Genetic Variation in the Chromosome 17q23 Amplicon and Breast Cancer Risk

Linda E. Kelemen,1 Xianshu Wang,2 Zachary S. Fredericksen,3 V. Shane Pankratz,3 Paul D.P. Pharoah,4 Shahana Ahmed,4 Alison M. Dunning,4 Douglas F. Easton,5 Robert A. Vierkant,3 James R. Cerhan,3 Ellen L. Goode,3 Janet E. Olson,3 and Fergus J. Couch2

1Division of Population Health, Alberta Cancer Board, Calgary, Alberta, Canada; Departments of Laboratory Medicine and Pathology and Health Sciences Research, Mayo Clinic, Rochester, Minnesota; and Departments of Oncology and 5Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom

Abstract

Background: Gene amplification leading to overexpression is a common event in breast tumors that is linked to tumor development and progression. The 17q23 region is amplified in >40% of breast tumors and contains several candidate oncogenes. Because common genetic variation in several oncogenes has been associated with cancer risk, we assessed the relevance of common variants in the 17q23 candidate oncogenes to breast cancer.

Methods: We investigated 60 polymorphisms in the TUBD1, SEPT4, PRKCA, TBX2, TBX4, TEX14, TLK2, YPEL2, and PPM1E genes from this amplicon for association with breast cancer risk among 798 Caucasian breast cancer cases and 843 unaffected Caucasian controls from the Mayo Clinic.

Results: Eight polymorphisms in PRKCA, TBX4, TLK2, and YPEL2 displayed significant dose-response associations with breast cancer risk (P_trend < 0.05). Of these, PRKCA rs7342847 and TLK2 rs2245092 and rs733025 were also associated with hormone receptor–positive breast cancer: PRKCA rs7342847 (odds ratio, 0.7; 95% confidence interval, 0.6-0.9; P_trend = 0.002) and TLK2 rs733025 and rs2245092 (both: odds ratio, 0.8; 95% confidence interval, 0.7-1.0; P_trend = 0.03). Interactions between SEPT4 rs758377 and TEX14 rs302864 (PInteraction = 0.0003) and between TLK2 rs733025 and YPEL2 rs16943468 (PInteraction = 0.05) for risk of breast cancer were also observed.

Conclusion: These findings suggest that single polymorphisms and combinations of polymorphisms within candidate oncogenes from the 17q23 amplicon may influence risk of breast cancer overall and possibly specific molecular subtypes of breast tumors. The findings are discussed within the context of the results from two independent data sets. (Cancer Epidemiol Biomarkers Prev 2009;18(6):1864–8)

Introduction

Molecular cytogenetic studies have revealed common regions of DNA amplification in breast tumors (1). These regions contain oncogenes such as Erb2 and c-Myc that are frequently overexpressed as a result of the amplification (2) and contribute to tumor progression. Recently, common genetic variation in oncogenes such as FGFR2 (3) and AURKA (4), which are amplified and overexpressed in primary breast tumors (5-7), has been associated with an increased risk of breast cancer. This suggests that genes contributing to breast tumor progression as a result of amplification and overexpression may also influence breast cancer susceptibility through inherited genetic variation.

The chromosome 17q23 region is gained or amplified in >40% of breast tumors (8-10). We (11) and others (12) characterized the 4-Mb amplicon on chromosome 17q23 and identified several independent peaks of amplification-containing genes such as T-box transcription factor (TBX2; ref. 13), Peanut-like 2 (SEPT4; ref. 14), and protein phosphatase 1E (PPM1E; ref. 15).

In view of the possible link between genetic variation and amplification of oncogenes in breast cancer (3), we hypothesized that common genetic variation in selected amplified genes from the 17q23 amplicon (11), TUBD1, SEPT4, PRKCA, TBX2, TBX4, TEX14, TLK2, and YPEL2, which are overexpressed in breast tumors, is associated with breast cancer risk. Interestingly, these proteins are thought to influence the process of cell division and might be associated with cancer through induction of aneuploidy and/or polyploidy (14, 16-21) much like AURKA (22). In addition, because of a recent report suggesting an association between expression of mitotic protein kinases and luminal breast cancer (23), we examined associations between genetic variation in the kinase-encoding genes, PRKCA and TLK2, and estrogen receptor–positive (ER+) and progesterone receptor–positive (PR+) breast cancer.

