Pathway Analyses Identify TGFBR2 as Potential Breast Cancer Susceptibility Gene: Results from a Consortium Study among Asians

Xiangyu Ma1,2, Alicia Beeghly-Fadiel1, Wei Lu3, Jiajun Shi1, Yong-Bing Xiang4, Qiuyin Cai1, Hongbing Shen5, Chen-Yang Shen6,7, Zefang Ren8, Keitaro Matsuo9, Motoki Iwasaki11, Jirom Long1, Ben Zhang1, Bu-Tian Ji12, Ying Zheng1, Wenjing Wang3, Zhibin Hu5, Yao Liu5, Pei-Ei Wu6, Ya-Lan Shieh6, Shenming Wang13, Xiaoming Xie14, Hidemi Ito9, Yoshio Kasuga15, Kelvin Y.K. Chan10,16, Hiroji Iwata17, Shoichiro Tsugane11, Yu-Tang Gao4, Xiao Ou Shu1, Harold L. Moses18, and Wei Zheng1

Abstract

Background: The TGF-β signaling pathway plays a significant role in the carcinogenic process of breast cancer. Methods: We systematically evaluated associations of common variants in TGF-β signaling pathway genes with breast cancer risk using a multistage, case-control study among Asian women. Results: In the first stage, 341 single-nucleotide polymorphisms with minor allele frequencies ≥ 0.05 across 11 genes were evaluated among 2,926 cases and 2,380 controls recruited as a part of the Shanghai Breast Cancer Genetics Study (SBCGS). In the second stage, 20 SNPs with promising associations were evaluated among an additional 1,890 cases and 2,000 controls from the SBCGS. One variant, TGFBR2 rs1078985, had highly consistent and significant associations with breast cancer risk among participants in both study stages, as well as promising results from in silico analysis. Additional genotyping was carried out among 2,475 cases and 2,343 controls from the SBCGS, as well as among 5,077 cases and 5,384 controls from six studies in the Asian Breast Cancer Consortium (stage III). Pooled analysis of all data indicated that minor allele homozygotes (GG) of TGFBR2 rs1078985 had a 24% reduced risk of breast cancer compared with major allele carriers (AG or AA; OR, 0.76; 95% CI, 0.65–0.89; P = 8.42 × 10−4). Conclusion: These findings support a role for common genetic variation in TGF-β signaling pathway genes, specifically in TGFBR2, in breast cancer susceptibility. Impact: These findings may provide new insights into the etiology of breast cancer as well as future potential therapeutic targets. Cancer Epidemiol Biomarkers Prev; 21(7); 1176–84. ©2012 AACR.

Introduction

The TGF-β signaling pathway is composed of several multifunctional cytokines and receptors that are involved in regulating various essential cellular processes including growth, differentiation, apoptosis, angiogenesis, and homeostasis (1, 2). This pathway also plays an important role in the development and progression of multiple human diseases, such as cancer, asthma, autoimmune, and cardiovascular diseases (2–6). In the context of cancer, the TGF-β signaling pathway has both tumor-suppressing and tumor-promoting functions depending upon the cellular context. In the early stages of cancer, TGF-β signaling...
can inhibit tumor growth, whereas in later stage cancers, tumor invasiveness, and metastasis are promoted by TGF-β signaling (1, 3).

Signal transduction of the TGF-β ligands [TGF-β1, TGF-β2, and TGF-β3] is mediated through interactions with their receptors. The ligands first bind to TGFBR2, a transmembrane serine/threonine protein kinase receptor; this interaction is sometimes mediated by TGFBR3. This complex then binds to and activates TGF-β receptor 1 (TGFBR1), which in turn phosphorylates SMAD family member 2 (SMAD2) and SMAD3. Phosphorylated SMAD2 and SMAD3, in association with SMAD4, form a complex which accumulates in the nucleus and acts as a transcription factor to regulate target genes. SMAD7 can block the activation of SMAD2 and SMAD3, while the SMAD anchor for receptor activation (SARA, also known as ZFYVE9) stabilizes the SMAD4–TGFBR1 interaction (7, 8). Genetic variation in constituents of the TGF-β signaling pathway may result in altered protein function, increased or decreased target gene transcription, and therefore, the development and progression of breast cancer.

