Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

Cancer Epidemiology, Biomarkers & Prevention

Research Article

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

Xiangyu Ma[1,2], Alicia Beeghly-Fadiel[1], Wei Lu[3], Jiajun Shi[1], Yong-Bing Xiang[4], Qiuyin Cai[1], Hongbing Shen[5], Chen-Yang Shen[6,7], Zefang Ren[8], Keitaro Matsuo[9], Ui Soon Khoo[10], Motoki Iwasaki[11], Jirong Long[1], Ben Zhang[1], Bu-Tian Ji[12], Ying Zheng[3], Wenjing Wang[3], Zhibin Hu[5], Yao Liu[5], Pei-Ei Wu[6], Ya-Lan Shieh[6], Shenming Wang[13], Xiaoming Xie[14], Hidemi Ito[9], Yoshio Kasuga[15], Kelvin Y.K. Chan[10,16], Hiroji Iwata[17], Shoichiro Tsugane[11], Yu-Tang Gao[4], Xiao Ou Shu[1], Harold L. Moses[18], and Wei Zheng[1]

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; $P = 8.42 \times 10^{-4}$).

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

Authors' Affiliations: [1]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee; [2]Visiting from Department of Epidemiology, College of Preventive Medicine, Third Military Medical University, Chongqing; [3]Shanghai Center for Disease Control and Prevention, Shanghai, China; [4]Department of Epidemiology, Shanghai Cancer Institute, Shanghai, China; [5]Department of Epidemiology and Biostatistics, Nanjing Medical University, Nanjing, China; [6]Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan; [7]College of Public Health, China Medical University, Taichung, Taiwan; [8]School of Public Health, Sun Yat-sen University, Guangzhou, China; [9]Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya, Japan; [10]Department of Pathology, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; [11]Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan; [12]Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Department of Health and Human Services, Bethesda, Maryland; [13]First Affiliated Hospital, Sun Yat-sen University, Guangzhou, China; [14]Cancer Center, Sun Yat-sen University, Guangzhou, China; [15]Department of Surgery, Nagano Matsushiro General Hospital, Nagano, Japan; [16]Department of Obstetrics & Gynecology, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong SAR, China; [17]Department of Breast Oncology, Aichi Cancer Center Central Hospital, Nagoya, Japan; [18]Department of Cancer Biology, Medicine and Pathology, Vanderbilt University School of Medicine, Nashville, Tennessee

Note: Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Corresponding Author: Alicia Beeghly-Fadiel, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, 2525 West End Avenue, 8th Floor, Nashville, TN 37203. Phone: 615-936-3512; Fax: 615-322-0502; E-mail: alicia.beeghly@vanderbilt.edu

doi: 10.1158/1055-9965.EPI-12-0118

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

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 *TGFB1* gene (*T29C*, also known as *rs1982073*, which merged into *rs1800470*; refs. 7, 9–13). Although this SNP was shown to be associated with increased TGFB1 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 *TGFB1* 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 [*TGFB1*, *TGFB2*, *TGFB3*, *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

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

tagging 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 II 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 χ^2 or *t* test for categorical or continuous variables, respectively. Hardy–Weinberg equilibrium among controls was evaluated using χ^2 tests. ORs and corresponding CIs were determined by 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 $P \leq 0.10$ (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 $P \leq 0.05$ was interpreted as statistically significant unless otherwise indicated.

