New Breast Cancer Risk Variant Discovered at 10q25 in East Asian Women

Jiajun Shi1, Hyuna Sung2, Ben Zhang1, Wei Lu9, Ji-Yeob Cho2,5, Yong-Bing Xiang10, Mi Kyung Kim8,11, Motoki Iwasaki12, Jirong Long1, Bu-Tian Ji13, Sue K. Park2,3,5, Ying Zheng9, Shoichiro Tsugane12, Keun-Young Yoo6, Wenjing Wang6, Dong-Young Noh14,5, Wonshik Han14,5, Sung-Won Kim14, Min Hyuk Lee7, Jong Won Lee8, Jong-Young Lee14, Chen-Yang Shen15,16, Keitaro Matsuo17, Sei-Hyun Ahn8,11, Yu-Tang Gao10, Xiao Ou Shu1, Qiuyin Cai1, Daehee Kang2,3,5, and Wei Zheng1

Abstract

Background: Recently, 41 new genetic susceptibility loci for breast cancer risk were identified in a genome-wide association study (GWAS) conducted in European descendants. Most of these risk variants have not been directly replicated in Asian populations.

Methods: We evaluated nine of those nonreplication loci in East Asians to identify new risk variants for breast cancer in these regions. First, we analyzed single-nucleotide polymorphisms (SNP) in these regions using data from two GWAS conducted among Chinese and Korean women, including 5,083 cases and 4,376 controls (stage 1). In each region, we selected an SNP showing the strongest association with breast cancer risk for replication in an independent set of 7,294 cases and 9,404 controls of East Asian descents (stage 2). Logistic regression models were used to calculate adjusted ORs and 95% confidence intervals (CI) as a measure of the association of breast cancer risk and genetic variants.

Results: Two SNPs were replicated in stage 2 at P < 0.05: rs1419026 at 6q14 [per allele OR, 1.07; 95% confidence interval (CI), 1.03–1.12; P = 3.0 × 10⁻⁴] and rs941827 at 10q25 (OR, 0.92; 95% CI, 0.89–0.96; P = 5.3 × 10⁻⁵). The association with rs941827 remained highly statistically significant after adjusting for the risk variant identified initially in women of European ancestry (OR, 0.88; 95% CI, 0.82–0.97; P = 5.3 × 10⁻⁵).

Conclusion: We identified a new breast cancer risk variant at 10q25 in East Asian women.

Impact: Results from this study improve the understanding of the genetic basis for breast cancer. Cancer Epidemiol Biomarkers Prev; 22(7); 1297–303. ©2013 AACR.

Introduction

Genetic factors play a significant role in the etiology of breast cancer (1–4). To date, genome-wide association studies (GWAS) have identified approximately 67 genetic susceptibility risk loci for breast cancer (5–19). With a few exceptions (9, 13, 14, 16, 20), most susceptibility loci were initially identified in GWAS conducted in European ancestry populations. Most, if not all, of the initially reported risk variants, in the form of single-nucleotide polymorphisms (SNPs; termed as index SNPs in subsequent text), are tagging SNPs. These SNPs were identified likely through their linkage disequilibrium (LD) with disease variants. Because LD patterns differ across populations of different ancestries, some findings from GWAS conducted in European descendants cannot be directly extrapolated to other populations (20–27). We recently evaluated all breast cancer risk variants identified to date and found that approximately half of the risk variants identified initially in European descendants cannot be directly replicated in East Asians (28). In the present study, we investigated 9 regions where the index SNP...
has not been replicated in Asian samples in an attempt to identify other breast cancer risk variants for East Asian women.

Materials and Methods

Study populations

This study was conducted as part of the Asia Breast Cancer Consortium (ABCC), which has been described elsewhere (9, 13, 14, 18, 22, 28). Briefly, samples analyzed in this study were from 8 epidemiologic studies in the ABCC (Table 1). Samples were from 13,642 Chinese women, 11,713 Korean women, and 802 Japanese women. Chinese participants were selected from 4 studies: the Shanghai Breast Cancer Study (SBCS), the Shanghai Breast Cancer Survival Study (SBCSS), the Shanghai Endometrial Cancer Study (SECS, controls only), and the Shanghai Women’s Health Study (SWHS; refs. 9, 29, 30). Korean participants came from 4 studies: the Seoul Breast Cancer Study (SeBCS; ref. 20), Korea National Cancer Center (Korea-NCC), Korea Genome Epidemiology Study (KoGES; ref. 31), and Korean Hereditary Breast Cancer (KOHBRA; ref. 32). Japanese samples were from the Japan Nagano Breast Cancer Study (33). In total, 12,377 cases and 13,780 controls were analyzed in the present study.

