HHS Public Access

Author manuscript

Cancer Causes Control. Author manuscript; available in PMC 2015 March 01.

Published in final edited form as:

Cancer Causes Control. 2014 March ; 25(3): 293–307. doi:10.1007/s10552-013-0331-9.

The influence of genetic ancestry and ethnicity on breast cancer survival associated with genetic variation in the *TGF*-β-signaling pathway: The Breast Cancer Health Disparities Study

Martha L. Slattery[1], **Abbie Lundgreen**[1], **Marianna C. Stern**[2], **Lisa Hines**[3], **Roger K. Wolff**[1], **Anna R. Giuliano**[4], **Kathy B. Baumgartner**[5,4], and **Esther M. John**[6]

[1]University of Utah, Department of Medicine, 383 Colorow, Salt Lake City, Utah 84108

[2]Department of Preventive Medicine, Keck School of Medicine of USC, Norris Comprehensive Cancer Center, University of Southern California, Los Angeles, California 90089

[3]University of Colorado at Colorado Springs, Department of Biology, 1420 Austin Bluffs Parkway, Colorado Springs, Colorado 80918

[4]Moffitt Cancer Center and Research Institute, Tampa, Florida 33612

[5]Department of Epidemiology and Population Health, School of Public Health & Information Sciences, University of Louisville, Louisville, Kentucky 40292

[6]Cancer Prevention Institute of California, Fremont, California, USA, 94538 and Division of Epidemiology, Department of Health Research and Policy and Stanford Cancer Institute, Stanford University School of Medicine, Stanford, California 94305

Abstract

The TGF-β signaling pathway regulates cellular proliferation and differentiation. We evaluated genetic variation in this pathway, its association with breast cancer survival, and survival differences by genetic ancestry and self-reported ethnicity.

The Breast Cancer Health Disparities Study includes participants from the 4-Corners Breast Cancer Study (n = 1391 cases) and the San Francisco Bay Area Breast Cancer Study (n=946 cases) who have been followed for survival. We evaluated 28 genes in the TGF-β signaling pathway using a tagSNP approach. Adaptive rank truncated product (ARTP) was used to test the gene and pathway significance by Native American (NA) ancestry and by self-reported ethnicity (non-Hispanic white (NHW) and Hispanic/NA).

Genetic variation in the TGF-β signaling pathway was associated with overall breast cancer survival (P_{ARTP} = 0.05), especially for women with low NA ancestry (P_{ARTP} =0.007) and NHW women (P_{ARTP} =0.006). *BMP2*, *BMP4*, *RUNX1*. and *TGFBR3* were significantly associated with breast cancer survival overall (P_{ARTP}=0.04, 0.02, 0.002, and 0.04 respectively). Among women with low NA ancestry associations were: *BMP4* (P_{ARTP} = 0.007), *BMP6* (P_{ARTP} = 0.001), *GDF10* (P_{ARTP}=0.05), *RUNX1* (P_{ARTP}=0.002), *SMAD1* (P_{ARTP}=0.05), and *TGFBR2* (P_{ARTP}=0.02). A

Correspondence to: Martha L. Slattery.

The authors have no conflict of interest to report.

polygenic risk model showed that women with low NA ancestry and high numbers of at-risk alleles had twice the risk of dying from breast cancer as did women with high NA ancestry.

Our data suggest that genetic variation in the TGF-β signaling pathway influences breast cancer survival. Associations were similar when the analyses were stratified by genetic ancestry or by self-reported ethnicity.

Keywords

TGFβ; *BMP*; *RUNX*; breast cancer; survival; Hispanic

The TGF-β signaling pathway regulates cellular proliferation, differentiation, apoptosis, and extracellular matrix remodeling and is involved in angiogenesis and inflammatory response [1]. The TGF-β family can be divided into two signaling pathways: (1) the bone morphogenetic proteins (BMPs) and growth and differentiations factors (GDFs); and (2) the TGFβs, activins, and myostatin. Smad proteins mediate the cellular effects of the TGF-β protein family, with BMPs and GDFs acting through Smad1, Smad5, and Smad8, whereas other members of the TGF-β family act through Smad2 and Smad3 [2]. The Smad pathway is thought to be the major TGF-β signal transduction pathway [3]. The Runt-related transcription factors (RUNX), including RUNX1, RUNX2, and RUNX3, also are involved in the TGF-β signaling pathway. Studies in *RUNX3* knockout mice have shown apoptotic defects in response to TGF-β; *RUNX2* transgenic mice have been shown to be hypersensitive to TGF-β[4]. All three *RUNX* genes have been shown to bind Smads [5–7], thus further influencing the TGF-β signaling pathway.

It is biologically plausible that alterations of the TGF-β signaling pathway may influence breast cancer prognosis given its regulatory role in angiogenesis, inflammation, and tumor growth. Although in early stages of cancer TGF-β may exhibit tumor suppressive effects, in later stages of breast cancer it appears to be pro-tumorigenic by stimulating invasion[8]. Moreover, high serum levels and high levels of expression of TGF-β and its receptors have been linked to breast cancer prognosis [9] and presence of phosphorylated-Smad2 has been associated with positive node status [10]. A study by deKruijf and colleagues [8] showed that high levels of TGF-β receptor expression in conjunction with Smad expression conferred an unfavorable prognosis after breast cancer diagnosis. The RUNX transcription factors also have been proposed as influencing survival, with RUNX2 being highly expressed in cell lines that are metastatic to bone. Because of BMPs' role in bone formation, they have been examined for their involvement in metastasis to the bone after breast cancer diagnosis and disease progression [11]. Additionally, BMPs have been associated with estrogen-induced proliferation of breast cancer cells [12]. One study has shown that BMP-Smad activation is involved in the progression of estrogen receptor positive (ER+) breast cancers specifically [13].

Incidence and mortality rates of breast cancer have been shown to vary by race and ethnicity [14, 15]. Among women in the Southwestern United States, those who are Native American (NA) have breast cancer incidence rates that are roughly one quarter to one third of those observed for women who are classified as non-Hispanic white (NHW). Hispanic women

have breast cancer incidence rates between women who are NA and those of European descent. Differences in breast cancer risk factors, such as parity, do not account for these differences [16]. Exploration of differences in disease rates can utilize genetic ancestry under the assumption that biological differences stemming from genetic factors influence the carcinogenic process. On the other hand, consideration of self-reported race/ethnicity as a stratification tool can focus on unidentified cultural factors that may contribute to disparities in cancer rates and might be common across the population, irrespective of underlying genetic differences. We considered these two methods of stratification to help understand the biological and cultural contributions to breast cancer survival.

In this study we evaluated the associations between genetic variability in the TGF-β signaling pathway and survival after diagnosis with breast cancer. We evaluated *TGFβ1* and its receptors, *SMAD* genes, *BMP* genes and their receptors, *RUNX* genes, activins and their receptors (*ACVR1*, *ACVR2A*, *ACVR2B*, and *ACVRL1*), and GDFs (*GDF10* and myostatin). We evaluated associations in an admixed population of NHW and Hispanic and Native American (NA) women, giving us the capability to examine associations by genetic ancestry as well as by self-reported ethnicity. We also evaluated survival after diagnosis with breast cancer by estrogen receptor (ER) and progesterone receptor (PR) tumor type.

Methods

This analysis from the Breast Cancer Health Disparities Study includes participants with information on survival from two population-based case-control studies, the 4-Corners Breast Cancer Study (4-CBCS) that included women from Arizona, Colorado, New Mexico and Utah, and the San Francisco Bay Area Breast Cancer Study (SFBCS) [14] who completed an in-person interview and who had a blood or mouthwash sample available for DNA extraction. In the 4-CBCS, participants were between 25 and 79 years of age with a histologically confirmed diagnosis of first primary invasive breast cancer (n=1391) between October 1999 and May 2004[17]. The SFBCS included women aged 35 to 79 years from the San Francisco Bay Area diagnosed with a first primary histologically confirmed invasive breast cancer (n= 946) between April 1997 and April 2002 [18, 19]. All participants signed informed written consent prior to participation; this study was approved by the Institutional Review Boards for Human Subjects at the University of Utah and the Cancer Prevention Institute of California.

Data Harmonization

Data used as adjustment variables were harmonized across the study centers and study-specific questionnaires as previously described [14]. Women were asked to self-report race and, with the option to report multiple categories if appropriate, i.e., NHW and Hispanic, Hispanic and NA. Women who reported any Hispanic ethnicity or NA were broadly classified as being Hispanic. Women also were classified as either pre-menopausal or post-menopausal based on responses to questions on menstrual history. Women were considered as post-menopausal if they reported either a natural menopause or if they reported taking hormone therapy (HT) and were still having periods or were at or above the 95th percentile

of age for those who reported having a natural menopause (i.e. ≥ 12 months since their last period); other were classified as pre-menopausal.