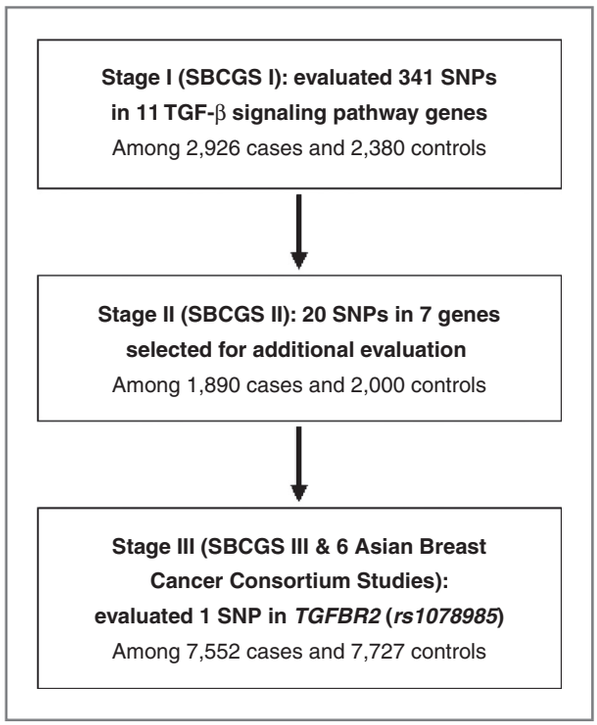

Figure 1. Study design.

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 *TGFB1*, *TGFB2*, *TGFB3*, *TGFBR1*, *TGFBR2*, *TGFBR3*, *SMAD2*, *SMAD3*, *SMAD4*, *SMAD7*, and *SARA* genes; of these, 341 had MAFs $\geq$ 5% among controls in our study population (Supplementary Table S1). Our coverage of the polymorphisms in these 11 genes (MAFs $\geq$ 5%) was estimated to be approximately 83.6% using an $r^2 = 0.8$ and a pairwise tagging approach.

Associations with breast cancer risk for the 341 TGF-β signaling pathway SNPs with MAFs $\geq$ 5% yielded significant *P* values from additive, dominant, or recessive models for 43 SNPs (Supplementary Table S2). When possible, consistency of associations between SBCS-I and SBCS-II study populations was assessed; 5 SNPs (*rs6696224*, *rs10493858*, *rs12132114*, *rs11165293*, and

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

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

Study stage	Ethnicity	Cases	Controls	Mean age[a]	% Postmenopausal[a]	ER^{+}[b]
Stage I						
SBCGS I	Chinese	2,926	2,380	**51.7/50.2**	42.9/41.4	1,581 (54.3%)
Stage II						
SBCGS II	Chinese	1,890	2,000	52.8/53.3	**48.8/55.0**	924 (59.9%)
Stage III						
SBCGS III	Chinese	2,475	2,343	**53.8/55.0**	50.4/52.5	1,542 (62.3%)
Taiwan	Chinese	1,049	1,046	**51.6/47.5**	**52.6/39.6**	634 (66.1%)
Hong Kong	Chinese	429	621	45.8/45.6	50.3/41.8	157 (70.4%)
Nanjing	Chinese	1,757	1,823	50.6/50.2	50.9/47.0	651 (54.9%)
Guangzhou	Chinese	804	853	49.0/49.2	**39.8/50.7**	156 (73.6%)
Nagoya	Japanese	640	644	51.4/51.1	48.4/48.5	353 (73.2%)
Nagano	Japanese	398	397	53.8/54.1	**54.8/65.7**	294 (74.4%)
Total		12,368	12,107	**51.8/51.3**	**47.8/47.0**	

[a]Cases/controls; bold values denote significant difference at $P \leq 0.01$.
[b]Number and percentage of estrogen receptor positive (ER^{+}) breast cancer cases among those with data available.

rs12562433) in 4 loci had inconsistent associations between the 2 SBCS study populations and were not further evaluated. Using an $r^2 \geq 0.3$, 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 2. Gene and SNP information for included TGF-β signaling pathway genes among women in Shanghai