Genotyping and quality control

Stage 1 testing was conducted using existing data from 2 GWAS, wherein 5,285 Chinese women and 4,777 Korean women were genotyped primarily using the Affymetrix Genome-wide Human SNP Array 6.0. Genotyping protocols have been described elsewhere (9, 18, 20). From the Chinese GWAS, we included 1 negative control and at least 3 positive quality control (QC) samples from the Coriell Cell Repositories in each of the 96-well plates for genotyping with Affymetrix SNP Array 6.0 chips. A total of 273 positive QC samples were successfully genotyped; the average concordance rate was 99.9% with a median value of 100%. Genetically identical and unexpected duplicate samples were excluded, as they were close relatives with a pair-wise proportion of identify-by-descent (IBD) estimate greater than 0.25. All samples with a call rate less than 95% were excluded. SNPs were excluded if: (i) minor allele frequency (MAF) < 1%; (ii) call rate < 95%; or (iii) genotyping concordance rate < 95% in QC samples. The final dataset included 2,918 cases and 2,324 controls for 690,947 markers. For the Korean GWAS, the Affymetrix SNP Array 6.0 was used (20). A total of 30 QC samples were successfully genotyped; the average concordance rate was 99.8%. SNPs were excluded if: (i) genotype call rate < 95%; (ii) MAF < 1% in either cases or controls; (iii) evidence for deviation from Hardy–Weinberg equilibrium (HWE) at \( P < 10^{-6} \); or (iv) poor genotyping cluster plot in either cases or controls. After QC filtering, the final dataset included 2,165 cases and 2,052 controls for 555,525 markers. All samples from both studies were genetically confirmed to be females.

We used the program MACH 1.0 (34) to impute genotypes for autosomal SNPs in HapMap Phase II release 22 for samples from the Chinese and Korean GWAS. Only SNPs with imputation quality score RSQR ≥ 0.3 were included in subsequent analyses. Dosage data for imputed SNPs in samples from each of the stage 1 studies were analyzed using the program mach2dat (34).

We genotyped 9 selected SNPs in stage 2 using the iPLEX MassARRAY platform (Sequenom). PCR primers

Table 1. Summary of selected characteristics of participants by studies

<table>
<thead>
<tr>
<th>Study</th>
<th>No. of cases</th>
<th>No. of controls</th>
<th>Population ethnicity</th>
<th>Study designa</th>
<th>Mean age, by</th>
<th>Menopause, c %</th>
<th>ER, d %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SBCGS1 (GWAS)</td>
<td>2,918</td>
<td>2,324</td>
<td>Chinese</td>
<td>Population-based</td>
<td>52/50</td>
<td>43/42</td>
<td>65</td>
</tr>
<tr>
<td>SBCGS2 (stage 2)</td>
<td>1,613</td>
<td>1,800</td>
<td>Chinese</td>
<td>Population-based</td>
<td>53/53</td>
<td>50/55</td>
<td>62</td>
</tr>
<tr>
<td>SBCGS3 (stage 3)</td>
<td>2,601</td>
<td>2,386</td>
<td>Chinese</td>
<td>Population-based</td>
<td>54/65</td>
<td>50/63</td>
<td>65</td>
</tr>
<tr>
<td>SeBCS1 (GWAS)</td>
<td>2,165</td>
<td>2,052</td>
<td>Korean</td>
<td>Hospital-based</td>
<td>48/51</td>
<td>36/55</td>
<td>63</td>
</tr>
<tr>
<td>SeBCS2 (stage 2)</td>
<td>777</td>
<td>1,104</td>
<td>Korean</td>
<td>Hospital-based</td>
<td>48/48</td>
<td>36/37</td>
<td>63</td>
</tr>
<tr>
<td>KOHBRA/KoGES</td>
<td>1,397</td>
<td>3,209</td>
<td>Korean</td>
<td>Hospital-based with community controls</td>
<td>40/50</td>
<td>23/NA</td>
<td>63</td>
</tr>
<tr>
<td>Korea-NCC</td>
<td>505</td>
<td>504</td>
<td>Korean</td>
<td>Hospital-based</td>
<td>49/49</td>
<td>50/45</td>
<td>65</td>
</tr>
<tr>
<td>Nagano</td>
<td>401</td>
<td>401</td>
<td>Japanese</td>
<td>Hospital-based</td>
<td>54/54</td>
<td>55/65</td>
<td>75</td>
</tr>
<tr>
<td>Total</td>
<td>12,377</td>
<td>13,780</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aCase–control study.

bMean age of cases–controls with available data.

cProportion of postmenopausal status of cases–controls with available data.

dAmong cases with ER data.