Genetic Data

DNA was extracted from either whole blood or mouthwash samples. Whole Genome Amplification (WGA) was applied to the mouthwash-derived DNA samples prior to genotyping. Quality control results were comparable for the two DNA sources. A tagSNP approach was used to characterize variation across candidate genes. TagSNPs were selected using the following parameters: linkage disequilibrium (LD) blocks were defined using a Caucasian LD map and an r^2=0.8; minor allele frequency (MAF) >0.1; range of −1500 bps from the initiation codon to +1500 bps from the termination codon; and 1 SNP/LD bin. Additionally, 104 Ancestry Informative Markers (AIMs) were used to distinguish European and NA ancestry in the study population [14]. All markers were genotyped using a multiplexed bead array assay format based on GoldenGate chemistry (Illumina, San Diego, California). In the current analysis, we evaluated the following genes in the TGF-β signaling pathway: *ACVR1* (16 SNPs), *ACVR2A* (6 SNPs), *ACVR2B* (3 SNPs), *ACVRL1* (4 SNPs), *BMP1* (10 SNPs), *BMP2* (6 SNPs), *BMP4* (4 SNPs), *BMP6* (23 SNPs), *BMP7* (24 SNPs), *BMPR1A* (9 SNPs), *BMPR1B* (18 SNPs), *BMPR2* (8 SNPs), *GDF10* (6 SNPs), *MSTN* (1 SNP), *RUNX1* (8 SNPs), *RUNX2* (17 SNPs), and *RUNX3* (8 SNPs), *SMAD1* (4 SNPs), *SMAD2* (5 SNPs), *SMAD3* (40 SNPs), *SMAD4* (2 SNPs), *SMAD6* (1 SNP), *SMAD7* (10 SNPs), *TGFβ1* (2 SNPs), *TGFβRI* (5 SNPs), *TGFβ2* (1 SNP), *TGFβR2* (1 SNP), and *TGFβR3* (5 SNPs). Supplemental Table 1 (online) details the genes and SNPs evaluated. Genotyping was completed for 933 women from the 4-CBCS who self-identified as NHW, 412 Hispanic, 8 American Indian, 14 NHW/Hispanic, 10 NHW/American Indian, 10 Hispanic/American Indian, and 4 NHW/Hispanic/American Indian and for 252 women from the SFBCS self-reported being NHW and 694 who reported being Hispanic.

Tumor Characteristics and Survival

Data on survival were available from local cancer registries that provided information on date of death or last follow-up (month and year), cause of death, and stage of disease at time of diagnosis. Survival (in months) was calculated as the difference between diagnosis date and date of death or last follow-up. Survival information was complete for each study through May of 2012. Information on cause of death was provided and was classified as breast cancer if either the primary or contributing cause of death was noted as breast cancer. SEER summary stage data were classified as local, regional, or distant. ER and PR tumor status was provided by local tumor registries which included the Utah Cancer Registry, the New Mexico Cancer Registry, the Arizona Cancer Registry, the Colorado Cancer Registry, and the Northern California Cancer Registry.

Statistical Methods

The program STRUCTURE was used to compute individual ancestry for each study participant assuming two founding populations [20, 21]. A three-founding population model was assessed, but did not fit the population structure with the same level of repeatability and correlation among runs as the two-founding population model. Participants were classified

by level of percent NA ancestry. Assessment across categories of ancestry was done using cut-points based on the distribution of genetic ancestry in the study control population (2597 Hispanic/NA and 1586 NHW controls) with the goal of creating distinct ancestry groups that had sufficient power to assess associations, especially when looking at menopausal status within admixture groups. Two strata of ≤28% and >28% were used to evaluate associations by level of NA ancestry. Genetic ancestry was used as a continuous variable when included in the models to adjust for possible confounding.

Associations between SNPs and risk of dying of breast cancer were evaluated using Cox proportional hazards models to obtain multivariate hazard ratios (HR) and 95% confidence intervals (CI) for all women and for women stratified by genetic ancestry, self-reported ethnicity (either NHW or Hispanic/NA) and by ER/PR status using SAS version 9.3 (SAS Institute, Cary, NC). We were not able to evaluate the category of ER−/PR+ tumors because there were too few women available for analysis. Individuals were censored when they died of causes other than breast cancer or were lost to follow-up. All SNPs were evaluated as a co-dominant model and if initial analysis suggested too few homozygous variant carriers a dominant model was used. However, in some instances the recessive model clearly fit the data and was used to calculate risk estimates. In addition to the minimal adjustments for age, study center, body mass index (BMI) in the referent year, parity, and genetic ancestry models also were adjusted for SEER summary stage. ER/PR tumor status was not adjusted in these models since these markers were not associated with tumor status and survival. Further adjustment for self-reported ethnicity did not alter associations (data not shown in tables). SNP p values were based on 1 degree of freedom (df) Wald chi-square tests, comparing the homozygote variants to the common genotypes when the co-dominant model was indicated. Interactions between genetic variants and genetic ancestry, self-reported ethnicity, and ER/PR status related to survival were assessed using p values from one and two degree of freedom Wald chi-square tests respectively.

We used the adaptive rank truncated product (ARTP) method (http://dceg.cancer.gov/bb/tools/artp) that utilizes a highly efficient permutation algorithm to determine the significance of each gene and of the TGF-β signaling pathway with survival after breast cancer diagnosis [22, 23]. We permuted the breast cancer survival outcome 10,000 times in R version 3.0.1 (R Foundation for Statistical Computing, Vienna, Austria). SNP associations were assessed among the observed and permuted data in R using p values from likelihood-ratio tests comparing full Cox proportional hazards models adjusted for age, BMI in referent year, disease stage, and genetic ancestry to reduced models excluding the SNP term. We report both gene and pathway p values (P_{ARTP}) based on five truncation points.

A polygenic risk summary score was created to estimate the risk of mortality associated with this pathway. SNPs included in the summary score were restricted to those located on genes with a P_{ARTP} of 0.10 or less for all women and/or for the specific ancestry strata; only those SNPs that contributed to the best fitting ARTP model were selected. The score for each SNP was based on the inheritance model with the co-dominant or additive model having a score of zero, one, or two based on the number of high-risk alleles, whereas scores of zero or two were assigned for the dominant and recessive models. At-risk alleles were assigned based on specific ancestry group risk. Risk estimates, based on varying numbers of at-risk alleles,

were estimated from proportional hazards models, taking into account confounding variables as described above. Cut-points were selected to maintain at least 10 deaths in each category. Women missing genotype data on two or more SNPs were excluded.

Results

Fifty percent of all deaths among NHW and 57% of all deaths among Hispanic/NA women were from breast cancer (Table 1). Of the 1152 women who self-reported being Hispanic or NA, 10.4% had NA ancestry levels at or below 18%, compared to 99% of NHW women. The majority of women who self-reported being Hispanic or NA had over 28% NA ancestry. Women were more likely to die if they were older at the time of diagnosis, had ER-/PR- tumors, or were diagnosed when their tumor was at a more advanced stage. There was no statistically significant difference in survival among women who self-reported being NHW or Hispanic/NA or when categorized by level of NA ancestry.