Although there is considerable biologic plausibility for the involvement of the TGF-β signaling pathway in the development of breast cancer, limited information is available about the impact of genetic variation on breast cancer risk. Most studies of genetic variation in TGF-β signaling pathway genes and breast cancer risk have focused on a few single-nucleotide polymorphisms (SNP), and findings have been inconsistent. The most extensively studied variant in the TGF-β pathway is a SNP located in exon 1 of the TGFBI gene (T29C, also known as rs1982073, which merged into rs1800470; refs. 7, 9–13). Although this SNP was shown to be associated with increased TGFBI protein secretion, its association with breast cancer risk has been inconsistent (7, 9–12). Our recent field synopsis included data from 32 studies for this variant; no significant association with breast cancer risk was found using allelic, dominant, or recessive models (10). Notably, results from a large study that evaluated 354 genetic variants in 17 TGF-β pathway genes for associations with breast cancer risk among women of European ancestry found that only this SNP (rs1982073, which merged into rs1800470) retained statistical significance after correction for multiple comparisons in analyses of progesterone receptor negative (PR−) breast cancer (14). Three other TGFBI variants (rs1800468, rs1800469, and rs1800471) and one TGFBR1 variant (rs11466445) have also been previously evaluated; meta-analyses for these SNPs have not found significant associations with breast cancer risk (10, 15, 16).

This study was undertaken to comprehensively evaluate the associations of genetic variants in the TGF-β signaling pathway with breast cancer risk among Asian women. In the discovery stage, 341 genetic variants in 11 TGF-β pathway genes [TGFBI, TGFBR2, TGFBR3, TGFBR1, TGF-β receptor 2 (TGFBR2), TGF-β receptor 3 (TGFBR3), SMAD2, SMAD3, SMAD4, SMAD7, and SARA] were evaluated among 2,926 cases and 2,380 controls from studies of Chinese women in Shanghai. Promising SNPs were then evaluated for replication of associations with breast cancer risk among an additional 1,890 cases and 2,000 controls from Shanghai. Finally, one SNP was further genotyped among 7,552 cases and 7,727 controls, comprised of 7 additional independent studies conducted among Chinese and Japanese women, as part of the Asian Breast Cancer Consortium.

Materials and Methods

Study population

The Shanghai Breast Cancer Genetics Study (SBCGS) includes data from 4 population-based studies conducted among Chinese women in urban Shanghai: the Shanghai Breast Cancer Study (SBCS; refs. 17, 18), the Shanghai Breast Cancer Survival Study (SBCSS; ref. 19), the Shanghai Endometrial Cancer Study (SECS; ref. 20), and the Shanghai Women’s Health Study (SWHS; ref. 21). Details of these studies have been described previously (17). Briefly, the SBCS is a 2-stage (SBCS-I and SBCS-II), population-based, case–control study. SBCS-I recruitment occurred between August 1996 and March 1998; SBCS-II recruitment occurred between April 2002 and February 2005. The SBCSS included newly diagnosed breast cancer cases ascertained via the population-based Shanghai Cancer Registry between April 2002 and December 2006. The SECS is a population-based, case–control study of endometrial cancer conducted between January 1997 and December 2003 using a protocol similar to the SBCS; only community controls from the SECS were included in the SBCGS. The SWHS is a population-based cohort study of women from urban communities in Shanghai who were recruited between 1996 and 2000. In this analysis, stage I (SBCGS I) included 2,926 cases and 2,380 controls from the SBCS, SBCSS, and SWHS; stage II (SBCGS II) included 1,890 cases and 2,000 controls from the SBCS, SBCSS, SWHS, and SECS. Stage III included 2,475 cases from the SBCSS and 2,343 controls from the SWHS and SECS (SBCGS III), as well as data from 6 collaboration studies, including 2,095 women from Taiwan (22, 23), 1,050 women from Hong Kong (24), 3,580 women from Nanjing, China (25, 26), 1,657 women from Guangzhou, China; 1,284 women from Nagoya, Japan (27); and 795 women from Nagano, Japan (28), participating in the Asian Breast Cancer Consortium. Appropriate approval was granted from all relevant Institutional Review Boards and all included participants provided informed consent.