TGFβ pathway genes	Genomic location	Gene span, kb	SNPS in HapMap[a]	SNPs genotyped		Genetic variation coverage (%)[b]
				All	MAF ≥ 5%	
Ligands						
TGFB1	19q13.1	23.2	16	14	13	93.8
TGFB2	1q41	95.1	106	44	29	81.1
TGFB3	14q24	23.1	19	13	6	68.4
Receptors						
TGFBR1	9q22	49.1	49	24	16	91.8
TGFBR2	3p22	87.6	118	102	80	97.5
TGFBR3	1p33-p32	203.7	298	112	86	81.5
Cofactors						
SMAD2	18q21	98.1	87	45	37	100.0
SMAD3	15q21-22	129.3	170	76	51	77.6
SMAD4	18q21.1	49.5	25	10	8	92.0
SMAD7	18q21.1	30.9	27	17	6	48.1
SARA (ZFYVE9)	1p32.3	204.3	50	22	9	88.0

[a]SNPs with a MAF ≥ 5% in HapMap (v2 R24), Han Chinese (CHB) data, ± 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.

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

Table 3. Associations with breast cancer risk for selected TGF-β signaling pathway variants, the SBCGS

Information[a]	*N* (cases/controls)	Breast cancer risk, additive model[b] AB OR (95% CI)	BB OR (95% CI)	*P*	Dominant model[c] AB/BB OR (95% CI)	*P*	Recessive model[d] BB OR (95% CI)	*P*
***TGFBR2 rs1078985* (A/G), 15.9%, intron 3**								
SBCGS I	2,909/2,316	0.87 (0.76–0.98)	0.79 (0.55–1.12)	**0.0117**	0.86 (0.76–0.97)	**0.0132**	0.82 (0.57–1.17)	0.2656
SBCGS II	1,543/1,746	0.81 (0.69–0.95)	0.64 (0.39–1.04)	**0.0024**	0.79 (0.68–0.93)	**0.0038**	0.67 (0.42–1.09)	0.1064
Combined	4,452/4,062	0.84 (0.76–0.93)	0.73 (0.55–0.97)	**1.15×10^{-4}**	0.84 (0.76–0.92)	**1.88×10^{-4}**	0.77 (0.58–1.02)	0.0605
***TGFB2 rs2799086* (C/T), 25.0%, intron 2**								
SBCGS I	2,764/2,177	1.04 (0.92–1.17)	1.34 (1.05–1.71)	0.0589	1.07 (0.96–1.20)	0.2247	1.32 (1.04–1.68)	**0.0228**
SBCGS II	1,613/1,800	1.06 (0.92–1.23)	1.18 (0.90–1.55)	0.1821	1.08 (0.94–1.24)	0.2583	1.16 (0.89–1.51)	0.2883
Combined	4,377/3,977	1.05 (0.96–1.15)	1.27 (1.06–1.52)	**0.0191**	1.08 (0.99–1.18)	0.0907	1.25 (1.04–1.49)	**0.0147**
***TGFB2 rs17047740* (C/T), 9.9%, intron 2**								
SBCGS I	2,771/2,176	1.04 (0.90–1.20)	2.22 (1.19–4.12)	0.1088	1.08 (0.94–1.24)	0.2772	2.20 (1.18–4.08)	**0.0126**
SBCGS II	1,601/1,784	1.12 (0.94–1.34)	1.08 (0.54–2.14)	0.2107	1.12 (0.95–1.33)	0.1892	1.05 (0.53–2.09)	0.8801