When considering all women, genetic variation across the entire TGF-β-signaling pathway was significantly associated with breast cancer mortality (Pathway $P_{ARTP} = 0.045$). When considering individual genes, we observed statistically significant associations for *BMP2, BMP4, RUNX1*, and *TGFBR3* (Gene $P_{ARTP} = 0.04$, 0.02, 0.002, and 0.04 respectively) (Table 2). There were a few SNPs in other genes that also showed statistically significant associations; however, the corresponding genes did not show overall association of statistical significance as summarized by ARTP. One SNP was significantly associated with breast cancer survival for *BMP1* (rs13257482), *BMP2* (rs7270163), *BMP6* (rs270413), *BMPR1B* (rs10049681), *RUNX2* (rs598953), *SMAD2* (rs1792658), *TBFβR1* (rs6478974) *andTBFβR3* (rs6678564) at the 0.05 level or less, two SNPs were associated with breast cancer survival for *GDF10* (rs7093975 and rs1902724) and *RUNX1* (rs1474479 and rs1883066), and three SNPs were associated for *BMP4* (rs17563, rs2761887, and rs4898820). There were no significant differences in mortality by ER/PR status for any of the genes evaluated (see online supplement table 2)

Women with low NA ancestry or who reported being NHW had a pathway P_{ARTP} of 0.007 and 0.006, respectively, while women with higher NA ancestry or who reported having any Hispanic/NA ethnicity pathway P_{ARTP} values of 0.18 and 0.51, respectively. A comparison of gene and SNP associations by genetic ancestry and by self-reported ethnicity is shown in Table 3. Several genes in the pathway were associated with breast cancer survival with findings being similar when comparing women with low NA ancestry and NHW women or when comparing those women with higher NA ancestry levels and those who self-reported being Hispanic or NA. Among women with low NA ancestry, P_{ARTP} values were significant at or below the 0.05 level for six of 12 pathway genes (Table 3): *BMP4, BMP6, GDF10, RUNX1, SMAD1*, and *TGFBR2*. Similar associations were observed in *BMP4, BMP6, GDF10, RUNX1*, and *SMAD1* in NHW women, although *TFGBR2* was not statistically significant. In general, individual SNP p values tended to be slightly stronger for NHW women and those with higher NA ancestry.

Evaluation of associations based on p values is influenced by sample size. In our study evaluation by self-reported ethnicity resulted in a slightly larger sample for those with low

NA ancestry vs. NHWs and women who self-reported being Hispanic /NA vs. women with high NA ancestry. This is because roughly 19% of those who self-reported being Hispanic or NA were in the lower NA ancestry group. Thus, we compared the associations for SNPs that were different either by genetic ancestry or by self-reported ethnicity (Table 4 shows all SNPs where the p for interaction was <0.05). For the most part, associations were strikingly similar between groups, especially when looking at risk estimates within strata.

The polygenic risk model showed increasing risk of dying from breast cancer with increasing number of at-risk alleles. This increased risk was observed for each NA ancestry group (Figure 1) although women with the highest risk of dying associated with increasing number of at-risk alleles were those with the lowest NA ancestry. It is of interest to note that among women in the lowest category of at-risk alleles, 4.5% of those with low NA ancestry and 5.3% of those with higher NA ancestry died from breast cancer. This is in contrast to 27% of women with low NA ancestry and 19.8% of those with higher NA ancestry who were in the highest category of at-risk alleles.

Discussion

In this study, we have taken a comprehensive gene and pathway approach to assess the association between genetic variations in the TGF-β signaling pathway and survival after breast cancer diagnosis. Examination of associations by both genetic ancestry and self-reported ethnicity provided insight into the potential biological basis for differences in associations. Stronger associations with the pathway were observed for those women with lower NA ancestry ($P_{ARTP} = 0.007$) compared with women with higher NA ancestry ($P_{ARTP} = 0.18$). Associations with genes that were most important also varied slightly by level of NA ancestry. Among those with low NA ancestry the most significant associations were with *BMPs*, *TGFB1* and its receptors, and *RUNX1*. *SMAD* and Activin genes had little influence on mortality overall or in specific NA ancestry groups.

Bone is the most common metastatic site for breast cancer [24]. BMPs are key factors in bone formation and thought to play a major role in bone metastasis[25]. However, the roles of BMPs are complex, with studies showing both growth promoting and inhibitory effects [26]. BMP2 has been associated with both decreased cell proliferation and promotion of invasiveness in MCF-7 breast cancer cells [27]. BMP2, BMP4, BMP6, and BMP7 have been shown to induce angiogenesis [24]. BMP7 has been shown to promote cell migration and invasion [28] and BMP7 protein expression has been associated with accelerated bone metastasis [28]. BMPRIA has been shown to prolong survival in mice by reducing invasiveness and bone metastasis [29], whereas BMPRIB has been associated with high tumor grade and poor prognosis [30]. Thus, there is support for the hypothesis that BMPs and their receptors influence survival after diagnosis with breast cancer. Genes that seemed most important with regard to survival in our study were *BMP4* and *GDF10* for all women and *BMP6* among women with low NA ancestry. *BMP4* has been suggested as a promoter of invasive behavior although it has been shown to reduce migration and invasion [31]. GDF10 expression, through its interaction with RUNX2, has been associated with lung cancer survival [32], although studies in breast cancer are lacking. BMP6 has been associated with estrogen induced breast cancer cell proliferation [33, 34] and has been

shown to inhibit apoptosis in breast cancer cells [35]. Associations between specific BMP-related SNPs and breast cancer prognosis have not been reported. Our data suggest the importance of genetic variation in these genes in survival after diagnosis with breast cancer, although little is known about functionality of specific SNPs associated with risk.

Genetic variation in *RUNX* 1 was highly associated with breast cancer survival overall and especially among women with low NA ancestry. Most studies to date have reported on RUNX2 which promotes cell migration and invasive properties leading to metastatic bone disease [36–38]. Inactivation of RUNX3 also has been associated with breast cancer progression [39]. However, a genome-wide association analysis of ER alpha showed that *RUNX1* is involved in ER regulation of genes [40]. *RUNX1* also has been shown to be mutated in breast cancers [41]. For the *RUNX1* SNPs that were associated with survival in our study, having a variant allele was for the most part associated with increased risk (in one instance it was protective). This could indicate that lower levels of *RUNX1* could subsequently impact regulation of other important genes associated with the ER or work through multiple mechanisms. While the exact mechanism can only be speculated, our finding merits replication in other studies.

TGF-β has been shown to have both tumor promoting and tumor inhibitory action [42]. Studies have shown that during tumor progression the tumor inhibitory effects are lost, whereas the tumor promoting effects remain intact [42]. Expression of TGF-βR2 has been associated with longer survival time among women with ER- tumors [42]. Some studies have shown that TGF-β1 protein levels are associated with shorter disease-free survival, especially among those with node-negative tumors[9], whereas others have reported associations between TGFβ1-expressing tumors and a greater likelihood of breast cancer recurrence[43, 44]. A study by Mu [45] found significantly higher TGF-β1 expression with the TT genotype of rs1982073, with an accompanying two-fold increase in risk of breast cancer death; the study by Zheng supported these findings [46]. We found a similar association with rs1800469 for women with higher levels of NA ancestry, but not for women with low NA ancestry. It is not clear why the association was seen only among women with higher NA ancestry. Because these women are more likely to have ER- tumors it is possible that they are more susceptible to the effects of TGF-β. However, we did not see differences in its association with survival according to ER status. Our findings suggest that among women with higher NA ancestry, alterations in the TGF-β signaling pathway might be more relevant for breast cancer progression, perhaps due to other genetic alterations more likely to be present in the NA ancestral background, or from other unidentified non-genetic risk factors that correlate with high NA ancestry and are associated with survival.

Many associations with breast cancer survival were observed, regardless of NA ancestry or ethnicity. However, as mentioned above, some associations differed by genetic ancestry with similar differences by self-reported ethnicity. While the reasons for these differences are not clear, it is known that at the population level breast cancer incidence and mortality rates vary by NA ancestry [14, 15], suggesting a possible biological underpinning. In general the pathway, gene, and SNP associations were more significant for women who reported being NHW or who were classified as having low NA ancestry. However, in some instances associations were different between self-reported race/ethnicity and NA ancestry

group. While unidentified factors associated with culture could contribute to these observations, they likely stem from genetic ancestry modifying the risk of breast cancer death associated with these genes. A logical explanation for these differences is misclassification stemming from arbitrary cut-points. Women who reported being only NHW have the lowest level of genetic ancestry which is slightly lower than that of the low NA ancestry group. Women within the highest NA ancestry group have been determined by AIMs to have the most NA ancestry. Half of the women who self-reported being Hispanic or NA in the low admixture group had admixture comparable to the NHW, however the other half had levels higher than most of the NHW but lower than the majority of women who reported being Hispanic or NA. When looking at genetic ancestry by case and control status, cases who self-reported being Hispanic or NA were significantly (p=0.006) more likely to be in the very low of the ancestry group than were controls. Our data suggest that evaluating the most extreme ends of the distribution of the data separates the effects associated with genetic ancestry. Although utilization of genetic ancestry markers to classify individuals is based on a set of markers measured in everyone, the cut-points selected were arbitrary and chosen to maximize power.