Genotyping and quality control

Over the past few years, we have genotyped TGF-β signaling pathway genes in several projects. To maximize our coverage of genetic variation for these genes, in the discovery stage, we included all genotyping data generated in these projects for this analysis. First, 88 haplotype
taging SNPs (htSNPs) in 11 genes were genotyped among 2,083 participants using a targeted genotyping system (Affymetrix Inc.). Second, 412 SNPs in 11 TGF-β pathway genes were genotyped as part of the Affymetrix Genome-Wide SNP Array 6.0 (Affymetrix Inc.) for 5,242 participants. Third, one SNP (rs1800469) was genotyped by TaqMan (Applied Biosystems) among 1,978 participants. Fourth, 3 SNPs (rs1800469, rs1800470, and rs1800471) were genotyped by RFLP among 2,277 participants. Finally, 2 SNPs (rs1461085 and rs2026811) were genotyped with the Sequenom iPLEX MassARRAY platform (Sequenom) among 1,978 participants. Twenty-seven SNPs were genotyped by more than one method, so that the total number of SNPs genotyped was 479. Analysis for stage I was conducted for 341 SNPs with a minimum minor allele frequency (MAF) of 5% among genotyped controls. Twenty promising SNPs were selected for additional stage I genotyping by the Sequenom iPLEX MassARRAY platform (stage II). One replicated SNP (rs1078985) was further evaluated among participants of SBCGS III and 6 Asian collaboration studies (stage III). Genotyping of these women was also carried out with the Sequenom iPLEX MassARRAY. For all genotyping methods, blinded duplicate samples and quality controls (QC) were included as described previously (17, 18). All included SNPs had call rates and concordance rates of at least 95% among duplicates within each platform, as well as across genotyping platforms. Laboratory personnel were blinded to the case-control and QC status of all samples.

**Statistical analysis**

All statistical analyses, except where noted, were conducted with SAS version 9.2 (SAS Institute Inc.) and Stata 11.0 (Stata Corporation). Characteristics of demographic data between cases and controls were compared with the χ² or t test for categorical or continuous variables, respectively. Hardy–Weinberg equilibrium among controls was evaluated using the χ² test. ORs and corresponding 95% CIs were calculated using logistic regression models that included adjustment for age. Additive, dominant, and recessive models of effect were used for all SNPs. Interactions were evaluated using likelihood ratio tests for nested models; case-only analyses were used to evaluate associations between SNPs and tumor characteristics. Heterogeneity between stages I and II results was evaluated using the Cochran’s Q statistic; significant heterogeneity was determined by P ≤ 0.10 (29). Pooled analysis was conducted with data from the SBCGS and the Asian Breast Cancer Consortium for the association between TGFBR2 rs1078985 and breast cancer risk. Linkage disequilibrium (LD) was assessed with SNAP (30). The Bonferroni adjustment was used to address the issue of false positive findings arising from multiple comparisons. Quanto was used for power calculations (31). All statistical tests were 2-tailed, and P ≤ 0.05 was interpreted as statistically significant unless otherwise indicated.

### Results

A 3-stage study design was used (Fig. 1). In total, 12,368 breast cancer cases and 12,107 controls were included in this analysis (Table 1). The 3 study stages included 2,926 cases and 2,380 controls from the SBCGS (stage I), 1,890 cases and 2,000 controls from the SBCGS (stage II), and 7,552 cases and 7,727 controls (stage III) from the SBCGS and the Asian Breast Cancer Consortium. Overall, cases were slightly older and more likely to be postmenopausal than controls. Information on breast cancer stage was available for the majority of the cases (91.8%) from the Shanghai studies; only 193 (3.0%) were in situ breast cancer (data not shown). Information about included SNPs and genetic coverage of 11 TGF-β signaling pathway genes is shown in Table 2. In stage I, a total of 479 SNPs were genotyped across the TGB1, TGB2, TGB3, TGBR1, TGBFR2, TGBFR3, SMAD2, SMAD3, SMAD4, SMAD7, and SARA genes; of these, 341 had MAFs ≥ 5% among controls in our study population. The total number of SNPs genotyped was 479. Analyses were evaluated using the Cochran’s Q statistic; significant heterogeneity was determined by P ≤ 0.10 (29). Pooled analysis was conducted with data from the SBCGS and the Asian Breast Cancer Consortium for the association between TGFBR2 rs1078985 and breast cancer risk. Linkage disequilibrium (LD) was assessed with SNAP (30). The Bonferroni adjustment was used to address the issue of false positive findings arising from multiple comparisons. Quanto was used for power calculations (31). All statistical tests were 2-tailed, and P ≤ 0.05 was interpreted as statistically significant unless otherwise indicated.