Materials and Methods

Mayo Clinic Study. The Mayo Clinic Breast Cancer study is an ongoing clinic-based case-control study...
initiated in February 2001 at Mayo Clinic (Rochester, MN). The study design has been presented previously (3). Consecutive cases were women ages ≥18 y, without a prior history of cancer, diagnosed with histologically confirmed primary invasive breast cancer within 6 mo of recruitment. Women with no history of cancer were frequency matched on age (5-y age category) and region of residence to cases. Controls were recruited from those seeking general medical examinations at the outpatient practices at Mayo Clinic. The 798 cases and 843 controls in this study were Caucasian women (99% of study participants) who completed a self-administered risk factor questionnaire and provided blood samples for genomic DNA extraction.

**Single Nucleotide Polymorphism Selection, Genotyping, and Quality Control.** Tag single nucleotide polymorphisms (SNP) and all putative functional SNPs for TUBD1, SEPT4, TBX2, TBX4, TEX14, TLK2, and YPEL2 were identified according to previously published criteria (24). Two functional SNPs were included from PRKCA in place of the large number of tagSNPs in this gene. SNPs were genotyped with the GoldenGate Assay (Illumina, Inc.; refs. 25-27). Only samples and SNPs with call rates >95% were included in analyses. Concordance between 100 duplicate samples was >99.99%. A single SNP (rs16943326) from the PPM1E gene in the 17q23 amplicon that displayed a significant inverse association with postmenopausal breast cancer in stage I of the Cancer Genetic Markers of Susceptibility (CGEMS) genome-wide association study (GWAS) rank of 11 (P = 0.00004; ref. 28) was genotyped with the Taqman assay in the Mayo Clinic Genotyping Shared Resource.

The Studies of Epidemicology and Risk Factors in Cancer Heredity (SEARCH), an ongoing population-based study of invasive breast cancer cases of European ancestry ascertained through the Eastern Cancer Registration and Information Centre in England (29), was used to further evaluate the association between TLK2 rs2245092, the SNP most significantly associated with risk in the Mayo Clinic sample, and breast cancer. Successful genotyping was achieved for 97.6% of 4,470 cases and 4,560 controls with the Taqman assay.

**Statistical Analysis.** Individual SNP associations for breast cancer risk were assessed using unconditional logistic regression to estimate odds ratios (OR) and 95% confidence intervals (95% CI) using ordinal (log additive) and codominant genetic models. Analyses were adjusted for age and region of residence in the Mayo Clinic sample. Observed OR values in the Mayo Clinic sample were compared with empirical OR value distributions using Monte Carlo simulation methods (30) that permitted case-control status using 2,000 iterations. The “corrected” OR value for the SNP with the smallest observed OR value was then calculated by comparing its value with the corresponding empirical distribution of permuted OR values. We compared ORs and 95% CIs observed in the Mayo Clinic sample with age-adjusted estimates in the SEARCH sample for TLK2 rs2245092 and with estimates adjusted for age, race, and postmenopausal hormone use among postmenopausal women in the CGEMS sample for PPM1E rs16943326.

In secondary analyses, interactions between pairwise combinations of the SNPs with statistically significant individual associations were evaluated by including into statistical models a product term of the coefficients for the log-additive genetic models of both SNPs and assessing significance with a likelihood ratio test. We also subset cases defined by ER+ and PR+ tumor status and compared SNP associations in the kinase-encoding genes, PRKCA and TLK2 (23), between cases and all controls using unconditional logistic regression.

**Results**

A total of 60 SNPs (Supplementary Table S1) were examined in Mayo Clinic cases and controls. Covariates that differed in distribution by case-control status (Table 1) did not affect risk estimates when evaluated in statistical models.