Combined	4,372/3,960	1.08 (0.97–1.20)	1.58 (1.01–2.47)	**0.0418**	1.10 (0.99–1.23)	0.0883	1.56 (1.00–2.44)	0.0511
***TGFBR1 rs2026811* (C/A), 46.9%, intron 1**								
SBCGS I	1,616/1,585	0.82 (0.70–0.96)	0.88 (0.72–1.07)	0.1340	0.84 (0.72–0.97)	**0.0221**	1.00 (0.84–1.18)	0.9834
SBCGS II	908/888	0.89 (0.71–1.10)	0.91 (0.70–1.18)	0.4278	0.89 (0.73–1.10)	0.2830	0.98 (0.78–1.23)	0.8552
Combined	2,524/2,473	0.84 (0.74–0.96)	0.89 (0.76–1.04)	0.0956	0.86 (0.76–0.97)	**0.0125**	0.99 (0.87–1.14)	0.9274
***TGFBR1 rs10733710* (G/A), 18.7%, intron 6**								
SBCGS I	947/889	1.28 (1.05–1.55)	1.42 (0.69–2.93)	**0.0111**	1.28 (1.06–1.55)	**0.0108**	1.31 (0.64–2.68)	0.4682
SBCGS II	1,609/1,799	1.10 (0.95–1.27)	1.04 (0.74–1.47)	0.2952	1.09 (0.95–1.25)	0.2262	1.01 (0.72–1.42)	0.9466
Combined	2,556/2,688	1.16 (1.03–1.30)	1.11 (0.82–1.52)	**0.0226**	1.15 (1.03–1.29)	**0.0132**	1.06 (0.78–1.44)	0.7092
***TGFBR2 rs304822* (C/T), 38.9%, 3′ flanking region**								
SBCGS I	2,853/2,255	0.84 (0.74–0.95)	0.92 (0.78–1.10)	0.0791	0.86 (0.76–0.96)	**0.0084**	1.02 (0.87–1.20)	0.8101
SBCGS II	1,598/1,780	0.91 (0.78–1.05)	0.88 (0.72–1.08)	0.1589	0.90 (0.78–1.03)	0.1382	0.93 (0.77–1.13)	0.4687
Combined	4,451/4,035	0.87 (0.79–0.95)	0.91 (0.80–1.04)	**0.0272**	0.88 (0.80–0.96)	**0.0034**	0.98 (0.87–1.11)	0.7868
***TGFBR3 rs284185* (T/A), 11.0%, intron 4**								
SBCGS I	2,763/2,167	1.00 (0.87–1.15)	1.71 (1.04–2.83)	0.3103	1.04 (0.90–1.19)	0.6176	1.71 (1.04–2.83)	**0.0352**
SBCGS II	1,609/1,791	0.86 (0.72–1.02)	1.72 (0.96–3.07)	0.5621	0.90 (0.76–1.07)	0.2207	1.77 (0.99–3.16)	0.0532
Combined	4,372/3,958	0.94 (0.84–1.05)	1.72 (1.17–2.51)	0.6682	0.98 (0.88–1.09)	0.7060	1.74 (1.19–2.54)	**0.0042**
***SMAD3 rs7178117* (G/C), 10.0%, intron 1**								
SBCGS I	2,771/2,177	1.14 (0.99–1.32)	1.57 (0.94–2.64)	**0.0190**	1.17 (1.01–1.34)	**0.0340**	1.53 (0.91–2.58)	0.1055
SBCGS II	1,600/1,781	1.11 (0.93–1.32)	0.82 (0.43–1.58)	0.4599	1.09 (0.92–1.29)	0.3293	0.81 (0.42–1.55)	0.5203
Combined	4,371/3,958	1.13 (1.01–1.27)	1.22 (0.82–1.82)	**0.0185**	1.14 (1.02–1.27)	**0.0195**	1.19 (0.81–1.77)	0.3761