![Figure 1. Study design.](image-url)
rs12562433) in 4 loci had inconsistent associations between the 2 SBCS study populations and were not further evaluated. Using an $r^2 > 0.3$, the remaining variants were found to represent 20 independent loci. One SNP from each of these loci was selected for additional genotyping in stage II. Design failed for one variant (rs12403389), so it was replaced with another SNP in high LD (rs10874915), despite not having a statistically significant association itself. Genotyping failed for one variant (rs745103), and so it could not be further analyzed.

Eight SNPs were found to have significant associations with breast cancer risk in the combined analysis of stages I and II data (Table 3); results from the 2 study stages for all 8 were not significantly heterogeneous ($P > 0.10$; data not shown). One SNP (TGFBR2 rs1078985) had significant associations with breast cancer risk in both study stages, as well as highly consistent risk estimates. When results from the 2 stages were combined, both heterozygotes (OR, 0.84; 95% CI, 0.765–0.93) and homozygotes (OR, 0.73; 95% CI: 0.55–0.97) had significantly lower risks of breast cancer

### Table 1. Selected characteristics of participants in the Asian Breast Cancer Consortium by study stage

<table>
<thead>
<tr>
<th>Study stage</th>
<th>Ethnicity</th>
<th>Cases</th>
<th>Controls</th>
<th>Mean age$^a$</th>
<th>% Postmenopausal$^b$</th>
<th>ER$^{1,b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stage I</td>
<td>SBCGS I</td>
<td>Chinese</td>
<td>2,926</td>
<td>2,380</td>
<td>51.7/50.2</td>
<td>42.9/41.4</td>
</tr>
<tr>
<td>Stage II</td>
<td>SBCGS II</td>
<td>Chinese</td>
<td>1,890</td>
<td>2,000</td>
<td>52.8/53.3</td>
<td>48.8/55.0</td>
</tr>
<tr>
<td>Stage III</td>
<td>SBCGS III</td>
<td>Chinese</td>
<td>2,475</td>
<td>2,343</td>
<td>53.8/55.0</td>
<td>50.4/52.5</td>
</tr>
<tr>
<td>Taiwan</td>
<td>Chinese</td>
<td>1,049</td>
<td>1,046</td>
<td>52.6/39.6</td>
<td>50.9/47.0</td>
<td>651 (54.9%)</td>
</tr>
<tr>
<td>Hong Kong</td>
<td>Chinese</td>
<td>429</td>
<td>621</td>
<td>45.8/45.6</td>
<td>50.3/41.8</td>
<td>157 (70.4%)</td>
</tr>
<tr>
<td>Nanjing</td>
<td>Chinese</td>
<td>1,757</td>
<td>1,823</td>
<td>50.6/50.2</td>
<td>50.9/47.0</td>
<td>651 (54.9%)</td>
</tr>
<tr>
<td>Guangzhou</td>
<td>Japanese</td>
<td>804</td>
<td>853</td>
<td>49.0/49.2</td>
<td>49.0/49.2</td>
<td>156 (73.6%)</td>
</tr>
<tr>
<td>Nagoya</td>
<td>Japanese</td>
<td>640</td>
<td>644</td>
<td>51.4/51.1</td>
<td>48.4/48.5</td>
<td>353 (73.2%)</td>
</tr>
<tr>
<td>Nagano</td>
<td>Japanese</td>
<td>398</td>
<td>397</td>
<td>53.8/54.1</td>
<td>50.3/54.1</td>
<td>294 (74.4%)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>12,368</td>
<td>12,107</td>
<td>51.8/51.3</td>
<td>47.8/47.0</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Cases/controls; bold values denote significant difference at $P < 0.01$.

$^b$Number and percentage of estrogen receptor positive (ER$^+$) breast cancer cases among those with data available.