Eight SNPs in four genes (PRKCA, TBX4, TLK2, and YPEL2) showed significant dose-responsive (P<0.05) associations with breast cancer risk in the Mayo Clinic sample (Table 2). Decreased risks were associated with each copy of the minor allele for PRKCA rs7342847 (P trend = 0.02), TBX4 rs11867179 (P trend = 0.04), and TLK2 rs733025, rs12451705, rs6504112, and rs2245092 (P trend ≤ 0.02), whereas increased risk was associated with each copy of the minor allele for TBX4 rs3744448 (P trend = 0.03) and YPEL2 rs16943468 (P trend = 0.03). Two other SNPs in SEPT4 and YPEL2 displayed associations with breast cancer risk in codominant genetic models (Supplementary Table S2). Overall, the number of observed significant associations (n = 10) was greater than the number expected by random variability (n = 3). When compared with the permuted P value distribution, none of the associations with breast cancer was statistically significant at P < 0.05. SNPs with observed nonsignificant associations with breast cancer are listed in Supplementary Table S3.

Given the similarity in ORs and the correlation between the four TLK2 SNPs (pairwise r² = 0.57-0.81; Table 2), we used backward elimination regression to evaluate the independent effects of TLK2 SNPs with breast cancer. Only TLK2 rs2245092 was retained by the model at P = 0.01, suggesting that rs2245092 or a variant in strong linkage disequilibrium accounted for these associations with breast cancer. We subsequently evaluated TLK2 rs2245092 for an association with breast cancer in the SEARCH study. The association was not confirmed: each allele was associated with an OR of 1.0 (95% CI, 0.9-1.1; P trend = 0.98).

In exploratory analyses driven by the hypothesis that coamplified genes may cooperate to promote tumor development, significant interactions were observed between SEPT4 rs758337 and TEX14 rs302864 (P interaction = 0.0003) and between TLK2 rs733025 and YPEL2 rs16943468 (P interaction = 0.05) under a log-additive genetic model.

We also evaluated associations between SNPs and hormone receptor–positive breast cancer in the Mayo Clinic sample. ER+/PR+ tumors were inversely related to PRKCA rs7342847 (OR, 0.7; 95% CI, 0.6-0.9; P trend = 0.002) and to TLK2 rs2245092 and rs2245092 (both: OR,
0.8; 95% CI, 0.7-1.0; $P_{\text{trend}} = 0.03$) under log-additive genetic models (Supplementary Table S4). None of the SNPs in either gene was associated with hormone receptor–negative breast cancer (data not shown). Data from 2,018 SEARCH cases with ER+ tumor status (PR+ status was unavailable) showed no association with $\text{TLK2 rs}2245092$ (OR, 1.0; 95% CI, 0.9-1.1; $P_{\text{trend}} = 0.98$).

Finally, the CGEMS study reported associations with postmenopausal breast cancer risk among women with one (OR, 0.96; 376 cases) or two (OR, 0.37; 2-degree-of-freedom $P = 0.00004$; 29 cases) copies of the minor allele of $\text{PPM1E rs}16943326$ (minor allele frequency = 0.23 among CGEMS controls). A nonsignificant trend toward increased risk of postmenopausal breast cancer was observed for this SNP among homozygote minor allele carriers (22 cases) in the Mayo Clinic sample (Supplementary Table S3). Significant associations of other SNPs in the Mayo Clinic sample by menopausal status are shown in Supplementary Table S5.

**Discussion**

Common genetic variation in amplified oncogenes such as $\text{FGFR2}$ has been associated recently with risk of breast cancer, suggesting that certain genes can both predispose

