 1.8–10.1), which in case subjects under the age of 50 years was 10.8 (3). We found that it was 0.69 (95% CI = 0.43–1.09). The difference between the U.S. and Australian OR estimates was significant (P < .001). In terms of log OR, the U.S. estimates had an SE of 0.44. Our estimate had an SE of 0.24, so that we had more than 80% power to detect effects equivalent to log OR > 0.6; i.e., OR > 1.8. The effect size previously reported (3) (OR = 4.3) is equivalent to log OR = 1.5. Therefore, our study had ample power to detect effects even half the magnitude of those reported previously in case subjects.

Roodi et al. (3) reported evidence for deviation from Hardy–Weinberg equilibrium in case subjects for the ER polymorphism in codon 325. We found no evidence of this in any subgroup or overall. We obtained this result despite having studied larger numbers of women, with the consequent ability of our study to detect smaller departures from equilibrium.

What could explain the differences between the two studies? First, we studied women under the age of 40 years, whereas the majority of women in the U.S. study (3) were older than 50 years. There was no evidence, however, that the allele frequency varied with age either in the U.S. case subjects or in the Australian case subjects and/or control subjects. The differences in allele frequency that we did find, in terms of city of residence and country of birth, were relatively small (ORs < 1.5), compared with the effect reported by Roodi et al. (3) (OR = 3.1), and were measured with precision.

Second, the U.S. study (3) used a clinic-based series of case subjects and provided no information about the racial background or other demographic features of the case subjects, with or without a family history of breast cancer. On the other hand, our subjects (both case and control) were derived from complete population listings; registration of cancers is a statutory requirement in Australia, as is registration of adults eligible to vote on the electoral roll. Although our response rates for participants who provided a blood sample were 60% in case subjects and 46% in control subjects, these participants did not differ from interviewed participants from whom a blood sample was not obtained in any of the factors shown to differ between case subjects and control subjects (9). Furthermore, when comparing case subjects and control subjects, we adjusted for these factors. Given that 40% of Australians live in either Melbourne or Sydney as well as the sampling strategies used, it seems reasonable to extrapolate our findings to the Australian female population, at least to those with both parents born in Australia.

Third, Roodi et al. (3) relied on self-reported family histories only, while we tried to validate cancers in relatives and found that this validation was possible in 78% of first-degree relatives and 54% of second-degree relatives (9). The analyses of allele frequency were not greatly influenced, however, by choosing reported or verified family histories or by restricting family history of breast cancer to only first-degree relatives. Therefore, it is unlikely that the first or third issue could explain the different findings, but the second issue may have played a part.

The study by Roodi et al. (3) must be considered as “an hypothesis-generating” report. The investigators identified six ER polymorphisms and sought associations between these polymorphisms and ER phenotypes and other clinicopathologic parameters, including tumor size, grade, and stage. Apparently, the only nominally significant association that they found was with the codon 325 polymorphism; i.e., a large number of tests were performed, so by chance alone at least one test could have given a nominally significant P value. Whether or not such an association was a reflection of multiple comparisons or was indicative of a “real” effect can be assessed only by attempts at replication such as ours.

Our study has demonstrated that it is important to try to replicate in large population-based samples of case subjects and control subjects the observations from
References


Notes

Supported by the National Health and Medical Research Council of Australia, VicHealth, the NSW Cancer Council, and the Peter MacCallum Cancer Institute.

Manuscript received August 18, 1997; revised January 26, 1998; accepted January 29, 1998.