### Table 2. Gene and SNP information for included TGF-β signaling pathway genes among women in Shanghai

<table>
<thead>
<tr>
<th>TGF-β pathway genes</th>
<th>Genomic location</th>
<th>Gene span, kb</th>
<th>SNPs in HapMap$^a$</th>
<th>SNPs genotyped</th>
<th>Genetic variation coverage (%)$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ligands</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGFβ1</td>
<td>19q13.1</td>
<td>23.2</td>
<td>16</td>
<td>14</td>
<td>93.8</td>
</tr>
<tr>
<td>TGFβ2</td>
<td>1q41</td>
<td>95.1</td>
<td>106</td>
<td>44</td>
<td>81.1</td>
</tr>
<tr>
<td>TGFβ3</td>
<td>14q24</td>
<td>23.1</td>
<td>19</td>
<td>13</td>
<td>68.4</td>
</tr>
<tr>
<td>Receptors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TGFBR1</td>
<td>9q22</td>
<td>49.1</td>
<td>49</td>
<td>24</td>
<td>91.8</td>
</tr>
<tr>
<td>TGFBR2</td>
<td>3p22</td>
<td>87.6</td>
<td>118</td>
<td>102</td>
<td>97.5</td>
</tr>
<tr>
<td>TGFBR3</td>
<td>1p33-p32</td>
<td>203.7</td>
<td>298</td>
<td>112</td>
<td>81.5</td>
</tr>
<tr>
<td>Cofactors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SMAD2</td>
<td>12q21</td>
<td>98.1</td>
<td>87</td>
<td>45</td>
<td>100.0</td>
</tr>
<tr>
<td>SMAD3</td>
<td>15q21-22</td>
<td>139.3</td>
<td>170</td>
<td>76</td>
<td>77.6</td>
</tr>
<tr>
<td>SMAD4</td>
<td>18q21.1</td>
<td>49.5</td>
<td>25</td>
<td>10</td>
<td>92.0</td>
</tr>
<tr>
<td>SMAD7</td>
<td>18q21.1</td>
<td>30.9</td>
<td>27</td>
<td>17</td>
<td>48.1</td>
</tr>
<tr>
<td>SARA (ZFYVE9)</td>
<td>1p32.3</td>
<td>204.3</td>
<td>50</td>
<td>22</td>
<td>88.0</td>
</tr>
</tbody>
</table>

$^a$SNPs with a MAF $\geq 5\%$ in HapMap (v2 R24), Han Chinese (CHB) data, $\pm 10$ kb for each gene.

$^b$Coverage of HapMap CHB SNPs by our genotyped SNPs for $r^2 = 0.8$, using a pairwise tagging approach in Tagger.
Table 3. Associations with breast cancer risk for selected TGF-β signaling pathway variants, the SBCGS

<table>
<thead>
<tr>
<th>Information&lt;sup&gt;a&lt;/sup&gt;</th>
<th>N (cases/controls)</th>
<th>AB OR (95% CI)</th>
<th>BB OR (95% CI)</th>
<th>P</th>
<th>AB/BB OR (95% CI)</th>
<th>P</th>
<th>BB OR (95% CI)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TGFBR2 rs1079895 (A/Q)</strong></td>
<td>947/899</td>
<td>1.28 (1.05–1.55)</td>
<td>1.42 (0.69–2.93)</td>
<td>0.0111</td>
<td>1.28 (1.06–1.55)</td>
<td>0.0108</td>
<td>1.31 (0.64–2.68)</td>
<td>0.4682</td>
</tr>
<tr>
<td>Combined</td>
<td>2,566/2,688</td>
<td>1.16 (1.03–1.30)</td>
<td>1.11 (0.82–1.52)</td>
<td>0.0226</td>
<td>1.15 (1.03–1.29)</td>
<td>0.0132</td>
<td>1.06 (0.78–1.44)</td>
<td>0.7092</td>
</tr>
<tr>
<td><strong>TGFBR2 rs304822 (C/T)</strong></td>
<td>947/899</td>
<td>1.28 (1.05–1.55)</td>
<td>1.42 (0.69–2.93)</td>
<td>0.0111</td>
<td>1.28 (1.06–1.55)</td>
<td>0.0108</td>
<td>1.31 (0.64–2.68)</td>
<td>0.4682</td>
</tr>
<tr>
<td>Combined</td>
<td>2,566/2,688</td>
<td>1.16 (1.03–1.30)</td>
<td>1.11 (0.82–1.52)</td>
<td>0.0226</td>
<td>1.15 (1.03–1.29)</td>
<td>0.0132</td>
<td>1.06 (0.78–1.44)</td>
<td>0.7092</td>
</tr>
<tr>
<td><strong>SMAD3 rs7178117 (G/C)</strong></td>
<td>947/899</td>
<td>1.28 (1.05–1.55)</td>
<td>1.42 (0.69–2.93)</td>
<td>0.0111</td>
<td>1.28 (1.06–1.55)</td>
<td>0.0108</td>
<td>1.31 (0.64–2.68)</td>
<td>0.4682</td>
</tr>
<tr>
<td>Combined</td>
<td>2,566/2,688</td>
<td>1.16 (1.03–1.30)</td>
<td>1.11 (0.82–1.52)</td>
<td>0.0226</td>
<td>1.15 (1.03–1.29)</td>
<td>0.0132</td>
<td>1.06 (0.78–1.44)</td>
<td>0.7092</td>
</tr>
</tbody>
</table>