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Level</th>
<th>Cases ($n = 798$)</th>
<th>Controls ($n = 843$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>20–39</td>
<td>56 7</td>
<td>48 6</td>
</tr>
<tr>
<td></td>
<td>40–49</td>
<td>192 24</td>
<td>166 20</td>
</tr>
<tr>
<td></td>
<td>50–59</td>
<td>224 28</td>
<td>274 32</td>
</tr>
<tr>
<td></td>
<td>60–69</td>
<td>195 24</td>
<td>207 25</td>
</tr>
<tr>
<td></td>
<td>70+</td>
<td>131 16</td>
<td>148 18</td>
</tr>
<tr>
<td>State of residence$^a$</td>
<td>MN</td>
<td>502 63</td>
<td>552 66</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>296 37</td>
<td>291 34</td>
</tr>
<tr>
<td>Family history$^b$</td>
<td>Yes</td>
<td>366 47</td>
<td>345 43</td>
</tr>
<tr>
<td>Postmenopausal status</td>
<td>Yes</td>
<td>480 64</td>
<td>579 72</td>
</tr>
<tr>
<td>Age at menarche (y)</td>
<td>&lt;12</td>
<td>132 18</td>
<td>122 16</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>224 31</td>
<td>184 24</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>218 30</td>
<td>238 32</td>
</tr>
<tr>
<td></td>
<td>≥14</td>
<td>154 21</td>
<td>209 28</td>
</tr>
<tr>
<td>Postmenopausal hormone use (mo)</td>
<td>0</td>
<td>430 58</td>
<td>366 49</td>
</tr>
<tr>
<td></td>
<td>1-60</td>
<td>131 18</td>
<td>160 22</td>
</tr>
<tr>
<td></td>
<td>60+</td>
<td>184 25</td>
<td>216 29</td>
</tr>
<tr>
<td>Smoking (pack-years)</td>
<td>None</td>
<td>467 63</td>
<td>500 66</td>
</tr>
<tr>
<td></td>
<td>≤4 y</td>
<td>46 6</td>
<td>67 9</td>
</tr>
<tr>
<td></td>
<td>&gt;4 y</td>
<td>231 31</td>
<td>192 25</td>
</tr>
<tr>
<td>Tumor status</td>
<td>ER Positive</td>
<td>486 82</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PR Positive</td>
<td>440 74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>PR Negative</td>
<td>152 26</td>
<td></td>
</tr>
</tbody>
</table>

NOTE: Numbers may not total 798 cases and 843 controls due to missing data. Percentages may not total 100 due to rounding.

$^a$Other refers to IA, IL, WI, ND, and SD.

$^b$Family history in first- or second-degree relative with breast or ovarian cancer.

Table 2. Selected adjusted ORs and 95% CIs between polymorphisms in genes on the chromosome 17q23 amplicon and breast cancer risk, Mayo Clinic, 2001-2005

<table>
<thead>
<tr>
<th>Gene/SNP rsID</th>
<th>Per allele</th>
<th>$P_{\text{trend}}$</th>
<th>Homozygotes common allele (referent)</th>
<th>Heterozygotes*</th>
<th>Homozygotes rare allele*</th>
<th>$P$ 2f$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{PRKCA}$ rs7342847</td>
<td>0.8 (0.7-1.0)</td>
<td>0.02</td>
<td>335 290</td>
<td>344 426</td>
<td>0.7 (0.6-0.9)</td>
<td>119 127</td>
</tr>
<tr>
<td>$\text{TBX4}$ rs11867179</td>
<td>0.8 (0.7-1.0)</td>
<td>0.04</td>
<td>445 451</td>
<td>310 319</td>
<td>1.0 (0.8-1.2)</td>
<td>43 73</td>
</tr>
<tr>
<td>$\text{TLK2}$ rs733025</td>
<td>0.8 (0.7-0.9)</td>
<td>0.006</td>
<td>378 345</td>
<td>339 395</td>
<td>0.8 (0.6-1.0)</td>
<td>81 103</td>
</tr>
<tr>
<td></td>
<td>0.8 (0.7-1.0)</td>
<td>0.02</td>
<td>260 226</td>
<td>374 427</td>
<td>0.7 (0.6-0.9)</td>
<td>163 189</td>
</tr>
<tr>
<td></td>
<td>0.8 (0.7-0.9)</td>
<td>0.008</td>
<td>314 285</td>
<td>365 404</td>
<td>0.8 (0.7-1.0)</td>
<td>119 154</td>
</tr>
<tr>
<td></td>
<td>0.8 (0.7-0.9)</td>
<td>0.005</td>
<td>356 325</td>
<td>357 405</td>
<td>0.8 (0.6-1.0)</td>
<td>85 113</td>
</tr>
<tr>
<td>$\text{YPEL2}$ rs16943468</td>
<td>1.3 (1.1-1.8)</td>
<td>0.03</td>
<td>679 745</td>
<td>113 95</td>
<td>1.3 (1.0-1.8)</td>
<td>6 3</td>
</tr>
</tbody>
</table>

NOTE: Adjusted for age and region of residence.

Abbreviations: Ca, cases; Co, controls.

*Referent is homozygous common allele group.