Misclassification by self-reported ethnicity is also possible given that some people report multiple races and may report ethnicity differentially for a myriad of reasons, especially individuals with mixed ancestry. Follow-up with some participants in the New Mexico center resulted in some people actually changing their self-reported ethnicity (personnel communication K. Baumgartner). As stated earlier, our data suggest that cases were more likely to have lower NA ancestry than controls, which corresponds to disease trends that show that women of European ancestry have higher breast cancer incidence rates than those with more NA ancestry. We have adjusted associations for known factors that could influence risk of breast cancer death and differ by race/ethnicity, i.e. BMI, although we acknowledge that other unidentified factors could be present. Given that our research is driven by the observed differences in breast cancer risk by NA ancestry, we believe that results obtained using genetic ancestry maximizes power, accurately discriminate the biological influences of NA ancestry, and are generally reinforced by our assessment associations stratified by ethnicity.

While the study has several strengths, including a large population-based sample of both NHW and Hispanic/NA women to evaluate associations with survival within a targeted candidate pathway, there are also limitations. We focused the analysis on genes within the pathway rather than individual tagSNPs. However, a logical next step is to examine the tagSNPs we identified as being important and the pathway in more depth in order to identify potentially functional SNPs that could be targeted for therapy. Additionally, since tagSNPs were not based on NA populations, some important SNPs could have been missed that influence risk among those with greater NA ancestry; further evaluation of tagSNPs in NA populations is warranted. The study included women who participated in the original case-control studies and we lack the capability to evaluate these associations among non-participants. Our analysis my not have included women with more advanced disease who may have been too sick to participate or were deceased before being contacted for the study. It is therefore possible that associations would have been stronger if women with more advanced disease had been included in the analysis. Additionally, there are limitations to our

polygenic risk score. Although we utilized ARTP permutated data to identify important genes and SNPs, another dataset for validation would be desirable. It should also be recognized that the risk estimates presented could be inflated, although setting an ARTP p value of <0.10 helps mitigate this effect. Detailed treatment data were not available, although we believe that adjustment for disease stage helps to overcome this limitation. Despite these limitations, we believe that the main messages are valid. First, risk increases with increasing “at-risk” alleles. We found that 5% of women in the lowest category of at-risk alleles died compared with almost 20% of women in the highest category of at-risk alleles. Second, the risk of dying for women with low NA ancestry in the highest category of at-risk alleles is almost twice that observed for women with higher NA ancestry.

In conclusion, our data suggest that genetic variation in the TGF-β signaling pathway influences survival after breast cancer. Associations were observed for both NHW and Hispanic women, although several genes were more strongly associated among women with low NA ancestry. Our data suggest that stratification that is able to best separate the effects of genetic ancestry is the most robust when evaluating genetic risk. Future studies that confirm these findings and determine functionality of SNPs within the pathway will enhance our understanding of the TGF-β signaling pathway and hopefully help identify potential drug targets to improve breast cancer prognosis.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would also like to acknowledge the contributions of the following individuals to the study: Dr. Juan Pablo Lewinger for his guidance on statistical issues; Sandra Edwards for data harmonization oversight; Jennifer Herrick of data management and harmonization; Erica Wolff and Michael Hoffman for laboratory support; Jocelyn Koo for data management for the San Francisco Bay Area Breast Cancer Study; Dr. Tim Byers for his contribution to the 4-Corners Breast Cancer Study; and Dr. Josh Galanter for assistance in selection of AIMs markers.

Grant Support: The Breast Cancer Health Disparities Study was funded by grant CA14002 from the National Cancer Institute to Dr. Slattery. The San Francisco Bay Area Breast Cancer Study was supported by grants CA63446 and CA77305 from the National Cancer Institute, grant DAMD17-96-1-6071 from the U.S. Department of Defense, and grant 7PB-0068 from the California Breast Cancer Research Program. The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000036C awarded to the Cancer Prevention Institute of California; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement #1U58 DP000807-01 awarded to the Public Health Institute. The 4-Corners Breast Cancer Study was funded by grants CA078682, CA078762, CA078552, and CA078802 from the National Cancer Institute. The research also was supported by the Utah Cancer Registry, which is funded by contract N01-PC-67000 from the National Cancer Institute, with additional support from the State of Utah Department of Health, the New Mexico Tumor Registry, and the Arizona and Colorado cancer registries, funded by the Centers for Disease Control and Prevention National Program of Cancer Registries and additional state support. The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official view of the National Cancer Institute or endorsement by the State of California Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors. The Mexico Breast Cancer Study was funded by Consejo Nacional de Ciencia y Tecnología (CONACyT) (SALUD-2002-C01-7462).

References

1. Gordon KJ, Blobe GC. Role of transforming growth factor-beta superfamily signaling pathways in human disease. Biochim Biophys Acta. 2008; 1782(4):197–228. [PubMed: 18313409]
2. Wharton K, Derynck R. TGFbeta family signaling: novel insights in development and disease. Development. 2009; 136(22):3691–3697. [PubMed: 19855012]
3. Buck MB, Knabbe C. TGF-beta signaling in breast cancer. Annals of the New York Academy of Sciences. 2006; 1089:119–126. [PubMed: 17261761]
4. Blyth K, Cameron ER, Neil JC. The RUNX genes: gain or loss of function in cancer. Nat Rev Cancer. 2005; 5(5):376–387. [PubMed: 15864279]
5. Lee KS, Hong SH, Bae SC. Both the Smad and p38 MAPK pathways play a crucial role in Runx2 expression following induction by transforming growth factor-beta and bone morphogenetic protein. Oncogene. 2002; 21(47):7156–7163. [PubMed: 12370805]
6. Cameron ER, Blyth K, Hanlon L, Kilbey A, Mackay N, Stewart M, Terry A, Vaillant F, Wotton S, Neil JC. The Runx genes as dominant oncogenes. Blood Cells Mol Dis. 2003; 30(2):194–200. [PubMed: 12732183]
7. Ito Y, Osato M, Ito K. RUNX and cancer. Ann Acad Med Singapore. 2003; 32(5 Suppl):S6–7. [PubMed: 14968716]
8. de Kruijf EM, Dekker TJ, Hawinkels LJ, Putter H, Smit VT, Kroep JR, Kuppen PJ, van de Velde CJ, Ten Dijke P, Tollenaar RA, et al. The prognostic role of TGF-beta signaling pathway in breast cancer patients. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2013; 24(2):384–390. [PubMed: 23022998]
9. Desruisseau S, Palmari J, Giusti C, Romain S, Martin PM, Berthois Y. Determination of TGFbeta1 protein level in human primary breast cancers and its relationship with survival. British journal of cancer. 2006; 94(2):239–246. [PubMed: 16404434]
10. Figueroa JD, Flanders KC, Garcia-Closas M, Anderson WF, Yang XR, Matsuno RK, Duggan MA, Pfeiffer RM, Ooshima A, Cornelison R, et al. Expression of TGF-beta signaling factors in invasive breast cancers: relationships with age at diagnosis and tumor characteristics. Breast Cancer Research and Treatment. 2010; 121(3):727–735. [PubMed: 19937272]
11. Jacobs EJ, Feigelson HS, Bain EB, Brady KA, Rodriguez C, Stevens VL, Patel AV, Thun MJ, Calle EE. Polymorphisms in the vascular endothelial growth factor gene and breast cancer in the Cancer Prevention Study II cohort. Breast cancer research : BCR. 2006; 8(2):R22. [PubMed: 16613616]
12. Schneider BP, Winer EP, Foulkes WD, Garber J, Perou CM, Richardson A, Sledge GW, Carey LA. Triple-negative breast cancer: risk factors to potential targets. Clinical cancer research : an official journal of the American Association for Cancer Research. 2008; 14(24):8010–8018. [PubMed: 19088017]
13. Kesler KA, Rieger KM, Hammoud ZT, Kruter LE, Perkins SM, Turrentine MW, Schneider BP, Einhorn LH, Brown JW. A 25-year single institution experience with surgery for primary mediastinal nonseminomatous germ cell tumors. The Annals of thoracic surgery. 2008; 85(2):371–378. [PubMed: 18222228]
14. Slattery ML, John EM, Torres-Mejia G, Lundgreen A, Herrick JS, Baumgartner KB, Hines LM, Stern MC, Wolff RK. Genetic variation in genes involved in hormones, inflammation and energetic factors and breast cancer risk in an admixed population. Carcinogenesis. 2012; 33(8): 1512–1521. [PubMed: 22562547]
15. Swan J, Edwards BK. Cancer rates among American Indians and Alaska Natives: is there a national perspective. Cancer. 2003; 98(6):1262–1272. [PubMed: 12973851]
16. Gilliland FD, Hunt WC, Baumgartner KB, Crumley D, Nicholson CS, Fetherolf J, Samet JM. Reproductive risk factors for breast cancer in Hispanic and non-Hispanic white women: the New Mexico Women's Health Study. American journal of epidemiology. 1998; 148(7):683–692. [PubMed: 9778175]
17. Slattery ML, Sweeney C, Edwards S, Herrick J, Baumgartner K, Wolff R, Murtaugh M, Baumgartner R, Giuliano A, Byers T. Body size, weight change, fat distribution and breast cancer

risk in Hispanic and non-Hispanic white women. Breast Cancer Res Treat. 2007; 102(1):85–101. [PubMed: 17080310]