**NOTE:** Estimates and P values in bold denote significance at P < 0.05.
<sup>a</sup>Information includes alleles (major or reference allele/minor allele) as determined by allele frequency among all genotyped controls, MAF among all genotyped controls, and region of the gene where the SNP is located.

<sup>b</sup>Breast cancer risk for heterozygotes (AB) and minor allele homozygotes (BB), compared with major allele homozygotes (AA), in models adjusted for age and genotyping stage when appropriate; P<sub>additive</sub>.

<sup>c</sup>Breast cancer risk for minor allele carriers (AB/BB) compared with major allele homozygotes (AA), in models adjusted for age and genotyping stage when appropriate; P value for dominant association.

<sup>d</sup>Breast cancer risk for minor allele homozygotes (BB) compared with major allele carriers (AA/AB), in models adjusted for age and genotyping stage when appropriate; P value for recessive association.

than major allele homozygotes (AA). Furthermore, both additive and dominant effect models were highly significant (P < 1.9 × 10<sup>−5</sup>). This surpassed a Bonferroni corrected significance threshold for the number of variants evaluated in stage II (P = 0.05/19 = 2.63 × 10<sup>−5</sup>). In addition, nominally significant associations with breast cancer risk were also found for TGFBR2 rs2799086, TGFBR2 rs17047740, TGFBR1 rs2026811, TGFBR1 rs10733710, TGFBR2 rs304822, TGFBR3 rs284185, and SMAD3 rs7178117 in the combined analysis, although none of these SNPs had significant associations in stage II. Regression models shown in Table 3 include adjustment for age and genotyping stage when appropriate; additional adjustment for education, age at menarche, age at menopause, age at first live birth, menstrual status, a first-degree relative with breast cancer, use of hormone replacement therapy, previous history of fibroadenoma, physical activity, body mass index, and waist-to-hip ratio.
**Discussion**

In this multistage study, we comprehensively evaluated genetic variation of 11 genes in the TGF-β signaling pathway with breast cancer risk among Asian women. One SNP (rs1078985) in intron 3 of the TGFBR2 gene, showed a consistent association in all 3 stages. Pooled analysis revealed a significantly reduced risk of breast cancer in a recessive genetic model (OR, 0.76; 95% CI, 0.65–0.89; $P = 8.42 \times 10^{-4}$). This novel finding provides support for a role of TGF-β signaling pathway in the etiology of breast cancer. Although the association of rs1078985 with breast cancer risk was identified in our study initially under the additive model (stages I and II), after evaluating data from 7 additional studies, a recessive model seemed to best explain the association. This may be due to the reduced power for detecting recessive associations in stages I and II. Using Quanto, we found that the power to find an association for an SNP with an MAF of 16% was less than 42% for a recessive model in the analysis of stages I and II combined. After including data from the SBCGS III and the 6 collaboration studies, however, we had >87% power to detect such an association.

Results from *in silico* analysis were supportive of an association between TGFBR2 rs1078985 and breast cancer risk. Using TFSEARCH (32), a web-based program that searches for transcription factor binding sites, an Nkx-2.5 binding site was found to be present when the major A allele was present, but not when the minor G allele was.