$^a$Two-degree-of-freedom $P$ value.
to (3) and contribute to progression of (6) breast cancer through independent mechanisms. We observed that breast cancer risk in the Mayo Clinic sample was associated with eight common polymorphisms in the selected candidate oncogenes PRKCA, TBX4, TLK2, and YPEL2 on chromosome 17q23, a region we previously identified as amplified frequently in breast tumors (11). Moreover, the SNP by SNP interactions observed between SEPT4 and TEX14, in particular, are intriguing given that both genes participate in the separation of the dividing cell at the end of mitosis (19, 20). This supports the notion that interdependent mechanisms involving genes from the 17q23 amplicon may contribute to breast cancer.

We also observed significant associations between PRKCA kinase rs7342847 and TLK2 kinase rs733025 and rs2245092 and hormone receptor–positive tumors in the Mayo Clinic study. This observation raises the possibility that these variants predispose to specific subtypes of breast cancer, as has been observed for rs2981582 in FGFR2, rs13281615 in 8q24, and rs3803662 in TNRC9, which have recently been verified as genetic risk factors for breast cancer (31). Our findings are also intriguing given the recent suggestion that certain mitotic kinases contribute to luminal breast cancer (23). A larger sample of ER+/PR+ tumors, however, is needed to confirm the absence of association observed in our study with these SNPs.

We were unable to confirm the association between TLK2 rs2245092 and breast cancer using the SEARCH sample. Possibly, one of the other TLK2 SNPs with which TLK2 rs2245092 was in strong linkage disequilibrium in the Mayo Clinic sample may have shown an association with breast cancer had it been genotyped in SEARCH. We were also unable to confirm the association reported in CGEMS between PPMIE rs16943326 and postmenopausal breast cancer using the Mayo Clinic data. In post hoc evaluation of the CGEMS data, SNPs in TUBD1, SEPT4, TBX2, and YPEL2 were represented in both CGEMS and the Mayo Clinic samples; however, only YPEL2 rs16943468 was significantly associated with postmenopausal breast cancer in both studies, and similar to PPMIE rs16943326, the ORs were in opposite directions. At a minor allele frequency of 0.06 (YPEL2 rs16943468) or 0.22 (PPMIE rs16943326), if the true OR among homozygotes is in the range of 1.2 to 1.5, then a much larger study than Mayo Clinic or CGEMS would be needed to have sufficient power to confirm the association at either the conventional P < 0.05 or genome-wide (P ≤ 10−7) level of significance (32). It also remains possible that the distribution of other exposures may explain the differences in genetic associations with CGEMS (e.g., 60% of Mayo Clinic cases did not use postmenopausal hormones versus <30% in the Nurses’ Health Study; ref. 33), emphasizing the need for careful interpretation of results between study populations.

The observed significant P values for SNPs shown to be associated with breast cancer in the Mayo Clinic sample were not supported by distributions obtained using mathematical permutation methods. However, the biological plausibility of our hypothesis and the support from published literature of examples of oncogenes within amplons that contribute to cancer at the level of commonly inherited variation (e.g., FGFR2; ref. 3) suggest that the findings are at least as plausible as the explanation of random variability (34). Replication in similar populations of adequate sample size, therefore, should be the standard that determines the significance of other potentially promising associations cited in this report.

It is plausible that a proportion of women at risk of breast cancer due to chromosome 17q23 polymorphisms may also have gene amplification in this chromosomal region or that other biological or clinical features may modify the genetic associations. Although tissues and clinical data were unavailable for evaluation in the present report, we are pursuing these hypotheses for future investigation.

In conclusion, this study may suggest a role for genetic variation within candidate oncogenes from the 17q23 amplicon in predisposition to breast cancer. Further studies will be critical to clarify the importance of these findings.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

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.

References

16. Morse-Gaudio M, Connolly JM, Rose DP. Protein kinase C and its
Genetic Variation in the Chromosome 17q23 Amplicon and Breast Cancer Risk

Linda E. Kelemen, Xianshu Wang, Zachary S. Fredericksen, et al.


Updated version
Access the most recent version of this article at:
http://cebp.aacrjournals.org/content/18/6/1864

Supplementary Material
Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2009/05/18/1055-9965.EPI-08-0486.DC1

Cited articles
This article cites 33 articles, 15 of which you can access for free at:
http://cebp.aacrjournals.org/content/18/6/1864.full.html#ref-list-1

Citing articles
This article has been cited by 5 HighWire-hosted articles. Access the articles at:
/content/18/6/1864.full.html#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.