18. John EM, Horn-Ross PL, Koo J. Lifetime physical activity and breast cancer risk in a multiethnic population: the San Francisco Bay area breast cancer study. Cancer Epidemiol Biomarkers Prev. 2003; 12(11 Pt 1):1143–1152. [PubMed: 14652273]
19. John EM, Phipps AI, Davis A, Koo J. Migration history, acculturation, and breast cancer risk in Hispanic women. Cancer Epidemiol Biomarkers Prev. 2005; 14(12):2905–2913. [PubMed: 16365008]
20. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics. 2003; 164(4):1567–1587. [PubMed: 12930761]
21. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. Genetics. 2000; 155(2):945–959. [PubMed: 10835412]
22. Yu K, Li Q, Bergen AW, Pfeiffer RM, Rosenberg PS, Caporaso N, Kraft P, Chatterjee N. Pathway analysis by adaptive combination of P-values. Genetic epidemiology. 2009; 33(8):700–709. [PubMed: 19333968]
23. Kai Yu, OL.; Wheeler, William. ARTP Gene and Pathway p-values computed using the Adaptive Rank Truncated Product. 2.0.0 edn. R package; 2011.
24. Ye L, Bokobza SM, Jiang WG. Bone morphogenetic proteins in development and progression of breast cancer and therapeutic potential (review). International journal of molecular medicine. 2009; 24(5):591–597. [PubMed: 19787192]
25. Ye L, Lewis-Russell JM, Kyanaston HG, Jiang WG. Bone morphogenetic proteins and their receptor signaling in prostate cancer. Histology and histopathology. 2007; 22(10):1129–1147. [PubMed: 17616940]
26. Kim M, Choe S. BMPs and their clinical potentials. BMB Rep. 2011; 44(10):619–634. [PubMed: 22026995]
27. Clement JH, Raida M, Sanger J, Bicknell R, Liu J, Naumann A, Geyer A, Waldau A, Hortschansky P, Schmidt A, et al. Bone morphogenetic protein 2 (BMP-2) induces in vitro invasion and in vivo hormone independent growth of breast carcinoma cells. International journal of oncology. 2005; 27(2):401–407. [PubMed: 16010421]
28. Alarmo EL, Parssinen J, Ketolainen JM, Savinainen K, Karhu R, Kallioniemi A. BMP7 influences proliferation, migration, and invasion of breast cancer cells. Cancer Letters. 2009; 275(1):35–43. [PubMed: 18980801]
29. Katsuno Y, Hanyu A, Kanda H, Ishikawa Y, Akiyama F, Iwase T, Ogata E, Ehata S, Miyazono K, Imamura T. Bone morphogenetic protein signaling enhances invasion and bone metastasis of breast cancer cells through Smad pathway. Oncogene. 2008; 27(49):6322–6333. [PubMed: 18663362]
30. Bokobza SM, Ye L, Kynaston HE, Mansel RE, Jiang WG. Reduced expression of BMPR-IB correlates with poor prognosis and increased proliferation of breast cancer cells. Cancer genomics & proteomics. 2009; 6(2):101–108. [PubMed: 19451094]
31. Alarmo EL, Kallioniemi A. Bone morphogenetic proteins in breast cancer: dual role in tumourigenesis? Endocrine-related cancer. 2010; 17(2):R123–139. [PubMed: 20335308]
32. Tandon M, Gokul K, Ali SA, Chen Z, Lian J, Stein GS, Pratap J. Runx2 mediates epigenetic silencing of the bone morphogenetic protein-3B (BMP-3B/GDF10) in lung cancer cells. Molecular cancer. 2012; 11:27. [PubMed: 22537242]
33. Takahashi M, Otsuka F, Miyoshi T, Otani H, Goto J, Yamashita M, Ogura T, Makino H, Doihara H. Bone morphogenetic protein 6 (BMP6) and BMP7 inhibit estrogen-induced proliferation of breast cancer cells by suppressing p38 mitogen-activated protein kinase activation. The Journal of endocrinology. 2008; 199(3):445–455. [PubMed: 18780779]
34. Zhang M, Yan JD, Zhang L, Wang Q, Lu SJ, Zhang J, Zhu TH. Activation of bone morphogenetic protein-6 gene transcription in MCF-7 cells by estrogen. Chinese medical journal. 2005; 118(19): 1629–1636. [PubMed: 16232348]

35. Du J, Yang S, Wang Z, Zhai C, Yuan W, Lei R, Zhang J, Zhu T. Bone morphogenetic protein 6 inhibit stress-induced breast cancer cells apoptosis via both Smad and p38 pathways. Journal of cellular biochemistry. 2008; 103(5):1584–1597. [PubMed: 17879955]

36. Pratap J, Lian JB, Stein GS. Metastatic bone disease: Role of transcription factors and future targets. Bone.

37. Pratap J, Lian JB, Javed A, Barnes GL, van Wijnen AJ, Stein JL, Stein GS. Regulatory roles of Runx2 in metastatic tumor and cancer cell interactions with bone. Cancer Metastasis Rev. 2006; 25(4):589–600. [PubMed: 17165130]

38. Mendoza-Villanueva D, Deng W, Lopez-Camacho C, Shore P. The Runx transcriptional co-activator, CBFbeta, is essential for invasion of breast cancer cells. Molecular cancer. 9:171. [PubMed: 20591170]

39. Subramaniam MM, Chan JY, Soong R, Ito K, Ito Y, Yeoh KG, Salto-Tellez M, Putti TC. RUNX3 inactivation by frequent promoter hypermethylation and protein mislocalization constitute an early event in breast cancer progression. Breast Cancer Res Treat. 2009; 113(1):113–121. [PubMed: 18256927]

40. Stender JD, Kim K, Charn TH, Komm B, Chang KC, Kraus WL, Benner C, Glass CK, Katzenellenbogen BS. Genome-wide analysis of estrogen receptor alpha DNA binding and tethering mechanisms identifies Runx1 as a novel tethering factor in receptor-mediated transcriptional activation. Mol Cell Biol. 30(16):3943–3955. [PubMed: 20547749]

41. Ellis MJ, Ding L, Shen D, Luo J, Suman VJ, Wallis JW, Van Tine BA, Hoog J, Goiffon RJ, Goldstein TC, et al. Whole-genome analysis informs breast cancer response to aromatase inhibition. Nature. 2012; 486(7403):353–360. [PubMed: 22722193]

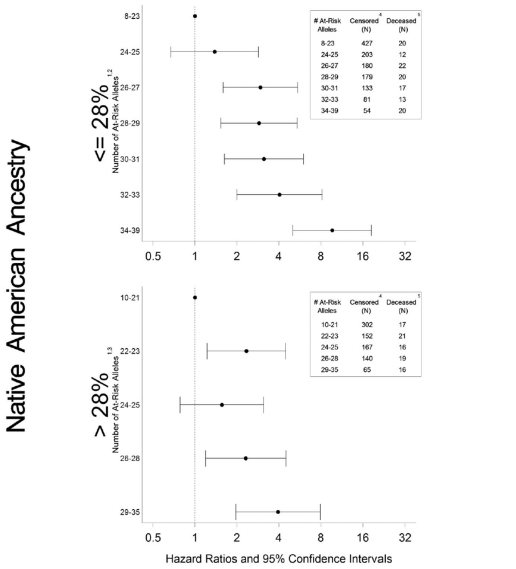
42. Buck MB, Fritz P, Dippon J, Zugmaier G, Knabbe C. Prognostic significance of transforming growth factor beta receptor II in estrogen receptor-negative breast cancer patients. Clinical cancer research : an official journal of the American Association for Cancer Research. 2004; 10(2):491–498. [PubMed: 14760070]

43. Koumoundourou D, Kassimatis T, Zolota V, Tzorakoeleftherakis E, Ravazoula P, Vassiliou V, Kardamakis D, Varakis J. Prognostic significance of TGFbeta-1 and pSmad2/3 in breast cancer patients with T1-2,N0 tumours. Anticancer research. 2007; 27(4C):2613–2620. [PubMed: 17695423]

44. Ivanovic V, Demajo M, Krtolica K, Krajnovic M, Konstantinovic M, Baltic V, Prtenjak G, Stojiljkovic B, Breberina M, Neskovic-Konstantinovic Z, et al. Elevated plasma TGF-beta1 levels correlate with decreased survival of metastatic breast cancer patients. Clinica chimica acta; international journal of clinical chemistry. 2006; 371(1–2):191–193.