*Including cases and controls from four studies conducted in Shanghai.

*Significant at \( p = 0.01 \) level.

Abbreviations: NA, not available; Nagano, Japan Nagano Breast Cancer Study; SBCGS, Shanghai Breast Cancer Genetic Study.
and allelic-specific extension primers were designed with the MassARRAY Assay Design 4.0 software, and alleles of each SNP were detected through matrix-assisted laser desorption/ionization–time-of-flight mass spectrometry on the MassARRAY Analyzer 4 System (Sequenom). In each 96-well plate, 5 QC samples were used in the Sequenom assay, including 1 negative control (water), 2 blinded duplicates, and 2 samples from the HapMap project. Mean concordance rate was 99.7% for the blind duplicates and 98.9% for the HapMap samples.

Statistical analyses

PLINK version 1.06 was used to analyze the genomewide data obtained in stage 1 (35). Population structure was investigated by analyzing 690,947 SNPs that passed QC in Chinese women (2,918 cases and 2,324 controls) and 555,525 SNPs that passed QC in Korean women (2,165 cases and 2,052 controls), showing an estimated inflation factor λ of 1.042 for Chinese (18) and 1.043 for Koreans (20). Therefore, population substructure, if present, should not substantially affect the results of this study. Logistic regression was used to estimate breast cancer risk under a log-additive model and adjusted for age and study site, when appropriate ORs associated with each SNP and 95% confidence intervals (CI) were estimated for cancer risk. Conditional analyses were conducted by adjusting the index SNP in each locus to evaluate possible independent association of breast cancer with the SNP under study. All analyses were conducted using SAS version 9.3 (SAS Institute Inc.) unless noted otherwise.

Results

Table 2 presents the associations of 9 index SNPs with breast cancer risk from the Chinese and Korean GWAS (stage 1). Results from the original GWAS of European-ancestry populations are shown, also. None of these index SNPs were associated with breast cancer risk at \( P < 0.05 \) in our study. To search for other possible variants in these regions which could be associated with breast cancer risk, we selected an SNP from each of these regions for further evaluation. In each region, we selected an SNP that showed the most significant association with breast cancer risk and is located within ±500kb of the index SNP of that region (Table 3).

Stage 1 and 2 results for the 9 selected SNPs are presented in Table 3. SNPs rs1419026 at 6p23 and rs941827 at 10q25 showed an association with breast cancer risk at \( P < 0.05 \) in the same direction in both stages. In the combined analysis of 12,377 breast cancer cases and 13,780 control women, per allele ORs were 1.07 (95% CI, 1.03–1.10) and 1.07 (1.04–1.10) for rs1419026 and rs941827, respectively. The association of rs1419026 with breast cancer risk was, in general, consistent across participating studies and heterogeneity test was not statistically significant (\( P = 0.675 \)). The association of breast cancer risk with rs941827 was substantially stronger in the Nagano study (OR, 0.72; 95% CI, 0.58–0.89; \( P = 0.002 \)) than other 7 studies combined (OR, 0.93; 95% CI, 0.89–0.97; \( P = 3.8 \times 10^{-5} \); \( P_{\text{heterogeneity}} = 0.027 \)). After excluding the Nagano study, the heterogeneity test was no longer statistically significant (\( P = 0.103 \)).
Table 3. Associations of breast cancer risk with SNPs selected for this study