NOTE: Estimates and *P* values in bold denote significance at $P \leq 0.05$.

[a]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.

[b]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_{trend}.

[c]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.

[d]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 ($P < 1.9 \times 10^{-4}$). This surpassed a Bonferroni corrected significance threshold for the number of variants evaluated in stage II (P, $0.05/19 = 2.63 \times 10^{-3}$). In addition, nominally significant associations with breast cancer risk were also found for *TGFB2 rs2799086*, *TGFB2 rs17047740*, *TGFBR1 rs2026811*, *TGFBR1 rs10733710*, *TGFBR2 rs304822*, *TGFBR3 rs284185*, and *SMAD3 rs7178117* in the combined analyze, 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, menopausal 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

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

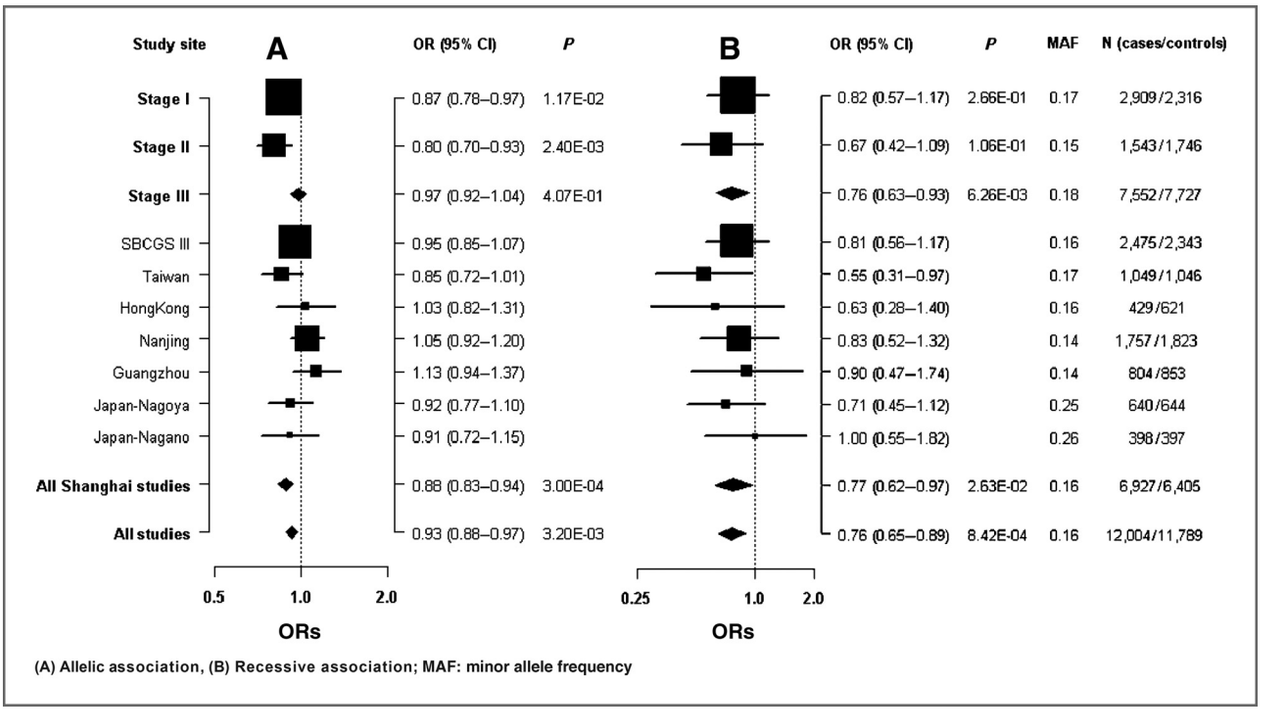

Figure 2 Forest plots for associations of breast cancer risk with *TGFBR2 rs1078985* by study site, the Asian Breast Cancer Consortium.

did not materially alter these findings (data not shown). Results for the remaining 11 SNPs evaluated in stage 2 are shown in Supplementary Table S3; results from analyses of the 2 stages combined were either nonsignificant or else had significant heterogeneity of associations with breast cancer risk.

TGFBR2 rs1078985 was then evaluated among an additional 4,818 subjects from the SBCGS III, as well as among 10,461 participants of 6 collaboration studies (Fig. 2). Pooled analysis of all data indicated a highly significant recessive effect (OR, 0.76; 95% CI, 0.65–0.89, $P = 8.42 \times 10^{-4}$). This was driven by the results of stage III, in which a 24% reduced risk of breast cancer (OR, 0.76; 95% CI, 0.63–0.93; P value $= 6.26 \times 10^{-3}$) was observed for minor allele homozygotes compared with major allele carriers. To evaluate a potential dual rule of the *TGFBR2 rs1078985* SNP with breast cancer risk, further analysis was conducted using data from the SBCGS by tumor stage (Table 4). Strong additive trends were seen among women with early and midstage disease, while the association with breast cancer risk was attenuated among women with advanced stage cancer (TNM stages III and IV). Heterogeneity tests, however, were not statistically significant ($P > 0.10$).