45. Mu L, Katsaros D, Lu L, Preti M, Durando A, Arisio R, Yu H. TGF-beta1 genotype and phenotype in breast cancer and their associations with IGFs and patient survival. British journal of cancer. 2008; 99(8):1357–1363. [PubMed: 18827819]

46. Zheng W. Genetic polymorphisms in the transforming growth factor-beta signaling pathways and breast cancer risk and survival. Methods in molecular biology. 2009; 472:265–277. [PubMed: 19107437]



[1]Includes SNPs associated with survival among all women (at-risk allele in parenthesis): *BMP2* rs7270163(A); *BMP4* rs17563(C), rs762642(T), rs2761887(C), rs4898820(T); *GDF10* rs7093975(C), rs762454(A), rs11598444(G), rs1902725(G), rs1902724(A); *RUNX1* rs2268288(C), rs2252585(T), rs8127225(T), rs1474479(A), rs1883066(G); *TGFB1* rs1800469(T); *TGFBR3* rs6678564(G).

[2]Includes SNPs uniquely associated among women with 0-28% NA ancestry (at-risk allele in parenthesis): *BMP6* rs10498671(C), rs11964227(G); *SMAD1* rs714195(G), rs12505085(A); *TGFBR2* rs3773644(T). Among women with non-missing covariate data, 108 women missing data on one genotype are included and 15 women missing genotype data on two (12 women) or three genotypes (2 women) are excluded.

[3]Includes SNPs uniquely associated among women with >28% NA ancestry (at-risk allele in parenthesis): *BMPR2* rs12621870(T); *SMAD6* rs2439385(C). Among women with non-missing covariate data, 205 women missing data on one genotype are included and 20 women missing genotype data on two (19 women) or three genotypes (1 woman) are excluded.

[4]Includes women alive as date of last follow-up, those lost to follow up, and those that died of causes other than breast cancer.

[5]Breast cancer mortality.

Hazard Ratios (HR) and 95% Confidence Intervals (CI) among primary invasive cases adjusted for age, study, BMI during referent year, parity, genetic ancestry, and SEER summary stage.

Figure 1.

Hazard ratios associated with number of at-risk alleles in genes as determined by the polygenic risk score in the TGF-β-Signaling Pathway.

Cancer Causes Control. Author manuscript; available in PMC 2015 March 01.

Table 1

Description of the study population[1]

	NHW						Hispanic/Native American					
	Deceased	Alive[2]	Log-Rank	Survival Months		% Surviving	Deceased	Alive[2]	Log-Rank	Survival Months		% Surviving
	N (%)	N (%)	P Value	Median	Min, Max	5 years	N (%)	N (%)	P Value	Median	Min, Max	5 years
Vital Status												
Deceased	202	-	NA	73	13,157	62	202	-	NA	71.5	11,154	60
Alive	-	983		110	1,171	100	-	950		118	4,171	100
Cause of Death												
Breast Cancer	102	-	NA	59.5	13,144	50	115	-	NA	59	11,154	50
Other	100	-		89	19,157	75	87	-		86	14,140	75
Study Site[3]												
4-CBCS	129 (14)	804 (86)	0.12	101	1,147	94	63 (14)	395 (86)	0.56	101	4,145	93
SFBCS	73 (29)	179 (71)		148	19,171	91	139 (20)	555 (80)		124	10,171	93
Age at Diagnosis (years)												
<40	15 (19)	63 (81)	<.01	97	1,164	90	14 (14)	86 (86)	<.01	109.5	11,167	91
40–49	41 (13)	281 (87)		104.5	1,168	94	48 (13)	331 (87)		114	4,168	93
50–59	39 (12)	291 (88)		108	2,170	96	56 (18)	264 (83)		115	4,171	94
60–69	49 (17)	241 (83)		104.5	1,169	92	45 (19)	196 (81)		115	11,170	94
≥70	58 (35)	107 (65)		106	1,171	90	39 (35)	73 (65)		109	26,160	88
Menopausal Status												
Pre-menopausal	56 (14)	338 (86)	0.14	103	1,168	93	59 (13)	383 (87)	<.01	113	4,168	92
Post-menopausal	142 (18)	630 (82)		106	1,171	93	133 (21)	515 (79)		113	4,171	93
Estimated Native American Ancestry												
≤18%	198 (17)	972 (83)	0.57	106	1,171	93	19 (16)	101 (84)	0.08	119.5	23,169	95
>18 – 28%	1 (17)	5 (83)		105.5	25,151	100	25 (24)	78 (76)		112	4,170	91
>28 – 70%	3 (43)	4 (57)		108	92,158	100	141 (16)	717 (84)		112	4,171	93
> 70%	0 (0)	2 (100)		99	94,104	100	17 (24)	54 (76)		113	8,170	89
ER/PR Status												
ER+/PR+	98 (15)	548 (85)	0.02	106.5	1,170	95	91 (16)	489 (84)	0.04	115	4,171	95
ER+/PR−	23 (21)	86 (79)		101	9,171	93	22 (20)	88 (80)		115.5	12,167	95

	NHW						Hispanic/Native American					
	Deceased	Alive[2]	Log-Rank	Survival Months		% Surviving	Deceased	Alive[2]	Log-Rank	Survival Months		% Surviving
	N (%)	N (%)	P Value	Median	Min, Max	5 years	N (%)	N (%)	P Value	Median	Min, Max	5 years
ER−/PR+	1 (7)	14 (93)		105	11,161	93	4 (14)	24 (86)		121	37,164	96
ER−/PR−	39 (23)	132 (77)		103	1,168	86	51 (23)	168 (77)		111	11,170	85
SEER Summary Stage												
Local	114 (14)	715 (86)	<.01	108	1,171	96	84 (13)	564 (87)	<.01	117	4,171	96
Regional	75 (23)	247 (77)		100	10,170	88	102 (24)	328 (76)		111	4,170	89
Distant	11 (73)	4 (27)		55	21,122	40	6 (67)	3 (33)		89	31,125	67

[1] Primary invasive breast cancer cases.

[2] Vital status is through May 2012; includes those lost to follow-up

[3] 4-CBCS=4 Corners Breast Cancer Study; SFBCS = San Francisco Bay Area Breast Cancer Study

Table 2

Associations between TGF-β signaling pathway genes and breast-specific mortality cancer in all women

GENE		Deaths/Person Years	HR[2]	(95% CI)	Gene P_{ARTP}[1]	Pathway P_{ARTP}
BMP1					0.21	0.045
rs13257482	GG	139 / 11706	1.00			
	GA/AA	75 / 8593	0.73	(0.55, 0.97)		
BMP2					0.04	
rs7270163	AA	179 / 15236	1.00			
	AG/GG	38 / 5164	0.60	(0.42, 0.86)		
BMP4					0.02	
rs17563	TT	48 / 4833	1.00			
	TC	87 / 8790	1.06	(0.74, 1.52)		
	CC	49 / 4002	1.59	(1.04, 2.44)		
rs2761887	AA	53 / 6048	1.00			
	AC	110 / 10406	1.33	(0.96, 1.85)		
	CC	54 / 3960	1.70	(1.16, 2.49)		
rs4898820	TT	71 / 5599	1.00			
	TG/GG	146 / 14805	0.75	(0.56, 1.00)		
BMP6					0.14	
rs270413	TT/TC	183 / 16040	1.00			
	CC	34 / 4377	0.66	(0.46, 0.96)		
BMPR1B					0.44	
rs10049681	TT	92 / 7878	1.00			
	TC/CC	125 / 12551	0.73	(0.55, 0.96)		
GDF10					0.06	
rs7093975	CC	130 / 11311	1.00			
	CT	79 / 7526	0.95	(0.71, 1.26)		
	TT	8 / 1554	0.46	(0.22, 0.93)		
rs1902724	AA	113 / 8909	1.00			
	AC/CC	104 / 11497	0.74	(0.57, 0.97)		
RUNX1					0.002	
rs1474479	GG	106 / 10705	1.00			
	GA	86 / 7916	1.09	(0.81, 1.46)		
	AA	25 / 1798	1.67	(1.07, 2.62)		
rs1883066	GG	194 / 16047	1.00			
	GC/CC	23 / 4383	0.46	(0.29, 0.71)		
RUNX2					0.20	
rs598953	TT	91 / 7644	1.00			
	TA	107 / 9851	0.85	(0.64, 1.12)		
	AA	19 / 2934	0.51	(0.31, 0.83)		
SMAD2					0.16	
rs1792658	AA	125 / 10762	1.00			

GENE		Deaths/Person Years	HR[2]	(95% CI)	Gene P_{ARTP}[1]	Pathway P_{ARTP}
	AC	79 / 7766	0.84	(0.63, 1.13)		
	CC	13 / 1902	0.54	(0.30, 0.96)		
TGFβR1					0.21	
rs6478974	TT	58 / 6815	1.00			
	TA/AA	159 / 13601	1.35	(1.00, 1.84)		
TGFβR3					0.04	
rs6678564	GG	195 / 17037	1.00			
	GC/CC	22 / 3392	0.56	(0.36, 0.87)		

[1] Gene P_{ARTP}=0.08 for *TGFβ1* because of strong association in highest NA ancestry group.