Nkx-2.5 is a transcriptional regulator of iodide transport in thyroid and mammary cells; a role in cancer has been implicated, as Nkx-2.5 has been shown to be expressed in breast cancer cell lines, as well as in mammary glands during lactation (33). The ENCODE transcription factor chromatin immunoprecipitation (ChIP) track of the UCSC Genome Browser (Build 36 assembly, hg18; ref. 34), was also evaluated; this track shows regions where...
transcription factors have been shown to bind by ChIP with specific antibodies followed by DNA sequencing (35). SNP rs1078985 was found to be within one experimentally verified transcription factor binding region (NF-kB) and adjacent to 7 additional regions, the strongest signal of which was found for PU.1. NF-kB regulates genes involved in the immune and inflammatory responses, as well as genes important for cell proliferation, apoptosis, angiogenesis, invasion, and therefore carcinogenesis (36, 37). Overexpression of NF-kB regulates genes involved in the immune and inflammatory responses, as well as genes important for cell proliferation, apoptosis, angiogenesis, invasion, and therefore carcinogenesis (36, 37). Overexpression of NF-kB and NF-κB has been shown in breast cancer cell lines and breast carcinomas (36), and many cancer cells show aberrant constitutive NF-kB activation, which mediates resistance to chemotherapy and radiotherapy (36, 37). PU.1 is an erythroleukemia transformation specific–domain transcription factor that binds purine-rich sequences and can regulate alternative splicing of target genes; it has been postulated that PU.1 can reduce the transcriptional activity of the p53 tumor suppressor family, thereby altering cell-cycle regulation and apoptosis (38). Together, these data provide considerable biologic plausibility for a role for this TGFBR2 SNP in breast cancer etiology.

Four SNPs in TGFBR1 (rs1800468, rs1800469, rs1800470, and rs1800471), and 1 SNP in TGFBR2 (rs11466445) have been reported to be associated with breast cancer risk in previous studies (7, 9–13, 15, 16). Among them, rs1800470 (also known as T29C or rs1982073) has been the most frequently investigated (7, 9–13). Similar to other previously reported SNPs, the association of rs1800470 with breast cancer risk has not been robustly replicated (10). In our study, none of these SNPs were significantly associated with breast cancer risk. Instead, in addition to the one replicated association (rs1078985), 7 additional SNPs (TGFBR2 rs2799086, TGFBR2 rs17047740, TGFBR1 rs2026811, TGFBR1 rs10733710, TGFBR2 rs304822, TGFBR3 rs284185, and SMAD3 rs7178117) had some evidence for possible associations with breast cancer risk. However, none of these marginally significant associations remained significant after adjusting for multiple comparisons.

Major strengths of this study include a multistage study design with a large sample size, the population-based design of the SBCGS, and a comprehensive and systematic analysis of genetic variants in 11 TGF-β signaling pathway genes. Limitations to be considered include that only the SMAD-mediated TGF-β signaling pathway was evaluated. Although this is the best characterized mechanism of TGF-β signaling, many studies have also shown that TGF-βs can exert their effects through SMAD-independent pathways, such as phosphoinositide 3-kinase, mitogen–activated protein kinase, protein phosphatase 2, PKB, extracellular signal–regulated kinase, and c-jun-NH2–kinase (7). A further limitation of our study was the relatively low coverage of variants in the TGFBR3 and SMAD7 genes. However, our coverage was above 75% for 9 included genes (using an $r^2$ of 0.8) and was on average very good (83.6%).

In conclusion, our finding of an association between TGFBR2 rs1078985 and a reduced risk of breast cancer among Asian women was not only replicated within the SBCGS, but was also evident among data from 6 collaboration studies. Together, our results support an important role for SNPs in the TGF-β signaling pathway genes in breast cancer susceptibility. These findings may provide new insights into the etiology of breast cancer as well as future potential therapeutic targets.

**Disclosure of Potential Conflicts of Interest**

The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agents. U.S. Khoo is a consultant and an advisory board member of Vanderbilt University. No potential conflicts of interest were disclosed by the other authors.