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

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

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

	Stage 0 or I		Stage II		Stage III or IV		Total[a]		
	N_{cases}	OR (95% CI)	N_{cases}	OR (95% CI)	N_{cases}	OR (95% CI)	N_{cases}	OR (95% CI)	P^d
TGFBR2 rs1078985									
AA	1,661	1.00 (reference)	2,502	1.00 (reference)	495	1.00 (reference)	5,079	1.00 (reference)	0.7738
AG	568	**0.91 (0.82–1.02)**	823	**0.87 (0.79–0.96)**	181	0.98 (0.82–1.17)	1,704	**0.89 (0.82–0.96)**	
GG	45	**0.72 (0.52–1.01)**	72	**0.76 (0.58–1.01)**	15	0.81 (0.47–1.38)	144	**0.75 (0.60–0.94)**	
P^b		**0.0198**		**0.0011**		0.5294		**0.0003**	
AA/AG	2,229	1.00 (reference)	3,325	1.00 (reference)	676	1.00 (reference)	6,783	1.00 (reference)	0.6869
GG	45	0.74 (0.53–1.03)	72	0.79 (0.60–1.05)	15	0.81 (0.48–1.39)	144	**0.77 (0.62–0.97)**	
P^c		0.0769		0.1026		0.4502		**0.0263**	

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

[a]Includes 565 women without information on tumor stage.

[b]P_{trend}.

[c]P value for recessive association.

[d]P value for heterogeneity test.

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-κB) and adjacent to 7 additional regions, the strongest signal of which was found for PU.1. NF-κB 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-κB1 and NF-κB2 has been shown in breast cancer cell lines and breast carcinomas (36), and many cancer cells show aberrant or constitutive NF-κB activation, which mediates resistance to chemotherapy and radiotherapy (36, 37). PU.1 is an erythroblast 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 *TGFB1* (*rs1800468*, *rs1800469*, *rs1800470*, and *rs1800471*), and 1 SNP in *TGFBR1* (*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 (*TGFB2 rs2799086*, *TGFB2 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*–NH_2–kinase (7). A further limitation of our study was the relatively low coverage of variants in the *TGFB3* 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.

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

Authors' Contributions

Conception and design: X. Ma, W. Lu, C.-Y. Shen, B. Ji, X.-O. Shu, H.L. Moses, W. Zheng
Development of methodology: X. Ma, C.-Y. Shen, B. Ji, W. Zheng
Acquisition of data: J. Shi, Y.-B. Xiang, Q. Cai, H. Shen, C.-Y. Shen, Z. Ren, M. Keitaro, U.S. Khoo, M. Iwasaki, Y. Zheng, Z. Hu, Y. Liu, S. Wang, H. Ito, Y. Kasuga, K.Y.-K. Chan, H. Iwata, Y.-T. Gao, X.-O. Shu,
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. Iwasaki, B. Ji, Z. Hu, Y.-T. Gao, X.-O. Shu, W. Zheng
Administrative, technical, or material support: Y.-B. Xiang, M. Keitaro, M. Iwasaki, B. Ji, W. Wang, P.-E. Wu, X. Xie, K.Y.-K. Chan, Y.-T. Gao
Study supervision: W. Lu, C.-Y. Shen, B. Ji, Y.-T. Gao, W. Zheng

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 CA148667, R01 CA064277, R01 CA090899, and R37 CA070867 to W. Zheng, and R01 CA092585 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