[2] Hazard Ratios (HR) and 95% Confidence Intervals (CI) among primary invasive cases adjusted for age, study, BMI during referent year, parity, genetic ancestry, and SEER summary stage.

Table 3

Comparison of associations with breast cancer survival by genetic ancestry and self-reported race/ethnicity.

		Percent Native American Ancestry				Self-Reported Race/Ethnicity			
		≤28		>28		NHW		Hispanic/Native American	
Gene	**SNP**	**SNP P**	**Gene P_{ARTP}**	**SNP P**	**Gene P_{ARTP}**	**SNP P**	**Gene P_{ARTP}**	**SNP P**	**Gene P_{ARTP}**
BMP1	rs7812993 (D)	0.06	0.11	0.02	0.24	0.05	0.06	0.05	0.24
	rs3924231 (D)	0.05		0.85		0.02		0.69	
	rs3924229 (D)	0.07		0.27		0.03		0.25	
BMP4	rs17563	0.02	0.007	0.92	0.8	0.005	0.003	0.88	0.83
	rs762642	0.02		0.94		0.01		0.94	
	rs2761887	0.005		0.40		0.001		0.45	
	rs4898820 (D)[1]	0.03		0.58		0.01		0.68	
BMP6	rs10498671 (D)	<0.0001	0.001	0.30	0.12	0.0002	0.005	0.82	0.22
	rs267806 (D)	0.08		0.31		0.04		0.25	
	rs11243204 (D)	0.84		0.008		0.82		0.05	
	rs6910759	0.49		0.01		0.15		0.003	
	rs2068361	0.86		0.007		0.81		0.03	
	rs911749 (D)	0.03		0.40		0.02		0.62	
	rs11964227	0.0006		0.47		0.001		0.79	
BMPR1B	rs7698964 (D)	0.40	0.13	0.05	0.24	0.79	0.15	0.23	0.72
	rs4145993 (D)	0.05		0.33		0.08		0.53	
	rs7694043 (D)	0.007		0.54		0.01		0.79	
	rs3796442 (D)	0.12		0.05		0.14		0.12	
BMPR2	rs1980153 (D)	1.00	0.82	0.30	0.06	0.18	0.54	0.04	0.06
	rs4675278 (D)	0.51		0.04		0.89		0.19	
	rs12621870 (D)	0.88		0.007		0.79		0.03	
	rs1199496 (D)	0.92		0.07		0.59		0.01	
GDF10	rs7093975	0.04	0.05	0.42	0.33	0.10	0.56	0.24	0.05
	rs762454 (R)[1]	0.02		0.78		0.03		0.65	
	rs11598444 (D)[1]	0.50		0.27		0.67		0.06	
	rs1902725 (D)	0.43		0.15		0.61		0.02	
	rs1902724 (D)[1]	0.13		0.08		0.74		0.01	
RUNX1	rs2268288 (D)[1]	0.01	0.002	0.50	0.17	0.03	0.0004	0.84	0.18
	rs1474479	0.004		0.96		0.003		0.90	
	rs1883066 (D)[1]	0.003		0.04		0.0004		0.15	
SMAD1	rs714195 (R)[1]	0.03	0.05	0.83	0.72	0.02	0.03	0.62	0.77
	rs12505085 (D)[1]	0.04		0.57		0.06		0.72	
SMAD3	rs12708492 (D)	0.05	0.4	0.02	0.82	0.08	0.68	0.10	0.85
TGFB1	rs1800469	0.81	0.53	0.002	0.004	0.51	0.82	0.02	0.06

		Percent Native American Ancestry				Self-Reported Race/Ethnicity			
		≤28		>28		NHW		Hispanic/Native American	
Gene	**SNP**	**SNP P**	**Gene P_{ARTP}**	**SNP P**	**Gene P_{ARTP}**	**SNP P**	**Gene P_{ARTP}**	**SNP P**	**Gene P_{ARTP}**
TGFBR2	rs3773644 (D)[1]	0.02	0.02	0.71	0.81	0.23	0.58	0.39	0.38
TGFBR3	rs6678564 (D)[1]	0.04	0.16	0.09	0.11	0.17	0.48	0.02	0.05

[1] Interaction p value > 0.05 for genetic ancestry and race/ethnicity but SNP contributes to significant strata-specific gene P_{ARTP}. Model selection is D=dominant; R = recessive; all others are co-dominant.

[2] SNP p values based on Cox proportional hazard models adjusted for age, study, BMI during referent year, parity, genetic ancestry, and SEER summary stage among primary invasive breast cancer cases.

Table 4

Associations between genes in TGF-β signaling pathway and breast cancer survival by genetic ancestry and self-reported race/ethnicity

	Percent Native American Ancestry					Self-Reported Race/Ethnicity				
	≤28		>28			NHW		Any Hispanic or Native American		
	Deaths/Person Years	HR (95% CI)	Deaths/Person Years	HR (95% CI)	P_{INT}	Deaths/Person Years	HR (95% CI)	Deaths/Person Years	HR (95% CI)	P_{INT}
BMP1 (rs7812993)					0.004					0.005
AA	86 / 7296	1.00	50 / 5492	1.00		71 / 6118	1.00	65 / 6670	1.00	
AG/GG	39 / 4777	0.69 (0.47, 1.01)	42 / 2851	1.61 (1.07, 2.44)		31 / 3961	0.66 (0.43, 1.01)	50 / 3667	1.45 (1.00, 2.10)	
BMP1 (rs3924231)					0.132					0.029
TT	96 / 8310	1.00	63 / 5657	1.00		80 / 6839	1.00	79 / 7127	1.00	
TC/CC	29 / 3763	0.66 (0.43, 1.01)	29 / 2686	1.04 (0.67, 1.63)		22 / 3240	0.56 (0.34, 0.90)	36 / 3209	1.08 (0.73, 1.61)	
BMP1 (rs3924229)					0.041					0.013
TT	106 / 9500	1.00	78 / 7256	1.00		88 / 7879	1.00	96 / 8878	1.00	
TC/CC	19 / 2587	0.63 (0.39, 1.03)	14 / 1073	1.39 (0.77, 2.49)		14 / 2214	0.54 (0.30, 0.95)	19 / 1445	1.34 (0.81, 2.21)	
BMP4 (rs762642)					0.079					0.036
TT	55 / 4009	1.00	32 / 3063	1.00		48 / 3353	1.00	39 / 3719	1.00	
TG	56 / 6100	0.66 (0.46, 0.97)	47 / 4080	1.02 (0.65, 1.61)		43 / 5055	0.61 (0.40, 0.93)	60 / 5125	1.03 (0.69, 1.55)	
GG	14 / 1979	0.51 (0.28, 0.91)	13 / 1199	1.03 (0.53, 1.97)		11 / 1685	0.42 (0.22, 0.81)	16 / 1493	1.02 (0.57, 1.85)	
BMP6 (rs10498671)					0.001					0.008
TT	71 / 8562	1.00	70 / 5982	1.00		57 / 7103	1.00	84 / 7441	1.00	
TC/CC	54 / 3525	2.17 (1.50, 3.13)	21 / 2356	0.77 (0.47, 1.26)		45 / 2990	2.19 (1.46, 3.30)	30 / 2891	0.95 (0.63, 1.45)	
BMP6 (rs267806)					0.055					0.020
CC	52 / 6255	1.00	28 / 1991	1.00		41 / 5320	1.00	39 / 2925	1.00	
CT/TT	73 / 5832	1.38 (0.96, 1.98)	64 / 6352	0.79 (0.50, 1.25)		61 / 4773	1.51 (1.01, 2.26)	76 / 7412	0.79 (0.53, 1.18)	
BMP6 (rs11243204)					0.038					0.265
AA	80 / 7703	1.00	45 / 5104	1.00		64 / 6477	1.00	61 / 6330	1.00	
AG/GG	45 / 4378	0.96 (0.66, 1.40)	47 / 3238	1.77 (1.16, 2.69)		38 / 3610	1.05 (0.70, 1.58)	54 / 4006	1.44 (1.00, 2.09)	
BMP6 (rs6910759)					0.024					0.003
AA	38 / 3360	1.00	42 / 4462	1.00		31 / 2659	1.00	49 / 5163	1.00	
AG	65 / 6204	0.95 (0.63, 1.42)	36 / 3178	1.31 (0.83, 2.06)		55 / 5226	0.88 (0.56, 1.37)	46 / 4156	1.29 (0.85, 1.94)	
GG	22 / 2523	0.83 (0.49, 1.42)	14 / 691	2.25 (1.20, 4.21)		16 / 2208	0.64 (0.35, 1.18)	20 / 1006	2.26 (1.32, 3.88)	