<table>
<thead>
<tr>
<th>Tested SNP</th>
<th>Chr./gene</th>
<th>Position, bp</th>
<th>Alleles</th>
<th>EAF</th>
<th>Index SNP</th>
<th>CEU Imputed</th>
<th>OR (95% CI)</th>
<th>P ( P_{\text{trend}} )</th>
<th>CHB Imputed</th>
<th>OR (95% CI)</th>
<th>P ( P_{\text{trend}} )</th>
<th>Combined Imputed</th>
<th>OR (95% CI)</th>
<th>P ( P_{\text{trend}} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs788166</td>
<td>2q31/DLX2</td>
<td>172623735 C/G</td>
<td>0.62</td>
<td>rs2016394</td>
<td>0.01</td>
<td>0.06</td>
<td>Yes</td>
<td>1.09 (1.02–1.15)</td>
<td>0.006</td>
<td>1.01 (0.97–1.06)</td>
<td>0.605</td>
<td>1.04 (1.00–1.08)</td>
<td>0.037</td>
<td></td>
</tr>
<tr>
<td>rs821287</td>
<td>6p23/RANBP9</td>
<td>13464145 G/A</td>
<td>0.68</td>
<td>rs204247</td>
<td>0.01</td>
<td>0.05</td>
<td>Yes</td>
<td>0.86 (0.77–0.95)</td>
<td>0.003</td>
<td>1.07 (1.02–1.12)</td>
<td>0.005</td>
<td>1.03 (0.99–1.08)</td>
<td>0.164</td>
<td></td>
</tr>
<tr>
<td>rs1419026</td>
<td>6q14/FAM46A</td>
<td>82145898 T/C</td>
<td>0.29</td>
<td>rs17529111</td>
<td>0.29</td>
<td>0.56</td>
<td>Yes</td>
<td>1.07 (1.00–1.13)</td>
<td>0.048</td>
<td>1.08 (1.02–1.13)</td>
<td>0.003</td>
<td>1.07 (1.03–1.12)</td>
<td>3.0 ( 10^{-4} )</td>
<td></td>
</tr>
<tr>
<td>rs10278902</td>
<td>7q33/ARHGEF5</td>
<td>143320582 G/A</td>
<td>0.59</td>
<td>rs720475</td>
<td>0</td>
<td>0</td>
<td>Yes</td>
<td>0.94 (0.88–1.00)</td>
<td>0.034</td>
<td>1.06 (1.01–1.11)</td>
<td>0.014</td>
<td>1.01 (0.98–1.05)</td>
<td>0.434</td>
<td></td>
</tr>
<tr>
<td>rs2608036</td>
<td>8q24/MYC</td>
<td>129194161 G/A</td>
<td>0.82</td>
<td>rs11780156</td>
<td>0.07</td>
<td>0.05</td>
<td>No</td>
<td>1.10 (1.02–1.19)</td>
<td>0.014</td>
<td>1.04 (0.98–1.10)</td>
<td>0.192</td>
<td>1.06 (1.01–1.11)</td>
<td>0.011</td>
<td></td>
</tr>
<tr>
<td>rs16921849</td>
<td>10p12/DNAJC1</td>
<td>21873162 G/A</td>
<td>0.34</td>
<td>rs11814448</td>
<td>0.01</td>
<td>0.01</td>
<td>Yes</td>
<td>1.07 (1.00–1.14)</td>
<td>0.036</td>
<td>0.98 (0.93–1.02)</td>
<td>0.342</td>
<td>1.01 (0.97–1.05)</td>
<td>0.621</td>
<td></td>
</tr>
<tr>
<td>rs941827</td>
<td>10q25/TCF7L2</td>
<td>114548877 C/T</td>
<td>0.29</td>
<td>rs7904519</td>
<td>0.01</td>
<td>0.02</td>
<td>No</td>
<td>0.88 (0.82–0.93)</td>
<td>5.0 ( 10^{-5} )</td>
<td>0.95 (0.90–1.00)</td>
<td>0.049</td>
<td>0.92 (0.89–0.96)</td>
<td>5.3 ( 10^{-6} )</td>
<td></td>
</tr>
<tr>
<td>rs500161</td>
<td>11q22/OVOL1</td>
<td>65452014 T/C</td>
<td>0.51</td>
<td>rs12575663</td>
<td>0.56</td>
<td>0.11</td>
<td>Yes</td>
<td>1.05 (0.99–1.12)</td>
<td>0.084</td>
<td>0.99 (0.95–1.04)</td>
<td>0.761</td>
<td>1.01 (0.98–1.05)</td>
<td>0.425</td>
<td></td>
</tr>
<tr>
<td>rs138019</td>
<td>22q13/MKL1</td>
<td>38941574 G/A</td>
<td>0.92</td>
<td>rs6001930</td>
<td>0.02</td>
<td>0.03</td>
<td>Yes</td>
<td>0.86 (0.77–0.97)</td>
<td>0.009</td>
<td>1.04 (0.96–1.14)</td>
<td>0.334</td>
<td>0.98 (0.91–1.04)</td>
<td>0.464</td>
<td></td>
</tr>
</tbody>
</table>