1. Ikushima H, Miyazono K. TGFbeta signalling: a complex web in cancer progression. Nat Rev Cancer 2010;10:415–24.
2. Moustakas A, Heldin CH. The regulation of TGFbeta signal transduction. Development 2009;136:3699–714.
3. Meulmeester E, Ten Dijke P. The dynamic roles of TGF-beta in cancer. J Pathol 2011;223:205–18.
4. Bottema RW, Kerkhof M, Reijmerink NE, Thijs C, Smit HA, van Schayck CP, et al. Gene-gene interaction in regulatory T-cell function in atopy and asthma development in childhood. J Allergy Clin Immunol 2010;126:338–46, 46 e1–10.
5. Tandon A, Tovey JC, Sharma A, Gupta R, Mohan RR. Role of transforming growth factor Beta in corneal function, biology and pathology. Curr Mol Med 2010;10:565–78.
6. Slattery ML, Herrick JS, Lundgreen A, Wolff RK. Genetic variation in the TGF-beta signaling pathway and colon and rectal cancer risk. Cancer Epidemiol Biomarkers Prev 2011;20:57–69.
7. Zheng W. Genetic polymorphisms in the transforming growth factor-beta signaling pathways and breast cancer risk and survival. Methods Mol Biol 2009;472:265–77.
8. Tang WB, Ling GH, Sun L, Liu FY. Smad anchor for receptor activation (SARA) in TGF-beta signaling. Front Biosci (Elite Ed) 2010;2:857–60.
9. Dunning AM, Ellis PD, McBride S, Kirschenlohr HL, Healey CS, Kemp PR, et al. A transforming growth factorbeta1 signal peptide variant increases secretion *in vitro* and is associated with increased incidence of invasive breast cancer. Cancer Res 2003;63:2610–5.
10. Zhang B, Beehly-Fadiel A, Long J, Zheng W. Genetic variants associated with breast-cancer risk: comprehensive research synopsis, meta-analysis, and epidemiological evidence. Lancet Oncol 2011;12:477–88.
11. Gu D, Zhuang L, Huang H, Cao P, Wang D, Tang J, et al. TGFB1 T29C polymorphism and breast cancer risk: a meta-analysis based on 10,417 cases and 11,455 controls. Breast Cancer Res Treat 2010;123:857–61.
12. Ma X, Chen C, Xiong H, Li Y. Transforming growth factorbeta1 L10P variant plays an active role on the breast cancer susceptibility in Caucasian: evidence from 10,392 cases and 11,697 controls. Breast Cancer Res Treat 2010;124:453–7.
13. Shin A, Shu XO, Cai Q, Gao YT, Zheng W. Genetic polymorphisms of the transforming growth factor-beta1 gene and breast cancer risk: a possible dual role at different cancer stages. Cancer Epidemiol Biomarkers Prev 2005;14:1567–70.
14. Scollen S, Luccarini C, Baynes C, Driver K, Humphreys MK, Garcia-Closas M, et al. TGF-{beta} signaling pathway and breast cancer susceptibility. Cancer Epidemiol Biomarkers Prev 2011;20:1112–9.
15. Woo SU, Park KH, Woo OH, Yang DS, Kim AR, Lee ES, et al. Association of a TGF-beta1 gene -509 C/T polymorphism with breast cancer risk: a meta-analysis. Breast Cancer Res Treat 2010;124:481–5.
16. Huang Y, Li B, Qian J, Xie J, Yu L. TGF-beta1 29T/C polymorphism and breast cancer risk: a meta-analysis involving 25,996 subjects. Breast Cancer Res Treat 2010;123:863–8.
17. Zheng W, Long JR, Gao YT, Li C, Zheng Y, Xiang YB, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. Nature genetics 2009;41:324–8.
18. Long JR, Cai QY, Shu XO, Qu SM, Li C, Zheng Y, et al. Identification of a functional genetic variant at 16q12.1 for breast cancer risk: results from the Asia Breast Cancer Consortium. PLoS Genetics 2010;6:e1001002.
19. Shu XO, Zheng Y, Cai H, Gu K, Chen Z, Zheng W, et al. Soy food intake and breast cancer survival. JAMA 2009;302:2437–43.
20. Cai H, Xiang YB, Qu S, Long J, Cai Q, Gao J, et al. Association of genetic polymorphisms in cell-cycle control genes and susceptibility to endometrial cancer among chinese women. Am J Epidemiol 2011;173:1263–71.
21. Zheng W, Chow WH, Yang G, Jin F, Rothman N, Blair A, et al. The Shanghai Women's Health Study: rationale, study design, and baseline characteristics. Am J Epidemiol 2005;162:1123–31.
22. Ding SL, Yu JC, Chen ST, Hsu GC, Shen CY. Genetic variation in the premature aging gene WRN: a case-control study on breast cancer susceptibility. Cancer Epidemiol Biomarkers Prev 2007;16:263–9.
23. Ding SL, Yu JC, Chen ST, Hsu GC, Kuo SJ, Lin YH, et al. Genetic variants of BLM interact with RAD51 to increase breast cancer susceptibility. Carcinogenesis 2009;30:43–9.
24. Chan KY, Liu W, Long JR, Yip SP, Chan SY, Shu XO, et al. Functional polymorphisms in the BRCA1 promoter influence transcription and are associated with decreased risk for breast cancer in Chinese women. J Med Genet 2009;46:32–9.
25. Wang Y, Hu Z, Liang J, Wang Z, Tang J, Wang S, et al. A tandem repeat of human telomerase reverse transcriptase (hTERT) and risk of breast cancer development and metastasis in Chinese women.Carcinogenesis 2008;29:1197–201.
26. Liang J, Chen P, Hu Z, Zhou X, Chen L, Li M, et al. Genetic variants in fibroblast growth factor receptor 2 (FGFR2) contribute to susceptibility of breast cancer in Chinese women. Carcinogenesis 2008;29:2341–6.
27. Hamajima N, Matsuo K, Saito T, Hirose K, Inoue M, Takezaki T, et al. Gene-environment interactions and polymorphism studies of cancer risk in the hospital-based epidemiologic research program at Aichi Cancer Center II (HERPACC-II). Asian Pac J Cancer Prev 2001;2:99–107.
28. Itoh H, Iwasaki M, Hanaoka T, Kasuga Y, Yokoyama S, Onuma H, et al. Serum organochlorines and breast cancer risk in Japanese women: a case-control study. Cancer Causes Control 2009;20:567–80.
29. Cochran WG. The combination of estimates from different experiments. Biometrics 1954;10:101–29.
30. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. Bioinformatics 2008;24:2938–9.