	Percent Native American Ancestry					Self-Reported Race/Ethnicity				
	≤28		>28			NHW		Any Hispanic or Native American		
	Deaths/Person Years	HR (95% CI)	Deaths/Person Years	HR (95% CI)	P_{INT}	Deaths/Person Years	HR (95% CI)	Deaths/Person Years	HR (95% CI)	P_{INT}
BMP6 (rs2068361)					0.045					0.208
GG	73 / 6645	1.00	48 / 3393	1.00		63 / 5640	1.00	58 / 4397	1.00	
GA	43 / 4672	0.85 (0.58, 1.26)	37 / 3805	0.63 (0.40, 0.97)		33 / 3810	0.88 (0.57, 1.35)	47 / 4668	0.71 (0.48, 1.05)	
AA	9 / 760	1.07 (0.53, 2.16)	7 / 1144	0.32 (0.14, 0.73)		6 / 633	0.90 (0.37, 2.18)	10 / 1272	0.47 (0.24, 0.94)	
BMP6 (rs911749)					0.035					0.039
GG	67 / 7486	1.00	49 / 4241	1.00		55 / 6237	1.00	61 / 5491	1.00	
GA/AA	58 / 4601	1.47 (1.03, 2.10)	43 / 4101	0.83 (0.55, 1.28)		47 / 3856	1.64 (1.09, 2.45)	54 / 4846	0.91 (0.62, 1.33)	
BMP6 (rs11964227)					0.003					0.010
GG	50 / 3706	1.00	38 / 3920	1.00		40 / 3129	1.00	48 / 4497	1.00	
GA	67 / 6088	0.80 (0.55, 1.15)	42 / 3558	1.20 (0.77, 1.87)		57 / 5053	0.82 (0.54, 1.24)	52 / 4593	1.07 (0.72, 1.59)	
AA	8 / 2294	0.27 (0.13, 0.57)	12 / 865	1.27 (0.66, 2.46)		5 / 1911	0.21 (0.08, 0.54)	15 / 1247	1.08 (0.60, 1.95)	
BMPR1B (rs7698964)					0.047					0.318
GG	104 / 9356	1.00	71 / 7098	1.00		84 / 7838	1.00	91 / 8616	1.00	
GA/AA	21 / 2731	0.81 (0.51, 1.31)	21 / 1245	1.65 (1.00, 2.72)		18 / 2255	0.93 (0.56, 1.56)	24 / 1721	1.32 (0.84, 2.10)	
BMPR1B (rs4145993)					0.040					0.084
CC	91 / 7949	1.00	60 / 5796	1.00		76 / 6799	1.00	75 / 6947	1.00	
CT/TT	34 / 4138	0.67 (0.45, 1.00)	32 / 2546	1.24 (0.81, 1.91)		26 / 3294	0.67 (0.43, 1.05)	40 / 3390	1.13 (0.77, 1.67)	
BMPR1B (rs7694043)					0.026					0.042
CC	40 / 5094	1.00	55 / 4776	1.00		29 / 4083	1.00	66 / 5787	1.00	
CT/TT	85 / 6968	1.69 (1.16, 2.47)	37 / 3566	0.88 (0.57, 1.34)		73 / 5998	1.76 (1.14, 2.73)	49 / 4536	0.95 (0.66, 1.38)	
BMPR1B (rs3796442)					0.013					0.037
CC	110 / 10067	1.00	59 / 6147	1.00		89 / 8270	1.00	80 / 7944	1.00	
CA/AA	15 / 2020	0.65 (0.38, 1.12)	33 / 2196	1.54 (1.00, 2.38)		13 / 1823	0.65 (0.36, 1.16)	35 / 2393	1.38 (0.92, 2.07)	
BMPR2 (rs1980153)					0.464					0.024
AA	99 / 9705	1.00	65 / 6191	1.00		86 / 8128	1.00	78 / 7768	1.00	
AT/TT	26 / 2376	1.00 (0.65, 1.55)	27 / 2151	1.27 (0.81, 2.01)		16 / 1958	0.69 (0.40, 1.18)	37 / 2569	1.53 (1.03, 2.27)	
BMPR2 (rs4675278)					0.040					0.266
GG	59 / 6190	1.00	48 / 3682	1.00		50 / 5217	1.00	57 / 4656	1.00	
GA/AA	66 / 5886	1.13 (0.79, 1.61)	44 / 4661	0.65 (0.43, 0.99)		52 / 4876	1.03 (0.69, 1.53)	58 / 5671	0.78 (0.54, 1.13)	

	Percent Native American Ancestry					Self-Reported Race/Ethnicity				
	≤28		>28			NHW		Any Hispanic or Native American		
	Deaths/Person Years	HR (95% CI)	Deaths/Person Years	HR (95% CI)	P_{INT}	Deaths/Person Years	HR (95% CI)	Deaths/Person Years	HR (95% CI)	P_{INT}
BMPR2 (rs12621870)					0.021					0.194
TT	76 / 7217	1.00	68 / 5212	1.00		63 / 6049	1.00	81 / 6380	1.00	
TC/CC	49 / 4808	1.03 (0.72, 1.48)	23 / 3119	0.52 (0.32, 0.83)		38 / 3987	0.95 (0.63, 1.42)	34 / 3940	0.65 (0.43, 0.97)	
BMPR2 (rs1199496)					0.176					0.049
AA	62 / 6181	1.00	32 / 3746	1.00		55 / 5265	1.00	39 / 4662	1.00	
AT/TT	63 / 5906	1.02 (0.71, 1.45)	60 / 4580	1.50 (0.97, 2.31)		47 / 4828	0.90 (0.60, 1.34)	76 / 5658	1.63 (1.10, 2.40)	
GDF10 (rs1902725)					0.539					0.038
GG	82 / 7477	1.00	69 / 5574	1.00		62 / 6219	1.00	89 / 6831	1.00	
GA/AA	43 / 4610	0.86 (0.59, 1.25)	23 / 2769	0.71 (0.44, 1.14)		40 / 3873	1.11 (0.74, 1.66)	26 / 3506	0.59 (0.38, 0.92)	
SMAD3 (rs12708492)					0.003					0.017
TT	21 / 2716	1.00	34 / 2311	1.00		16 / 2208	1.00	39 / 2820	1.00	
TC/CC	104 / 9355	1.61 (1.00, 2.57)	58 / 6032	0.60 (0.39, 0.92)		86 / 7869	1.63 (0.95, 2.81)	76 / 7517	0.72 (0.49, 1.06)	
TGFB1 (rs1800469)					0.026					0.303
CC	58 / 5489	1.00	17 / 2345	1.00		45 / 4684	1.00	30 / 3150	1.00	
CT	48 / 4862	0.92 (0.62, 1.36)	41 / 4079	1.52 (0.86, 2.69)		42 / 4064	1.03 (0.67, 1.58)	47 / 4877	1.08 (0.68, 1.71)	
TT	17 / 1388	1.07 (0.62, 1.85)	32 / 1783	2.54 (1.40, 4.61)		13 / 1062	1.24 (0.66, 2.33)	36 / 2110	1.81 (1.11, 2.97)	

[1] Hazard Ratios (HR) and 95% Confidence Intervals (CI) among primary invasive cases adjusted for age, study, BMI during referent year, parity, genetic ancestry, and SEER summary stage.