\( ^{a} \text{One SNP was selected in each locus for the study.} \\
^{b} \text{The closest gene to corresponding index SNP in each region.} \\
^{c} \text{Location based on NCBI Human Genome Build 36.3.} \\
^{d} \text{Effect/reference alleles were defined on the basis of NCBI Human Genome Build 36.3, dbSNP 126 forward strand.} \\
^{e} \text{Effect allele frequency in controls of stage 2 studies.} \\
^{f} \text{On the basis of HapMap genotype data release 27, dbSNP building 126.} \\
^{g} \text{Adjusted for age and study site.} \\
\text{Abbreviations: Chr., chromosome; EAF, effect allele frequency.} \)
Conditional analyses for rs1419026 and rs941827 were conducted by adjusting the index SNP in each of these loci. These analyses were conducted using stage 1 samples with data available for both new and index SNPs. The association with rs941827 at 10q25 remained statistically significant after adjusting for index SNP rs7904519 (OR, 0.88; 95% CI, 0.82–0.93; \( P = 5.3 \times 10^{-5} \); data now shown in tables). However, the significant association with rs1419026 disappeared after adjusting for its index SNP rs17529111 (OR, 1.05; 95% CI, 0.96–1.15; \( P = 0.293 \)). SNPs rs1419026 and rs941827 were associated with both estrogen receptor (ER)+ and ER– cancer (Table 4).

Nominally significant associations were also observed for 2 other SNPs (rs821287 and rs10278902) in stage 2. However, the direction of the association for these 2 SNPs was inconsistent in stages 1 and 2, and thus these 2 SNPs were considered not being replicated in this study.

Discussion

In this large study conducted in East Asian women, we identified a new genetic risk variant for breast cancer at 10q25, a breast cancer susceptibility locus identified recently in a GWAS of European descendants (19). We also found an SNP (rs1419026) at 6q14 that showed a stronger association with breast cancer risk in East Asians than the index SNP (rs17529111) initially identified in this locus in a GWAS conducted among European descendants. Our study has expanded the list of breast cancer risk variants identified for East Asian women and provides data that might be useful in fine-mapping GWAS-identified regions to identify causal variants for this common malignancy.

The index SNP rs7904519 at 10q25 was not replicated in our study. The risk allele frequency is very low in East Asian women (0.045) compared with European descendants (0.405). In the present study, we found a significant association of breast cancer risk with SNP rs941827, with an effective allele frequency of 0.26 in East Asian women and 0.29 in Europeans. These 2 SNPs are not correlated (\( r^2 < 0.03 \) in either CEU or CHB + JPT samples). The index SNP (rs7904519) is located in intron 3 of the TCF7L2 gene (NM_000756). The SNP identified in our study (rs941827) is located in intron 7 of the vesicle transport through interaction with t-SNAREs homolog 1A (yeast) (VTI1A) gene (NM_145206), approximately 215 kb upstream of rs7904519. The VTI1A gene encodes a soluble N-ethylmaleimide-sensitive fusion protein attachment protein receptor (SNARE) that mediates the transport of vesicles between the Golgi apparatus and the plasma membrane (36, 37). The potential role of the VTI1A protein in breast carcinogenesis remains unknown. Intriguingly, SNPs rs941827 and rs7904519 are included in a 420- to 540-kb region that has been found to be deleted in some breast and colorectal cancer samples (38, 39). This deletion causes a VTI1A-TCF7L2 fusion, which may affect the regulatory function of the TCF/β-catenin complex on the Wnt signaling (40).

Further studies are needed to clarify the mechanism of the association of VTI1A variants and breast cancer risk identified in this study.

In our study, some imputed SNPs could not be investigated properly because of their low imputation quality in stage 1. Some of the index SNPs evaluated in stage 1 showed an association in the same direction as reported previously in the European-ancestry study, although the association was not statistically significant perhaps due to a small sample size. In addition, we selected only one SNP per locus for stage 2 replication because of budget constraints. It is possible that additional risk variants may be located in some of these regions and can be further investigated in the future. Nevertheless, using data from East Asian women, we identified one new genetic risk variant at 10q25 for breast cancer. We also identified a risk variant at 6q14 that showed a stronger association with breast cancer in East Asians than the index SNP initially discovered in this region in a GWAS conducted in European descendants. These results are new and should be helpful for future studies to understand the genetic basis for breast cancer.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Disclaimer

The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Authors’ Contributions


www.aacrjournals.org

Cancer Epidemiol Biomarkers Prev; 22(7) July 2013

1301


Acknowledgments

The authors thank participants and research staff for this study; and Regina Courtney, Jie Wu, Jing He, Mary Jo Daly, and Bethanie Rammer for their help with sample preparation and technical support for the project at Vanderbilt.

References


New Breast Cancer Risk Variant Discovered at 10q25 in East Asian Women


Updated version

Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-12-1393

Cited articles

This article cites 40 articles, 4 of which you can access for free at:
http://cebp.aacrjournals.org/content/22/7/1297.full#ref-list-1

Citing articles

This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/22/7/1297.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, use this link http://cebp.aacrjournals.org/content/22/7/1297.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.