### Table 4. Association of TGFBR2 rs1078985 and breast cancer risk by tumor stage, the SBCGS

<table>
<thead>
<tr>
<th></th>
<th>Stage 0 or I</th>
<th>Stage II</th>
<th>Stage III or IV</th>
<th>Total*</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_{cases}$</td>
<td>OR (95% CI)</td>
<td>$N_{cases}$</td>
<td>OR (95% CI)</td>
<td>$N_{cases}$</td>
</tr>
<tr>
<td>TGFBR2 rs1078985</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>1.661</td>
<td>1.00 (reference)</td>
<td>2.502</td>
<td>1.00 (reference)</td>
</tr>
<tr>
<td>AG</td>
<td>0.91 (0.82–1.02)</td>
<td>0.87 (0.79–0.96)</td>
<td>0.98 (0.82–1.17)</td>
<td>0.89 (0.82–0.96)</td>
</tr>
<tr>
<td>GG</td>
<td>0.72 (0.52–1.01)</td>
<td>0.76 (0.56–1.01)</td>
<td>0.81 (0.47–1.38)</td>
<td>0.75 (0.60–0.94)</td>
</tr>
<tr>
<td>$P^b$</td>
<td>0.0011</td>
<td>0.5294</td>
<td>0.0001</td>
<td>0.0769</td>
</tr>
<tr>
<td>AA/AG</td>
<td>2.229</td>
<td>1.00 (reference)</td>
<td>3.325</td>
<td>1.00 (reference)</td>
</tr>
<tr>
<td>GG</td>
<td>0.74 (0.53–1.03)</td>
<td>0.79 (0.60–1.05)</td>
<td>0.81 (0.48–1.39)</td>
<td>0.77 (0.62–0.97)</td>
</tr>
<tr>
<td>$P^c$</td>
<td>0.1026</td>
<td>0.4502</td>
<td>0.0263</td>
<td>0.0263</td>
</tr>
</tbody>
</table>

NOTE: Bold values denote significance at $P \leq 0.05$.

*Includes 565 women without information on tumor stage.

$P^b$ value for recessive association.

$P^c$ value for heterogeneity test.
TGF-β Pathway Variants and Breast Cancer Risk

Authors' Contributions
Analysis and interpretation of data: X. Ma, A. Beeghly-Fadiel, J. Shi, C.-Y. Shen, J. Long, B. Ji,
Writing, review, and/or revision of the manuscript: X. Ma, A. Beeghly-Fadiel, Y.-B. Xiang, H. Shen, Z. Ren, M. Keitaro, M. Ivasaki, B. Ji, Z. Hu, Y.-T. Gao, X.-O. Shu, W. Zheng
Administrative, technical, or material support: Y.-B. Xiang, M. Keitaro, M. Ivasaki, B. Ji, W. Wang, P.-E. Wu, X. Xie, K.-Y. Chan, Y.-T. Gao

Acknowledgments
The authors thank the study participants and research staff for their contributions and commitment to this project, Regina Courtney for DNA preparation, Jing He for data processing and analyses, and Bethanie Rammer for clerical support in the preparation of this manuscript.

Grant Support
This work was supported in part by U.S. NIH USPHS grants R01 CA124558, R01 CA145667, R01 CA06277, R01 CA090899, and R37 CA070867 to W. Zheng, and R01 CA92585 and R01 CA118229 to X. O. Shu. X. Ma (from the Third Military Medical University in Chongqing, China) is supported by the China Scholarship Council (CSC) while the author is visiting the Vanderbilt Epidemiology Center in the United States.

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.

Received January 30, 2012; revised March 21, 2012; accepted April 16, 2012; published OnlineFirst April 26, 2012.

References


Pathway Analyses Identify *TGFBR2* as Potential Breast Cancer Susceptibility Gene: Results from a Consortium Study among Asians

Xiangyu Ma, Alicia Beeghly-Fadiel, Wei Lu, et al.


Updated version  Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-12-0118

Supplementary Material  Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2012/04/24/1055-9965.EPI-12-0118.DC1

Cited articles  This article cites 37 articles, 16 of which you can access for free at:
http://cebp.aacrjournals.org/content/21/7/1176.full#ref-list-1

Citing articles  This article has been cited by 3 HighWire-hosted articles. Access the articles at:
http://cebp.aacrjournals.org/content/21/7/1176.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, contact the AACR Publications Department at permissions@aacr.org.