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

31. Gauderman W, Morrison J. QUANTO 1.1: a computer program for power and sample size calculations for genetic-epidemiology studies. Available from: http://hydra.usc.edu/gxe/. 2006.
32. Heinemeyer T, Wingender E, Reuter I, Hermjakob H, Kel AE, Kel OV, et al. Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucleic Acids Res 1998;26:362–7.
33. Dentice M, Luongo C, Elefante A, Romino R, Ambrosio R, Vitale M, et al. Transcription factor Nkx-2.5 induces sodium/iodide symporter gene expression and participates in retinoic acid- and lactation-induced transcription in mammary cells. Mol Cell Biol 2004;24:7863–77.
34. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The human genome browser at UCSC. Genome Res 2002;12:996–1006.
35. Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature 2007;447:799–816.
36. Rayet B, Gelinas C. Aberrant rel/nfkb genes and activity in human cancer. Oncogene 1999;18:6938–47.
37. Lee CH, Jeon YT, Kim SH, Song YS. NF-kappaB as a potential molecular target for cancer therapy. Biofactors 2007;29: 19–35.
38. Tschan MP, Reddy VA, Ress A, Arvidsson G, Fey MF, Torbett BE. PU.1 binding to the p53 family of tumor suppressors impairs their transcriptional activity. Oncogene 2008;27:3489–93.

Published OnlineFirst April 26, 2012; DOI: 10.1158/1055-9965.EPI-12-0118

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

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.

Cancer Epidemiol Biomarkers Prev 2012;21:1176-1184. Published OnlineFirst April 26, 2012.